Challenges and Opportunities in Grid Connected Commercial Scale PV and Wind Farms

Professor Syed Islam, PhD, FIET, FIEAust, SMIEEE

Centre for Smart Grid and Sustainable Power Systems
Curtin University
Perth, WA 6845
AUSTRALIA
s.islam@curtin.edu.au

***Abstract*-Wind power generation and Solar PV installations continues to grow and are the two key contributors to the large scale variable power generation mix. Utility transmission system operators impose stringent grid codes internationally. Grid code regulations are defined by system operators to outline the rights and responsibilities of all the generators and loads that are connected to the transmission/distribution system. In the past, grid codes did not include any regulations for Wind Power Plants (WPPs) and PV plants as the penetration levels were small. However, the situation has now radically changed as many countries have installed many GW worth of variable power generations. The increased reliance on renewable power generation has raised some serious concerns regarding the negative impacts of large WPPs and PV on the stability of power networks. Modern grid codes require WPPS and PVs not only withstand various grid disturbances but also to contribute to network stability support and ancillary services provision, as do conventional generating units. Solar shading can cause PV output to drop by as much as 70% in a space of few seconds. This requires strategies for solar smoothing. In this plenary paper, status review, research challenges and opportunities in developing enabling technologies in grid code compliances, solar smoothing, storage and congestion management to maximise renewable energy penetration into the grid are presented.**

Index Terms- Wind Power, Variable Power Generation, PV, Congestion Management, Grid Codes.

I. Introduction

World Wind power generation reached 432 GW in 2015 and cumulative PV output reached 178 GW in 2014. Even though in 2015, the wind and PV power generation only contributes to 3.7% and 1.2% respectively to World total power generation, the growth forecast for both wind and PV is very high i.e. we could see 500 GW worth of PV by the end of 2050 and wind power generation could double to 800 GW by the end of 2020 according to Global Wind Energy Council. In some countries the penetration levels are already quite significant. For example, in Germany, the total PV installed was 40 GW approximately 7.5% of electricity consumption in 2015 and in Denmark wind power now contributes to approximately 20 percent of their electricity consumption, in South Australia wind power installed capacity is around 40% of their peak demand. The above data shows that, unless proper strategies are in place, there could be significant impact on system voltage and frequency due to the intermittent nature of solar and wind. To address this issue, Electric utilities and standard bodies around the World have introduced various grid codes for variable power generation to connect to their grid. Grid code regulations are defined by system operators to outline the rights and responsibilities of all the generators and loads that are connected to the transmission/distribution system. In the past, grid codes did not include any regulations for PVs or WPPs because the penetration level of solar and wind power was extremely small compared to the conventional generation systems. However, the situation has radically changed during the last few years as many countries witnessed tremendous increase in the number and capacity of PV and WPPs integrated into their electric networks. China installed 22 GW of grid connected PV in the first six months of 2016 and have already 50 GW worth of PV installed. Some plants are large and any outage could be catastrophic. In China, the Minsheng PV plant is 2 GW composed of six million solar panels spanning over 4607 hectares of land. The shift from conventional to the renewable energy resources has raised serious concerns regarding the negative impacts of large PV and WPPs, as intermittent generation units, on the stability of existing power networks. On the 28th of September, 2016, wind gust up to 140 km/hr brought down 23 transmission towers resulting in major generator to be disconnected, wind turbines stalled and the interconnector between SA and Victoria isolated causing the entire state to go into darkness. It took almost 3 days to restore power. Voltage and frequency drops were found to be the major cause of this outage. To safeguard the network against these threats, system operators in many countries have enforced stringent technical requirements on large WPPs and now PVs. Modern grid codes require WPPs to not only withstand various grid disturbances, but also contribute to the network stability support and ancillary service provision, as do conventional generation units [1]. For PVs solar smoothing is increasingly becoming a major discussion point. The issue of grid code requirements for large wind farms has been already explored in the literature. The US federal and regional grid codes were presented in [2]. Technical regulations of Canada, Germany, Ireland, UK and Spain were studied in[3],[4],[5],[6] and [7], respectively. Comparative study of the international grid codes was first reported in [8], evaluating Danish, British, German, and Irish Grid Codes. Similar comparative studies were reported in [9] and [10]. In Australia and New Zealand, AS/NZ 4777.2:2015 outlines various requirements for PV to connect to distribution grid.

This paper is organized as follows. In Section 2, some of the most common grid code requirements included in majority of international grid codes are explored. Section 3 reviews the wind energy conversion technologies for grid code compliances. Section 4 discusses some advances in technology in meeting voltage sags. Section 5 presents some of the key points in AS/NZ grid codes for grid connected PV. Section 6 presents a discussion on solar smoothing and some

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

of the emerging standards. Section 7 discusses variable power overgeneration and congestion management issues. Finally, in section 8 some emerging grid compliant solutions for different wind power generator technologies are given.

II. REVIEW OF DANISH AND AUSTRALIAN GRID CODES

A review of the international grid codes shows that the technical requirements enforced on large WPPs can be broadly classified into five groups: (1) fault ride-through requirements, (2) active and reactive power responses following disturbances, (3) extended variation range for the voltage-frequency, (4) active power control or frequency regulation support, and (5) reactive power control or voltage regulation capability. When a fault occurs at some points in the electric network, the supply voltage drops to the lower levels until protection devices detect the faulty area and isolate it from the rest of the network. During this interval, wind generators, like other system components, experience a voltage sag condition at their terminals depending on the type and location of the fault. As a result of this disturbance, wind farms (specifically those employing modern variable-speed technologies) may disconnect from the grid due to potential stability problems. A practical example of this incident was experienced in the European outage on 4 November 2006, which caused losing 4,892 MW of wind power generation in Western-Europe.

The disconnection of WPPs under voltage disturbances is unacceptable when wind power constitutes a significant part of the total network generation. Thus, modern grid codes require WPPs to continue their uninterrupted operation under various fault conditions according to given voltage-time profiles. Such requirements are usually referred to as low voltage ride through (LVRT) capability. Fig. 1 shows a practical example of the LVRT curve defined by the Danish system operator for wind generators connected to the transmission system (V > 110 kV): "a wind power plant shall remain connected to the transmission system under fault conditions when the voltage measured at the HV terminals of the grid connected transformer (point of common coupling – PCC) remains in *Area B* shown in Fig. 1" [11]. That is, WPPs are only allowed to disconnect from the grid when the voltage profile falls into *Area C*.

Some international grid codes also require large WPPs to continuously operate under small deviations of the voltage magnitude or frequency. This range of variations is usually referred to as 'the normal operation area', as shown in Fig. 2 for Danish Grid Code. From Fig. 2, it is clear that WPPs must retain at least 60% of their normal production for 15 minutes when the frequency varies between 50.2 Hz to 52 Hz and the PCC voltage changes in the range of ±10%.

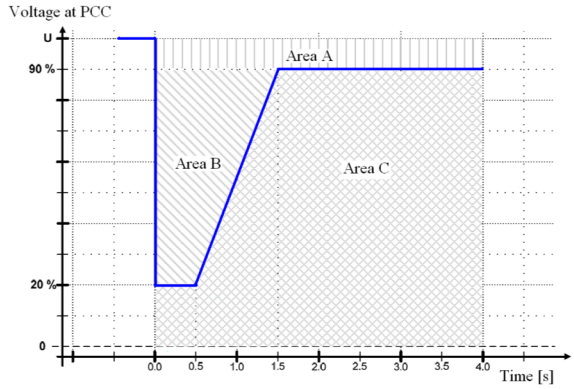


Fig. 1 Danish Grid Code for LVRT.

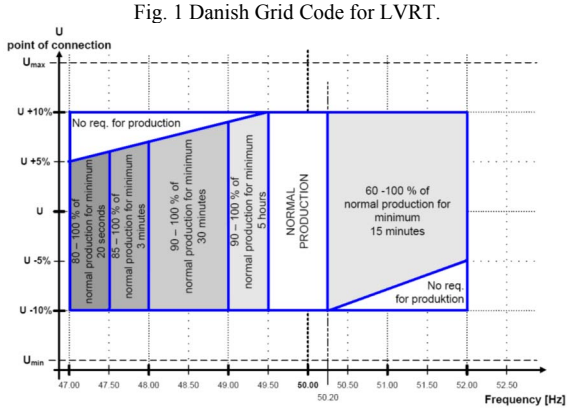


Fig. 2 Normal operation area of Danish Grid Code.

In the very recent years, grid codes of some countries like Australia, Denmark and Spain have enforced demanding voltage time profiles for voltage swell conditions. Voltage swells could be initiated by switching off large loads, energizing capacitor banks, or occurring faults in the grid. This requirement is commonly referred to as high voltage ride-through (HVRT) capability. Modern grid codes may also require large WPPs to withstand frequency excursions that can be observed under grid disturbances with large mismatch between the generation and consumption in the system.

Besides the LVRT and HVRT requirements, some international grid codes require large WPPs to provide reactive power support during the fault period and/or exhibit fast active power recovery during the supply voltage restoration. These regulations are enforced to fully exploit the control capability of modern wind farms and contribute to the system stability support following various grid disturbances. For instance, Danish Grid Code requires large WPPs to support the transient voltage stability of the network through providing the reactive current according to the characteristic shown in Fig. 3: "as the PCC voltage drops by more than 10% (i.e., operating within *Area B* in Fig. 1), the reactive current output of the WPP must follow the characteristic shown in Fig. 4 • with a tolerance of ±20% in 100 ms". It can be seen that the I_Q component of the WPP must increase by 2.5% (on a per unit basis) for each 1% drop in the PCC voltage. The injection of reactive power has the highest priority in *Area B*, but the free capacity of current must also be utilized to retain the active power production in proportion to the voltage sag magnitude. Furthermore, the Danish TSO stipulates active power restoration to the pre-fault value immediately after the fault clearance.

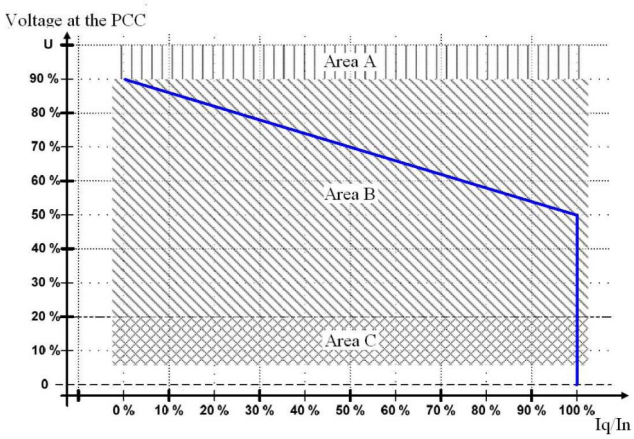


Fig. 3 Reactive Power as defined in Danish Grid Code.

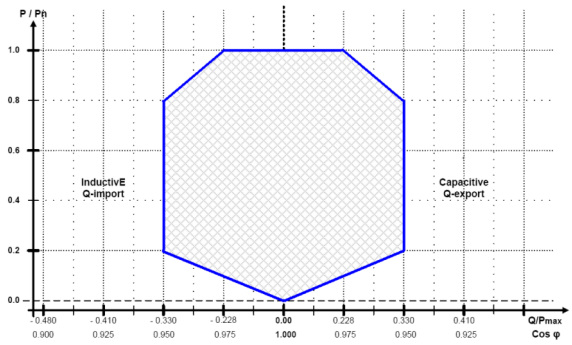


Fig. 4 Reactive Power and Power Factor Requirement of Danish Grid Code.

III. WIND POWER COMPLIANT TECHNOLOGIES

The ability of WPPs to meet the international grid codes is greatly influenced by the technology used in the wind farm. Three topologies have been widely used in wind generation systems. The simplest technology (type A) is constituted of a squirrel cage induction machine directly connected to the grid. This technology is referred to as a fixed-speed wind turbine because the machine slip can change in the very limited range of ±1%. Soft-starter is needed to avoid large inrush current at the machine starting instant. Also, reactive power compensation devices, e.g. capacitor banks or STATCOM, must be installed to locally provide the machine magnetization. The aerodynamics of fixed-speed wind turbine can be controlled by stall, active-stall or pitch control approaches [1].

In type B, a wound rotor induction machine is used and an adjustable resistance is connected to the rotor winding. The torque characteristic of the machine can be altered through the variable rotor resistance. As a result, the machine speed can vary in a limited range of ±10% around the synchronous speed; however, the slip power will LVRT be dissipated in the resistance as heat Therefore, the overall system performance is low and sizeable heat sinks are needed to avoid thermal instability problems inside the machine structure. This technology was popular in 90's but with the recent advances in power electronic converters, it has becomes obsolete. Type B wind generator is not studied in this paper. In type C and type D, back-to-back converters have been used so that the generator speed can vary widely according to the instantaneous wind speed at the site. These technologies are referred to as variable-speed wind generators. Type C is called doubly fed induction generator (DFIG). In DFIG-based wind generators, a wound rotor induction machine is mechanically coupled to the turbine through shaft and gear-box system. The three-phase stator winding of the machine is directly connected to the grid, whereas the rotor winding is connected via back-to-back voltage source converters (VSCs) rated at 30%-35% of the generator size. This gives a rotor speed variation range of about ± 25%. The rotor-side converter (RSC) controls the active and reactive power outputs from the machine, while the grid-side converter (GSC) regulates the dc-link voltage at the reference value. The DFIG-based wind generators have recently become the most dominant technology in the market as they offer many technical and economical advantages, including maximized power capture, reduced mechanical stresses imposed on the turbine,improved power quality, and reduced acoustical noise. In type D, the stator winding of a synchronous or induction generator is connected to the grid through fully rated back-to-backVSCs. The machine-side converter (MSC) controls the active and reactive power outputs of the machine. The GSC, on the other hand, maintains the dc-link voltage at its set point and controls the reactive power exchange with the grid, i.e., the GSC transfers the active power extracted from the wind turbine to the grid at an adjustable power factor. Cost-benefit analysis of wind turbines shows that the gear-box costs about 15% of the total wind turbine price and it reportedly has the highest failure rate between the electrical and mechanical components of a wind generation system. Therefore, multi-pole permanent magnet synchronous machines have recently become popular as they avoid gear-box. Moreover, since the machine excitation is not provided via the slip-ring system, the permanent magnet machine needs less monitoring and maintenance operations. This can be a major advantage for offshore wind farms, where access to the wind generators is extremely laborious.

IV. VOLTAGE SAGS

Most common type of Wind Power Generators arethe Doubly Fed Induction Generators (DFIGs). Voltage sags that occur at point of common couplings are symmetrical and asymmetrical types. Further phase angle jumps must also be considered. Due to varied network characteristics and constraints, the voltage sag conditions experienced at the wind generator terminals can be significantly different from the conditions at the PCC. Therefore, it is very important to identify the voltage sags that can practically affect the operation of wind generators.

Extensive simulation studies are carried out inMATLAB/Simulink to investigate the transient overshoots and ripples that appear in the rotor current and dc-link voltage when the DFIG is subjected to various types of (a)symmetrical faults [12]. For the first time, the impacts of phase-angle jump and operational constraints of circuit breakers are examined. Furthermore, the influences of sagparameters including type, initial point-on-wave instant, depth,and impedance angle are also investigated.

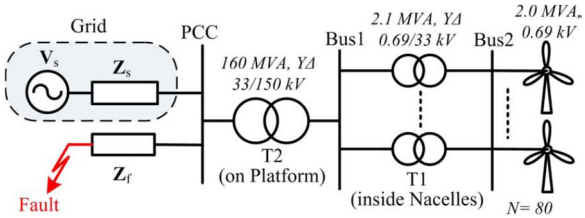


Fig. 6 Simulated Wind Power Systems

For symmetrical sags (type A), the natural response of the stator flux generates large transient overshoots in the rotor current and dc-link voltage at the beginning and the end of the three-phase faults. However, if the operational constraints of the circuit breakers are taken into account, the full voltage recovery will take place in two or three consecutive steps and smaller overshoots will be experienced
at the fault clearance instant. For wind generators connected to the distribution system or located offshore, overshoots during the recovery process are expected to increase as the fault current angle decreases. For asymmetrical sags (fault types B to G), steady-state oscillations as well as transient overshoots are observed in the rotor current and the dc-link voltage. The negative-sequence component of the grid voltage increases when the voltage sag is associated with large phase-angle jumps. This has significant detrimental impacts on wind turbines located offshore, since the transient overshoots in the rotor current and dc-link voltage can increase by 45%–60% and 20%–40%, respectively Table 1 shows that under identical sag conditions, the largest overshoots in the dc-link voltage are experienced for three-phase faults (type A), while the highest peaks for the rotor current are observed under phase-to-phase faults (types C and D). Also, a linear relationship is obtained between the sag depth and peaks of the rotor current or dc-link voltage, i.e., higher peaks will be expected as the depth increases.

TABLE I
SUMMARY OF SIMULATED RESULTS [12]

Fault (Sag Type)	α= 0°			α= –20°			α= –60°		
	θ_{peak} (°)	$I_{r,p}$	$V_{dc,p}$	θ_{peak} (°)	$I_{r,peak}$	$V_{dc,peak}$	θ_{peak} (°)	$I_{r,peak}$	$V_{dc,peak}$
Φ–g (C*&D*)	90&180	1.75	1.13	83&173	1.93	1.16	77&167	2.53	1.56*
Φ–Φ (C&D)	90&180	2.2	1.21	84&174	2.65	1.28	80&170	3.5**	1.46
2Φ–g (G&F)	90&180	2.15	1.2	85&175	2.45	1.26	82&172	3.25	1.41
3Φ (A)	–	1.9	1.23	–	2.1	1.3	–	2.95	1.61

*Most severe case in terms of the overshoot in the dc-link voltage.
**Most severe case in terms of the overshoot in the rotor current.

V. GRID CONNECTED PV CODES

In the case of PV plants, the grid codes are primarily targeting the inverters and their functionalities. The grid codes are either for small scale PVs connected to distribution systems at voltage less than 415 V, medium scale PVs connected to 6.6 kV to 35 kV sub-transmission and for large scale PVs connected to greater than 66kV transmission systems. There are several grid codes specifying regulations such as in China (GB/T 19964, GB/T 29319), USA (IEEE 1547 and IEEE 519), Germany (VDE-AR-N4105), and Australia/NZ (AS/NZS 4777-2:2015).
These grid codes mainly focussing on power quality, harmonics, flickers, voltage deviations, frequency, active/reactive power, anti-islanding, DC injection etc. Some of the details of these codes as given in AS/NZS4777-2:2015 are given below.

Power Quality

The AS/NZS 4777-2:2015 specify the power quality response modes:

- Contribute to power quality at point of customer installation
- Provide functions which provides support to grid : volt response modes, fixed power factor or reactive power response mode and power rate limit.

The volt response mode ensures that voltage can be maintained with increased number of systems connected at the point. The volt-watt mode varies the output power of inverter in response to voltage changes as shown in Figure 7.

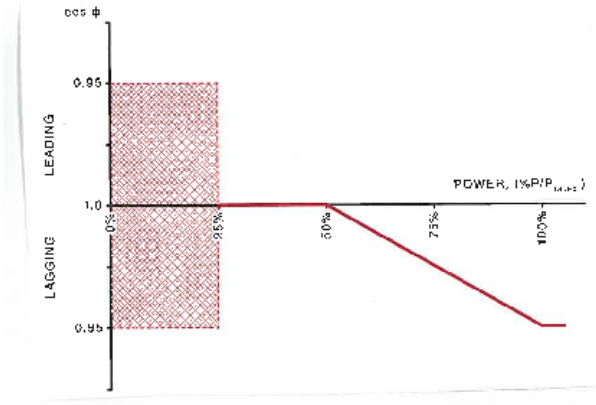


Fig. 7 Volt-Watt Response Mode (AS/NZS 4777-2) .

Power Factor

The AS/NZS4777-2:2015 states that for all current outputs from 25% to 100% of rated current, the inverter operate at unity power factor within the range 0.95 lagging to 0.95 leading. However, if the inverter operates in the power quality response mode as described above, then operation other than at unity power factor is acceptable.

Harmonics and Transients

The Australian grid code limits the THD to less than 5% when considering the current harmonic up to 50th. The individual even harmonics are restricted to 1% and the individual odd harmonics are restricted to 4%. For both single phase and three phase inverters, the dc injection current may not exceed 0.5% of the inverter rated current or 5 mA whichever is greater.

Volt-Var and Volt-Watt Response Modes

The volt-var and volt-watt response modes vary the reactive/active power output of the inverter in response to the voltage at the grid-interactive port. Figure 8 shows the various voltage ranges for the volt-var requirements. The inverter may provide a range up to 100% leading or lagging vars. The default values are 30% leading, 0%, 0% and 30% lagging in Figure 8 and 100%, 100%, 100% and 20% for voltage settings 1-4 for volt-watt requirements.

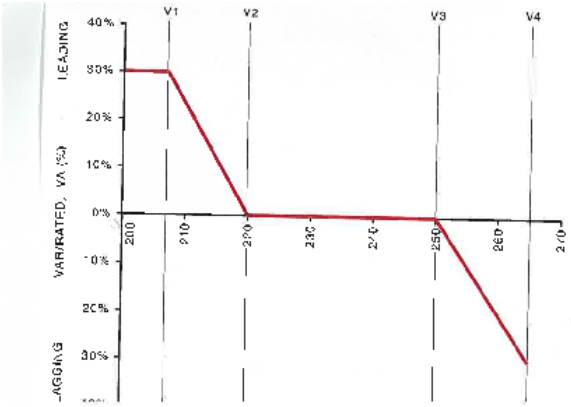


Fig. 8 Volt-Var Control Mode (AS/NZS 4777-2).

If the inverters are required to operate at reactive power mode, the ratio of reactive power to apparent power should be 100%. If the inverter is required to operate at fixed power factor mode, then minimum range of settings should be 0.8 leading to 0.8 lagging.

VI. SOLAR PV SMOOTHING – AN EXAMPLE

The continued increase in PV and Wind power generation displaces significant energy consumption that would have normally come from conventional generators. While the renewables are welcome, unfortunately they are intermittent in nature. Figure 9 below shows the so called duck curve for the California ISO from 2012 to 2020.

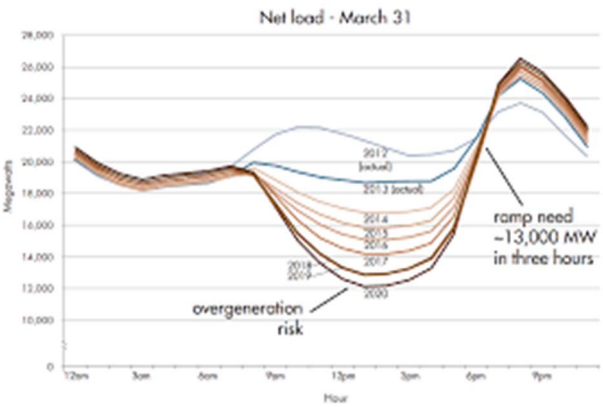


Fig. 9 The Duck Curve for the CAISO System [Source: NREL Report].

Figure 9 shows that the load demand during the mid- day could drop from 21 GW to just 12 GW by the year 2020 while the ramp need could be as high as 13,000 MW in three hours. This situation put enormous challenges to the system operator in balancing load and generation and to avoid system failure/blackouts.

To complicate the matter further, the intermittent nature of solar generation is such that when a cloud event or system disturbance occurs, the energy output of many solar PV systems can drop sharply, and this manifests itself as a rapid transfer of load away from the PV generation and onto the generation equipment.

This rapid transfer of load has the ability to affect the operation of the remaining power plants. The situation is even more serious for small islanded power systems. The ability of the power system to meet this sudden increase in load is dependent on the characteristics of the generator sets and how many are online at the time of the event.

Similarly when the cloud event finishes, the sometimes rapid reappearance of the sun manifests itself as a spike in PV system production and a sudden excess of energy flowing into the network. Both these scenarios could have detrimental impact on voltage and frequency. Figure 10 below shows how rapidly the PV output power could fluctuate on a system. The PV output drops from 250 kW to just 40 kW in 20 seconds. The lack of rotating inertia in the system impact the frequency and the lack of generation leads to drop in voltages as shown in Figures 11 and 12. A flywheel energy storage system and 2 back-up generators eventually restore the frequency and voltages [13].

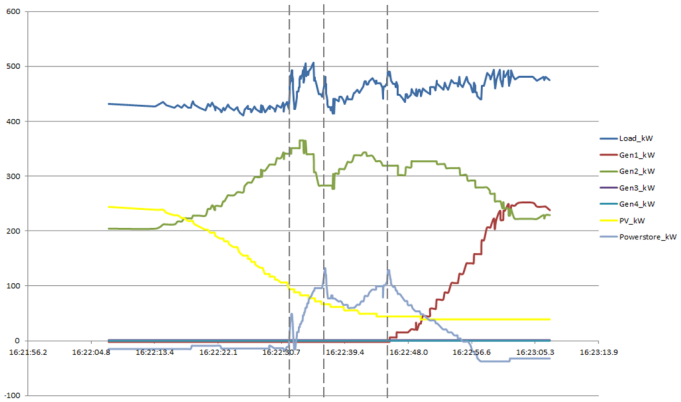


Fig. 10 Real Time Recording for a Cloud Event in a PV-Diesel-Flywheel system.

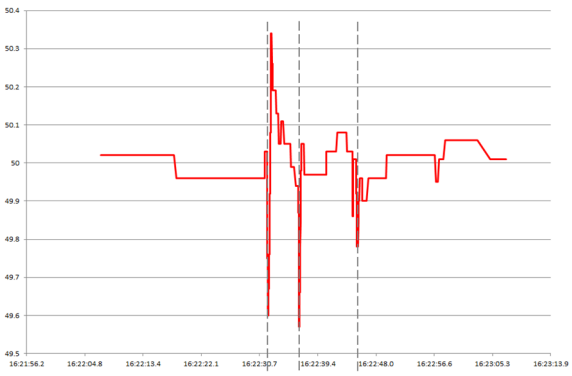


Fig. 11 Frequency Response during the Cloud Event.

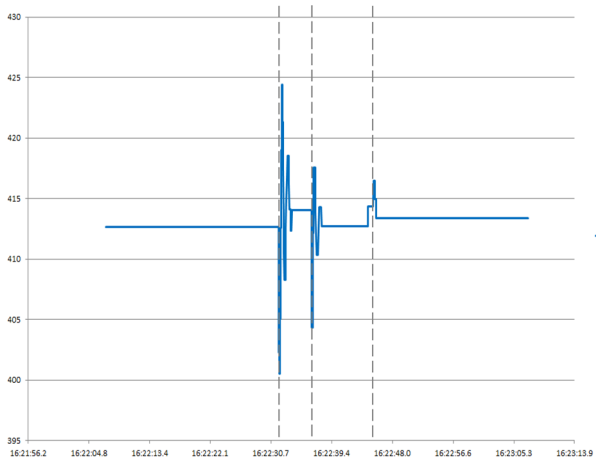


Fig. 12 Voltage Response during the Cloud Event.

VII. OVEGENERATION AND CONGESTION MANAGEMENT

Overgenerationis a term to describe the situation when output from renewable energy reaches so high that the system could not accommodate it. The CAISO duck curve shows that during the day there is potential for overgeneration due to increased PV penetration. There is a limit to how much can be taken off from spinning conventional generators due to technical limits. Under such circumstances, the PV inverters has to reduce output or disconnect and wind generators reduce output by blade pitching or face disconnection. The system operators carefully balance the load and generation including variable power generation together with control of PV and Wind and any storage system. Another consequence of overgeneration could be when the transmission lines are unable to carry the excess power generated. New transmission lines are preferable but cost prohibitive. New lines have the ability to smooth out renewable power generation uncertainties particularly if geographical issues are considered. Interconnection with neighboring countries (Europe) or adjacent states (Australia) could improve congestion as well as reliability. However when new lines are not a solution, in such cases renewable power curtailment could be avoided by using technology such as congestion management [14].

Risk of network congestion is the product of likelihood and the severity of network congestion. The severity of congestion may be indicated by the volatility in Localised Marginal Node Price (LMP) and the amount of wind power curtailment.

Volatility in LMP is most commonly used as an indicator of network congestion as congestion cost is a significant component of LMP in transmission systems.Pricing signals have been proposed as a control mechanism for renewable energy integration.

In [14], the approach taken first establishes a base case for LMP without incorporating network constraints. For each outage scenario, the LMP at each bus is compared to the base case LMP, weighted by the load at that bus and the overall weighted variation in LMP is found. To compare the LMP profile of a specific case to the base case, the term LMP_V is defined as:

$$LMP_V = \sqrt{\frac{1}{\sum_{i=1} P_{D,i}} \left(\sum_{i=1} \left[P_{D,i} \left(\frac{LMP_i - LMP_{i,base}}{LMP_{i,base}} \right)^2 \right] \right)}$$

Stochastic Dynamic Line Rating (DLR) could be used to minimize wind power curtailment [14]. In Figure 13, risk profiles in terms of (LMP-v) severity and likelihood for a number of N-1 and N-2 outages/contingencies are shown. From the Figure, it is clear that the risk of DLR lowers risk of congestion by lowering the severity of congestion for both N-1 and N-2 outages.

Transmission Lines in Europe are rated at 36 degrees C ambient and 0.6 m/s wind speed. Opportunities for DLR for many large wind farms producing close to capacity during high wind could reduce valuable wind power curtailment due to transmission congestion.

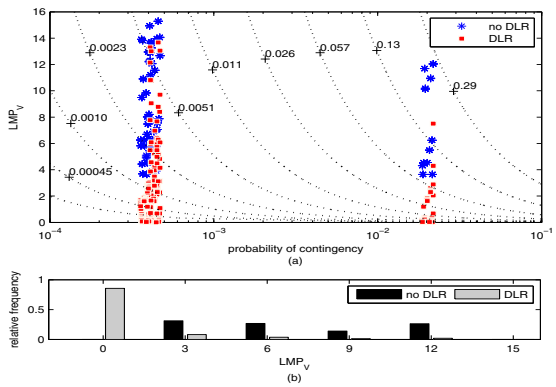


Fig. 13 Risk Profiles against Contingencies and DLR Improvement.

VIII. ACTIVE AND REACTIVE POWER SUPPORT – AN EXAMPLE

A Case Study to support voltage sag in controlling active and reactive power using a SMES shunt active/reactive power controller is presented below. The test system is shown in Figure 14 [15].

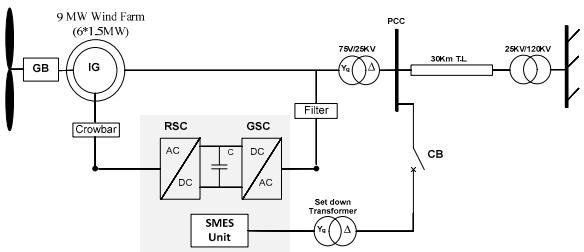


Fig.14 SMES Power Conditioner at PCC.

A voltage sag depth of 0.3 pu lasting for 0.1 s is applied at t = 2 s at the grid side of the system under study (Fig. 14). Without the SMES unit, the real power produced by the DFIG will drop to 0.45 pu, as shown in Fig.15(a). However, with the SMES unit connected to the system, the DFIG output power will drop only to 0.55 pu. Fig. 15(b) implies that, with the connection of the SMES unit and during the event of voltage sag, the reactive power support by the DFIG is reduced, and the steady-state condition is reached faster, compared to the system without SMES.

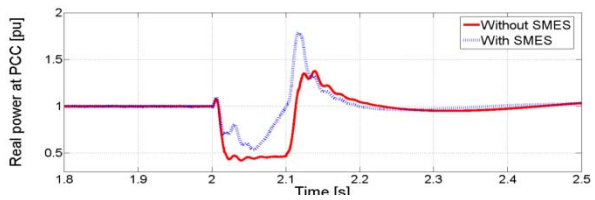


Fig.15a Active Power Regulation at PCC.

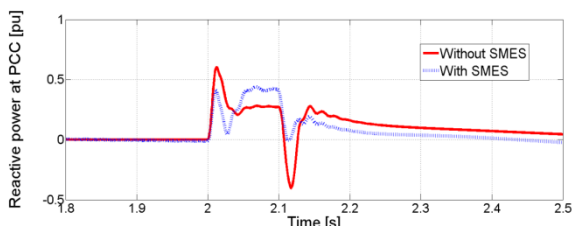


Fig.15b Reactive Power Regulation at PCC.

The voltage at the PCC is shown in Fig. 15(c), where without SMES, voltage will drop to 0.38 pu violating most of the low voltage ride through grid codes. However, by connecting the SMES unit, voltage drop at the PCC will be rectified to 0.68 pu, bringing the voltage level to a safety margin specified by wind turbine manufacturers. With the SMES connected to the system and due to active power compensation, the shaft speed oscillation and electromechanically torque along with settling time will also be substantially reduced.

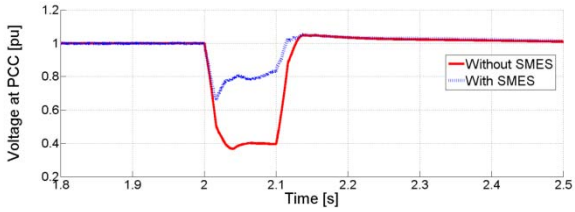


Fig. 15c Voltage Regulation at PCC.

IX. Conclusions

So far the grid code developments have been heavily influenced by instincts of utility engineers and isolated anecdotal experiences. A more consolidated approach in the area of grid connection regulations is required in future, addressing factors such as penetration level of wind power, network redundancy and robustness, short circuit level at the PCC etc. New emerging grid codes are expected to include stringent requirements on reactive power injection, inertia response and power oscillation damping. Grid side technologies using innovative power electronic technologies are required to meet current and emerging grid codes. This area of research and development requires a collective effort from researchers both from academia and industry. In this paper some international grid codes are presented, challenges and opportunities are put forward, and some enabling technologies are shown in order to solve some of these challenges.

Acknowledgment

The author acknowledges the contribution made by his PhD and Master degree students namely, Mansour Mohseni, Binayak Banerjee, Yousef Khaimiara, and David Stephens.

References

[1] O. Anaya-Lara, N. Jenkins, J. Ekanayake, P. Cartwright, and M. Hughes, "*Wind energy generation modeling and contro"l*,Wiley, 2009.

[2] R.M. Zavadil, C. J. Smith, "Status of wind-related US national and regional grid code activities," in Proc. *IEEE Power Engineering Society General Meeting*, San Francisco, 2005.

[3] S. Bernard, Beaulied, G. Trudel, "Hydro-Quebec grid code for wind farm interconnection," in Proc. *IEEE Power Engineering Society General Meeting*, San Francisco, 2005.

[4] I. Erlich, W. Winter, A Dittrich, "Advanced grid requirements for the integration of wind turbines into the German transmission system," in Proc. *IEEE Power Engineering Society General Meeting*, Montreal, 2006.

[5] I. M. Dudurych, M. Holly, M. Power, "Integration of wind power generation in the Irish grid," in Proc. *IEEE Power Engineering Society General Meeting*, Montreal, 2006.

[6] SKM, *New generation technologies and GB grid codes*, Dec. 2004. Available at< http://www.ofgem.gov.uk/>.

[7] A. Morales, X. Robe, J. L. R. Amenedo, S. Arnaltes, R. Zubiaur, Z. Torbado, "Advanced grid requirements for the integration of wind farms into the Spanish transmission system," *IET Renew. Power Gener*., vol. 2, no. 1, pp. 47–59, 2007.

[8] C. Jauch, J. Matevosyan, T. Ackermann, and S. Bolik, "International comparison of requirements for connection of wind turbines to power systems", *Wind Energy*, vol. 8, no. 8, pp. 295–306, Jul. 2005.

[9] W. Christiansen, D.T. Johnsen, *Analysis of requirements in selected grid codes*, Technical Report, Section of Electric Power Engineering, Technical University of Denmark (DTU), 2006.

[10] I. M. de Alegrıa, J. Andreu, J. L. Martın, P. Ibanez, J. L Villate, and H. Camblong, "Connection requirements for wind farms: A survey on technical requierements and regulation," *Renewable and Sustainable Energy Review*., vol. 11, pp. 1858–1872, Nov. 2007.

[11] Energinet, Technical regulation 3.2.5 for wind power plants with a power output greater than 11 kW, Sept. 2010. Available at http://www.energinet.dk>.

[12] M Mohseni, "Enhanced Control of DFIG Baesd Wind Power Plants to Comply with International Grid Codes" *PhD Thesis*, Curtin University, 2011.

[13] David Stephens, "Incrasing Renewable Eergy Hosting Capacity at Marble Bar", *Master Thesis*, Curtin University, 2015.

[14] Binayak Banerjee, D Jayaweera and S Islam, "Congestion management with Dynamic Line Ratings Considering Network Imbalance", *in the Proceedings of the 2015 IEEE PES General Meeting*, IEEE, 26-30 July 2015.

[15]A. Abu-Siada, S. Islam, "Application of SMES Unit in Improving the Performance of an AC/DC Power System", *IEEE Transactions on Sustainable Energy*, Vol. 2, No. 2, pp. 109-121, 2011.

2-D to 3-D Conversion of Videos using Fixed Point Learning Approach

Nidhi Chahal,[1,*] and Santanu Chaudhury[2†]
[1]Electrical Department, Indian Institute of Technology
New Delhi, India
[2]Electrical Department, Indian Institute of Technology, New Delhi, and Central Electronics Engineering Research Institute
Pilani, India
*nidhi.vce@gmail.com, †schaudhury@gmail.com

***Abstract*—The depth cues from multiple images are useful in accurate depth extraction while monocular cues from single still image are more versatile. In our paper, monocular cue which gives useful information about single frame and depth from motion using optical flow estimated from consecutive video frames are used to produce final depth maps. The machine learning approach is promising and new research direction in the field of depth estimation and thus 2-D to 3-D conversion. A fast automatic technique is proposed which utilizes a fixed point learning framework for the accurate estimation of depth maps of test images. For this task, a contextual prediction function is generated using training database of 2-D color and ground truth depth images. The depth maps obtained from monocular and motion depth cues of input video frames are used as input features for learning process. The depths generated from fixed point model are more accurate and reliable than MRF fusion of these depth cues. The stereo pairs are generated using depth maps predicted from fixed point learning. These final stereo pairs are converted to 3-D output video which is displayed on 3-DTV. For subjective evaluation, MOS score is calculated by showing final 3-D video to different viewers using 3-D glasses.**

***Index Terms*—Monocular, ZPS, Artifacts, MOS.**

I. Introduction

THE world of 3-D includes the third dimension of depth which can be perceived in form of binocular disparity by human vision. The different views of real world are perceived by both eyes of human as these are located at different positions. From these different views, the brain is able to reconstruct the depth information. A 3-DTV display presents two slightly different images of every scene to the individual eyes and thus 3-D perception can be realized. The better visualization of 3-D views demands accurate and denser depth map estimation. There is great impact of depth map quality on overall 3-D output.

In general, depth cues of multiple images gives more accurate depth extraction. And monocular depth cues which are utilized for depth estimation from single still image are more versatile. For the accurate and high quality conversion of 2-D images to 3-D models, a single solution does not exist. To enhance the accuracy of results, various depth cues should be combined in such a way that more dense and fine depth maps can be obtained. The depth consistency should also be maintained for the accurate depth estimation. The principle of depth consistency states that if the color values or intensities of the neighboring pixels are similar, these also should have similar depth values. In our paper, machine learning framework is introduced which gives faster and more accurate 2-D to 3-D conversion.

In training phase, the video frames are extracted first from training video and their appearance features are utilized; the depth values of neighboring blocks of a given image block are used as contextual features. The initial depths are obtained using monocular depth cue of input video frames considering single frame at a time and motion depth cue using more than one image. The depths from monocular and motion cues are used as input features. These features and ground truth depth values of neighboring blocks or patches are used in learning process. The generated contextual prediction function gives labeling i.e. assigning depth values of image blocks or individual pixels as output while input being both its features and depth values of rest of the image blocks or pixels.

Thus, the trained model is used in testing phase where new testing video frames are used as inputs and corresponding depth values are predicted using this learning framework. Finally, stereo pair is generated from test image and its corresponding predicted depth map using image warping technique. A 3-D structure of testing video frames is obtained from a training database of video frames and their depths. Thus final 3-D output images/video is obtained which can be displayed on 3-DTV display and 3-D output can be viewed using 3-D glasses.

II. Prior Work

In the last decade, a new family of machine learning approaches are introduced for the depth estimation of different 2-D images. In these techniques, training database of color and depth images are used to estimate the depth map of a query color image. In this way, the information is transferred from the structure correlations of the color and depth images in the database to the query color image and its 3-D structure is obtained. Some 3-D conversion systems have used monocular depth cues only for the depth extraction from single still image which has their own limitations. Most of the prior work have focused on obtaining the better quality of estimated depth maps, but at the expense of high computational cost algorithms.

A depth is estimated in [1] from a single color image using MRF framework and image parsing processing. The machine learning approach has also been adopted in [2] that exchanges transference of labels by directly depth map data.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

The semantic labels and a higher complex supervised model is incorporated in [3] to achieve more accurate depth maps. Konard in [4] has not used image registration step to reduce the computational burden of previous approaches and used a matching based search framework based on HOG features to find similar structured images. This approach has less computational cost than previous approaches, but for many practical applications, it is still too high. The depth maps are combined using MRF model which incorporates contextual constraints into fusion model described in [7], but depth learning is not done in this method. MRF and CRF are often limited to capturing less neighborhood interactions due to heavy computational burden in training and testing stages and thus, their modeling capabilities are also limited.

In our paper, fixed point learning framework provides depth learning from training database of RGB images and their ground truth depth values. The initial depth values from defocus and motion cues are used as training features. And we are using ground truth depth values which are obtained after manual labeling of input images for which ground truth depth maps are not available. These ground truth depths have been used as training labels which results to fixed point iteration. The learned fixed point function captures rich contextual/structural information and is easy to train and test and much faster also which balances the performance and learning time. For the testing images, reliable depth maps can be obtained using this trained model. The proposed learning framework provides automatic and fast 2-D to 3-D conversion method which converges at very less number of iterations with good accuracy in results.

Fixed point model [8] is a simple and effective solution to the structured labeling problem. We compared the results with MRF fusion as described in [7] of depth cues. Depth maps from focus and motion are fused using MRF approach. But the depths obtained from fixed point model have higher accuracy than MRF fused depths. The results are also obtained with other fusion methods like weighted averaging, least squares, maximizing approach, window technique. All the fusion depths have less accuracy as compared to fixed point model results. The error of depth maps are calculated by taking their difference with ground truth depths, it varies between [0 1]. The accuracy of depths can be obtained by subtracting error from 1 and multiplying with 100 to get the values in percentage.

III. Initial Depth Maps

The HVS (Human Visual System) exploits a set of visual depth cues to perceive 3-D scenes. The extraction of depth information from single view is known as monocular cue. There are various monocular depth cues as shading, texture gradient, accommodation, linear perspective and others. The motion depth cue provides depth extraction from multiple images. Both methods are used in the paper to utilize the benefits of both depth cues. The two important depth cues for depth extraction are focus/defocus and motion parallax on the basis of their relative importance in the human brain. The importance depends on the distance (JND-Just noticeable difference) from the observer. A depth cue with larger JND means that it is harder to be detected in the human vision system.

A. Depth From Defocus

The defocusing factor is used as monocular depth cue as discussed above. The principle that stuff which is more blurry is further away is used in depth estimation from focus depth cue. The method used here recovers depth from a single defocused image captured by uncalibrated conventional camera as described in [5]. From the blur amount at each edge location, a sparse depth map is estimated. Then, propagate the depth estimates to the whole image and obtain full depth map. The defocus or blur pattern depends on the shape of aperture and is often called the circle of confusion (CoC).

First, the amount of spatially varying defocus blur at edge locations is estimated. The input image is re-blurred using Gaussian kernel and the gradient ratio between input and re-blurred images is computed. At the edge location, this ratio has maximum value. Thus amount of defocus blur can be obtained from the gradient ratio using this maximum value.

$$b = \frac{1}{\sqrt{R_{max}^2 - 1}} b_0 \tag{1}$$

We used depth from defocus method in which blur estimation method is described for 1-D case and then 2-D image. Given maximum value R_{max} at the edge locations, the unknown blur amount b can be calculated using equation 1, where b_0 is the standard deviation of the re-blur Gaussian kernel. This is blur estimation method for 1-D case which can be easily extended for 2-D images, explained in paper referred here. The Fast PD which is an optimal solution and faster than conventional graph cut techniques is used to propagate depth estimates from edge locations to other non edge regions of an image.

B. Depth Using Optical Flow

The other depth cue used is depth from motion. There are two ways to calculate depth from motion parallax and these are motion blur and optical flow. The drawback of depth extraction using motion blur is the cost due to the multiple of high quality cameras, image processing programming and a big hardware background. In our experiments, optical flow is used for depth extraction and it is calculated between two consecutive frames taken from the scene as described in [6]. The length of the optical flow vectors will be inverse proportional to the distance of the projected point. The optical flow can also be used to recover depth from motion by using:

$$Z = v_c \cdot \frac{D}{V} \tag{2}$$

where v_c is the velocity of camera,
D is the distance of the point on image plane from focus of expansion,
V is the amplitude of flow and
Z is the depth of the point in the scene projected.

IV. Final Depths using Fixed Point Learning

There are infinite number of possible solutions for recovering 3-D geometry from single 2-D projection. The problem can be solved using learning based methods in vision. Initially a set of images and their corresponding depth maps are gathered in supervised learning. Then suitable features are extracted from the images. Based on the features and the ground truth depth maps, learning is done using learning algorithms. The process is shown step by step in Fig. 1. The depths of new images are predicted from this learned algorithm. The local as well as global features are used in a supervised learning algorithm which predicts depth map as a function of image.

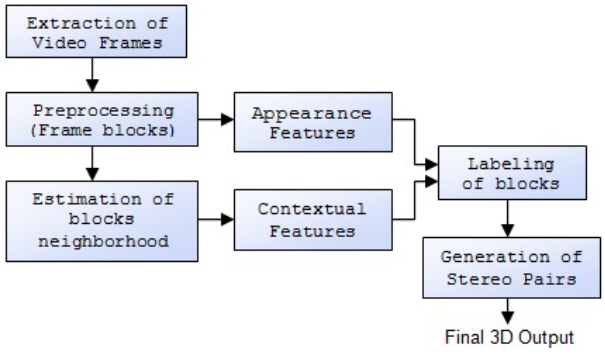


Fig. 1. 2-D video to 3-D output block diagram

If we take examples of monocular depth cues, local information such as variation in texture and color of a patch can give some information about its depth, these are insufficient to determine depth accurately and thus global properties have to be used. For example, just by looking at a blue patch, it is difficult to tell whether this patch is of a sky or a part of a blue object. Due to these difficulties, one needs to look at both the local and global properties of an image to determine depth.

The entire image is initially divided into small rectangular patches which are arranged in a uniform grid. And a single depth value for each patch is estimated. The absolute depth features are used to determine absolute depth of a patch which captures local feature processing. To capture additional global features, the features used to predict the depth of a particular patch are computed from that patch as well as the neighboring patches which is repeated at each of the multiple scales, so that the feature vector of a patch includes features of its immediate neighbors at large spatial scale, its far neighbors at a larger spatial scale, and again its very far neighbors at an even larger spatial scale.

The fixed point model approach is an effective and simple solution to the problem of structured labeling in image processing. A fixed point function is obtained with the labeling of the structure being both the output and the input. The overall fixed point function is a vector form of contextual prediction function of the nodes. The contextual prediction function gives labels of an individual node as output while inputs are both its features and labeling of rest of the nodes or neighboring nodes. The learned function captures rich contextual/structural information and is easy to train and test.

In image processing, a structured input is an image of all pixels and the structured outputs are the corresponding labels of these pixels. The structured input as shown in Fig. 2, can be represented as a graph, denoted as χ = (**I** , **J**). The block $b_i \in$ **I** represents i^{th} block with its features denoted as f_i. And **J** represents the neighborhood of each block. M denotes the number of neighbors a block can have in its neighborhood specification. That is, M number of blocks specifies the neighborhood N_i of block b_i in four directions: left, right, top and bottom.

The labeling of the neighborhood of b_i is denoted as $\mathbf{q}N_i$ where **q** denotes labeling of all the blocks in χ. The labels for all blocks can be represented as **p** = (p_i : 1...ϕ), where $p_i \in \phi$ and ϕ is the label space. For each block b_i, a context prediction function Ψ takes both b_i's appearance features f_i and contextual features $\mathbf{q}N_i$ as inputs. And this prediction function predicts the labeling $p_i \in \phi$ for the blocks of image. The contextual prediction function defined in [8] can be represented as follows:

$$q_i = \psi(f_i, \mathbf{q}N_i; \delta) \tag{3}$$

where ψ is a regression function within range [0,1] and δ is the parameter of the function. The labeling **q** of all image blocks can be written as

$$\mathbf{q} = \psi(f_1, f_2, ...f_n, \mathbf{q}; \delta) \tag{4}$$

Given the depth values **q** and features f_1, f_2,...f_n of training data, the parameter δ is learned. In experiments, SVM is used as contextual prediction function. The function is trained for each image block on the basis of appearance features f_i and contextual features $\mathbf{q}N_i$. The contraction mapping, when learned is applied iteratively to the new structured inputs.

A. Local Features

The frames are extracted first from the 2-D input video and converted to gray scale. These frames are divided into 8x8 blocks to reduce the computations. The appearance features of these input frames/images is extracted then. The examples of appearance features of image are maximum height of the block, width of the block, aspect ratio and mean wise averaging of the image blocks. The RGB values can also be taken as appearance features of images.

B. Global Features

The neighborhood of each image block is identified then. The parameter M defines the span of neighborhood of image blocks. The neighboring blocks of ith block is defined by all the adjacent blocks to its left, right, top and bottom. In our experiments, we used M = 1 to 6 where M = 6 gives the best results, this value is obtained empirically. A normalized histogram of neighboring blocks is used as contextual feature $\mathbf{q}N_i$. For each block b_i, a 4 x M x C dimensional contextual feature vector is created. In this vector, 4 specifies the neighborhood in four directions (upper, lower, left and right), M is the span of the context and C is the number of class labels. In implementations, following values are used.

- M = 6 = Number of neighbors a node can have in the neighborhood specification.

- C = 255 = No of classes = No of histogram bins
- Number of iterations in prediction = 3

C. Depth Prediction

All the image blocks are labeled in this step. Labeling means assigning depth values to each block of an image. The ground truth depth values are used as training labels resulting in fixed point iteration. At each iteration, the depth value for which convergence is achieved, is assigned to the corresponding block during testing process. The fixed-point model converges very quickly at the testing stage with only 2 to 3 iterations. We have used three number of iterations in our experiments as it gives good accuracy with less implementation time. We experimented with L1 regularized support vector machine (SVM-L1), provided in the Lib linear software package as the classifier. The SVM parameter is taken as '-s 5 -c 1', KLR can also be used for this application.

V. Stereo Generation

The stereo pairs are generated at the end from 2-D original images and final obtained depth maps. Consider original 2-D image as intermediate image and generate left and right views from this image. The right and the left images will be half shifted or warped toward the respective direction and all the artifacts introduced by the processing will be part of both images but halved visible. So, less artifacts are produced in this method. We applied this technique [9] for final stereo generation.

First, pre-processing of depth maps is done by shifting depth map by ZPS i.e. Zero Parallax Setting:

$$Z_c = \frac{Z_n - Z_f}{2} \tag{5}$$

where Z_n and Z_f are the nearest clipping plane and the farthest clipping plane of the depth map. In an 8-bit depth map, Z_n = 255 and Z_f = 0. After that, the depth map is further normalized with the factor of 255, so that the values of the depth map lie in the interval of [-0.5, 0.5], values that are required by the image warping algorithm. Then, Gaussian filter is used for depth smoothing.

- w = 3 x σ where w is the filter window size
- σ is standard deviation
- Depth smoothing strength = Baseline/4

The stereo pair is obtained by transferring the original points at location (P_m,y) to left points (P_l,y) and right points (P_r,y):

$$P_l = P_m + \frac{Bf}{2Z} \tag{6}$$

$$P_r = P_m - \frac{Bf}{2Z} \tag{7}$$

where B is the baseline distance between two virtual cameras.
B = 5 % of width of depth image
f is the focal length of camera
f is chosen as one in our experiments without any loss of generality.

(a) Frames of Input Video 1

(b) Frames of Input Video 2

(c) Frames of Input Video 3

Fig. 2. Examples of 2-D video frames from input videos

TABLE I
Accuracy of Predicted Depths with respect to Ground Truth

Test	Fixed Point	MRF Model
1	83.57%	73.6%
2	83.45%	73.72%
3	82.71%	74.90%
4	82.39%	74.00%
5	83.00%	74.72%
6	82.38%	73.73%

VI. Implementation Results

In Fig. 2, examples of few input frames are shown for different videos which are converted to 3-D video in our work. These are cartoon videos of famous Mickey mouse and Donald. The first step is generating initial depth images from defocus and motion. These are shown in Fig. 3, in which frames of natural scene video are also considered apart from animated video. The depths from defocus and motion are shown in second and third rows respectively. The depth from motion is generated using two consecutive frames, thus, two depth images from motion are shown for three input frames in Fig. 3.

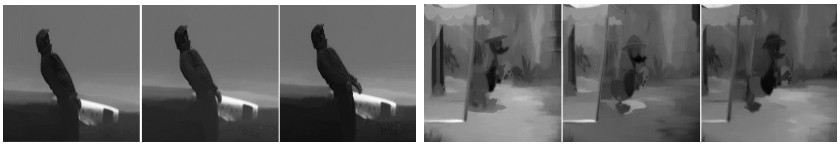
The final 3-D output is shown in Fig. 4, in which input color video of Mickey mouse is considered. The video is in wmv format and the resolution of video frames is 960×1080. The resolution of final 3-D image is 1920 x 1080 where final left and right frames are put side by side. The 3-D results for more frames are shown in Fig. 5. The results of fixed point learning and MRF model for some video frames in form of accuracy (percentage) is shown in Table I. The error of depth maps are calculated by taking their difference with ground

TABLE II
Subjective Quality Measure

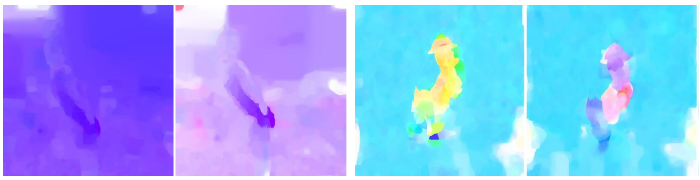
Video	Viewer's Score	MOS
Mickey	4, 4, 5, 4, 4	4.2
Donald	4, 4, 4, 4, 5	4.2
Human	4, 3, 5, 4, 4	4

MOS: Mean Opinion Score

(a) Input frames from different videos



(b) Corresponding depths from defocus depth cue



(c) Corresponding depths from motion depth cue

Fig. 3. Initial depths from defocus and motion. Two depth images from motion cue for corresponding three input frames.

(a) Input frame 1 (b) Input frame 2 (c) Depth from defocus (d) Optical Flow (f) Final 3-D

Fig. 4. Conversion of 2-D video frames to 3-D using fixed point learning

truth depths, it varies between [0 1]. The accuracy of depths can be obtained by subtracting error from 1 and multiplying with 100 to get the values in percentage. Example, if error between predicted depth and ground truth depth is 0.2, then accuracy is $1 - 0.2 = 0.8$ or 80%. The accuracy of depth maps is higher using fixed point learning process as compared to results with MRF model.

The subjective test is conducted to show the quality of 3-D images, that is, left and right image pair for every 3-D image on 3-DTV display by wearing 3-D glasses. Five participants were involved in the experiment to evaluate 3-D quality. Their individual scores are mentioned in Table II using which MOS i.e. mean opinion score is calculated. The MOS is the arithmetic mean of all the individual scores and can vary from 1 (worst) to 5 (best).

VII. Conclusion

The depth estimation is the key step for conversion of 2-D images or videos to 3-D. In this paper, final depths are obtained using depth cue of single still image and motion cues which use more than one image. Fixed point model uses ground truth depth maps and provides learning framework for accurate estimation of depths which gives higher accuracy than MRF and other fusion methods. Finally, stereo pairs are generated to get final 3-D output which can be viewed on any 3-DTV display. The quality of 3-D images is also evaluated using MOS which indicates good and reliable depth extraction. MRF has limitations in capturing neighborhood interactions which limits their modeling capabilities while fixed point learned function captures rich contextual/structural information and is easy to train and test. The model is much faster to train and balances the performance and learning time giving higher accuracy than MRF depth fusion.

References

[1] A. Saxena, M. Sun, and A.Y. Ng: Make3-D: Learning 3-D Scene Structure from a Single Still Image. Trans. on Pattern Anal. and Mach. Intell. vol. 31, pp. 824-840. IEEE (2009).

[2] J. Konrad, M. Wang, and P. Ishwar: 2-D-to-3-D Image Conversion by Learning Depth from Examples. Comput. Soc. Conf. on Comput. Vis. and Pattern Recognit. Workshops (CVPRW), pp. 16–22. IEEE (2012)

[3] C. Liu, J. Yuen, and A. Torralba.: Nonparametric scene parsing: Label transfer via dense scene alignment. on Comput. Vis. and Pattern Recognit, pp 1972-1979, IEEE (2009).

[4] J. Konrad, M.Wang, P. Ishwar, C.Wu, and D. Mukherjee.: Learning-based, automatic 2-D-to-3-D image and video conversion. Trans. on Image Process, vol. 22, no. 9, pp 34853496, IEEE (2013).

[5] Shaojie Zhuo and Terence Sim.: Defocus Map Estimation from a Single Defocused Image: Pattern Recognition, vol. 44, no. 9, pp-1852-1858, (2011)

[6] Deqing Sun , Stefan Roth , Michael J. Black.: Secrets of Optical Flow Estimation and their Principles. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2432-2439 (2010)

[7] M. Xu, H. Chen and P. K. Varshney. An Image Fusion Approach Based on Markov Random Fields: TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, vol. 49, no. 12, (2011).

[8] Q. Li, J. Wang and Z. Tu: Fixed-Point Model For Structured Labeling: Proceedings of the 30th International Conference on Machine Learning, Atlanta, Georgia, USA, vol. 28, (2013).

[9] Y. Kang, Y. Lai and Y. Chen: An Effective Hybrid Depth-Generation Algorithm for 2-D-to-3-D Conversion in 3-D Displays. Journal of Display Technology. vol. 9, (2013)

[10] A. Torralba and A. Oliva, *Depth Estimation from Image Structure*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 9, pp. 1226-1238, September 2002.

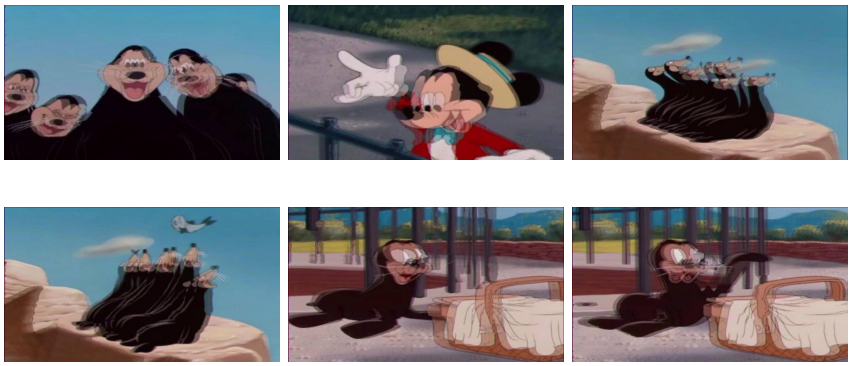


Fig. 5. Final 3-D output for few video frames

[11] Y. Salih, A. S. Malik and Z. May, *Depth Estimation Using Monocular Cues from Single Image*, National Postgraduate Conference (NPC), pp. 1-4, September 2011.

[12] F. Guo, J. Tang and H. Peng, *Automatic 2D-to-3D Image Conversion based on Depth Map Estimation*, International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 8, no. 4, pp. 99-112, 2015.

[13] D. C. Lee, A. Gupta, M. Hebert, and T. Kanade, *Estimating Spatial Layout of Rooms using Volumetric Reasoning about Object and Surfaces*, In NIPS, pp. 1288-1296, 2010.

[14] A. G. Schwing and R. Urtasun, *Efficient Exact Inference for 3D Indoor Scene Understanding*, In ECCV, pp. 299-313, 2012.

[15] A. Gupta, A. Efros, and M. Hebert, *Blocks world Revisited: Image Understanding using Qualitative Geometry and Mechanics*, In ECCV, vol. 6314, Lecture Notes in Computer Science, pp. 482-496, Springer, 2010.

Recent Advances in Direct Torque and Flux Control of IPMSM Drives

M. Nasir Uddin and Md. Mizanur Rahman
Department of Electrical Engineering
Lakehead University, Thunder Bay, Canada
E-mails: muddin@lakeheadu.ca, mrahman8@lakeheadu.ca

***Abstract*– This paper presents the latest advancements in direct torque and flux control (DTFC) schemes of interior permanent magnet synchronous motor (IPMSM) drives. A novel eighteen-sector based DTFC scheme incorporating a model based loss minimization algorithm is proposed to mitigate the torque ripples and achieve high efficiency as compared to the conventional six-sector based DTFC scheme. Finally, in order to have direct and better control of reducing the torque/flux ripples further, a nonlinear control incorporating motor electromagnetic developed torque and stator air-gap flux linkage as virtual control variables is developed. In conventional nonlinear controller d-q axis currents (i_d, i_q) are considered as virtual control variables that indirectly controls the torque/flux which may not be suitable for high performance drives. Thus, the proposed work overcomes the major drawback (i.e., torque ripple) of the conventional DTFC based IPMSM drive. Feasibility of the developed DTFC schemes is verified through both simulation and experimental results.**

***Index Terms*— Interior permanent magnet motor, loss minimization algorithm, direct torque and flux control, nonlinear torque/flux control, torque ripple minimization and real time implementation.**

NOMENCLATURE

v_d, v_q	Stator voltages; *d*-and *q*- axis components.
i_d, i_q	Stator currents; *d*-and *q*- axis components.
L_d, L_q	Stator inductances; *d*-and *q*- axis components.
R_s	Stator per phase resistance.
Ψ	Rotor magnetic-flux linkage.
ψ_s	Stator flux linkage.
P	Motor pole pairs number.
ω_r	Actual rotor speed.
ω	Electrical speed of the rotor (ω=Pω_r)
i_{cd},i_{cq}	d-q axis core-loss armature current.
i_{oq}	q-axis torque-generating armature current.
P_{cu},P_{Fe}	Copper and iron loss of the motor.
k_ω, k_T	Speed error and torque error.
k_ψ	Stator flux linkage error.
ψ_d,ψ_q	Stator flux linkages; *d*-and *q*- axis components.
T_e, T_L	Motor electromagnetic developed and load torque.
J	Rotor inertia constant.
B_m	Friction damping coefficient.
T_e^*, ψ^*	Command torque and stator flux linkage.

I. INTRODUCTION

Advent and expanded accessibility of high energy rare earth magnetic material such as, neodymium-iron-boron and their alloys put high attention for researchers in the control of interior permanent magnet synchronous motor (IPMSM) as it has the capability of producing reluctance torque which offers high competence and low turbulence with reliability [1-2]. Researchers normally concentrate on vector control (VC) and direct torque and flux control (DTFC) in the field of drives. The DTFC offers fast transient response, less parameter dependence and it does not need the pulse-width modulation (PWM) which is a complex field oriented block as compared to field oriented control (FOC) [3]. On the other hand, DTFC suffers from variable switching frequency, high ripples in torque, flux distortion, hysteresis bound which degrade its performance of drive systems at low speed operation [4]. The variable switching frequency can generate mechanical resonance and due to flux distortion current harmonics generates which affect DTFC performance [5]. To overcome the torque/flux ripples of DTFC researchers mainly concentrate on the inverter control algorithms [6-7]. In DTFC the motor actual torque and air gap flux linkage values are compared with the corresponding reference values. Corresponding error signals are passed through the torque and flux hysteresis comparators to generate the logic signals of the voltage vector look up table [8]. Generally, high torque ripples from conventional six-sector based DTFC scheme restricts better dynamic performance of industrial drives. Therefore, to overcome the unbalanced voltage effects in any sectors and to reduce motor torque ripples, conventional six sectors are replaced with the proposed eighteen-sector. Again, most of the researchers considered reference air-gap flux as constant based on the trial-and-error method to simplify the control task which may not be acceptable for high performance drives [9]. However, it is not possible to minimize the motor losses with constant air-gap flux as the air-gap flux changes with the operating conditions [10,11].

Therefore, a novel eighteen-sector based DTFC scheme is developed to achieve high dynamic performance with reduced torque ripples as compared to the conventional 6-sector based DTFC. In addition, a model based loss minimization algorithm (LMA) is integrated with the proposed DTFC scheme in order to optimize the efficiency

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

along with high dynamic performance. Eighteen sectors are developed to overcome the unbalanced voltage vector selection of conventional six-sector based system that minimizes the torque ripples. Further, a nonlinear controller with virtual torque and flux controls is also developed for IPMSM drive to minimize the drive torque ripples. The complete IPMSM drives incorporating the developed control techniques are successfully implemented in real-time using digital signal processor (DSP) board-DS1104 for laboratory 5-hp motor. The effectiveness of the proposed control techniques are verified in both simulation and experiment at different operating conditions. It is found that the nonlinear controller based IPMSM drive provides the best performance in terms of torque ripples among all the DTFC schemes. It is found that the proposed nonlinear/18-sector based DTFC scheme could be the potential candidate for real-time industrial drives.

II. Proposed 18-Sector Based DTFC Scheme

Fig. 1 shows the proposed IPMSM drive incorporating eighteen-sector based DTFC scheme and model based LMA. The traditional DTFC divides stator flux into six sectors where unbalanced voltage effect in one sector can result the unbalanced change of flux. Speed oscillations as well as torque ripples can be reduced by adding more sectors in classical DTFC as switching frequency increases with the increase of sector numbers. Therefore, twelve new sectors are added to the original six sectors and the spaces are divided such as, (-10^0, 10^0), (10^0, 30^0)……and (330^0, 350^0). The command air gap flux reference and motor reference torque magnitudes are compared with their respective actual values and the errors are processed through hysteresis band controllers. Based on the outputs of torque and flux hysteresis controllers and stator flux-linkage space vector position number, the DTFC switching table produces the logic signals for three-phase voltage source inverter [13]. The lookup table to select the proper voltage vector for the proposed DTFC scheme is shown in Table-I that results in tracking the actual torque/flux to follow the reference torque/flux. A three level torque hysteresis comparator is utilized for the proposed DTFC so that the torque ripple remains at the minimum level.

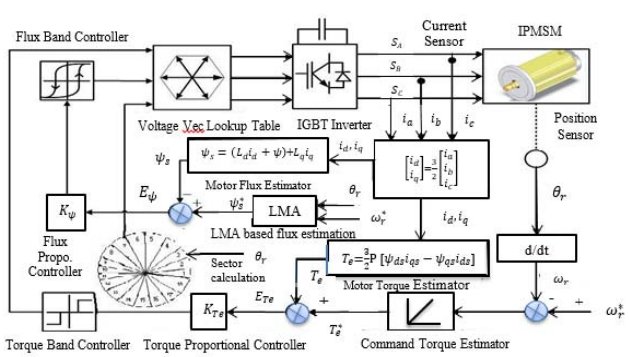


Fig. 1: 18-sector based DTFC scheme incorporating LMA of IPMSM drive.

For torque hysteresis controller output H_T=+1 means that the reference torque is less than the estimated one and a voltage vector should be selected in a way to increase the torque. H_T=-1 means that torque should be decreased. The exactly same control principle is used for the amplitude of stator flux. For example, if the stator flux is in sector 1, to decrease flux (H_ψ=-1) and to increase the torque (H_T=+1), voltage vector is applied two steps ahead of the current sector voltage vector. Thus, the voltage vector V_3 is applied. When H_T =-1 and H_ψ=-1 and flux linkage space vector is in sector 1, then voltage vector is applied that is 2 steps backward of the current sector number. That's why voltage vector V_5 is applied, which is shown in Table-I.

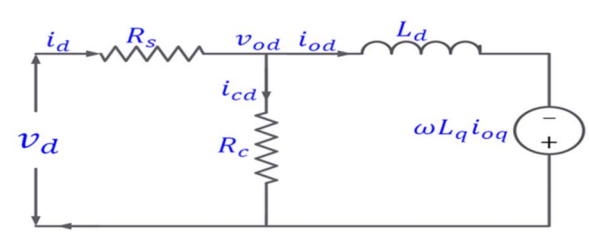
Table-I
DTFC Look-up table for voltage vectors

H_ψ	H_T	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9
+1	+1	V_2	V_2	V_3	V_3	V_3	V_4	V_4	V_4	V_5
	0	V_0	V_0	V_0	V_7	V_7	V_7	V_0	V_0	V_0
	-1	V_6	V_1	V_1	V_1	V_2	V_2	V_2	V_3	V_3
-1	+1	V_3	V_3	V_3	V_4	V_4	V_4	V_5	V_5	V_5
	0	V_7	V_7	V_7	V_0	V_0	V_0	V_7	V_7	V_7
	-1	V_5	V_6	V_6	V_6	V_1	V_1	V_1	V_2	V_2

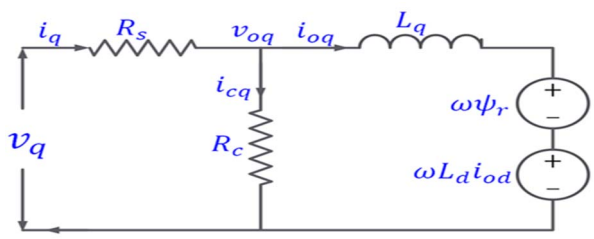
H_ψ	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}	S_{16}	S_{17}	S_{18}
+1	V_5	V_5	V_6	V_6	V_6	V_1	V_1	V_1	V_2
	V_7	V_7	V_7	V_0	V_0	V_0	V_7	V_7	V_7
	V_3	V_4	V_4	V_5	V_5	V_5	V_5	V_6	V_6
-1	V_6	V_6	V_6	V_1	V_1	V_1	V_2	V_2	V_2
	V_0	V_0	V_0	V_7	V_7	V_7	V_0	V_0	V_0
	V_2	V_3	V_3	V_4	V_4	V_4	V_2	V_5	V_5

A. Motor Model Based Loss Minimization Algorithm

For the eighteen-sector based DTFC scheme, a model-based LMA is developed to obtain the optimum flux, which will act as a command flux for the IPMSM drive. The model based LMA is used to calculate the optimum i_d^* which ensures the minimum power losses. The d-q axis model of IPMSM drives incorporating the iron loss is shown in Fig. 2.



(a)



(b)

Fig. 2: IPMSM model incorporating iron loss; (a) d-axis, (b) q-axis.

Core loss armature currents are obtained from Fig. 2 as:

$$i_{cd} = \frac{L_d\frac{d(i_{od})}{dt} - - L_q i_{oq}\omega}{R_c} = \frac{-L_q i_{oq}\omega}{R_c} \quad (1)$$

$$i_{cq} = \frac{L_q\frac{d(i_{oq})}{dt} + \psi_r\omega + L_d i_{od}\omega}{R_c} = \frac{\psi_r\omega + L_d i_{od}\omega}{R_c} \quad (2)$$

Total electric power loss can be written as:

$$P_E = \frac{3}{2}\left(R_s {i_d}^2 + R_s {i_q}^2 + R_c {i_{cd}}^2 + R_c {i_{cq}}^2\right) \quad (3)$$

Higher efficiency is achieved when the total power losses are minimum and minimum power loss condition is obtained by differentiating the total power loss with respect to i_d and then set it to zero to get optimum i_d as:

$$i_d^* = \frac{-[R_s i_q\frac{di_q}{di_d} + R_c i_{cd}\frac{di_{cd}}{di_d} + R_c i_{cq}\frac{di_{cq}}{di_d}]}{R_s} \quad (4)$$

To minimize the electrical losses and to maximize the efficiency steady state value of the core loss component of d- and q- axis current is calculated from Fig. 4 (a) and 4(b) as:

$$i_{cd} = \frac{-L_q\omega(i_q - i_{cq})}{R_c} = -\frac{L_q\omega i_q}{R_c} + \frac{L_q\Psi_r\omega^2 + L_d L_q\omega^2(i_d - i_{cd})}{{R_c}^2} \quad (5)$$

$$i_{cq} = \frac{L_d\omega\,(i_d - i_{cd}) + \Psi_r\omega}{R_c} - \frac{\Psi_r\omega}{R_c} + \frac{L_d\omega\, i_d}{R_c} - \frac{L_d\omega}{R_c}\left[\frac{-L_q\omega(i_q - i_{cq})}{R_c}\right] \quad (6)$$

After some manipulation, one can obtain:

$$i_{cd} = \frac{-L_q i_q\omega R_c + L_q\Psi_r\omega^2 + L_d L_q\omega^2 i_d}{{R_c}^2 + L_d L_q\omega^2} \quad (7)$$

$$i_{cq} = \frac{\Psi_r\omega R_c + L_d\omega\, i_d R_c + L_d L_q i_q\omega^2}{{R_c}^2 + L_d L_q\omega^2} \quad (8)$$

Differentiating d-q axis core loss current i_{cd} and i_{cq} with respect to i_d:

$$\frac{di_{cd}}{di_d} = \frac{L_d L_q\omega^2}{{R_c}^2 + L_d L_q\omega^2} \quad (9)$$

$$\frac{di_{cq}}{di_d} = \frac{L_d\omega R_c}{{R_c}^2 + L_d L_q\omega^2} \quad (10)$$

In steady state, $\frac{di_q}{di_d}$=0. Substituting the values of $\frac{di_{cd}}{di_d}$, $\frac{di_{cq}}{di_d}$, $\frac{di_q}{di_d}$ into (4) yields:

$$i_d = {i_d}^* = -\frac{R_c L_d\omega[i_{cd} L_q\omega + i_{cq}\omega]}{R_s({R_c}^2 + L_d L_q\omega^2)} \quad (11)$$

Using equation (6) and (11):

$$i_q = {i_q}^* = \frac{L_d\omega\, i_{0d} + \Psi_r\omega}{R_c} + i_{0q} \quad (12)$$

For DTFC, reference flux is estimated from the reference d-q axis current as:

$$\Psi_s^* = \psi_d^* + j\psi_q^* = (\Psi_r + L_d i_d^*) + jL_q i_q^* \quad (13)$$

The steady-state output power, P_{out}, total loss P_L, and efficiency η of the IPMSM can be calculated as:

$$P_E = P_{cu} + P_{Fe} \text{ and } P_L = P_E + P_{mechanical} \quad (14)$$

$P_{mech.}$= Mechanical frictional constant* ω (15)

The steady-state output power, $P_{out} = T_L\omega$ (16)

Motor efficiency is given by;$\eta = \frac{P_{out}}{P_{total}} * 100\%$ (17)

III. Virtual Torque-Flux Control Based Nonlinear Controller

Considering sinusoidal induced EMF with no damper winding on rotor and negligible core losses, state model of an IPMSM drive is derived from synchronous machine can be obtained by [14] as follows:

$$\frac{di_d}{dt} = \frac{1}{L_d}[v_d - R_s i_d + P\omega_r L_q i_q] \quad (18)$$

$$\frac{di_q}{dt} = \frac{1}{L_q}[v_q - R_s i_q - P\omega_r L_d i_d - P\omega_r\psi_m] \quad (19)$$

$$\frac{d\omega_r}{dt} = \frac{1}{J}[T_e - T_L - B_m\omega_r] \quad (20)$$

$$T_e = \frac{3}{2}P[\psi_d i_q - \psi_q i_d] \quad (21)$$

$$\Psi = \psi_d^2 + \psi_q^2 \quad (22)$$

Motor actual speed converge to the command speed for speed tracking and error dynamics (speed) is specified as:

$$e_\omega = \omega_r^* - \omega_r \quad (23)$$

where: e_ω= speed error and ω_r^*= speed reference.

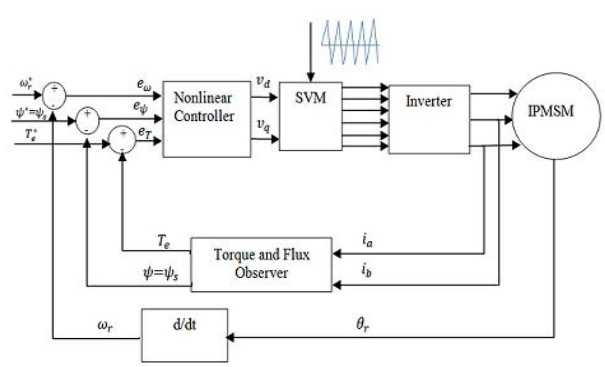


Fig. 3: Proposed nonlinear torque/flux control scheme of IPMSM drive.

Then, from (20), and (23) one can get:

$$\dot{e_\omega}=-\dot{\omega_r}=\frac{1}{J}(B_m\omega_r+T_L-T_e) \quad (24)$$

where: $\dot{e_\omega}$= rate of change of speed error (error dynamic). To fulfill the speed tracking objectives, derivative of the initial Lyapunov function, $V=\frac{1}{2}e_\omega^2$, given by:

$$\dot{V}=e_\omega\dot{e_\omega}=\frac{e_\omega}{J}(B_m\omega_r+T_L-T_e) \quad (25)$$

To ensure the stability of (25), T_L and ψ_s is identified as virtual control variable such that (25) becomes negative semi-definite. Command motor electromagnetic developed torque and stator air-gap flux are chosen as:

$$T_e^*=B_m\omega_r+T_L+k_\omega Je_\omega \text{ and } \psi^*=\psi_{ref} \quad (26)$$

where k_ω is closed-loop feedback gain, replacing (26) into (25):

$$\dot{V}=-k_\omega e_\omega^2 \quad (27)$$

Speed tracking requirements of (27) are fulfilled assuming $k_\omega>0$. Corresponding derivatives of the error dynamics:

$$\dot{e_T}=\dot{T_e^*}-\dot{T_e} \quad (28)$$

$$\dot{e_\psi}=\dot{\psi}^*-\dot{\psi} \quad (29)$$

From the derivative of motor command electromagnetic developed torque:

$$\dot{T_e^*}=\frac{1}{J}(B_m-k_\omega J)(T_e-T_L-B_m\omega_r) \quad (30)$$

From the derivative of motor actual electromagnetic developed torque:

$$\dot{T_e}=\frac{3}{2}P[\psi_d\frac{d}{dt}(i_q)+\dot{\psi_d}i_q-\psi_q\frac{d}{dt}(i_d)-\dot{\psi_q}i_d] \quad (31)$$

Substituting the values of $\dot{T_e^*}$ and $\dot{T_e}$ in (28):

$$\dot{e_T}=\frac{(B_m-k_\omega J)}{J}(T_e-T_L-B_m\omega_r)-\frac{3}{2}P\psi_d[\frac{1}{L_q}(-R_si_q-P\omega_rL_di_d-P\omega_r\psi_m)]+\frac{3}{2}P(i_d-\frac{\psi_d}{L_q})v_q+\frac{3}{2}P\psi_q[\frac{1}{L_d}(-R_si_d+P\omega_rL_qi_q)]-\frac{3}{2}P(i_q-\frac{\psi_q}{L_d})v_d \quad (32)$$

Stator air-gap flux linkage error dynamics are derived as:

$$\dot{e_\psi}=\dot{\psi}^*-\dot{\psi}=2\psi_dR_si_d+2\psi_qR_si_q-2\psi_dv_d-2\psi_qv_q \quad (33)$$

New Lyapunov function including the speed, torque and stator air-gap flux linkage variables can be defined as:

$$V_1=\frac{1}{2}(e_\omega^2+e_T^2+e_\psi^2) \quad (34)$$

The derivative of this Lyapunov function is:

$$\dot{V_1}=e_\omega\dot{e_\omega}+e_T\dot{e_T}+e_\psi\dot{e_\psi}$$

$$=-k_\omega e_\omega^2+e_T\left\{\frac{(B_m-k_\omega J)}{J}(T_e-T_L-B_m\omega_r)-\frac{3}{2}P\psi_d\left[\frac{1}{L_q}(-R_si_q-P\omega_rL_di_d-P\omega_r\psi_m)\right]+\frac{3}{2}P(i_d-\frac{\psi_d}{L_q})v_q+\frac{3}{2}P\psi_q[\frac{1}{L_d}(-R_si_d+P\omega_rL_qi_q)]-\frac{3}{2}P(i_q-\frac{\psi_q}{L_d})v_d\right\}+e_\psi[2\psi_dR_si_d+2\psi_qR_si_q-2\psi_dv_d-2\psi_qv_q] \quad (35)$$

To establish the global asymptotic stability of (35), voltages v_d and v_q are chosen as:

$$v_d=\frac{1}{\psi_q\left(i_q-\frac{\psi_q}{L_d}\right)+\psi_d\left(i_d-\frac{\psi_d}{L_q}\right)}\left\{\frac{2\psi_q(B_m-k_\omega J)}{3JP}(T_e-T_L-B_m\omega_r)-\psi_d\psi_q\left[\frac{1}{L_q}(-R_si_q-P\omega_rL_di_d-P\omega_r\psi_m)\right]+\psi_q^2\left[\frac{1}{L_d}(-R_si_d+P\omega_rL_qi_q)\right]+\left(i_d-\frac{\psi_d}{L_q}\right)\left[\psi_dR_si_d+\psi_qR_si_q+\frac{1}{2}k_\psi e_\psi\right]+\frac{2\psi_q}{3P}k_Te_T\right\} \quad (36)$$

$$v_q=\frac{1}{-[\psi_q\left(i_q-\frac{\psi_q}{L_d}\right)+\psi_d\left(i_d-\frac{\psi_d}{L_q}\right)]}\left\{\frac{2\psi_d(B_m-k_\omega J)}{3JP}(T_e-T_L-B_m\omega_r)-\psi_d^2\left[\frac{1}{L_q}(-R_si_q-P\omega_rL_di_d-P\omega_r\psi_m)\right]+\psi_d\psi_q\left[\frac{1}{L_d}(-R_si_d+P\omega_rL_qi_q)\right]-\left(i_q-\frac{\psi_q}{L_d}\right)\left[\psi_dR_si_d+\psi_qR_si_q+\frac{1}{2}k_\psi e_\psi\right]+\frac{2\psi_d}{3P}k_Te_T\right\} \quad (37)$$

where, k_ω, k_ψ and k_T are closed-loop feedback constants. Final Lyapunov's derivative in (35) can be simplified as:

$$\dot{V_1}=-k_\omega e_\omega^2-k_Te_T^2-k_\psi e_\psi^2 \quad (38)$$

If k_ω, k_T and k_ψ>0, then tracking requirements are met and guarantees asymptotic stability in the system.

IV. Real-Time Implementation

The proposed torque and flux control based IPMSM drive is experimentally implemented using digital-signal processor (DSP) board DS1104 for a prototype 5-hp motor. The snapshot of experimental set up is shown in Fig. 4. A personal computer is used where DSP board is equipped and supplied by a software development system which provides the feature of changing command speeds and index pulses from the keyboard with editing and downloading capabilities. The DS1104 board is mainly 64-bit floating point DSP and uses a PPC603e processor with a clock pulse of 250 MHz supported by 32 MB of global and 8 MB of flash memory. A 16-bit TMS320F240 operates as a slave processor which provides input/output port configuration. An optical incremental encoder is used to sense the rotor position and fed to the DSP board. Two phase currents i_a and i_b are measured by Hall-effect current sensors. A/D (analog to digital) converters and interface circuits are used to feed the currents into the DSP board. Finally, firing pulses are achieved and these six pulses are fed to the IGBT based driver circuit. A dc-generator is used to as loading machine to the motor. A real-time Simulink model of the proposed drive system is built and downloaded loaded to the DSP board for real-time implementation. A sampling frequency of 10 kHz is found to be suitable for the overall drive system.

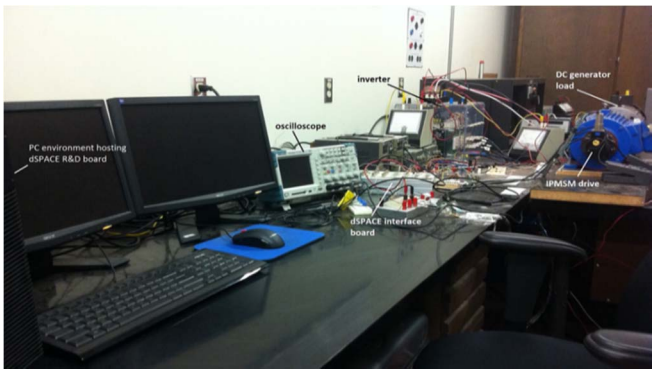


Fig. 4: Snapshot of the prototype drive system.

V. Sample Simulation and Experimental Results

A. Simulation Results

The performance of the proposed developed control techniques is tested in simulation at different operating conditions using MATLAB/Simulink software. Figures 5a and 5b show the starting responses of the conventional 6-sector based DTFC with LMA and proposed 18-sector based DTFC with same LMA of IPMSM drive, respectively at rated conditions (183 rad/s and 19.1 Nm). It is clearly seen from Figs. 5a and 5b that, motor torque ripples are reduced significantly with proposed DTFC scheme while maintaining similar speed and efficiency responses as compared to the conventional 6-sector based DTFC scheme.

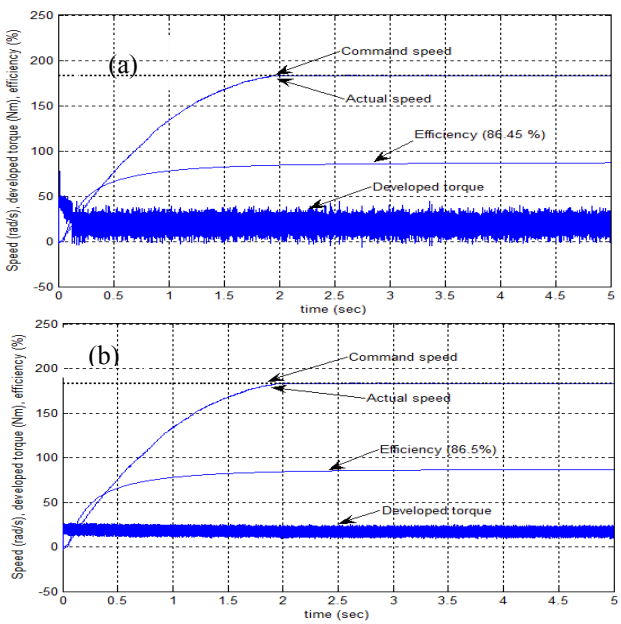


Fig. 5: Speed, torque and efficiency responses of LMA based DTFC scheme at rated conditions: (a) conventional six-sector based DTFC; (b) proposed eighteen-sector based DTFC.

Simulated speed and torque responses of the conventional and nonlinear torque/flux control scheme of IPMSM drive at rated condition are presented in Figs. 6a and 6b, respectively. In this case LMA was not integrated with any of the control techniques for the sake of only testing the torque ripple. It is seen from these figures that the nonlinear torque/flux control scheme exhibits significant reduction in torque ripple than that of the traditional DTFC scheme.

B. Experimental Results

The experimental speed and efficiency responses of the proposed DTFC scheme at rated speed (183 rad/s) with light load (25% rated load) conditions is shown in Fig. 7. It is seen that at rated speed condition proposed DTFC based IPMSM drive can follow the rated command speed smoothly and

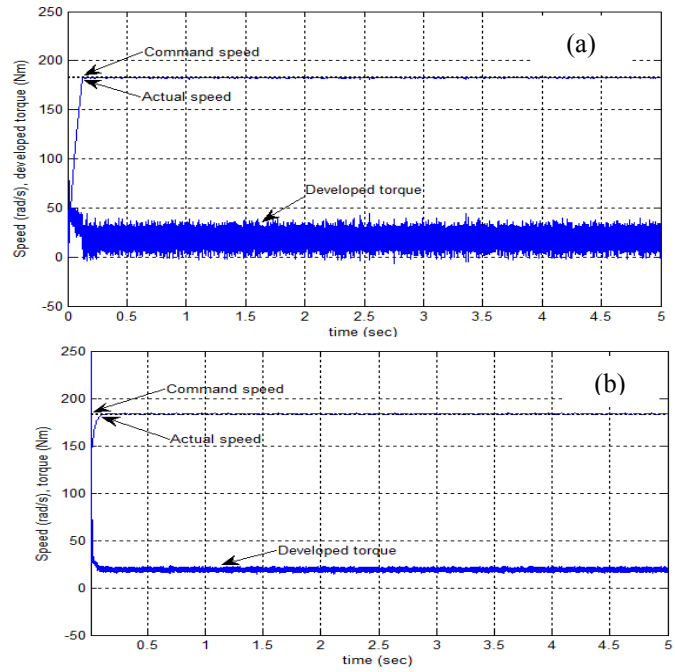


Fig. 6: Simulated speed and developed torque responses at rated conditions: (a) conventional; (b) nonlinear torque/flux control scheme.

quickly while maintaining high efficiency (85%) of the drive. The motor was tested at light-load condition as the loss increases relatively at light-load condition. Figure 8 demonstrates the comparative torque ripple analysis of the conventional 6-sector based DTFC with LMA, 18-sector based DTFC incorporating the same LMA and nonlinear torque/flux scheme at rated speed and 50% rated load. It is found from Fig. 8a that the classical DTFC has higher torque ripples which degrades the performance of the drive system, whereas 18-sector based DTFC scheme maintains lower torque ripples as shown in Fig. 8b. The nonlinear torque/flux control scheme maintains the lowest torque ripple among all the control schemes as shown in Fig. 8c.

VI. Conclusion

A novel 18-sector based DTFC scheme incorporating a model based loss minimization algorithm has been developed to mitigate the torque ripple while maintaining high efficiency as compared to the conventional 6-sector based DTFC scheme. A nonlinear controller incorporating virtual and torque and flux control has also been developed. Comparative results showed that the proposed 18-sector based DTFC scheme performed better than that of the conventional 6-sector based DTFC. The nonlinear based torque/flux control scheme exhibited the lowest torque triple than that of the traditional DTFC schemes. Because of high computational requirements of nonlinear scheme LMA was not integrated with this controller. Thus, this paper would be helpful to select the appropriate DTFC scheme for real-time

industrial IPMSM drives based on the requirements (e.g., lowest torque, reasonable torque ripple with high efficiency).

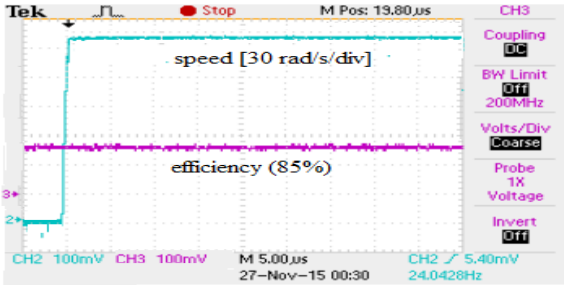


Fig. 7: Experimental speed and efficiency responses of the proposed DTFC at rated speed and no load.

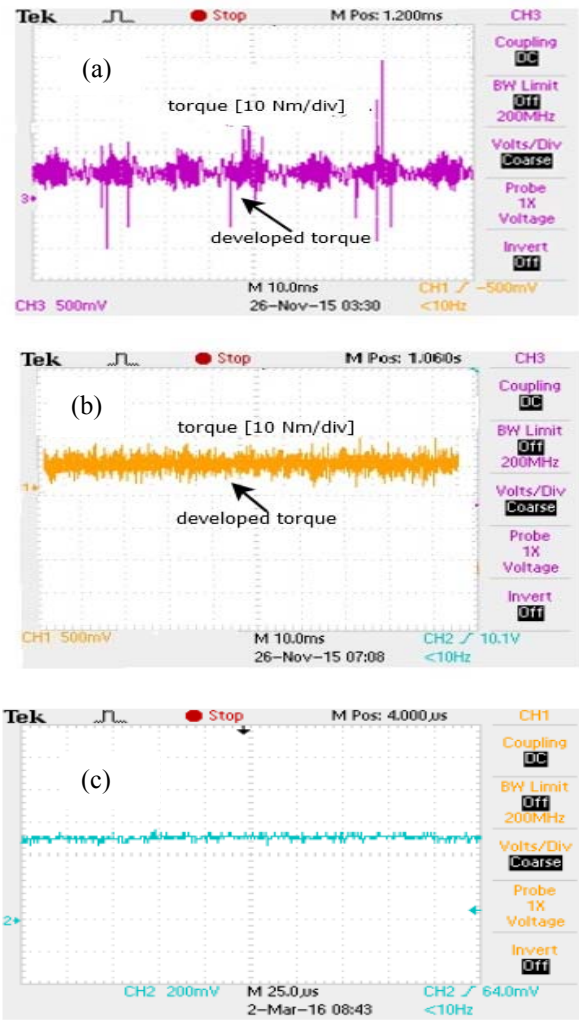


Fig. 8: Torque responses at rated speed and 50% rated load: a) conventional six-sector based DTFC with LMA; (b) eighteen-sector based DTFC with LMA; (c) nonlinear DTFC scheme.

APPENDIX – MOTOR PARAMETERS

Label	Parameter	Value with units
P_{rated}	IPMSM rated power	3.7 kW
V_{rated}	Rated rms voltage	183 v
I_{rated}	Rated current	14.2 A
f_{rated}	Rated frequency	87 Hz
ω_r	Rated rotor speed	183 rad/s
P	Pole pair numbers	3
L_d, L_q	d- and q- axis inductance	.00506, .00642 H
R_s	Stator resistance	0.242 ohm
J	Rotor inertia constant	0.0133 Kg.m^2
B_m	Damping coefficient	0.001 Nm/rad/s
ψ_m	Magnetic flux constant	.2449 v/rad/s

REFERENCES

[1] K. J. Tseng and S. B. Wee, ‘‘Analysis of flux distribution and core losses in interior permanent magnet motor,’’ IEEE Transactions on Energy Conversion, vol. 14, no. 4, pp. 969-975, December 1999.

[2] R. S. Rebeiro and M. N. Uddin, ‘‘Performance Analysis of an FLC-based online adaptation of both hysteresis and PI controllers for IPMSM drives,’’ IEEE Transactions on Industry Applications, vol. 48, no. 1, pp. 12-19, January/February 2012.

[3] M. N. Uddin and M. Hafeez, ‘‘FLC-based DTC scheme to improve the dynamic performance of an IM drive,’’ IEEE Transactions on Industry Applications, vol. 48, no. 2, pp. 823-831, March/April 2012.

[4] M. F. Rahman, L. Zhong and Khiang Wee Lim ‘‘A direct torque-controlled interior permanent magnet synchronous motor drive incorporating field weakening,’’ IEEE Transactions on Industry Applications, vol. 34, no. 6, pp. 1246-1253, November/December 1998.

[5] A. Lokriti, I. Salhi and S. Doubabi ‘‘IM direct torque control with no flux distortion and no static torque error, ISA Transaction, vol. 59, pp. 256-267, 2015.

[6] Kyo-Beum Lee, Joong-Ho Song, Ick Choy and Ji-Yoon Yoo ‘‘Torque ripple reduction in DTC of induction motor driven by three-level inverter with low switching frequency,’’ IEEE Transactions on Power Electronics, vol. 17, no. 2, pp. 255-264, March 2002.

[7] Lixin Tang, Limin Zhong, M. F. Rahman and Y. Hu, ‘‘A novel direct torque controlled interior permanent magnet synchronous machine drive with low ripple in flux and torque and fixed switching frequency’’ IEEE Transactions on Power Electronics, vol. 19, no. 2, pp. 346-354, March 2004.

[8] M. Nasir Uddin, H. B. Zou, and F. Azevedo, "Online Loss Minimization Based Adaptive Flux Observer for Direct Torque and Flux Control of PMSM Drive" IEEE Transactions on Industry Applications, vol. 52, no.1, pp.425-431, Jan/Feb 2016.

[9] L. Tang, L. Zhong, M. F. Rahman and Y. Hu, ‘‘A novel direct torque controlled interior permanent magnet synchronous machine drive with low ripple in flux and torque and fixed switching frequency’’ IEEE Transactions on Power Electronics, vol. 19, no.2, pp.346-354, 2004.

[10] Z. Ibrahim and E. Levi, ‘‘A comparative analysis of fuzzy logic and PI speed control in high performance AC drives using experimental approach’’ IEEE Transactions on Industry Applications, vol. 38, no.5, pp.1210-1218, September/October 2002.

[11] Shigeo Morimoto, Yoji Takeda and Takao Hirasa, ‘‘Loss minimization control of permanent magnet synchronous motor drives’’ IEEE Transactions on Industrial Electronics, vol. 41, no.5, pp.511-517, October 1994.

[12] Yong Liu, Z. Q. Zhu and D. Howe, “Direct torque control of brushless dc drives with reduced torque ripple,” IEEE Transactions on Industry Applications, vol. 41, no. 2, pp. 599–608, March/April 2005.

[13] H. F. Abdul Wahab and H. Sanusi, “Simulink model of direct torque control of induction machine,” American Journal of Applied Sciences, vol. 5, issue. 8, pp. 1083– 1090, 2008

[14] Tianfu Sun, Jiabin Wang and Xiao Chen ‘‘Maximum torque per ampere (MTPA) control for interior permanent magnet synchronous machine drives based on virtual signal injection,’’ IEEE Transactions on Power Electronics, vol. 30, no. 9, pp. 5036-5045, September 2015.

[15] Data sheets of SEMIKRON, SKHI 22 AR inverter manual.

[16] M. N. Uddin and J. Lau "Adaptive backstepping based nonlinear control of an IPMSM drive," IEEE 35th Power Electronics Specialists Conference (PESC), vol. 5, pp. 3451-3457, June 2004.

[17] M. A. Rahman, D. M. Vilathgamuwa, M. N. Uddin and King-Jet Tseng "Nonlinear control of interior permanent magnet synchronous motor," IEEE Transactions on Industry Applications, vol. 39, no. 2, pp. 408-416, March/April 2003.

[18] M. Nasir Uddin and Md. Muminul Islam Chy "Online parameter-estimation-based speed control of PM AC motor drive in flux-weakening region," IEEE Transactions on Industry Applications, vol. 44, no. 5, pp. 1486-1494, September/October 2008.

Multiscale–Multiphysics Modeling of Nonclassical Semiconductor Devices

Shaikh S. Ahmed, Mohammad Z. Rashid, Saad M. Al-Qahtani, Abdulmuin M. Abdullah, Md Rezaul Karim Nishat, Khadija A. Khair, Ye Wu, Mayada M. Taher, Sameer Al-Sibiani, and Abdussamad A. Muntahi

Department of Electrical and Computer Engineering, Southern Illinois University at Carbondale
1230 Lincoln Drive, Carbondale, IL 62901, USA.
ahmed@siu.edu

***Abstract*—This work describes our on-going efforts to develop a multiscale *Quantum Atomistic Device Simulator* (QuADS 3-D) where: a) material parameters are obtained atomistically using first-principles, b) structural relaxation and phonon dispersions are studied via molecular mechanics/dynamics, c) a variety of tight-binding models (*s*, sp^3s^*, $sp^3d^5s^*$) are used for the calculation of electronic bandstructure and interband transition rates, and d) coupled charge-phonon transport is simulated using a combined Monte Carlo-NEGF framework. The atom-by-atom simulation capability in QuADS 3-D exposes new degrees-of-freedom at nanoscale (such as engineering the stress, hybrid crystal cuts, composition, surface polarization, and electrostatics) and creates transformative design routes for boosting performance and reliability of novel nanoelectronic devices. Application of QuADS 3-D is demonstrated by *four examples*: 1) quantum and coulomb effects in nanoscale FETs; 2) correlation of structural modifications and reliability in AlGaN HEMTs; 3) effects of contact resistances in nanostructured thermoelectric coolers; and 4) efficiency *droop* in nanostructured III-N LEDs.**

***Index Terms*—QuADS 3-D, multiscale modeling, nanoscale devices, tight-binding, Monte Carlo, coupled transport.**

I. Introduction

As semiconductor devices shrink into the nanoscale regime and new classes of nanodevices emerge, performance and reliability of nanosystems engineered from these devices are fundamentally governed by an intricate interplay of *long-range* structural and built-in electrostatic fields (strain, polarization, applied and environmental biases), effects of *atomistic* entities (defects, disorder, surface relaxation) on the bandstructure (symmetry lowering, anisotropy, non-degeneracy), quantum *size-quantization*, and nonlinear, highly stochastic and dynamically-coupled *charge and phonon transport processes*. This suggests that the modeling of nanoscale devices must involve new theories and approaches on complexity of concurrent phenomena *at different length and time scales*. This type of research is highly interdisciplinary, covering the domains of chemistry, materials science, physics, and engineering.

On the computational side, for low-dimensionality devices, modeling approaches based on a *continuum* representation of the underlying material become invalid. On the other hand, a correct quantum description is well achieved by means of *ab initio* methods. A manifold of flavors and codes exist (such as ABINIT, SIESTA, PWscf) each of which is adapted to the specific needs of their users. Unfortunately, while offering intellectual appeal, these *ab initio* materials science methods can only model very small and *closed* systems having ~100 atoms *near or at equilibrium*. This fundamentally originates from the complexity of the quantum representation of electrons, possessing several internal degrees of freedom. The requirement of *self-consistency* further slows down such methodology. Note that, today's nanoscale devices exploit physical processes at a scale where the number of atoms in the *active region* is on the order of 10,000 to more than 100 million! While computational system sizes of millions of atoms appear at first sight huge and wasteful, it has been well demonstrated [1][2] that the underlying physical problems do require such large scale analysis. Therefore, modeling of nanodevices, to be adequately accurate yet computationally efficient, must employ a *multiscale and modular* approach (both hierarchical and handshake). With an objective to bridge the gap between atomistic and continuum modeling paradigms, for the past few years, our Group at SIU has been developing a multiscale *Quantum Atomistic Device Simulator* (QuADS 3-D) [3]. The key features of QuADS 3-D are described in Section II.

II. QuADS 3-D Simulator

QuADS 3-D enables quantum-corrected atomistic numerical modeling of non-equilibrium charge and phonon transport phenomena in realistically-sized systems containing more than 100 million atoms! QuADS 3-D is primarily being built upon three modules: i) Open source LAMMPS molecular dynamics code [4] for geometry construction and modeling structural relaxations. LAMMPS can model atomic, polymeric, biological, metallic, granular, and coarse-grained systems using a variety of force fields and boundary conditions. To improve accuracy, *ab initio* PWscf tool is used for parameterization of force and polarization coefficients and model bandstructure calculations; ii) *Extended version* of the open source NEMO 3-D tool for the calculation of excitonic and phonon spectra and optical transition rates [5][6]. NEMO 3-D enables the computation of electronic structure using a variety of *tight-binding models* (*s*, sp^3s^*, $sp^3d^5s^*$) that are optimized with a genetic algorithm tool. Whereas, for the calculation of *atomistic* (non-linear) strain relaxation, NEMO 3-D currently employs the atomistic valence-force field (VFF) with strain-dependent Keating potentials. NEMO 3-D includes spin in its fundamental atomistic tight binding representation; and iii) A quantum-corrected (benchmarked against the non-equilibrium Green function, NEGF, formalism) 3-D Monte Carlo electron-phonon transport kernel [7][8]. The QuADS simulation strategy, which is divided into different computational phases, spanning from the molecular structure of the constituting elements, to the electron and phonon band structure, transport and optical coupling, is depicted in Fig. 1. The Figure also shows the various length and time scales and the associated observables and how one passes between them, and the codes used along with their interdependencies. These core modules have successfully been benchmarked on several supercomputers. For example, recent

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

NEMO 3-D benchmarks show [9] that 3-D domain decomposition scheme can be utilized exceeding 32,000 cores on realistic electronic structures comprised of one billion atoms. On the other hand, in our recent efforts of modeling thermal conductivity, our group at SIU has successfully used LAMMPS on more than 120,000 cores in then Jaguar XT5, largest number of atoms simulated being ~3.8 billion!

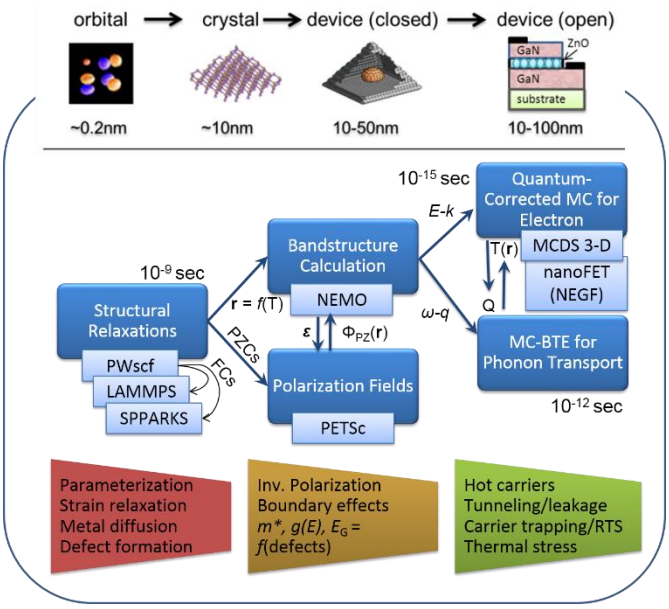


Fig. 1. QuADS 3-D simulation platform at SIU. Shown are the program flow (in bold arrows), the data interfaces, codes used, and relevant time scales. PWscf: *ab initio* module, FCs: force constants, PZCs: polarization coefficients, LAMMPS: massively parallel molecular dynamics code, SPPARKS: kinetic lattice MC (KLMC), NEMO: nanoelectronics modeling tool, MC: Monte Carlo kernels, NEGF: non-equilibrium Green function, **r**: atom positions, *E-k*: bandstructure, *ω-q*: phonon modes, Q: heat, Φ(**r**): potential.

Models and Algorithms—The tools used in QuADS 3-D exploit several recently developed novel, memory-miserly, and fast algorithms, are portable, and incorporate state-of-the-art fault-tolerant software design approaches. For example, the *eigen solvers* available in NEMO 3-D include the PARPACK library, a custom implementation of the Lanczos method, Block Lanczos method, the spectrum folding method and the Tracemin method. Regarding charge transport, while the NEGF approach is considered as the state-of-the-art, modeling realistically-extended devices within the dissipative regime (with scattering) using NEGF is prohibitively expensive. In the past, we have worked on the development of novel algorithms for the calculations of Green's functions and associated charge densities from large sparse matrices describing the underlying nanoscale systems. Among these algorithms, FIND (fast inverse using nested dissection), improved the computation speed significantly especially when the cross-section of the nanostructure becomes large [10]. As for the Monte Carlo transport kernel, in QuADS 3-D, quantum mechanical size-quantization effects have been accounted for via a *parameter-free effective potential* scheme [7]. The approach is based on a perturbation theory around thermodynamic equilibrium and derived from the idea that the semiclassical Boltzmann equation with the quantum corrected potential and the Wigner equation should possess the same steady state. The effective potential possesses no fitting parameters, as the size of the electron (wavepacket) is determined from its energy. The resultant quantum potential is, in general, two-degrees smoother than the original Coulomb and barrier potentials of the device, i.e. possesses two more classical derivatives, which essentially eliminates the problem of the statistical noise. Also, in this kernel, to treat full Coulomb (electron-ion and electron-electron) interactions properly, the simulator implements three real-space *electron dynamics* (ED) schemes: the particle-particle-particle-mesh (P^3M) method, the corrected Coulomb approach, and the fast multipole method (FMM) [11][12][13]. The effective force on an electron is computed as a combination of the short-range molecular dynamics force and the long-range Poisson force. For an *N*-particle system (with *M* grids), in contrast to the well-known P^3M algorithm with $O\left(N + M\log M\right)$ operation count, FMM's complexity is only ~*N*. Also, necessary *event-biasing algorithms* are used in the simulator that enhance the carrier statistics and result in a faster convergence of the channel current [14]. Finally, to properly treat the charge transport and the *heating effects* without any approximations, one in principle has to solve the *coupled* electron/hole−phonons–heat bath problem, where each sub-process has to be addressed in a somewhat individual manner and included in the global picture via a self-consistent loop [15][16][17]. The system is nonlinear, as the transition probabilities depend on the product of the electron and phonon distribution functions, and poses a multi-scale problem since the sub-processes involve different time scales: the velocity of the phonons is two orders of magnitude lower than the velocity of the electrons. In the coupled electron-phonon BTE formalism, as implemented in QuADS 3-D, phonons are treated, on an equal basis with electrons, as superparticles with group velocity dependent on the type and mode of phonons. The output of the thermal simulations is the temperature distribution, which is then used to calculate thermal conductivity and as an *input* to the charge transport kernel of the integrated simulator.

Below, in Section III, we delineate applications of QuADS 3-D in modeling *four* types of nanoscale devices: 1) quantum and coulomb effects in nanoscale FETs; 2) correlation of structural modifications and reliability in AlGaN HEMTs; 3) effects of contact resistances in nanostructured thermoelectric coolers; and 4) efficiency *droop* in nanostructured III-N LEDs.

III. Benchmark Simulations with QuADS 3-D

A. Quantum and Coulomb Effects in Nanoscale FETs

Quantum Effects in III-V FETs—Because of limited benefits of strain engineering in extremely scaled silicon devices and lack of demonstration of a performance gain at the product level with nanowires, nanotubes, graphene, and other exotic channel materials, there is a strong motivation to continue device scaling using high-transport III-V (such as InGaAs and InAsSb) channel materials beyond the year 2020. As for device *architecture*, the multi-gate FETs have been hailed as one of the biggest breakthroughs in transistor technology in the last 20 years. As compared to their single-gate counterparts, the multi-gate FETs offer faster switching, low power, improved short channel effects and improved linearity. However, there are several challenges with these novel FETs. In this Section, within a *dissipative* transport formalism, we investigate the performance of the multi-gate III-V FETs as compared to the planar single-gate counterpart, and show how *quantum size quantization* and *random dopant fluctuations* (RDF) affect their characteristics.

First, the device bandstructure parameters (bandgap, effective masses, and density of states) were computed within a 20-band nearest-neighbor $sp^3d^5s^*$ tight binding scheme. Fig. 2(a) shows the variation of Γ-valley effective mass as a function of

device diameter considering three crystallographic directions. Regarding the Monte Carlo transport kernel, the scattering mechanisms activated in the calculations include: acoustic intravalley, intervalley (both *g*- and *f*- phonon processes), inter-sub-band, piezoelectric, and surface roughness. As for size-quantization effects, the effective potential method was validated (for ballistic transport) using our fully quantum mechanical simulator, *nanoFET*, available on www.nanoHUB.org [18]. Fig. 2(b) shows the I_D-V_{DS} characteristics for two FETs having channel length of 18 nm and 9 nm, respectively. Noticeable is the large degradation in the drain current when size-quantization is included in the calculation. This can be attributed to the *very small* electron effective masses in $InAs_{0.7}Sb_{0.3}$. Also, one can see that the degradation is less in the 9 nm device. This is due to the fact that for a smaller channel length, electrons can gain higher kinetic energy, especially at the drain end where they behave more as *classical* particles.

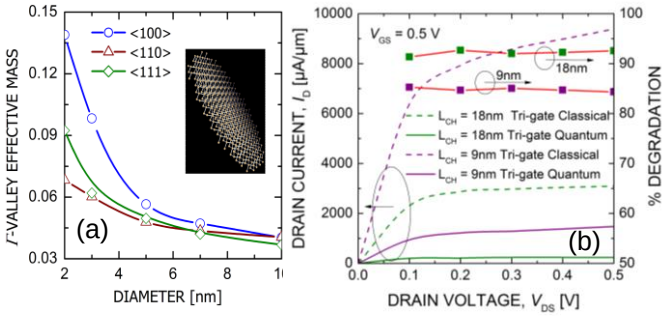


Fig. 2. (a) Effective mass of $InAs_{0.7}Sb_{0.3}$ nanowire plotted against diameter/base length. (b) I_D-V_{DS} characteristics and percentage current degradation.

In recent years, the study of discrete charge induced random telegraphic signal (RTS) in emerging devices has attracted much attention. RTS results from the capture and emission of charged particles from defect states/traps and leads to modulation in carrier density (electrostatics) and mobility (dynamics), dominates the low frequency noise (LFN) performance around and below threshold, and significantly changes the ON-state current in nanoscale FETs. Experiments on nanowire FET having gate width of ~100 nm and length of ~200 nm show that the on current fluctuation can be as high as 25% [19]. Here, first, we study the effects of random dopant fluctuations (RDF) on the drive capability of single- and tri- gate FETs. Fig. 3(a) shows the fluctuations of drain current for 25 samples from each category. Two features stem out from this Figure: a) The trigate offers a higher current (faster gate speed); and b) RDF is weaker in trigate, compared to the single gate device. Finally, the effect of the presence of a *single negative charge* located in various spots within the channel is studied. As can be seen from Fig. 3(b), of all the positions considered in the calculation,

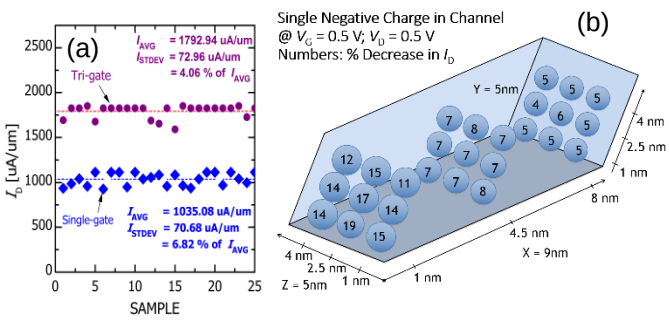


Fig. 3. (a) RDF in single and trigate $InAs_{0.7}Sb_{0.3}$ FETs. (b) Effects of a single localized negative charge on drain current degradation.

the impurity when placed at the source end of the device affects (degrades) the drain current most. This comes as a direct consequence of the reduction in the carrier *injection* velocity, which never gets recovered as the carriers move along the channel region.

Monolayer MoS_2—2D MoS_2 with a finite bandgap looks promising as channel material for FETs. Recently, MoS_2 transistors operating at gigahertz frequencies (with a cut-off frequency of ~6 GHz) have been reported [20]. In monolayer MoS_2, the *K*-point direct bandgap is ~1.8 eV, the electron effective mass at the K point, $m_e^* = 0.48m_0$ and is isotropic in nature. Also, the two degenerate valleys at the *K* point are almost parabolic [21]. First principles based calculations [22] have demonstrated that, for infinite layer separation, MoS_2 has a static dielectric constant of 1 whereas the previously reported values were in the range of 4.2−7.6. Nevertheless, electron mobility in monolayer MoS_2 is degraded by various scattering mechanisms. In this Section, using the recently reported value for the dielectric constant, we calculate the scattering rates (and transport parameters) in monolayer MoS_2 due to different phonon and Coulomb processes. The simulated MoS_2 FET has a gate length of 14 nm, source-drain length of 20 nm and a width of 10 nm. The channel thickness is ~0.65 nm (thickness of the MoS_2 layer) and the channel is undoped. We have used HfO_2 as top gate oxide (Top OX) with an equivalent oxide thickness (EOT) of ~2 nm. The MoS_2 channel sits on a SiO_2 buried oxide (BOX) layer. The following scattering mechanisms have been included in the Monte Carlo transport kernel: acoustic TA, acoustic LA, optical 0th, optical 1st, polar optical, 0th and 1st order surface/remote optical in top gate oxide and buried oxide layers, and remote Coulomb in top gate oxide and buried oxide layers. A pen picture of current degradation due to various scattering processes is depicted in Fig. 4. It was found that the total scattering rate is strongly dominated by *remote coulomb scattering*, which, when compared to the ballistic regime, degrades the drain current by ~46%.

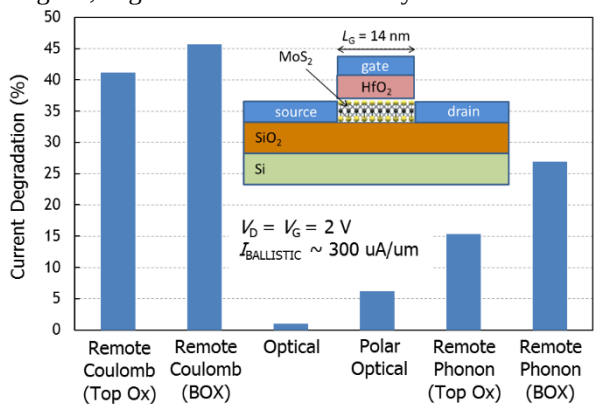


Fig. 4. Drain current degradation (with respect to ballistic value) due to various scattering processes in an MoS_2 based MOSFET (Inset: the device).

B. *Structural Modification and Reliability of AlGaN HEMTs*

AlGaN based nanoscale high-electron-mobility transistors (HEMTs) are the next generation of transistor technology. However, as evidenced by recent stress tests, *reliability* (a gradual decrease in the output current) of these devices remains a major concern [23][24][25][26]. Although, in these tests, physical damages were clearly visible in the device, the root cause, nature, and impact of these damages have not yet been fully assessed experimentally. This makes the extrapolated MTTF less meaningful and poses a significant challenge to predicting device lifetime, a task which is critical for devices

used in space applications. In this Section, we first develop a conceptual framework that enables fine-tuning of some of the existing models on hot-carrier stressing in AlGaN/GaN HEMTs. Next, *atomistic* simulations were carried out to study and quantify the distributions of long-range internal fields, generation of hot electrons in the active region, effects of localization/delocalization of trapped electrons in the gate dielectric layer, and the metal diffusion phenomenon.

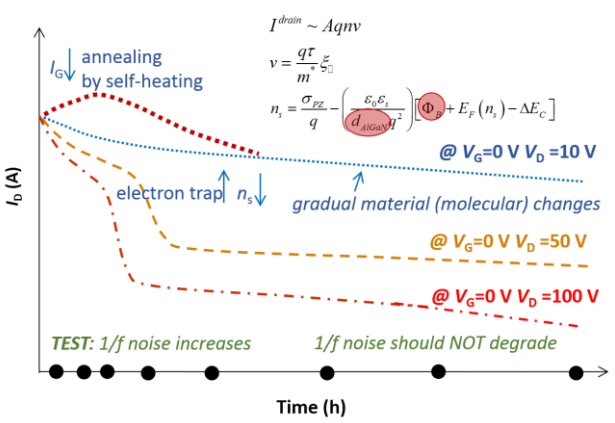


Fig. 5. *Schematic* of the evolution of I_D in a hot-electron stress test.

To illustrate, refer to the schematic (Fig. 5) of a typical evolution of drain current, I_D, during a *hot electron* ON-state stress test. Our first understanding is that, as the hot electrons, in an increasing number, get trapped at the defect states. The gradual drop in the drain current is mainly due to the reduction in the channel charge density, n_s, as well as pronounced Coulomb scattering. Note that, as shown by the red dotted curve, and as observed in some stress experiments, the drain current could rise in the very beginning. This can be attributed to the reduction in the gate leakage current. Also, for the case where gate current is due to the electrons coming from the gate end, a rise in the drain current could be due to self-heating induced annealing of traps around the gate-drain access region. At even higher drain voltages, the hot electrons could be *distributed across the entire channel region*, resulting in a *steeper* reduction in the drain current.

Nevertheless, with an increase of stress time, the drain current continues to decrease (gradually), even though the gate current remains more or less constant (not shown). *Why does this happen?* To understand this, one needs to consider modifications in the molecular/atomic structure of the gate stack and look at the equation of the sheet electron density, n_s. Here, it is clear that any reduction in the effective thickness of the gate dielectric, d_{AlGaN} , or an increase in the barrier height, Φ_B , due to molecular changes (changes in alloy composition) of the gate contact material (for example, Ni/Pt/Au) can result in a gradual decrease in the drain current. To validate these claims, we simulated an AlGaN HEMT [25] consisting, from bottom to top, of 1.5 µm of GaN, followed by 1 nm of $In_{0.1}Ga_{0.9}N$ capped by 11 nm of GaN and 13 nm of recessed $Al_{0.32}Ga_{0.64}N$. The source-gate and gate-drain separations are both 0.75 µm, while the gate length is 0.1 µm. The channel electron density in an AlGaN HEMT is primarily created via the presence of a strong internal field. Here, first, the thermal conductivity of the AlGaN/GaN system was computed using the LAMMPS software. This was then used to obtain the temperature distribution (self-heating) in the active region, which is then again used in LAMMPS to study the structural relaxation and an effective thinning of the barrier layer (Fig. 6a). Using NEMO 3-D and with relaxed atomic coordinates, we then calculated and plotted the charge density as a function of the thickness of the AlGaN barrier layer (Fig. 6b). The reduction in the induced charged density for thinner barrier layers, when integrated into the transport kernel, resulted in an increase in the threshold voltage and thereby a drop in the drain current (Fig. 6c).

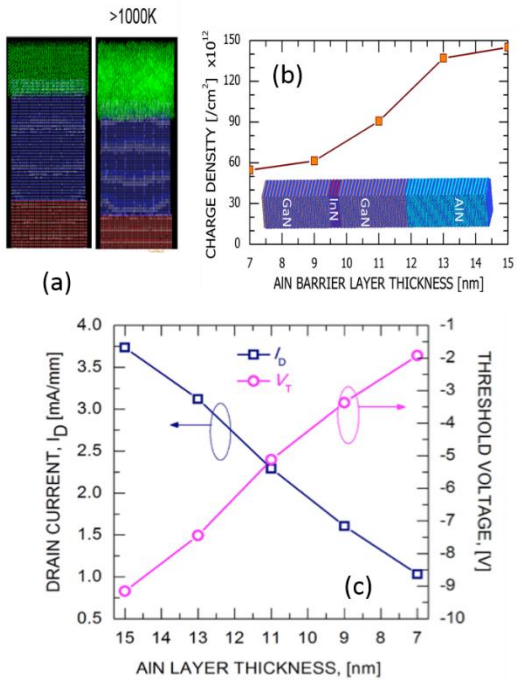


Fig. 6. (a) Atomic diffusion in AlGaN; (b) Charge density and (c) I_D and V_{TH} variation with a reduction in d_{AlGaN}.

C. *Effects of Contact Resistances in Thermoelectric Coolers*

Recently, integration of Bi_2Te_3 superlattice based thin-film thermoelectric coolers (TECs) into state-of-the-art electronic packages has been experimentally demonstrated and active cooling of ~15 ºC at the *hot spot* on a silicon chip with a high heat flux (~1,300 W/cm^2) has been reported [27]. In this Section, we go beyond the thin-film paradigm and numerically study the performance of a Bi_2Te_3 *nanowire* based TEC for hot-spot cooling. The nanowire in configured in the *z*-direction and the effective mass of the wire was calculated using a 20-band spin-corrected $sp^3d^5s^*$ tight-binding scheme. Next, using LAMMPS, the lattice thermal conductivity of Bi_2Te_3 nanowires was calculated as a function of diameter. Finally, the *ZT* of Bi_2Te_3 nanowires was obtained, as described in Ref. [28], and shown in Fig. 7(a).

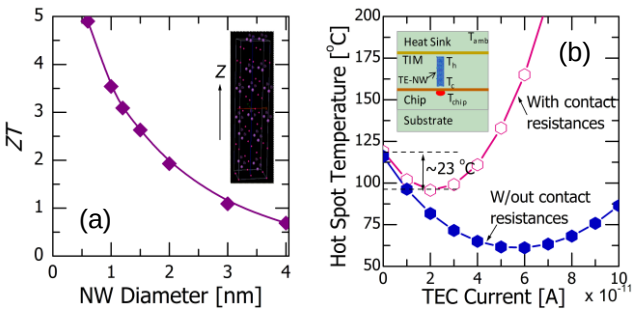


Fig. 7. (a) *ZT* of Bi_2Te_3 nanowire as a function of nanowire width (inset shows the nanowire in the *z* direction); (b) Temperature of the hot spot plotted against the TEC current. NW diameter is 0.6 nm. The reference values of electrical and thermal contact resistances at the interface of super lattice–metal layer (10^{-11} Ω-m^2 and 1×10^{-6} m^{2}K/W, respectively) and at the interface of TEC module-heat spreader layer (10^{-10} Ω-m^2 and 8×10^{-6} m^{2}K/W, respectively) are obtained from Ref. [27]

For modeling a *complete* on-chip nanowire-based embedded thermoelectric cooler (eTEC) unit, as shown in the inset of Fig. 7(b), a thermal circuit model has been used with heat current and thermal resistances. This Figure shows the variation of hot-spot temperature (such as in a microprocessor) as a function of the drive current of the TEC for two different cases: *without* and *with* the contact electrical and thermal resistances. The active cooling for these two cases is 56 ºC and 23 ºC, respectively [29]. The results suggest that, while TECs could offer significant active cooling, to be efficient, appropriate measures must be taken to reduce/minimize the effects of thermal and electrical contact resistances before these devices can be deployed in HPC/data-center cooling applications.

D. Efficiency Droop in Nanostructured III-N LEDs

Nano-engineered InGaN heterostructures are promising candidates for efficient and full-spectrum optical emitters for use in conventional (e.g. general illumination, lasers) as well as emerging (e.g. near-field photolithography, quantum cryptography, diagnostic medicine and biological imaging) applications. To illustrate, according to the U.S. Department of Energy (DOE), it is estimated that, by 2035, switching to solid-state lighting (SSL) could reduce national lighting electricity consumption by nearly 75%, the equivalent to 5.1 quadrillion BTUs, as well as spare the atmosphere 180 million metric tons of greenhouse gas emissions. SSL devices could, theoretically, be over 90% efficient, vibration and shock resistant, have exceptionally long-life (over 50,000 hours), and offer a wide-range of tunability of the emission characteristics.

Droop in SSL devices is the loss of efficiency at high operating current density causing them to burn more energy per lumen. This is the *most pressing* problem in solid-state lighting technology today. Here, using QuADS 3-D, we first investigate how *droop* and degradation in SSL devices are governed by an intricate interplay of crystal atomicity, built-in structural fields, and charge and phonon transport processes at different length and time scales, and then explore new design space for boosting SSL efficiency.

InN/GaN nanostructures exhibit atomically inhomogeneous and *long-range* internal electromechanical fields originating mainly from: a) the fundamental atomicity and interface discontinuity, b) epitaxial strain, c) piezoelectricity, and d) spontaneous (pyroelectric) polarization. The simulated GaN/InGaN/GaN disk-in-wire LED structures have height and base length of 100 nm and 14 nm, respectively, and contain ~10M atoms. The undoped $In_{0.25}Ga_{0.75}N$ quantum disk is positioned at the center of the GaN nanowire and has a thickness of ~3 nm. The distribution of 3-D atomistic strain fields is shown in Fig. 8a. The induced piezoelectric (PZ) and the pyroelectric (PY) potential distributions along the growth direction and through the center of the disk are plotted in Fig. 8b using different hierarchical models as parameters. Compared to the pyroelectric potential, the piezoelectric counterpart is significantly large. The side panel shows the 2-D potential distribution projected in the *x*-*z* plane featuring long-range and a dipole characteristic. The internal fields affect the electronic and optical properties of the devices significantly. Fig. 8c shows the polar plots of interband optical transition rates between the HOMO (ground hole) and the LUMO (ground electron) states in the *x*-*y* plane in the InGaN quantum disk revealing pronounced *anisotropy* and suppression of optical transition. The energy of the optical transition (bandgap) may lie well below the *bulk* bandgap (so called quantum confined Stark effect) and can exhibit strong variation as a function of disk composition (indium mole fraction, Fig. 8d). Next, to improve the optical characteristics, a semi-polar *m*-plane InGaN LED [30][31] is simulated and the spontaneous emission rates

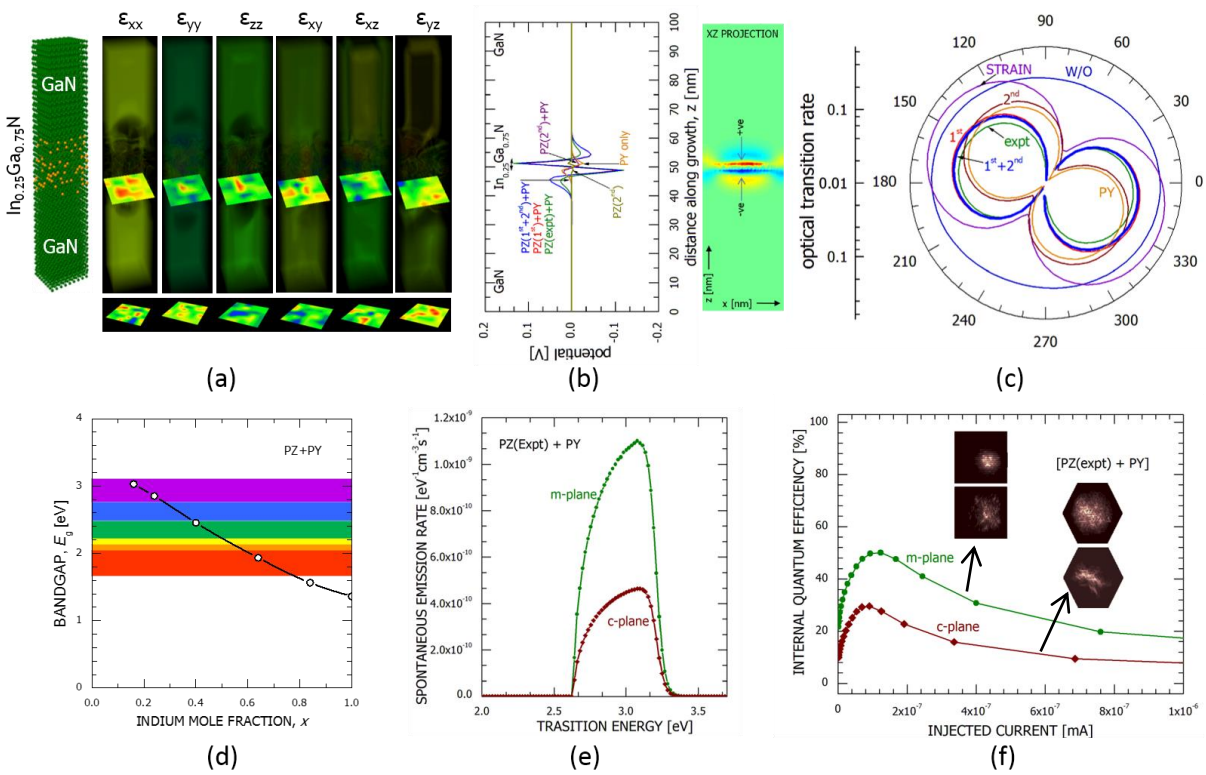


Fig. 8. (a) Atomistic strain distributions. (b) Internal polarization fields. (c) Optical transition rates. (d) Variation of bandgap with mole fraction. (e) Spontaneous emission rates, and (f) IQE for *c*- and *m*- plane devices; inset shows HOMO (bottom) and LUMO (top) wavefunctions.

for these two structures are plotted in Fig. 8e. Finally, the microscopically determined parameters are incorporated in the off-the-platform Sentaurus TCAD toolkit and used to compare the internal quantum efficiency (IQE) for *m*-plane and *c*-plane disk-in-wire LEDs (Fig. 8f). It is found that for a disk thickness of ~3 nm the average interband optical transition rates is higher in the *m*-plane quantum disk which results in improved internal quantum efficiency.

IV. Conclusions

QuADS 3-D allows the calculation of electronic states, optical response, and transport in various semiconductor devices that use nanostructured active region including quantum dots, quantum wires, quantum wells and nanocrystals. Using QuADS 3-D, nanoelectronic device designers will be able to address many challenging issues including crystal atomicity, defects, interfaces and surfaces, strain relaxation, piezoelectric and pyroelectric polarization, quantum confinement, highly-interacting and dissipative current and phonon paths, and performance in harsh environments—all on an equal footing. With the multi-million atom handling capability, the simulator creates new engineering routes for optimizing the efficiency and reliability of nanoelectronic and optoelectronic devices that were previously infeasible. A web-based online interactive version of the simulator for educational use will also be available on nanoHUB.

Acknowledgment

This work was supported by the US National Science Foundation Grant Nos. ECCS-1102192 and CCF-1218839. Part of the computational work was performed on the XSEDE and the nanoHUB HPC platforms. Access to Synopsys TCAD toolset via the SIU site is also acknowledged.

References

[1] Shaikh Ahmed, Sharnali Islam, and Shareef Mohammed, "Electronic Structure of InN/GaN Quantum Dots: Multimillion Atom Tight-Binding Simulations," *IEEE Transactions on Electron Devices*, vol. 57, issue 1, pp. 164–173, January 2010.

[2] Muhammad Usman, Yui H. Matthias Tan, Hoon Ryu, Shaikh Ahmed, and Gerhard Klimeck, "Importance of Piezoelectric Effect in Matching an Experimental Excitonic Spectrum of a Single Quantum Dot Molecule", *Nanotechnology*, vol. 22, 315709, 2011.

[3] Shaikh Ahmed, Krishnakumari Yalavarthi, Vamsi Gaddipati, Abdussamad Muntahi, Sasi Sundaresan, Shareef Mohammed, Sharnali Islam, Ramya Hindupur, Dylan John, And Joshua Ogden, "Quantum Atomistic Simulations of Nanoelectronic Devices using QuADS," in *Nano-Electronic Devices: Semiclassical and Quantum Transport Modeling*, book edited by D. Vasileska and S. M. Goodnick, Springer, pp. 405–441, 2011.

[4] Freely available at http://lammps.sandia.gov/

[5] Shaikh Ahmed, Neerav Kharche, Rajib Rahman, Muhammad Usman, Sunhee Lee, Hoon Ryu, Hansang Bae, Steve Clark, Benjamin Haley, Maxim Naumov, Faisal Saied, Marek Korkusinski, Rick Kennel, Michael Mclennan, Timothy B. Boykin, and Gerhard Klimeck, "Multimillion Atom Simulations with NEMO 3-D", In Meyers, Robert (Ed.) *Encyclopedia of Complexity and Systems Science*, vol. 6, pp 5745–5783. Springer New York 2009.

[6] Gerhard Klimeck, Shaikh Ahmed, Neerav Kharche, Hansang Bae, Steve Clark, Benjamin Haley, Sunhee Lee, Maxim Naumov, Hoon Ryu, Faisal Saied, Marta Prada, Marek Korkusinski, and Timothy B. Boykin, "Atomistic Simulation of Realistically Sized Nanodevices Using NEMO 3-D", *IEEE Trans. Electron Devices*, Vol. 54, 9, pp. 2079–99, 2007.

[7] S. Ahmed, C. Ringhofer, D. Vasileska, "Parameter-Free Effective Potential Method for Use in Particle-Based Device Simulations," *IEEE Trans. Nanotechnology*, vol. 4, pp. 465–471, July 2005.

[8] D. Vasileska and S. Ahmed, "Narrow-Width SOI Devices: The Role of Quantum Mechanical Size Quantization Effect and the Unintentional Doping on the Device Operation," *IEEE Trans. Electron Devices*, vol. 52, pp. 227–236, 2005.

[9] Benjamin P. Haley, Sunhee Lee, Mathieu Luisier, Hoon Ryu, Faisal Saied, Steve Clark, Hansang Bae, and Gerhard Klimeck, "Advancing nanoelectronic device modeling through peta-scale computing and deployment on nanoHUB," *Journal of Physics: Conference Series*, vol. 180, 012075, 2009.
Also, http://cobweb.ecn.purdue.edu/~gekco/nemo3D/index.html

[10] Song Li, Shaikh Ahmed, Eric Darve, and Gerhard Klimeck, "Compute the Diagonal of Sparse Matrix Inverse using FIND Algorithm," *Journal of Computational Physics*, 227, pp. 9408–9427, 2008.

[11] C. J. Wordelman, U. Ravaioli, "Integration of a particle-particle-particle-mesh algorithm with the ensemble Monte Carlo method for the simulation of ultra-small semiconductor devices," *IEEE Trans. Electron Devices*, 47, pp. 410–416, 2000.

[12] L. Greengard and V. Rokhlin, "A fast algorithm for particle simulations," *J. Comput. Phys*., 135, 2, pp. 280–292, 1997.

[13] V. Rokhlin, "Rapid solution of integral equations of classical potential theory," *J. Comp. Phys.*, 60, pp. 187–207, 1985.

[14] M. Nedjalkov, S. Ahmed, and D. Vasileska, "A self-consistent event biasing scheme for statistical enhancement," *Journal of Computational Electronics*, 3, pp. 305–309, 2004.

[15] D. Vasileska, K. Raleva, and S. M. Goodnick, "Electrothermal Studies of FD SOI Devices that Utilize a New Theoretical Model for the Temperature and Thickness Dependence of the Thermal Conductivity," *IEEE Trans. Electron Devices*, **57**, 3, 726-728, 2010.

[16] C. Jacoboni and L. Reggiani, 'The Monte Carlo Method for the Solution of Charge Transport in Semiconductors with Applications to Covalent Materials," *Rev. Modern Phys*., vol. 55, pp. 645–705, 1983.

[17] M. Fischetti, and S. Laux, "Monte Carlo study of electron transport in silicon inversion layers," *Phys. Rev. B*, vol. 48, pp. 2244–2274, 1993.

[18] URL: https://nanohub.org/tools/nanofet/ (accessed on 12/04/2016)

[19] C. Wei, Y. Jiang, Y. Z. Xiong, X. Zhou, N. Singh, S. C. Rustagi, G. Q. Lo, and D. Lee Kwong, "Impact of Gate Electrodes on 1/f Noise of Gate-All-Around Silicon Nanowire Transistors," *IEEE Electron Device Letters*, vol. 30, 10, pp. 1081–1083, October 2009.

[20] Daria Krasnozhon, Dominik Lembke, Clemens Nyffeler, Yusuf Leblebici, and Andras Kis., "MoS_2 Transistors Operating at Gigahertz Frequencies," *Nano Lett*., 14, p. 5905−5911, 2014.

[21] E. S. Kadantsev and P. Hawrylak, "Electronic structure of a single MoS_2 monolayer," *Solid State Comm*, 152, pp. 909−913, 2012.

[22] F. Hüser, T. Olsen, and K. S. Thygesen., "How dielectric screening in two-dimensional crystals affects the convergence of excited-state calculations: Monolayer MoS_2," *Phys. Rev. B*, 88, 245309, 2013.

[23] Jungwoo Joh, Jesús A. del Alamo, Kurt Langworthy, Sujing Xie, Tsvetanka Zheleva, "Role of stress voltage on structural degradation of GaN high-electron-mobility transistors," *Microelectronics Reliability*, 51, pp. 201–206, 2011.

[24] D. Marcon, G. Meneghesso, T.-L. Wu, S. Stoffels, M. Meneghini, E. Zanoni, and S. Decoutere, "Reliability Analysis of Permanent Degradations on AlGaN/GaN HEMTs," *IEEE Trans. Electron Devices*, 60, 3132, 2013.

[25] Fabio Alessio Marino, Nicolas Faralli, Tomás Palacios, David K. Ferry, Stephen M. Goodnick, and Marco Saraniti, "Effects of Threading Dislocations on AlGaN/GaN High-Electron Mobility Transistors," *IEEE Trans. Electron Devices*, 57, 1, pp. 353–360, Jan 2010.

[26] S. Choi, G. Woodruff, E. Heller, D. Dorsey, R. Vetury, and S. Graham, "The Impact of Bias Conditions on Self-Heating in AlGaN/GaN HEMTs," *IEEE Trans. Electron Devices*, 60, 1, pp. 159–62, 2012.

[27] I. Chowdhury, R. Prasher, K. Lofgreen, G. Chrysler, S. Narasimhan, R. Mahajan, D. Koester, R. Alley and R. Venkatasubramanian, "On-chip cooling by superlattice-based thin-film thermoelectrics," *Nature Nanotechnology*, vol. 4, pp. 235–238, 2009.

[28] L. D. Hicks, M. S. Dresselhaus, "Effect of quantum-well structures on the thermoelectric figure of merit," *Phys. Rev. B*, vol. 47, 12727, 1993.

[29] Afsana Sharmin, et al., "Multiscale Design of Nanostructured Thermoelectric Coolers: Effects of Contact Resistances," *Journal of Electronic Materials*, vol. 44, no. 6, pp. 1697–1703, 2015.

[30] Md. R. Nishat, S. Al-Qahtani, Y. Wu, V. Chimalgi, and S. Ahmed, "GaN/InGaN/GaN Disk-in-Wire Light Emitters: Polar *vs.* Nonpolar Orientations," *in proceedings of 18th International Workshop on Computational Electronics (IWCE) 2015*, West Lafayette, IN, USA, pp. 1-2, 2-4 Sept. 2015.

[31] Vinay Chimalgi, Md. R. K. Nishat, and Shaikh Ahmed, "Nonlinear Piezoelectricity and Efficiency *Droop* in Hexagonal In(Ga)N/GaN Disk-in-Wire LEDs," *Superlatt. Microstructures*, vol. 84, pp. 91–98, 2015.

Freeform Electronics for Advanced Healthcare

Muhammad M. Hussain,[1,*] Aftab M. Hussain,[1] Joanna M. Nassar,[1] Arwa T. Kutbee,[1] Abdurrahman Gumus,[1] and Amir N. Hanna[1]

[1]Integrated Nanotechnology Lab and Integrated Disruptive Electronic Applications (IDEA) Lab

Electrical Engineering, Computer Electrical Mathematical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST)
Thuwal 23955-6900, Saudi Arabia

[*]muhammad.hussain@kaust.edu.sa

Abstract—**Freeform (physically flexible, stretchable and reconfigurable) electronics can be critical enabler for advanced personalized healthcare. With increased global population and extended average lifetime of mankind, it is more important than ever to integrate advanced electronics into our daily life for advanced personalized healthcare. In this paper, we discuss some critical criteria to design such electronics with enabling applications.**

***Index Terms*—Freeform, electronics, healthcare.**

I. Introduction

We live in a complex world with rising global population. While human resource can be a great asset but we first have to ensure they are healthy and happy. Interestingly rising global population imposes a challenge on further distribution of limited resources to ensure such healthy and happy life. Specially with more aging population with extended lifetime, we must have to ensure everyone leads a healthy and happy life. In that regard, electronics may potentially be used to facilitate such prosperity. It is true that today's electronics are sufficiently powerful in context of data management (processing, storage, communication, display). However, they are often rigid, bulky, brittle and uniformly structured. Interestingly, human bodies have more irregularity in their shapes, asymmetry on their skins and inherent softness. Therefore, a critical barrier for seamless integration of today's advanced electronics with human body lies in their interfaces. One potential solution is to use soft curvilinear electronics which can be easily conformed to human body with suitable mechanical dexterity. From that perspective, in the last two decades, many interesting exciting materials have been emerged fuelled by innovation in chemistry and material science disciplines [1]. While organic and molecular electronics are spearheading and have penetrated in the high resolution display market, their inherent limited electron transport characteristics, thermal instability and ultra-large-scale-integration (ULSI) incompatibility, impede their future and wide ranging success. One dimensional material such as carbon nanotube (CNT) and various nanowires overcome the electron transport and thermal instability issue with great prowess but largely fail due to the ULSI incompatibility. While two dimensional graphene or other dichalcogenide atomic crystal structure materials are promising, still their manufacturability is a valid and sustained concern for the integrated device manufacturers (IDMs). Therefore, with due credits to all these emerging materials, traditional crystalline materials like silicon, silicon germanium, gallium nitride, III-V materials based electronics can still be the backbone of freeform electronics as long as we can make them physically flexible, stretchable and reconfigurable. Therefore, in this paper we will describe some important criteria for such electronic system design focusing on advanced healthcare for both *in-vivo* and *in-vitro* applications. We will focus on bulk monocrystalline silicon (100) which is used to fabricate ninety percent of the electronics today.

II. Design Criteria for Freeform Electronics Focusing on Advanced Healthcare Applications

First and the foremost challenge is the choice of biocompatible materials based electronics whose integration density and performance plus energy efficiency is strictly directed by thermal compatibility of human body. It is well known that human body (both *in-vivo* and *in-vitro*) should not experience more than 2°C/cm^2 heat increment. This dictates that the associated power consumption from body integrated electronics is restricted to 40 mW/cm^2. Now, take an example of brain mapping. According to the World Health Organization (WHO), twenty five percent of world's population suffer from some sort of neurological disorders. Examples can be from normally ignored depression to Alzheimer's disease, Parkinson's disease, etc. Therefore, it is important to study (map) our brain activities. Today's gold standard is micro-needle based Utah Array which is the most effective method as of today but after a yearlong constant use, it inflames the penetration area and needs immediate removal. Thus makes it an invasive system. Non-invasive systems such as Magnetic Resonance Imaging (MRI) can be fantastic, however, it's humongous size and weight with kilowatts range power consumption makes it unsuitable for constant use. Also, when an episode begins for a patient, unless we have a convenient wearable system, it becomes undeniably difficult to monitor brain activity. Now, electroencephalogram (EEG) can be used as non-invasive wearable system, however data acquisition through such system comes with very weak signal-to-noise ratio and thus making the overall reliable data crunch nearly impossible. Hence, if we can make a soft patch which will be placed on brain's soft cells and tissues in the intracranial space, then it can possibly send us invaluable information. Now, the challenge begins as human brains have 82 billion neurons. It is impossible that every single of them can be tracked in foreseeable future. At the same time, brain's soft exposed areas for exploration is limited to mere 8 × 8 cm^2 area. Therefore, it is important for us to use every bit of that area with soft sensors. However, sensors collecting analog signals need associated minimal electronics to convert the data into digital signals and then transmitting that to out of scalp. Hence, if we follow traditional integration schemes of lateral in-plane monolithic integration, it is impossible to maximize number of sensors within a limited space. Therefore, recently we have demonstrated a complementary metal oxide semiconductor (CMOS) based integration strategy which allows soft materials like polymer to be used as interposing base layer and

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

fabricating soft biocompatible sensors on one side and relevant data management electronics on the other side (Fig. 1) [2]. An effective analogy will be two sides of the same coin. However, it is easy to say than done. Although it sounds like 3D-IC design, in reality having a soft medium makes the integration sequence highly complex. Still, using electrochemical deposition we overcame the challenge with through-polymer-via (TPV) formation. These highly conductive copper TPVs ensure seamless flow of electrical signals from the sensors on one side to processing electronics on the other. We utilized state-of-the-art electronics fabrication processes to ensure compatibility with high-end electronics. Further, the scalable fabrication process ensures the successive miniaturization of sensor nodes with generation of the device, thus providing higher spatial resolution. We enhanced the integration scheme by integrating multiple layers of electronics on the sensor array to obtain a completely integrated area-efficient system for Electrocorticography (ECoG). Hence, we demonstrated the integration of sensor array within a soft material system (polyimide), with state-of-the-art silicon CMOS circuitry. Further, the integration scheme can be used to integrate layers different materials together making heterogeneous 3D integration possible. The resulting system is free-form (flexible and stretchable) if the component layers are free-form. With this integration strategy, going forward we will be able to design body integrated electronics (both *in-vivo* and *in-vitro*) in area efficient manner.

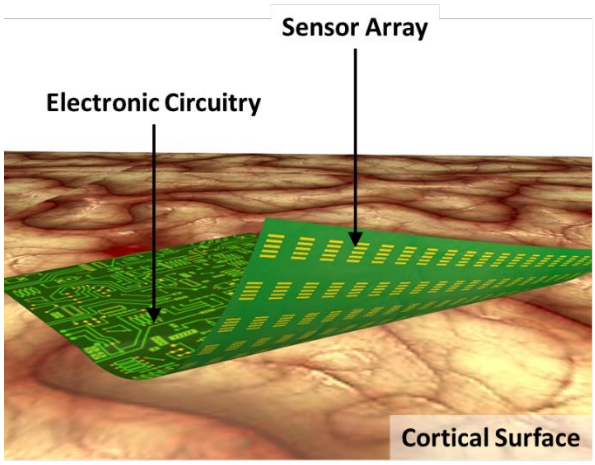


Fig. 1. Illustration of a flexible, 3D integrated brain machine interface.

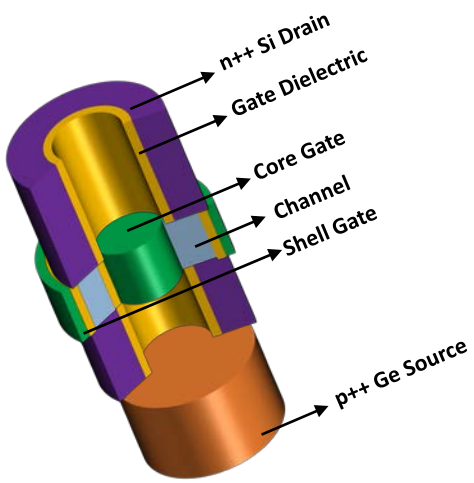


Fig. 2. Illustration of a hetero-structure nanotube TFET with core-shell gates.

Next challenge is thermal management. State-of-the-art planar and non-planar transistors have an established trend of area scaling resulting into higher power dissipation induced hotter electronics and systems. To address this critical issue, we have designed a nanotube architecture based heterogeneous tunnel field effect transistor (TFET) with core-shell gate stacks (Fig. 2) [3, 4]. Such device offers the highest performance (drive current) among all the reported TFETs with pinned sub-threshold slope lesser than 60 mV/dec. Such device overcomes issues of paradoxical relationship of limited performance and lower power dissipation in traditional devices. By using hetero-structure materials system we have taken advantage of higher injection from source side and thus higher drain current.

After integration strategy and thermal management, important consideration is the operating environment which is bio-fluidic or skin-mounted. In both cases, conventional radio frequency (RF) antenna based communication medium fails due to unintended frequency dispersion and attenuation. Therefore, from antenna design to their integration and choice of material has to sustain within bio-fluidic environment which is considered as lossy medium. For skin-mounted wearables, it if affected by anomalous physical deformation related to our daily activities such as flexing, stretching, twisting, etc. Therefore, not only we need to develop ultra-thin flexible antenna, but also a stretchable one. In the past several in-plane stretchable antennas have been demonstrated but due to stretching their length fundamentally changes which compromises their fixed frequency operation. Therefore, we adopted a design of helical architecture which stretches out-of-plane and its core body remains detached from the skin, lowering lossy skin medium's interference (Fig. 3) [5]. Additionally, using copper (which is biocompatible) as based metal for the antenna, which is not physically stretchable, we overcame the issue of permanent physical elongation based frequency dispersion and attenuation. However, due to fractal design, copper based helical architecture antenna allowed to have sufficient stretchability to withstand daily activities of human body. Using such novel but balanced integration strategy, we demonstrated wearable antennas which is operable even up to 390 meters at fixed frequency of 2.4 GHz.

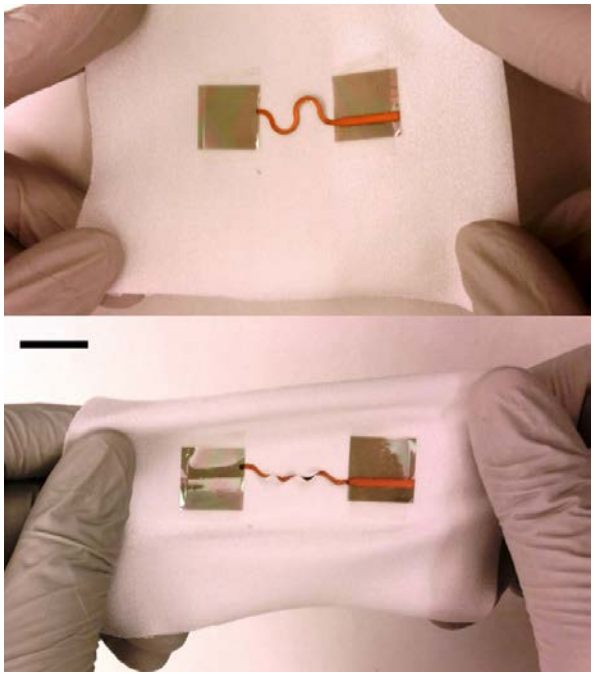


Fig. 3. Optical images of fabricated antenna attached to a stretchable fabric, before and after 30% straining. Scale bar is 2 cm.

Next important consideration is power management. While it is important to lower power consumption of any electronics in a world of growing number of electronics, it is equally important to have biocompatible energy harvesting and storage system. In that regard, we have unusual energy resource like bio-fluids such as sweat and saliva. We redesigned microbial fuel cells (MFCs) using CMOS compatible processes and achieved nW to mW level power generation. Now, what is MFC? Although nearly eighty percent of the earth has water, only a limited percentile is potable. Reverse osmosis is the leading technology among seven other corresponding technologies but each of them is excessively power consuming. Therefore, water-energy nexus is a critical challenge in today's world with rising population. MFC is a lab based interesting device whose electrogenic bacteria can consume waste from the waste water and in return releases electrons. If properly collected, such electrons can make the overall cell operation sustainable and energy independent. Thus a MFC can purify water while saving precious energy. We deployed nickel monosilicde based Ohmic contact and silicon based CMOS technology to develop micro-scale MFC which can generate mW level of power for on-chip power generation. By using saliva as fuel, we demonstrated nW scale power generation with further potential of scaling up [6]. More importantly, by doing so we have eliminated status-quo fuels such as chemically hazardous acetate based chemistry. Additionally, by removing traditional proton exchange membrane, we have made the overall device affordable. Therefore, such devices are important for *in-vivo* power generation with bio-compatibility.

Final criteria for any device for wider deployment to benefit the larger portion of the world's population is economy. While CMOS technology can certainly drive down the cost, but CMOS technology alone cannot do it for every person on earth. Therefore, recently we have demonstrated household materials based wearable multi-functional skin-type flexible integrated platform which can monitor the environment with out-of-plane arrangement and body vitals with in-plane integration [7]. A comparative study shows that such wearable gadgets provide the body vitals like body temperature, heartrate, blood pressure and skin hydration data same as that of obtained from smart watches (Apple Watch and Samsung Gear) within ±5% accuracy but significantly lower cost (< $25). Now, how is it even possible to use household materials for such sophisticated functionalities? Oddly enough, most physical sensing is dependent on resistive and/or capacitive sensing. Now, both of these sensing really depend on conductive and insulating materials. Our investigative study shows household materials such as aluminium foils, copper foils, plastic wraps, sponge, 3M papers, etc. with their conductive medium and micro-fibril architecture are effective conductive and resistive materials. Next, their capacitive and resistive properties can be further tuned based on their shapes and sizes and that is exactly how we built variety of sensors: pressure, proximity, temperature, humidity, flow, photo, strain sensors, etc. Another beauty of the overall scheme is paper has been in use for hundreds of years and such sensors can be made easily with record affordability. This also perfectly aligns with our vision of Do It Yourself (DIY) Electronics – simple to understand and learn and easy to implement and use electronics.

III. Key Advanced Healthcare Applications

We will briefly discuss three wearable freeform electronic systems which essentially disrupt the way today's healthcare technologies and ways. Our first example is smart thermal patch. Millions of world's population suffer from daily bruises, pains, arthritis and such. Most often, such suffering population use chemical based off-the-shelves patches which often are not suitable for children. For various locations of human body, we have to use different sizes raising the economic consideration. And finally, they can work at the most for ten hours only and then driving the cost up further. Therefore, we have demonstrated a copper (any biocompatible, corrosion resistant, economical material with suitable residual stress can be used) based CMOS compatible integration strategy to develop a smart digitized thermal patch integrating copper based heaters and temperature sensors which can conform to irregular body contours (Fig. 4) [8]. Moreover the whole system is stretchable up to 300% thus such a system can be used while wrapping our index fingers and then can be stretched out to cover larger space of other body locations. Such flexibility and stretchability is unprecedented. With appropriate integration of freeform silicon CMOS electronics, we have demonstrated App based smart gadget controlled pain management. In future, such system can be used for precision thermotherapy based targeted destruction of cancerous cell instead of conventional chemotherapy with negative side effects.

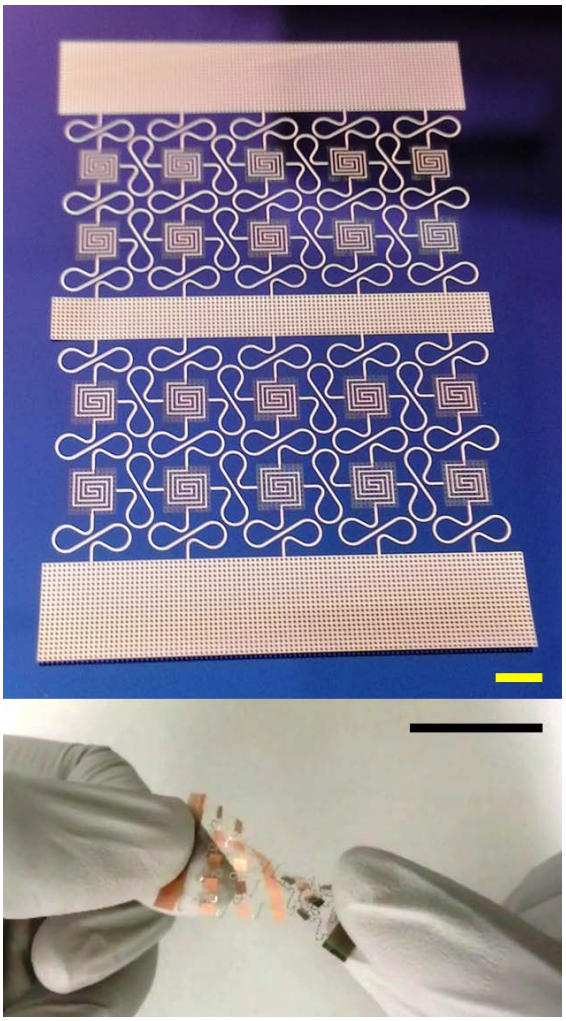


Fig. 4. Optical images of fabricated thermal patch on the wafer, and being handled with bare hands without the use of special handling equipment. Scale bars are 2 mm and 2 cm respectively.

Our next example is presumably world's first wearable *in-vivo* device. Today for orthodontic treatment, metallic one-size-fits-all dental braces are used for bon regeneration. It costs

thousands of dollars associated with its lingering usage time. With the advances in panoramic x-ray technology, we are now capable to personalize such dental braces by making them using additive manufacturing such as 3D printing (even using transparent materials). We have taken more advanced steps by integrating near-infrared lasing capability embedded on and in this personalized 3D printed transparent braces. Such capability enhances the bone regeneration significantly and reduces the overall cost and discomfort. However, it would not have been possible without an appropriate energy storage. Major challenge to integrate any traditional lithium based energy storage is their toxicity. Therefore, we have introduced non-toxic micro-scale flexible batteries to be used as on-demand power supply. In future we envision to use smart phone based remote (wireless) powering for such *in-vivo* devices for advanced healthcare applications.

Our final example is inspired by the vision, "What is after sensors?" Today more and more we are embracing Internet of Things (IoTs) which suggests trillions of sensors deployed worldwide to create cyber-physical system which eventually generates big data. Personalized healthcare initiatives are geared toward developing understanding about personalized healthcare need (instead of blanket globalized treatment) by analyzing genetic mapping by chip-based genome sequencing and physiological and psychological monitoring using wearables. However, going forward in future when we enter the era of Internet of Everything (IoE), the main question would be what we do with the analysed big data. To address such questions, we have taken steps to personalize nutrient formulation and eventually personalize medicines, by developing wearable personalized medicinal platform. Using expendable polymer based smart microfluidics architecture and freeform CMOS electronics, we have shown that based on an already established library of genetic mapping and instant collection of physiological and psychological data, we can dynamically formulate the required nutrition level or medicine of a user, prepare it *in-situ* and deliver on-demand. Such devices are unprecedented and shows the realization of the vision the future of affordable advanced personalized healthcare for everyone.

IV. Conclusions

In this paper, we have laid down our vision about freeform electronics for advanced personalized healthcare. We have shown five main design criteria for any such device (*in-vivo* and *in-vitro*): (i) integration strategy, (ii) thermal management, (iii) biocompatibility, (iv) power management and (v) economics. By using examples ranging from brain machine interface to low-cost paper-based wearable monitors we have shown criticality of such design criteria. Further, we have presented three advanced personalized healthcare freeform electronic systems which will disrupt the way we know and use healthcare today. Presented devices are smart digitized thermal patch for pain management, smart dental brace for accelerated orthodontic treatment and finally wearable personalized medicinal platform for *in-situ* dynamic formulation and on-demand delivery of nutrients and medicines. Immense possibilities exist going forward to positively change the way we lead our life today and multi-disciplinary effort with open minded engineering innovation, we can accelerate such growth to bring tomorrow's world to today.

Acknowledgment

This publication is based upon work supported by the King Abdullah University of Science and Technology (KAUST) and Boeing Company.

References

[1] A. M. Hussain and M. M. Hussain, "CMOS-Technology-Enabled Flexible and Stretchable Electronics for Internet of Everything Applications," *Adv. Mater.*, vol. 28(22), pp. 4219–4249, Jun. 2008.

[2] A. M. Hussain and M. M. Hussain, "Deterministic Integration of Out-of-Plane Sensor Arrays for Flexible Electronic Applications," *Small*, vol. 12(37), pp. 5141–5145, Oct. 2016.

[3] A. N. Hanna, H. M. Fahad and M. M. Hussain, "InAs/Si Hetero-Junction Nanotube Tunnel Transistors," *Sci. Rep.* vol. 5, 9843, doi:10.1038/srep09843, 2015.

[4] A. N. Hanna and M. M. Hussain, "Si/Ge Hetero-structure Nanotube Tunnel Field Effect Transistor," *J. Appl. Phys.* vol. 117, pp. 014310, Jan. 2015.

[5] A. M. Hussain, F. A. Ghaffar, S. I. Park, J. A. Rogers, A. Shamim and M. M. Hussain, "Metal/Polymer Based Stretchable Antenna for Constant Frequency Far-Field Communication in Wearable Electronics," *Adv. Funct. Mater.* vol. 25(42), pp. 6565–6575, Nov. 2015.

[6] J. E. Mink, R. Qaisi, B. E. Logan and M. M. Hussain, "Energy Harvesting from Saliva in Micro-Sized Microbial Fuel Cells," *NPG (Nature) Asia Mater.* vol. 6, pp. e89, Mar. 2014.

[7] J. M. Nassar, M. D. Cordero, A. T. Kutbee, M. A. Karimi, G. A. Torres Sevilla, A. M. Hussain, A. Shamim and M. M. Hussain, "Paper Skin Multi-Sensory Platform for Simultaneous Environmental Monitoring," *Adv. Mater. Tech.* vol. 1(1), Apr. 2016.

[8] A. M. Hussain, E. B. Lizardo, G. A. Torres Sevilla, J. M. Nassar and M. M. Hussain, "Ultra-Stretchable and Flexible Copper Interconnect Based Smart Patch for Adaptive Thermotherapy," *Adv. Healthcare Mater.* vol. 4(5), pp. 665–673, Apr. 2015.

Electrical Power Generation Through Speed Breaker

Kafi Mohammad Ullah[1], K. M. Ahsan-uz-Zaman[2], Sohag Hosen[3], Rakibul Hasan Khan[4], Sanjida Parvin[5]
[1,2,3,4] Dept. of EEE, Mymensingh Engineering College (University of Dhaka), Mymensingh
[5]Dept. of EECE, Military Institute of Science & Technology (MIST), Dhaka
[1]kafi.mec.07@gmail.com

***Abstract*— In the present situation power becomes basic need for human life. Energy is responsible for major developments of any country's economy. Conventional energy sources generate most of the energy of today's world. But the population is increasing day by day and the conventional energy sources are diminishing. Moreover, these conventional energy sources are polluting and responsible for global warming. So, non-conventional sources are needed to be developed for power generation which are clean, environment friendly and sustainable. In this research we propose a renewable non-conventional energy source based on speed breaker mechanism. Our project is to enlighten the streets utilizing the jerking pressure which is wasted during the vehicles passes over speed breaker in roadside. We can tap the energy generated by moving vehicles and produce power by using the speed breaker as power generating unit. The kinetic energy of the moving vehicles can be converted into mechanical energy through rack and pinion mechanism and this mechanical energy will be converted to electrical energy using generator which will be used for lighting the street lights. Therefore, by using this mechanism we can save lot of energy which can fulfill our future demands.**

Keywords*— *kinetic energy, speed breaker, rack & pinion, generator, non-conventional energy, street light.

I. Introduction

Now a day's power has become the major need for human life. Energy is an important input in all the sectors of any countries economy. The availability of regular conventional fossil fuels will be the main sources for power generation, but there is a fear that they will get exhausted eventually by the next few decades. Therefore, we have to investigate other types of renewable sources .The day-to-day increasing population and decreasing conventional sources for power generation, provides a need to think on non-conventional energy resources [1] [2].Another major problem, which is becoming the exiting topic for today is the pollution. Power stations and automobiles are the major pollution producing places. So non-conventional power source is needed to reduce this problem. We proposed a non-conventional power generating system based on speed breaker mechanism which generate electricity without using any commercial fossil fuels, which is not producing any polluting products [3]. In this paper, our aim is to conserve the kinetic energy which convert into electricity that gone wasted, while vehicles move.

II. Methodology

Power can be produced from conventional and nonconventional energy sources. In this paper we show energy conversion from kinetic energy to rotational energy and rotational energy to electrical energy respectively [4]. This project explains the mechanism of electricity generation from speed breakers.

It is a simple but optimum process to generate energy from speed breaker arrangements. There are a large number of vehicles running on the road. These vehicles are going over a number of speed breakers present on the road. We want to replace this traditional speed breakers with our proposed speed breaker. It is an Electro-Mechanical unit. This system utilizes both mechanical technologies and electrical techniques for the power generation and its storage. The generation will be proportional to the traffic density [5].

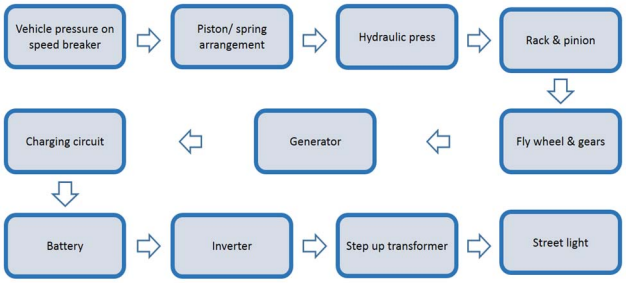


Fig. 1. Block diagram of

III. System Design & Function

In this research, vehicle pressure on the speed breaker which is converted into rotary energy through rack and pinion using hydraulic press. Consequently, this rotary energy rotates generator that generates electrical power which is being stored through battery using charging circuit [6]. The whole system is represented on Fig. 2 and Fig. 3.

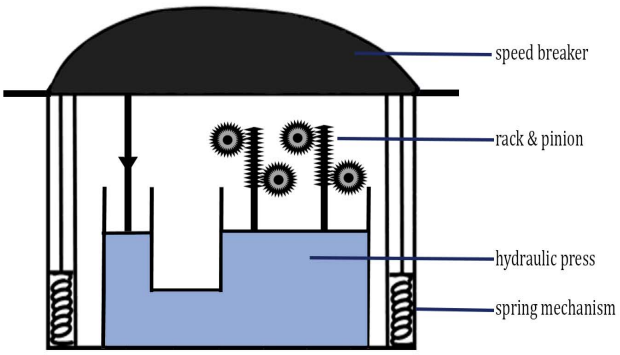


Fig. 2. Speed breaker mechanism

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

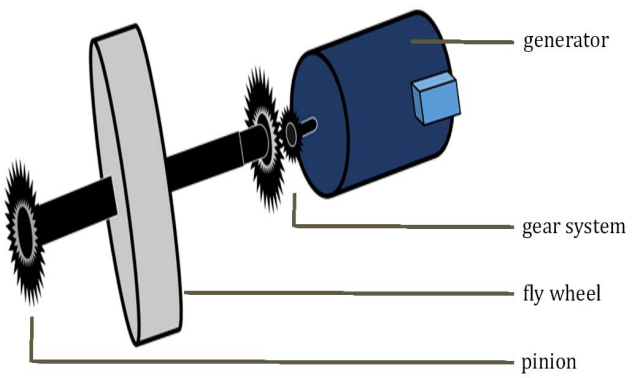


Fig. 3. Fly wheel and generator mechanism

When the vehicle load acted upon the speed breaker system is transmitted to rack and pinion arrangements. Because of the weight of vehicle the top portion of the speed breaker moves downwards [7]. The pressure rod which one side is attached to the bottom of the speed breaker and other side to small piston of hydraulic press. The hydraulic press convert the force into 4times from small piston to large piston. 3 rakes attached to large piston.

Every rack connected with 2 one way pinion. Pinions are getting rotation from the movement of rack backward and forward. Every pinion connect to a fly wheel which absorbs the energy when demand is less and releases the same when it is required. Fly wheel connected to a large gear and large gear connected to a small gear which attached to a generator. Generator convert rotational energy into electrical energy [8]. When the vehicle passes the speed breaker, the expansion of springs takes the speed breaker at previous state which done the system reverse.

The charging circuit charges a battery. The Inverter circuit converts this DC voltage into AC voltage and a step up transformer stepped up the AC voltage. A dark sensing circuit is used to sense the night so that the street light lit on.

IV. System Construction

Speed breaker, spring arrangement, hydraulic press, rack and pinion combination, freewheeling and gear combination, generator, charging circuit, battery, dark sensing & switching circuit, inverter circuit and step up transformer used to the system in this research work.

A. Speed breaker

It's the top portion of the system which is made of iron in curved shape. The main function of this speed breaker is to sustain the pressure of vehicle and squeezes it when vehicle passes through it.

B. Spring arrangement

A spring is an elastic body whose function is to distort when loaded and to recover its original shape when the load is removed. It cushions, absorbs or controls energy either due to shocks or due to vibrations [9]. There are four helical springs below the speed breaker which are squeezed when vehicle pressurizes upon it and it bring the speed breaker at previous state.

C. Hydraulic press

In a hydraulic press a small force applied on a column of liquid is converted into a much greater force available to another column of liquid. It is an application of Pascale law. In our proposed system it convert the force into 4times from small piston to large piston when the pressure created on speed breaker using this equation,

$$F_2 = \frac{F_1}{a} \times \mathrm{A}$$

F_1= created force on small piston by speed breaker
F_2= converted force by hydraulic press
a = area of small piston
A = area of large piston

D. Rack & Pinion

Rack and pinion can convert rotary to linear of from linear to rotary motion [10]. Rack is a linear gear and pinion is a circular gear. Applied force on rack is converted to rotation by pinion. The mechanical force is converted into rotational force.

E. Fly wheel

The primary function of flywheel is to act as an energy accumulator. It reduces the fluctuations in speed [11]. It absorbs the energy when demand is less and releases the same when it is required.

F. Generator

The device which converts mechanical energy into electrical energy is called generator. An AC generator is used for producing alternating current which contains an assembly of stationary (stator) and moving parts (rotor). The rotor is connected with the gear. The torque which generated by gear rotates the rotor of the generator. The rotor creates a moving magnetic field around the stator, which induces a voltage difference between windings of stator and produce the alternating current (AC) output of the generator [12].

G. Charging circuit

Charging circuit is used to charge the battery. Fig. 4. shows the charging circuit.

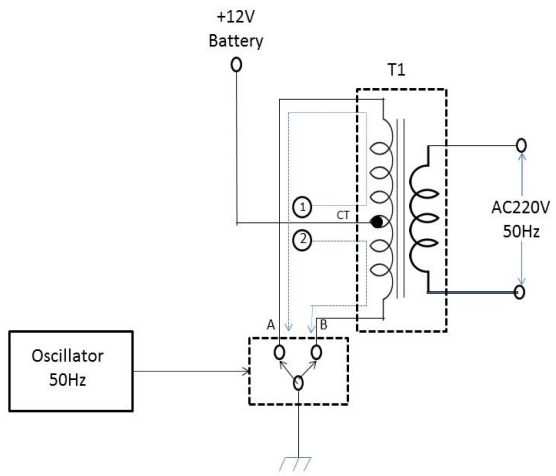


Fig. 4. Charging circuit

H. Dark sensing and switching circuit

Dark sensing circuit sense the dark and switch the light on. Fig. 5. shows the dark sensing & switching circuit.

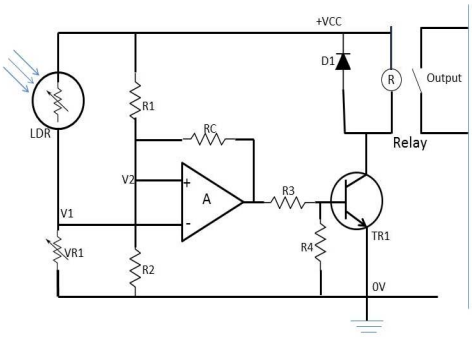


Fig. 5. Dark sensing & switching circuit.

I. Inverter circuit & step-up transformer

Inverter convert DC voltage to AC voltage and step up transformer is a type of transformer which stepped up the AC voltage. In this system inverter circuit converts 12V DC to 15V AC. Step up Transformer makes the voltage to 250 V AC from 15 V AC.

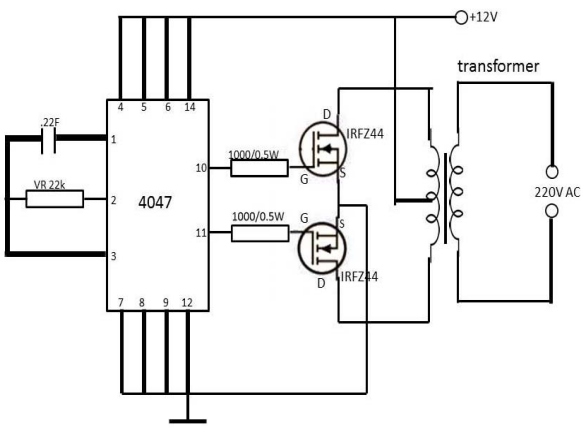


Fig. 6. Inverter circuit & step-up transformer.

V. Cost Analysis

The different essential equipment and its cost with quantities are given to Table I.

TABLE I. COST ANALYSIS

Equipment	No. of set	Price (BDT)
Rack	2	200
Pinion	4	400
Gear	2	200
Spring	4	400
Body frame	1	1000
Hydraulic press	1	3000
Battery	4	8000
Fly wheel	1	500
Wiring		200
Maintenance cost		1000
Installation cost		1000
Generator		3000
Total cost		18900

VI. Power Calculation & Result Analysis

Let's consider,

The mass of vehicle moving over the speed breaker = 350Kg (Approximately)

Height of speed breaker = 15 cm

Weight of the Body = 350 Kg × 9.8 = 3430 N

Distance traveled = Height of the speed breaker = 15cm

Work done = weight of the body × distance travelled by the pressure of vehicle

Power = Work done/Second = (3430×0.15)/60 = 8.58 Watts

Output Power developed for 1 vehicle passing over the speed

Power developed for 60 minutes (1 hr.) = 514.5 watts

Power developed for 24 hours = 12.35 Kw

Our proposed system can provide 250 v and 24 amp.

We are using CFL bulb (100 watt)

In one km 60 bulbs are needed.

Total watt = 60×100 = 6000 watt = 6 Kw

This power generated by vehicles is more than sufficient to run four street lights in the night time**.**

VII. Advantages

Using this proposed technology we can get the following advantages:

- Nonpolluting power generating source
- Power generation with low cost.
- Simple construction, mature technology.
- Less floor area required and no obstruction to traffic.
- No need of manual work during power generation.
- Low installation and maintenance cost
- Power generation using non-conventional energy sources.
- Easy for maintenance and no fuel transportation problem.
- Simple construction, mature technology, and easy maintenance.

VIII. Conclusion

The listed system is non-conventional and the way of power generation technique is also echo friendly. It has advantage that it does not utilize any external source [13]. By using this system we will able to reduce power crisis and load shedding. The stored electricity could satisfy the daily requirement for street lighting. We can also use it for signal system on road, tollbooth or any other useful work. So from the every angle we can see this system is very much effective to reducing the power crisis in Bangladesh.

ACKNOWLEDGMENT

The authors are grateful to Assistant Prof. Dr. Ratan Kumar Nandi, Dept. of Electrical & Electronic Engineering (EEE), Mymensingh Engineering College for his continuous encouragement and supervision throughout the course of the research.

REFERENCES

[1] Sharma, P.C., "Non-conventional power plants", Public Printing Service, New Delhi, 2003.

[2] Mukherjee, D. Chakrabarti, S., "Non-conventional power plants", 2005.

[3] Sharma.P.C , *Principle of renewable energy systems* (Public printing service, New Delhi, 2003).

[4] Watts,G., "Effects of speed distribution on the Hormonoise model predictions", Inter-noise Conference, Prague, 2004.

[5] Dr. Anders Brandt & MSc. John Granlund Swedish Road Administration. "Bus Drivers Exposure to Mechanical Shocks Due To Speed Bumps". Society for Experimental Mechanics, IMAC 25th Conference and Exposition on Structural Dynamics 2008.

[6] A. Padma Rao, A. Kiran Kumar and S. Suresh, "Power Generation from Speed Breaker by Rack and Ratchet Mechanism," *International Journal of Current Engineering and Technology*, sp. no. 2, February 2014.

[7] Abdul Razzak Pathan, Aniket Garate, Karthikeyan N and SonaliRetharekar, "Power generation through speed breaker,".

[8] Amanpreet Kaur, Shivansh Kumar Singh, Rajneesh, Parwez and Shashank, "Power Generation Using Speed Breaker with Auto Street Light," *International Journal of Engineering Science and Innovative Technology (IJESIT),* vol. 2, no. 2, March 2013.

[9] "Power System Stabilizers", by Mitsubishi Corporation-A release notes from Mitsubishi Co.

[10] P.M. Anderson and A.A. Fouad, „Power System Control and Stability", Galgotia Publications.

[11] "Power System Dynamics and Control", K.R.Padiyar, Interline Publishers Bangalore.

[12] D. Venkata Rao, K. Prasada Rao, Chiranjeeva Rao and R. Umamaheswara Rao, "Design and Fabrication of Power generation System using Speed Breaker," *International Journal of Current Engineering and Technology,* vol. 4, no. 4, August 2014.

[13] Nota, R., Barelds, R., "Engineering methods for road traffic and railway noise after validation and fine-tuning", Harmonies, 2005.

Improved Convergence of an Adaptive Identification Algorithm for a Large Sparse System

Ankit Deshmukh,[1,] K. Ramakrishnan,[2] and Goshaidas Ray[1]
[1]Department of Electrical Engineering, I.I.T Kharagpur
Kharagpur-721302, West Bengal, India.
[2]Department of Electrical Engineering, Pondicherry College,.
Pondicherry-605014, India
ankitdeshmukh52@gmail.com; ramakrishnan@pec.edu; gray@ee.iitkgp.ernet.in

1

***Abstract*—Estimation of small tap-coefficients of large sparse system using μ-law based proportionate normalized least square (MPNLMS) algorithm yields slow converges, since the proportionality of these coefficients is ignored in the updated process. The individual activation factor-MPNLMS (IAF-MPNLMS) algorithm solves this problem by assigning new gain distribution factor while updating the tap-coefficients. However, this algorithm also suffers from slow convergence if an input signal is highly correlated. In this paper, a modified activation factor is suggested to improve the convergence speed of the MPNLMS and IAF-MPNLMS algorithms for sparse system identification while the input signal is highly correlated. For correlated input signals, the simulation results show that the proposed algorithm acquires better convergence speed than the other two algorithms.**

***Index Terms*—Adaptive filters, sparse system, proportionate normalized least mean square, individual activation factor.**

I. INTRODUCTION

An impulse response can be qualitatively classified as sparse if a large fraction of its energy is concentrated within a small fraction of its duration. Adaptive system identification is a challenging task for sparse systems [1]-[2]. Sparse systems have gathered a lot of attention and they characterized by long impulse responses with only a few nonzero coefficients. In the literature, several adaptive schemes can be found to accelerate the convergence toward the sparse optimum solution and they have been applied in a wide range of applications such as data compression, echo cancellation, spectrum cartography, and medical signal processing. For example, proportionate adaptive filters assign to each tap co-efficient a different step size proportional to its amplitude [3]. It has been reported that direct application of normalized least mean square (NLMS) algorithm to system identification provides unsatisfactory performance when the impulse response is sparse. Firstly, the adaptive algorithm has to operate on a relatively long length filter, and secondly, most of the coefficients that will occur during adaptation are the almost zero-valued coefficients [4]-[5]. To address this problem, the concept of proportionate NLMS (PNLMS) was introduced in [6], where each coefficient is updated with step-size proportional to its magnitude. PNLMS algorithm has fast initial convergence speed but it is not maintained over the whole adaptation process, its performance also degrades as the sparseness decreases [7]. An essential role of individual-activation factor (IAF) lies, in fact, in an adaptive algorithm to prevent the adaptive coefficients from stalling when their magnitudes are significantly smaller than the largest coefficent. The IAF-PNLMS algorithm revisits the PNLMS algorithm for devising a new way to determine the activation factor for each tap-coefficient individually, consequently improving the convergence speed [8]-[9], whereas in PNLMS a common individual activation factor is used to all tap-coefficients. Recently, an improved version of optimal convergence is achieved while all coefficients attain in a vicinity of their optimal values in a same number of iterations. To achieve this, a non-linear relationship i.e. the μ-law conception is introduced to PNLMS algorithm (MPNLMS) to obtain fast initial convergence over the complete adaptation process in case of a sparse system at the expence of large computational burden [10]-[11]. However, for highly sparse system, the MPNLMS converges slowly during the update of small coefficients, since the proportionality of these coefficients is not taken into account in adaptation process. Thus, to achieve a fast convergence algorithm for highly sparse system identification, the μ-law conception is introduced into the IAF-PNLMS algorithm which is termed as IAF-MPNLMS algorithm [11]. However, the performance of IAF-MPNLMS algorithm goes on degrading as the input signal becomes highly correlated. It has been reported that the logarithmic activation factor to the inactive coefficient does not prove effective while the magnitude of these coefficients is very small. In this paper, we propose a new square-root strategy to increase the gain associated to the inactive coefficients of the IAF-MPNLMS when active coefficients approach to its converse value, which in turn speed up the global convergence. The versatility of the proposed approach has been demonstrated through numerical examples in a variety of applications and it outperforms over the recent results.

The remainder of the paper is organized as follows: Section II summarizes the PNLMS, MPNLMS, IAF-PNLMS, and IAF-MPNLMS algorithms. The modified IAF-MPNLMS algorithm is proposed in section III. Simulations results are presented in section IV. Section V draws conclusions.

II. PROBLEM FORMULATION

This section presents the problem of system identification followed by the brief description of the PNLMS, MPNLMS, IAF-PNLMS, and IAF-MPNLMS algorithms.

A. System Identification

The problem of system identification can be described with the figure 1 where $x(n)$ is the input or exciting signal, $y(n)$ is the output signal of the model, $e(t)$ is the error signal, $z(n)$ is

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

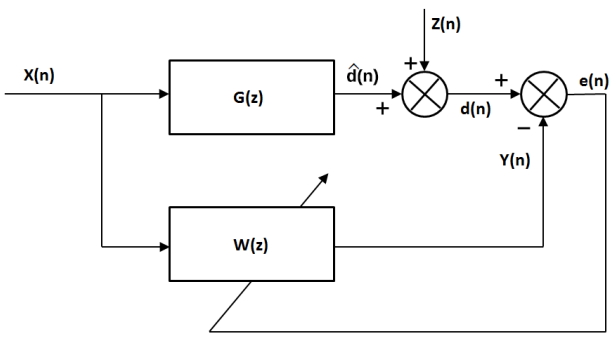


Fig. 1. System Identification Using Adaptive Filters

the measurement signal and $d(n)$ is the desired signal. Both plant $G(z)$ and adaptive filter $W(z)$ are excited by the same input signal $x(n)$.

Here, the objective is to estimate the parameters of the model $W(z)$ of the plant $G(z)$. Initially, the positions of unknown non-zero parameters of $W(z)$ are assumed to be known on the basis of some prior knowledge of the plant $G(z)$. These parameters of $W(z)$ are then estimated through an adaptive filtering algorithm such that the square of the difference between the plant output $d(n)$ and adaptive filter output $y(n)$ is minimized.

B. PNLMS, MPNLMS Algorithms and their variants.

The PNLMS algorithm was developed for specifically for sparse system identification. The main idea being the allotment of different step sizes to different coefficients proportional to the magnitude of the estimated filter coefficients. However, the tap coefficient magnitudes are not known a priori, therefore the tap coefficients are recursively updated as follows [10]:

$$W(n+1) = W(n) + \beta \frac{G(n)X(n)e(n)}{X^T(n)G(n)X(n) + \delta}, \quad (1)$$

where, β is step size, δ is regularization parameter that avoids the division by zero in the above expression, and the error $e(n)$ is given by,

$$e(n) = d(n) - W^T(n)X(n), \quad (2)$$

$$\begin{aligned} X(n) &= [x(n)\ x(n-1)\ \cdots\cdots x(n-L+1)]^T \\ &= \text{input vector}, \\ W(n) &= [w(n)\ w(n-1)\ \cdots\cdots w(n-L+1)]^T \\ &= \text{tap coefficient vector}, \end{aligned}$$

where L is the length of the filter and the gain distribution matrix associated with the filter tap coefficients and they are expressed as

$$G(n) = diag[g_1(n)\ g_2(n)\ .....g_L(n)] \quad (3)$$

and

$$g_i(n) = \frac{\phi_i(n)}{\frac{1}{L}\sum_{i=1}^{L}\phi_i(n)} \quad (4)$$

where $\phi_i(n)$ is the proportionality function and it is defined as

$$\phi_i(n) = max[f_i(n), F_i(n)] \quad (5)$$

with individual activation factor $f_i(n)$ as

$$f_i(n) = \rho \times max(\gamma, F_1(n), F_2(n)....., F_L(n)) \quad (6)$$

For PNLMS algorithm, the choice of F_i as

$$F_i(n) = |w_i(n)|, \quad for\ i = 1,\ 2\ \cdots\cdots L \quad (7)$$

The parameter γ is a constant and positive number (typical value is 0.01) prevents the filter coefficients from stalling during initialization and ρ with a typical value of 0.01, prevents coefficients from stalling when they are much smaller than the magnitude of largest coefficient. The equation (6) employs step size proportional to the magnitude of coefficients. The function $f_i(n)$ in equation (6) is called activation function. The PNLMS algorithm, has very fast initial convergence speed, however, after the initial period, it begins to slow down dramatically, even becoming slower than the NLMS algorithm.

The MPNLMS algorithm improves the convergence of PNLMS algorithm. It achieves this by computing the optimal proportionate step-size during the adaptation process. The MPNLMS algorithm was derived such that all coefficients attain a value within a ϵ-vicinity of their optimal values in the same number of iterations, $F_i(n)$ in (6) is modified as,

$$F_i(n) = \ln(1 + \mu|w_i(n)|), \quad for\ i = 1,\ 2\ \cdots\cdots L$$

where, $\mu = 1/\epsilon$ (ϵ is a small positive number). It is to be noted that in both PNLMS and MPNLMS algorithms, the activation factor $f_i(n)$ is same for all the coefficients which, in turn, is responsible for an imperfect performance in identification of highly sparse systems.

To overcome this difficulty Individual activation factors (IAF) were introduced in [8] and [11] for PNLMS and MPNLMS algorithms respectively. The new activation function was designed in such a way that it would converge to the coefficient magnitude as the number of iterations increase and it would be greater than zero to avoid freezing of the algorithm. Thus the following individual activation function formulation was adopted:

$$f_i(n) = \gamma F_i(n) + (1-\gamma)(\phi_i(n-1)) \quad (8)$$

where, $0 < \gamma < 1$ and

$$\phi_i(n) = max[f_i(n), F_i(n)] \quad (9)$$

In equation (8) the first term of the right-hand side helped $f_i(n)$ converge to the corresponding coefficient magnitude and the second term ensured it was always greater than zero. The activation factors were initialized with a small positive constant $f_i(n) = 10^{-4} > 0$. To have the estimated coefficients be proportional to the magnitudes of the corresponding system tap-coefficients over the adaptation process, the activation factors were given a learning period of L samples equal to filter length and $\gamma = 1/2$ was chosen, thus equation (8) was revised to: [8]

$$f_i(n) = \begin{cases} (1/2)F_i(n) + (1/2)(\phi_i(n-1)), & n = mL, \\ & m = 1, 2, 3.. \\ f_i(n-1), & \text{otherwise} \end{cases} \quad (10)$$

The equation (10) instead of equation (6) along with equation (7) gives the IAF-PNLMS algorithm. Replacing (7) by the expression

$$F_i(n) = \ln(1 + \mu|w_i(n)|)$$

in the IAF-PNLMS algorithm provides the IAF-MPNLMS algorithm. The IAF-MPNLMS algorithm was introduced in order to achieve the advantages of both the MPNLMS and the IAF-PNLMS algorithms

III. The Modified IAF-MPNLMS Algorithm

IAF-MPNLMS algorithm, discussed in the previous section comprised of two major features. The first being the role of activation factor in the adaptation of inactive tap coefficients and second, the logarithmic function $F_i(n)$ responsible for adjusting the gain control factors for active tap coefficients. It has been reported that the individual activation factor was given by equation (8) contains past information (some memory) of the magnitude of the ith coefficient, where factor γ can be seen as a forgetting factor related to the memory of the adaptive filter coefficient magnitude. The equation (8) consists of two terms on the right-hand side, the second term ensures that $f_i(n)$ be always greater than zero. But the activation factors are initialized with a small positive constant (typically, $f_i(0) = 10^{-2}/L$), such that $f_i(n) > 0$.

(A) Now if the initial value $f_i(0) = 10^{-2}/L > 0$ then it can be seen clearly that, in IAF-MPNLMS algorithm, $f_i(n) > 0$ always and this, in turn, implies that $F_i(n) > 0$ Hence, the second term in (8) looses its importance.

(B) It is noted from hypothesis **(A)** that the activation function now will have only one term that is $F_i(n)$ on the right hand side of (8). But for inactive coefficients, which have near zero magnitudes and the expression for $F_i(n)$ can be written as:

$$F_i(n) \approx \mu|w_i(n)|$$

In this case, the activation factor, which is instrumental in deciding the gain control factor for inactive coefficients does not have a non-linear relationship to the coefficient magnitude. To overcome this difficulty, the activation function is now modified in such a manner that does not change the gains for active tap-coefficients, but will provide better gain distribution for inactive tap coefficients while the nonlinear/square-root form of individual activation function is adopted during the complete adaptation process. For the proposed method, we will replace the activation function simply by the square root of the estimated filter tap coefficient. This is because the shape of a square root function gives better gain distribution than that of a logarithmic function for incredibly small magnitudes (inactive taps). The new activation function is also given a learning period of samples equal to the filter length as in IAF-MPNLMS algorithm. The Modified IAF-MPNLMS algorithm (MIAF-MPNLMS) can be formulated as:

$$f_i(n) = \begin{cases} \sqrt{|w_i(n)|} & n = mL, \\ & m = 1, 2, 3.... \\ f_i(n-1) & otherwise \end{cases} \tag{11}$$

and,

$$F_i(n) = \ln(1 + \mu|w_i(n)|)$$

with

$$\phi_i(n) = max[f_i(n), F_i(n)] \tag{12}$$

the elements of $G(n)$ are computed as:

$$g_i(n) = \frac{\phi_i(n)}{\frac{1}{L}\sum_{i=1}^{L}\phi_i(n)} \tag{13}$$

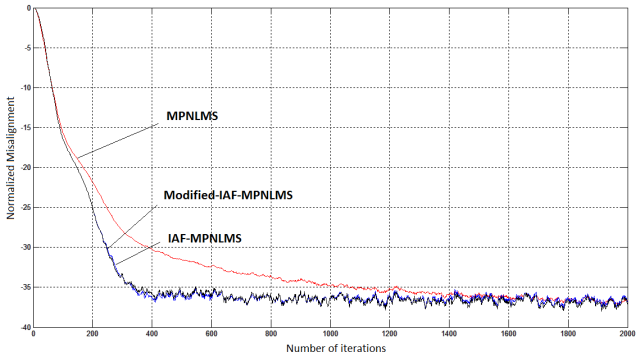


Fig. 2. Adaptation curves for a highly sparse system for uncorrelated input signal.

where,

$$G(n) = diag[g_1(n)\ g_2(n)\ .....g_L(n)] \tag{14}$$

The tap coefficients are updated as:

$$W(n+1) = W(n) + \beta\frac{G(n)X(n)e(n)}{X^T(n)G(n)X(n) + \delta} \tag{15}$$

In this algorithm, the activation factor and hence the gain of the inactive taps have been modified so as to achieve better convergence with respect to IAF-MPNLMS and MPNLMS algorithms.

IV. Simulation Results

In this section, Monte Carlo simulations (average of 100 independent runs) are carried out for the MPNLMS, IAF-MPNLMS, and the proposed Modified IAF-MPNLMS (MIAF-MPNLMS) algorithms for a highly sparse system identification problem. Normalized Misalignment is defined by

$$N_M = \left[\frac{\|P(n) - W(n)\|_2}{\|P(n)\|_2}\right]^2$$

and it is defined as the measure of algorithm performance, where $P(n)$ is the true value of system parameters and $W(n)$ is the estimated values.

- Example 1:

In this example, a sparse system having 200 taps out of which 4 taps are active, is considered. The active taps (0.02, 0.2, -0.1, 0.02) are considered at the positions (11, 30, 135, 185) respectively. The uncorrelated input signal is a zero mean with unity variance and the additive noise signal is also zero mean and the signal to noise ratio (SNR) is assumed to be 30 dB. The simulation results are obtained by considering the following data: step size β is 0.3 for all cases, $\rho = 0.01$, $\gamma = 0.01$ for MPNLMS, the value of μ is 250 for MIAF-MPNLMS (proposed), IAF-MPNLMS and 100 for MPNLMS, $f_i(0) = 0.0001$ for IAF MPNLMS and MIAF-MPLNMS. The results are represented in figure 2 for the case when the input signal is uncorrelated and it is observed that there is no improvement in the speed of convergence of the proposed scheme compared to the IAF MPNLMS algorithm.

- Example 2:

In this example, the same system and parameters as in example-1 have been considered only the input signal used to identify the system is correlated in the following manner: The input signal used is a correlated unity-variance ($AR(2)$)

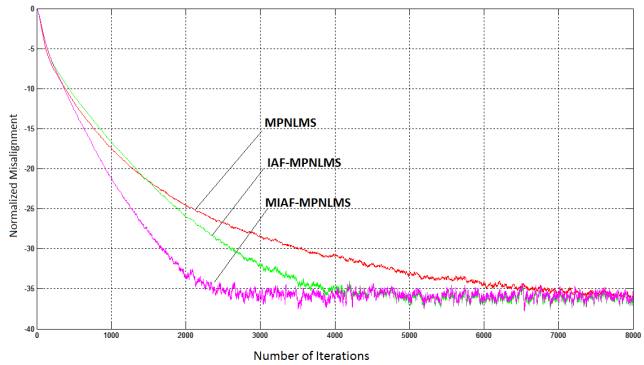


Fig. 3. Adaptation curves for a highly sparse system for correlated input signal.

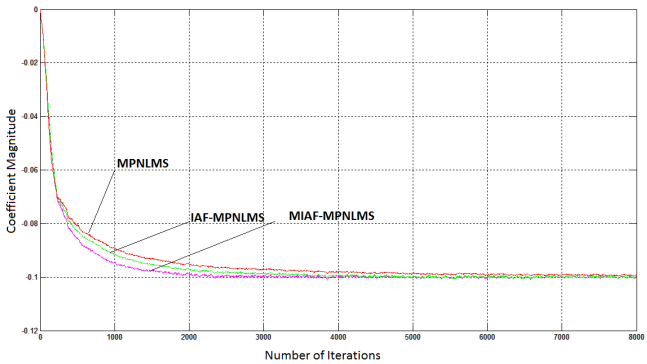


Fig. 4. Performance curves of the 135-th filter tap coefficient (actual value=-0.1) for correlated input.

process with $b_1 = 1.29$, $b_2 = -0.75$ and $v(n)$ is the white noise with variance $\sigma_v^2 = 0.20$. The input signal $x(n)$ is generated by the expression given below.

$$x(n) = b_1 x(n-1) + b_2 x(n-2) + v(n)$$

The results of the proposed algorithm are compared with existing algorithm and they are represented in figure 3 for the case when the input signal is correlated and with the parameters $\beta = 0.3$, $f_i(0) = 10^{-4}$, $\rho = 0.01$, $\gamma = 0.01$ with $\mu = 100$ for MPNLMS and $\mu = 250$ for the rest two algorithms to obtain the same misalignment.

The convergence of individual filter tap-coefficients for correlated input signals are also presented in figures 4 and 5 which correspond to the 135th and the 185th filter tap coefficients respectively. In case of highly correlated input signal, it may be noted that the proposed identification scheme based on modified activation function outperforms than the other two methods presented in this paper.

V. CONCLUSIONS

In this paper, a modification in the individual activation factor was proposed to provide new gain distribution for updating the inactive and active tap-coefficients of the sparse system. As a consequence, the performance and convergences of the proposed algorithm significantly improved than the IAF-MPNLMS and MPNLMS algorithms for identification of highly sparse system while inputs are highly corelated. In addition, the simulation results reveal that the proposed scheme although exhibits same convergence performance for

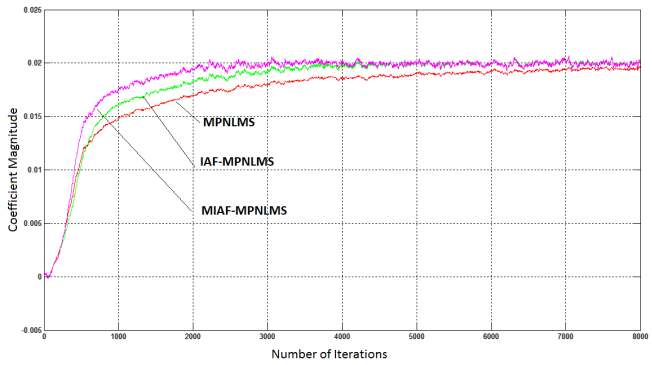


Fig. 5. Performance curves of the 185-th filter tap coefficient (actual value=0.02) for correlated input.

uncorrelated input signal as that of IAF-MPNLMS and MIAF-MPNLMS algorithms, but it outperforms over the recent algorithms for sparse system identification under correlated input signals until it reaches the steady state value.

REFERENCES

[1] K. Dogancay and P. Naylor, "Recent advances in partial update and sparse adaptive filters," *in Proc. European Signal Processing Conferences, Antalya, Turkey , Sept., EUSIPLO* 2005.

[2] S. Chouvardas, K. Slavakis, Y. Kopsinis, and S. Theodoridis, "A aparsity promoting adaptive algorithm for distributed learning," *IEEE Trans on Signal Processing*, vol. 60, no.10, pp. 5412–5425, 2012.

[3] R. Harris, D. Chabries, and F. Bishop, "A variable step size adaptive filter aigorithm," *IEEE Trans. Acoust. Speech, Signal Processing*, vol.34, no.2, pp.309-316, 1986.

[4] S. Haykin, *Adaptive filter theory*, 2nd ed. Prentice Hall, 2002.

[5] Y.F. Chen, and D.M. Etter, "Analisis of an adaptive technique for modeling sparse systems," *IEEE Trans on Acoustics, Speech and signal Processing*, vol. 37, no. 2, pp. 254-264, 1989. 508–518, 2008.

[6] D. Duttweiler, "Proportionate normalized least mean square adaptation in echo cancellers," *IEEE Trans on Speech Audio Processing*, vol. 8, pp. 508–518, 2008.

[7] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 1881–1884, 2002.

[8] F. C. de Souza, O.J.Tobias ,R. Seare, and D. R. Morgan, "A PNLMS algorithm with individual activation factors," *IEEE Trans on Signal Processing*, vol. 58, no. 4,2036-2047, 2010.

[9] F.C.de Souza, R. Sera and D.K. Morgan, "An enhanced IAF-PNLMS adaptive algorithm for sparse impulse response identification," *IEEE Trans on Signal Processing*, vol. 60,no.6, pp. 3301-3307-518, 2012.

[10] H. Deng and M. Doroslovacki, "Improving convergence of the PNLMS algorithm for sparse impulse response identification," *IEEE Signal Process. Lett.*, vol. 12, no. 3, pp. 181–184, 2005.

[11] Y Zhang, C. Wang, and N. Li, "A fast converging algorithm for systems with highly sparse impulse response," *IEEE Fifth International Joint Conference on Computational Sciences and Optimization*, vol. 12, no. 3, pp. 831–834, 2012.

Input Switched Single Phase SEPIC Controlled Rectifier With Improved Performances

Md. Shamsul Arifin*, Mohammad Jahangir Alam
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh
*shamsularifin126@gmail.com

Abstract—**DC power conditioning and processing is an essential challenge now a days. Rectifiers having bridge structures are considered to be nonlinear as they draw non sinusoidal current from line. This distorted current causes several problems related to power quality. In this paper, a new configuration of single phase single switch controlled rectifier using SEPIC topology is proposed. The input current of the proposed configuration is chopped at high frequency using a semiconductor switch to get step up/step down AC to DC conversion, instead of using conventional converters. The proposed circuit is compared with conventional one under same load and operating conditions. The proposed circuit shows better performance i.e. higher power factor at input side and lower distortion (THD) of input current than conventional one. The efficiency of proposed converter is also high. No other control circuitry is required to achieve these advantages.**

***Index Term*—Controlled Rectifier; SEPIC topology; DC-DC converter; THD**

I. Introduction

Telecommunications and computer industries have become essential part of modern society. Currently electric vehicles are occupying the place of conventional vehicles due to their environment friendly technology. DC power supply is necessary to run these devices, which consist of battery storage and charging system. Moreover renewable energy specially the solar energy is becoming a dominating power source in almost every country. DC power conditioning and processing is a major challenge in these sectors. The DC power supply is also required in operating sensitive electronic devices, separating magnets, particle acceleration, fusion, electrolyte cells, electro refining, chemical process, industrial cranes, protective and measuring instrument like relay and in DC motors. Most of the cases this DC power supply has come from AC supply. That's why power factor of AC side, AC current distortion and efficiency are the major issues to be concerned while converting AC to DC.

Uncontrolled bridge rectifiers draw non sinusoidal currents from line due to its non-linear characteristics. Several power quality problems arise at source side due to the non-linear characteristics. These problems include low power factor at input side, high distortion of input current, failure of transformers due to overheating, interference at communication channels, harmonic pollution at power system[1,2]etc. Due to harmonic pollution at power system power supply is interrupted, which cause malfunction or damage of sensitive electric device.

There are several methods to solve the problem of harmonic pollution in power system and new methods are still being investigated. One of these methods is to use a passive filter in input side. This filter consists of an inductor and a capacitor, which is bulky in size. This filter can improve THD but input power factor remains low. Moreover using bulky inductor and capacitor imposes a restriction on using it for high power application. To overcome these problems, switch mode regulator has been introduced. This method employs a single phase rectifier followed by a DC-DC converter. This method is also known as active filtering [3-11]. In this case the input filter required is relatively smaller than passive filter. Moreover improved input power factor is achieved in this method.

In this paper, a new configuration of single phase single switch controlled rectifier using SEPIC topology of high performance is reported. Very few remarkable works are done with SEPIC topology based controlled rectifier. SEPIC topology has advantage of less active components than conventional Boost followed by Buck converter. It also has other advantages like low noise operation, simple control scheme etc. Moreover SEPIC is more stable and produces less power ripples than Ĉuk converter, which frequently used in renewable energy sector [12-13].The proposed circuit chops input current by a single switch. This converter is different from conventional DC-DC converter regulated rectifiers because input current switching ensures input AC current to be almost in phase with input voltage without any control scheme. This would result almost sinusoidal input current with using of small filter and also ensure good power factor at AC side. The proposed converter provides step up/step down output DC voltage with the duty cycle control of switch. The efficiency is also high for both converters.

II. Proposed Circuit Configuration

The controlled rectifier circuit of proposed configuration is constructed with capacitors (C1, C2), inductors (L1, L2, L3, and L4), diodes and a switch (M1). As the voltage to be converted is AC, the conversion on both cycles of the input voltage is necessary. L1, L2and C1 are used for positive half cycle and L3, L4 and C2 are for negative half cycle. The

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

inductors L_pos, L_Neg, and capacitors C_pos, C_Neg form the input filter. C_Load and R_Load are the output capacitor and load respectively. Figure 1 shows the circuit construction of proposed single phase controlled rectifier.

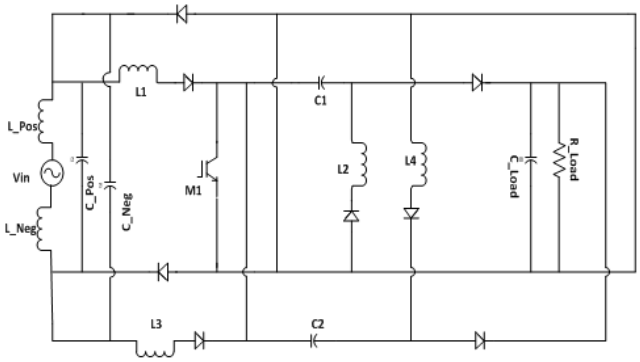


Fig. 1. Proposed SEPIC controlled Rectifier

A. Circuit Parameters

For simulation, the SEPIC inductors L1, L2, L3 and L4 have the inductance of 2mH each. The filter inductors L_pos and L_Neg have inductance of 2.5mH each. Each SEPIC capacitor C1 and C2 has capacitance of 1μF each. The filter capacitors C_pos and C_neg have the capacitance of 2μF. The capacitance of the capacitor at output C_Load is 100 μF and the load resistance R_Load is 100 Ω for reported controlled rectifier. An AC voltage having 300V peak and 50 Hz frequency is considered as supply source. An IGBT is used for switching purpose. All the diodes are considered to be ideal. The reported controlled rectifier is compared with the conventional single-phase SEPIC AC-DC configuration.

B. Simulated Waveforms

Simulated waveforms of current at input side, output DC voltage and supply AC voltage of the proposed controlled rectifier circuit at a switching frequency (f_s) of 6 kHz, duty cycle (D) of 0.3 and load resistance (R_Load) of 100 Ω are given in fig. 3.

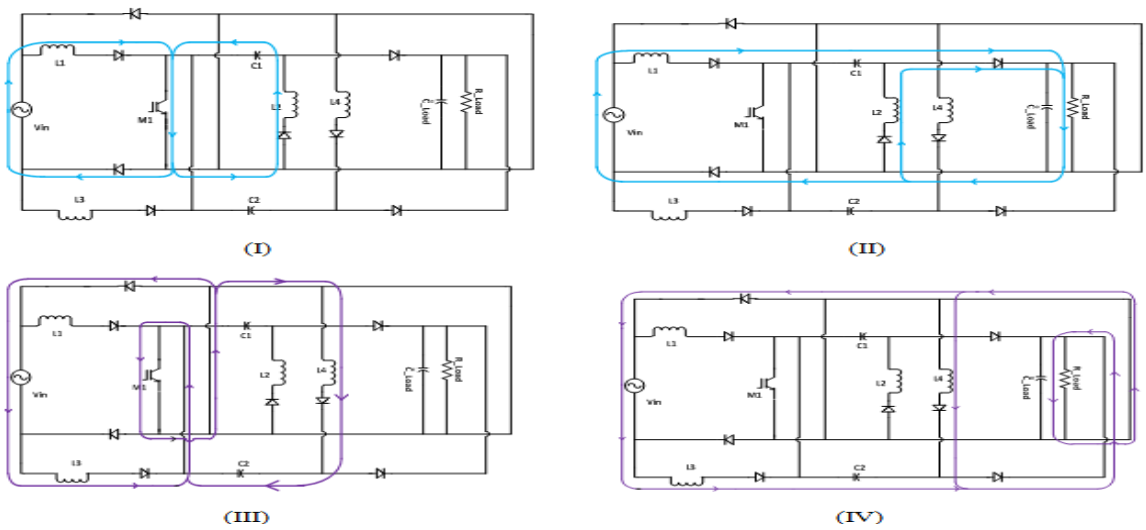


Fig. 2. Four modes of operation of proposed SEPIC controlled rectifier (I) Mode 1, when switch is ON during positive half cycle of the supply, (II) Mode 2, circuit when switch is OFF during positive half cycle of the supply,(III) Mode 3, when switch is ON during negative half cycle of the supply,(IV) Mode 4, circuit when switch is OFF during negative half cycle of the supply.

III. Modes of Operation

The proposed SEPIC controlled rectifier has four operating modes – positive and negative half cycles each with switch ON and OFF condition, as shown in Fig. 2. Modes 1 and 2 represent the positive half cycle operation with switch ON and OFF position respectively. Similarly, modes 3 and 4 are associated with the negative half cycle operation of the proposed circuit. In this proposed scheme energy transfer is associated with inductors and capacitors L1, L2, C1 and L3, L4 C2 for positive and negative half cycle of conversion respectively.

IV. Simulation Results

The proposed SEPIC topology based controlled rectifier circuit is simulated with PSIM professional version 9.0.3.400.

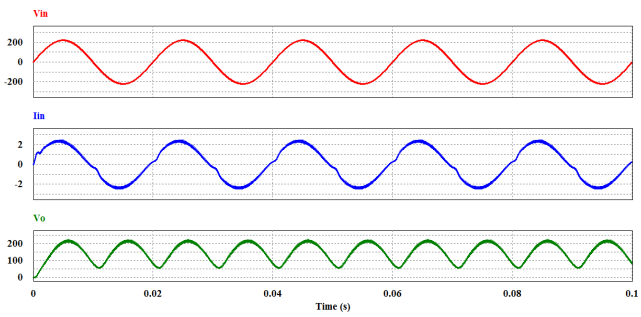


Fig. 3. Waveforms of AC Supply voltage, input AC current and output voltage of the proposed configuration at f_s = 6 kHz, D = 0.3 and R_Load = 100Ω.

Observing this figure it is found that, the current at input side has become nearly sinusoidal. Fig. 4 shows the corresponding spectrum of the input current for the proposed rectifier.

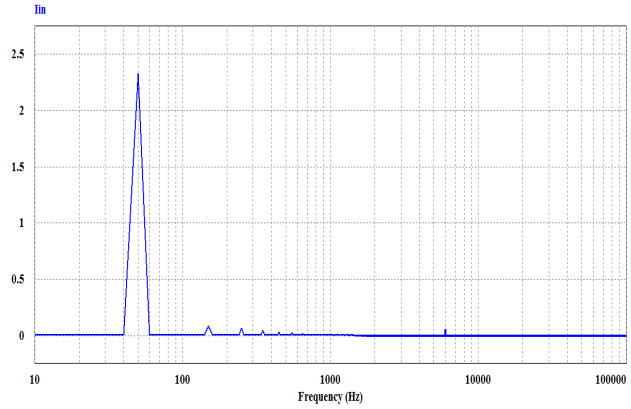


Fig. 4. Spectrum of AC Input Current of the reported configuration at $f_s = 6$ kHz, D = 0.3 and R_Load = 100Ω.

Fig. 5 represents the waveforms of supply AC voltage, current at input side and output DC voltage of the conventional SEPIC AC-DC converter. Fig.6 shows the spectrum of input AC current of conventional circuit at $f_s = 6$ kHz, D = 0.3, and R_Load = 100 Ω, under same operating condition to show the obtained improvement.

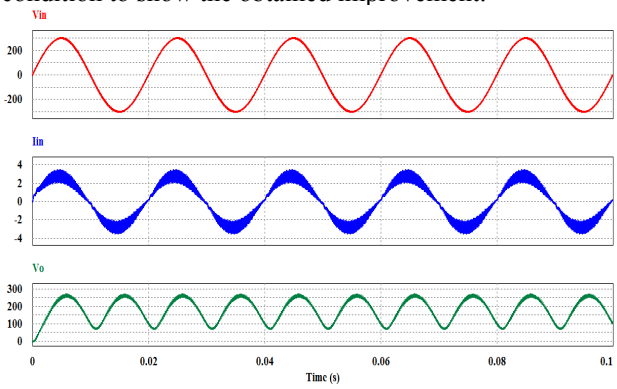


Fig. 5. Waveforms of AC Supply voltage, input AC current and output voltage of the conventional configuration at $f_s = 6$ kHz, D = 0.3 and R_Load = 100Ω.

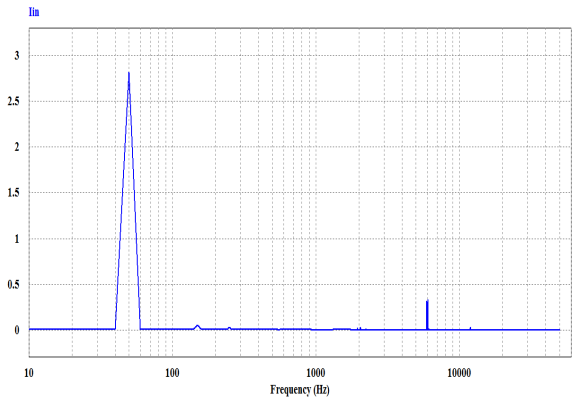


Fig. 6. Spectrum of AC Input Current of the conventional configuration at f_s=6 kHz, D = 0.3 and R_Load = 100Ω

C. *Quantitative Comparison*

The reported controlled rectifier scheme is compared with conventional one having SEPIC configuration at the output. Comparison is made in terms of THD (%) of input AC current, power factor at AC side and efficiency of conversion. The results of comparisons are shown in Fig. 7 to 9.

The bar chart shown in Fig. 7 indicates that the input current THD (%) of the proposed controlled rectifier is lesser than the conventional one for all duty cycles. The proposed rectifier exhibits higher power factor in AC side at all duty cycles compared to conventional one (Fig. 8). In terms of efficiency, the proposed controlled rectifier of SEPIC configuration has almost equal conversion efficiency as the conventional one (Fig 9).

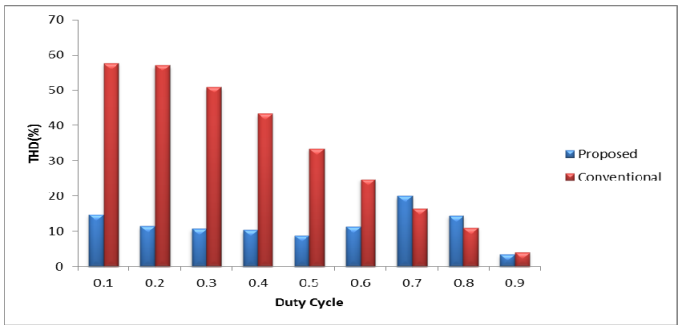


Fig. 7. Comparison of THD of AC current at input side between conventional and proposed SEPIC controlled rectifier at $f_s = 4$ kHz and R_Load = 100Ω.

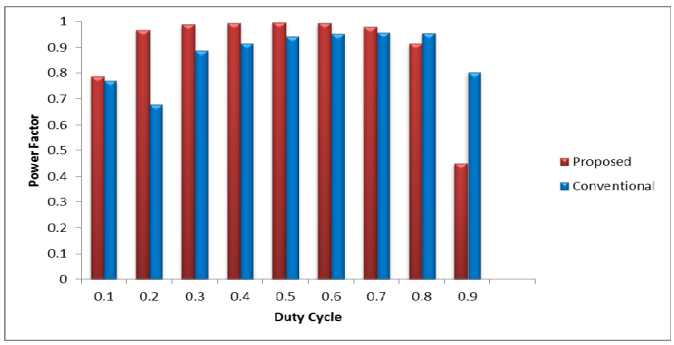


Fig. 8. Comparison of power factor of input AC side between conventional and proposed SEPIC controlled rectifier at $f_s = 4$ kHz and R_Load = 100Ω.

Fig. 9. Comparison of conversion efficiency between conventional and proposed SEPIC controlled rectifier at $f_s = 4$ kHz and R_Load = 100Ω.

D. *Performances under Variation of Load*

The proposed and conventional schemes are compared by varying load at frequency of 4 KHz and 50 % duty cycle. The load is altered from 50Ω to 500Ω and the results are presented in terms of power factor at input AC side and distortion of input AC current.

The comparative scenarios of THD (%) of input AC current, power factor at input side and conversion efficiency between the reported and the conventional scheme are represented in Fig. 10 to Fig. 12. The reported converter gives lower THD of input current with respect to the conventional circuit for all the loads. This reported circuit

also has higher power factor and conversion efficiency than the conventional single-phase rectifier for all the loads.

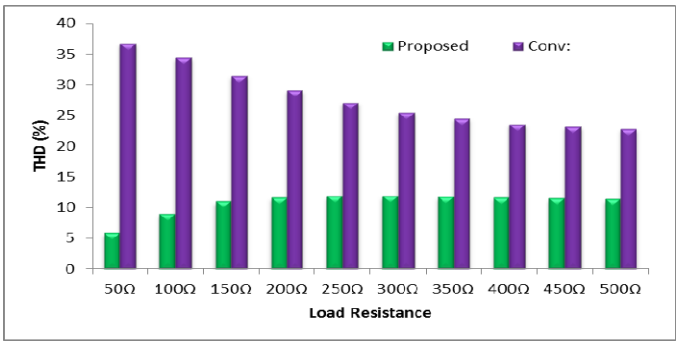


Fig. 10. Comparative scenarios of THD (%) of input AC current between reported and conventional scheme with variation of load at $f_s = 4$ kHz and D = 0.5.

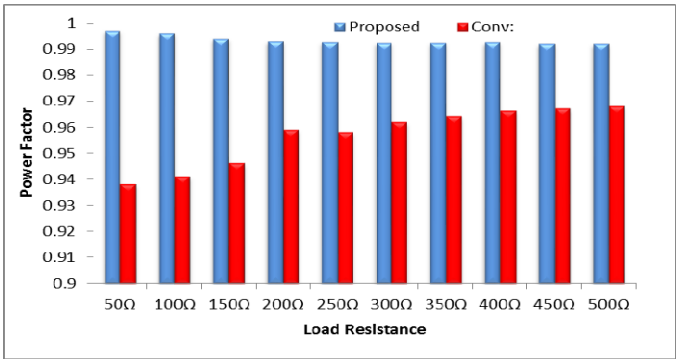


Fig. 11. Comparative scenarios of power factor at input AC side between reported and conventional scheme with variation of load at $f_s = 4$ kHz and D = 0.5.

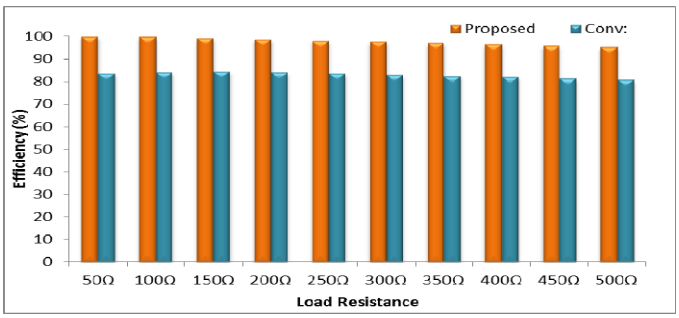


Fig. 12. Comparative scenarios of conversion efficiency between reported and conventional scheme with variation of load at $f_s = 4$ kHz and D = 0.5.

V. Conclusion

The work presented in this paper focus on the design and analysis of single phase SEPIC configuration based controlled rectifier with improved performance. The circuit chops the input current at the AC side in contrast to the conventional scheme, which chops the DC output. The fundamental advantage of the proposed circuit is that smaller sized filter is needed to achieve better performances in terms of input current THD and power factor at input side. The proposed converter provides control in terms of output voltage. Moreover single switch makes the control easier. Compared to conventional rectifier with dc-dc converter, this converter has lower distortion of input current and high power factor in AC side. This converter has only one switch which reduces the switching loss and increases efficiency. The proposed controlled rectifier is applicable for wider range of load voltage and switching frequency.

References

[1] M. M. S. Khan, M. S. Arifin, M. H. Rahaman, I. K. Amin, M. R. T. Hossain, A. H. Abedin, M. A. choudhury & M. N. Uddin, " Input switched high performance three phase Buck-Boost controlled rectifier" IEEE International Conference on Industrial Technology (ICIT), Cape town, pp. 557-562, February 2013

[2] Zixin Li,Yaohua Li,Ping wang,Hai Bin Zhu,Cong Wei Liu and Wei Xu "Control of Boost type PWM rectifier in stationary frame under unbalanced input voltage " *IEEE Trans on Power Electronics*,Vol. 25No.10,October 2010.

[3] B. Singh, B. N. Singh, A. Chandra, and K. Al-Haddad, A. Pandey, D.P. Kothari, "A review of single-phase improved power quality AC-DC converters", *IEEE Transactions on Industrial Electronics*, Vol. 50, No.5, pp 962 – 981, Oct. 2003.

[4] O. Garcia, J. A. Cobos, R. Prieto, P. Alou, and J. Uceda, "Single Phase Power Factor Correction: A Survey", *IEEE Transactions on Power Electronics,* Vol. 18, No. 3, pp. 749-755, May 2003.

[5] M. M. S. Khan, M. S. Arifin, M. R. T. Hossain, M. A. Kabir, A. H. Abedin, & M. A. Choudhury, " Input switched single phase buck and buck-boost AC-DC converter with improved power quality" 7th International Conference on Electrical & Computer Engineering (ICECE), Dhaka, pp. 189-192. December 2012

[6] J W KOLAR and T. Friedli "The essence of three phase PFC rectifier system" IEEE 33rd International Telecommunication Energy Conference(INTELEC),pp 1-27, Oct 2011.

[7] M. A. Kabir, A. H. Abedin, S. Islam & M. A. Choudhury, "Ćuk topology based single phase AC-DC converter wih low THD and high power factor" IECON 2011-37th Annual Conference on IEEE Industrial Electronics Society, Melbourne, pp. 1256-1261,November 2011.

[8] M. R. T. Hossain, M. H. Rahaman, M. S. Arifin, M. A. Choudhury, A. H. Abedin, & M. N. Uddin, "Input switched single phase high performance bridgeless AC-DC Zeta converter".IEEE International conference on Power Electronics, Drives and Energy Systems (PEDES), Bengaluru, pp. 1-5, December 2012

[9] M. S. Patil, and S. P. Patil, "Single-Phase Buck-Type Power Factor Corrector with Lower Harmonic Contents in Compliance with IEC 61000-3-2" International Journal of Engineering Science and Technology, Vol. 2(11), pp. 6122-6130,2010

[10] P. F. Melo, R. Gules, E. F. R. Romaneli, and R. C. Annunziato, "A Modified SEPIC Converter for High-Power-Factor Rectifier and Universal Input Voltage Applications", *IEEE Transactions on Power Electronics*, Vol. 25, No. 2,pp. 310-321,February 2010

[11] N. Hasan, M. Hossain, M. A. Kabir, M. A. Choudhury & M.N. Uddin, "AC side switched active filter based high performance three-phase boost rectification with energy recovery snubber", Industry Applications Society Annual Meeting (IAS)",pp. 1-8. ,2012

[12] Jacob, Tony, and S. Arun. "Maximum Power Point Tracking of hybrid PV and wind energy systems using a new converter topology." International Conference on Green Technologies (ICGT) , pp. 280-287, 2012.

[13] Betten, John. "Benefits of a coupled-inductor SEPIC converter." Power Management, 2009.

ASSOCIATIVE MEMORY ALGORITHM for VISUAL PATTERN RECOGNITION with MEMRISTOR ARRAY and CMOS NEURON

Anik Chowdhury,* Mrinmoy Sarkar, Aqeeb Iqbal Arka and A. B. M. Harun-ur Rashid
Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology
Dhaka-1205, Bangladesh
*anik.chy.113@gmail.com

***Abstract*—This paper presents the development of a neuromorphic system for visual pattern recognition. Associative memory algorithm has been used to recognize specific patterns and the method is implemented with discrete circuit elements that use memristors as the synapse. Weight of the synapses between the inputs and output neurons are adjusted by memristors. The amount of the neural block depends on how many patterns and each block contains 30 weighted synapses connected to the output. Each input corresponding to a pixel in a 6×5 pixel image generates voltage pulses according to the pixel value. The input voltage pulses weight are then maintained by the memristors and integrated by the output neurons. Each of the output of neurons is compared with a fixed threshold voltage which will determine whether the output bit is 0 or 1. Successfully demonstration of the system has been done by training and recognizing images of numbers from 1 to 8.**

***Index Terms*—Artificial neural network, CMOS neuron, memristor, multiple encoding mode, neuromorphic, pattern recognition, weighted synapse**

I. Introduction

Artificial neural networks are the most revered implementation of bio-inspired engineering which is capable of imitating information processing system [1]–[3]. Taking ideas from sentient species' neuron which is defined as nodes, ideas have been brought up for computational neural networks [4] that are able to simulate the inner function of a neuron. A class of statistical models may be defined as "neural" if they contain a set of adaptive weights which is numerical parameters tuned by certain learning algorithms ability to approximate non-linear function of their inputs. And the adaptive weights are to be considered as connection strengths between nodes. A neural network has three major parts: an input layer, hidden layer(s) and one output layer [5]. Synapses which are the main memory zone work as a connection between the biological neural network. Recently a passive device named memristor has been developed and successfully used as a synapse [6]–[8]. However, most of the neuromorphic systems [9] utilizing the memristor have been implemented in software and verified with only simulations.

The motivation of this work is to develop a memristor- based ANN for pattern recognition [10]–[12]. The algorithm is called associative memory [13]–[15]. At first weights of the ANN is converted into resistance values. Then the corresponding memristors are trained to certain resistance values. And then corresponding image pixels are fed into the input synapses to check whether it can make the desirable output or not. Synapses are implemented with memristors and the output neuron is implemented with CMOS circuits. As a demonstration, the system is tested with images of numbers from 1 to 8.

The rest of the paper is organized as follows. Section II introduces the proposed system with detailed explanations of different blocks in the system. The system operation, programming technique and proposed learning rule are shown in Section III. Experimental results are shown in Section IV and conclusions are drawn in Section V.

II. Proposed System

Overall system can be divided into two main functional blocks. The first block consisting of weighted synapses which are replaced by memristors and the second block is the output neurons which is replaced by the op-amp. In the following subsections ANN circuit realization, multiple training encoding mode, the memristor and the output neuron will be discussed in detail.

A. ANN Circuit Realization

For solving the pattern recognition problem a network should be developed. Larger implementation of the circuit described in [16], [17] is used in this paper. A simplified version of the circuit is shown in Fig. 1. Here P_1, P_2, P_n are the pixel inputs of the system, R_{M1}, R_{M2}, R_{Mn} are the weights of the synapses which are basically the memristors, and R_F is called the feed forward resistance. However, this

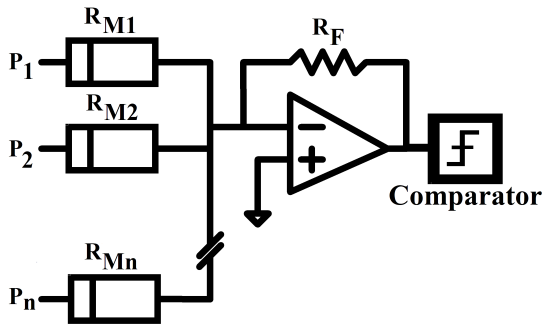


Fig. 1. Basic circuit of neural network

circuit can only use positive weights. All the ANNs can not be solved with positive weights. To overcome this issue, overall network is modified as described in [16], [17]. Fig. 2 shows the modified version of the circuit. As shown in Fig. 2, inputs are connected to the memristors and their corresponding inverted outputs are connected to R_N. This block can generate negative weights.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

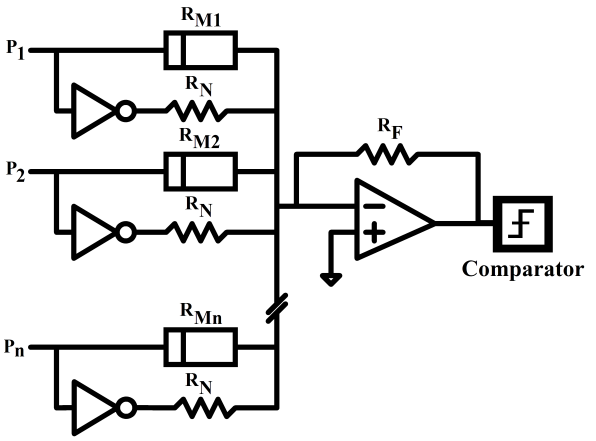


Fig. 2. Modified circuit of neural network

B. Multiple Training Encoding Mode

As described earlier, weighted synapses can be converted into resistance. For determining the resistance values, WANG et al.'s multiple training encoding strategy [18] is used in this paper. Using this strategy, the 2^N number of input patterns can be converted into a N number of output bits. If the input patterns are defined in a bidirectional array X_i and output patterns are defined in a bidirectional array Y_i then the correlation matrix G which elements are the weights of the input synapses can be defined as given in (1) where q_i are set of weight vectors. The overall strategy is given in [18].

$$G = \sum_{i=1}^{N} q_i X_i^T Y_i \tag{1}$$

C. Memristor a Synapse

A memristor is a two-terminal device that can change its resistance according to the integral of current flowing through it [19]. Various types of memristors have been found [20] recently which can model synaptic weights. Using the equations described in [17], values of R_M, R_N and R_F can be calculated as given in (2), (3), (4).

$$R_M = \frac{R_F R_N}{R_F - G \times R_N} \tag{2}$$

$$R_N = R_{MHigh} - \frac{R_{MHigh} G_{High} (R_{MHigh} - R_{MLow})}{R_{MHigh} G_{High} - R_{MLow} G_{Low}} \tag{3}$$

$$R_F = \frac{R_N (R_{MHigh} G_{High} - R_{MLow} G_{Low})}{R_{MHigh} - R_{MLow}} \tag{4}$$

Where G_{High} and G_{Low} are the highest and lowest weights of the ANN which are the highest and the lowest values of the correlation matrix G, while R_{MLow} and R_{MHigh} are the highest and the lowest optimal range of the memristors which are assigned as 10KΩ and 100KΩ. Following the equations given in (3) and (4), R_N and R_F are calculated as 18.182KΩ and 41.778KΩ respectively.

D. CMOS opamp as a Neuron

An op-amp is used as CMOS neuron [21] in this ANN circuit. Operational amplifiers are an integral part of the most analog circuits and systems. Designing amplification block to achieve high gain is not only the concerned area for the modern op-amp, it also includes proper biasing circuitry along with better compensation technique for stable operation.

III. System Operation

This section explains the overall system operation followed by the programming technique, proposed learning rule, and overall system operation.

A. Programming Technique

A lot of memristor models existed for simulation of the memristor. In this paper generalized memristor model [22] is used for simulation. There are two modes in programming a memristor which are writing which will decrease the device resistance and erasing which will do the opposite [22]. In order to make sure that the device is not under stress $1us$ pulse which amplitude is higher than the threshold value is used in this paper. An amplitude pulse which has a much smaller value than the threshold voltage is used for reading the current value of memristor. 100mV, $15us$ pulse is chosen to be sufficiently small in order to avoid disturbing the current state of the device while reading it.

B. Proposed Learning Rule

For programming a single memristor a MATLAB script has been written to provide the voltage inputs to automatically drive the memristor to the desired resistance. The script will keep generating programming pulses with amplitudes proportional to the difference between the current resistance R_i of the memristor and the desired value R_d. According to the difference value, for the larger difference a bigger amplitude pulse will be sent and for the smaller difference, a smaller pulse will be sent. When the memristor almost reaches the desired value amplitude will be getting smaller. For keeping the device safe programming pulse value is limited below V_{max}. In short, pulses with adapted amplitude will be sent to programme the memristor, and then read the memristor to check whether its resistance is within a certain tolerance which is 15Ω of the desired value. Overall algorithm is given in [17].

C. Overall System Operation

At first, the memristors should be trained to their desired resistance value. Then they can be added to the main network. For this reason, a switching block will be necessary that can

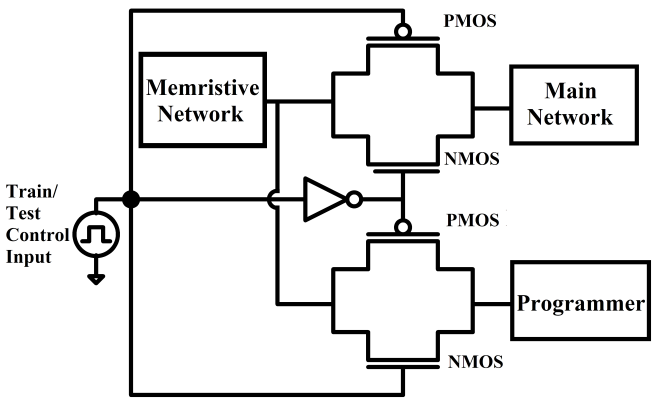


Fig. 3. Switching block using transmission gate

switch the memristors to the main circuit after the training period. Basically, switches are made from transmission gates [23]. The main function of the switching block is to switch the memristors either in programming network or in the main network like shown in Fig. 3.

IV. Experimental Results

The simulation is done in LTSpice. TSMC180nm process [24] has been used for NMOS and PMOS specifications to design the CMOS switch blocks. For solving the pattern recognition problem, images of numbers from 1 to 8 are used. As discussed earlier, eight input patterns can be encoded into three output bits. Thirty pixels and their corresponding inverted outputs are fed into the main circuit. All the input pulses have been generated using MATLAB. The overall circuit diagram is shown in Fig. 4. In Fig. 4 n varies from

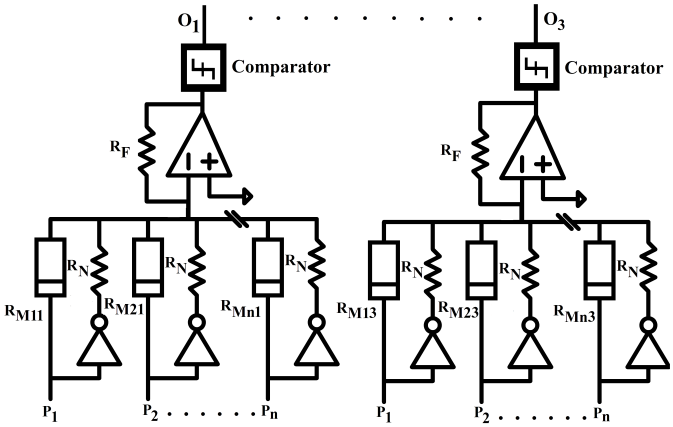


Fig. 4. Circuit realization of the system

1 to 30 for thirty pixels. Training images are shown in Fig. 5. As the algorithm works for bipolar input and output,

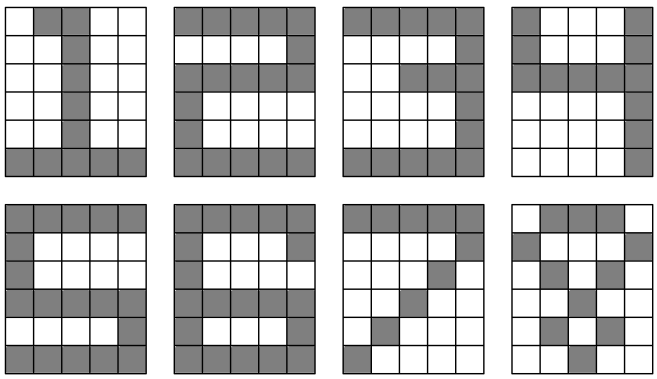


Fig. 5. Training images

a positive value 100mV is assigned for gray pixels and a negative value -100mV is assigned for white pixels. At first using the algorithm weights of the synapses are determined and then these weights are converted into resistance values as given in (2) with which the memristors will be trained. After training the memristor to their desired values, inputs, and their corresponding outputs are fed into the main circuit to see whether the circuits work or not. The Output from the integrators' is shown in Fig. 6. Three different output bits are used to express the eight different patterns as shown in Fig. 5. The first pattern which is "1" is expressed by binary output bit "000". The second pattern which is "2" is expressed by binary output bit "001". And thus, the last pattern which is "8" is expressed by binary output bit "111". For bipolar output bit, "1" is defined as positive voltage 4V and "0" is defined as negative voltage -4V. V_{th} is set at 0

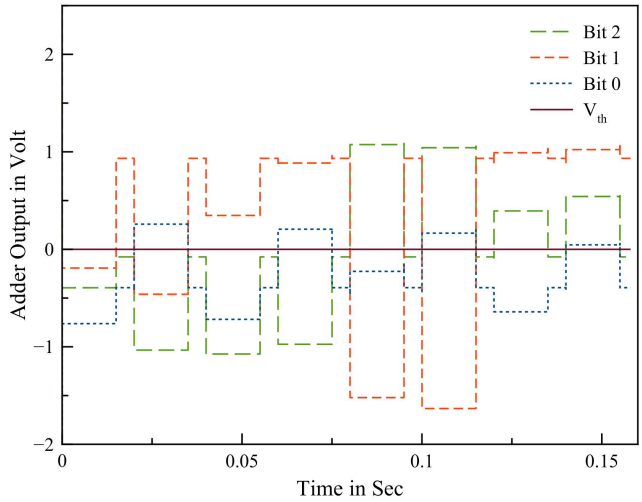


Fig. 6. Output signals from the integrators'

V to identify the output bits as shown in Fig. 6. At first

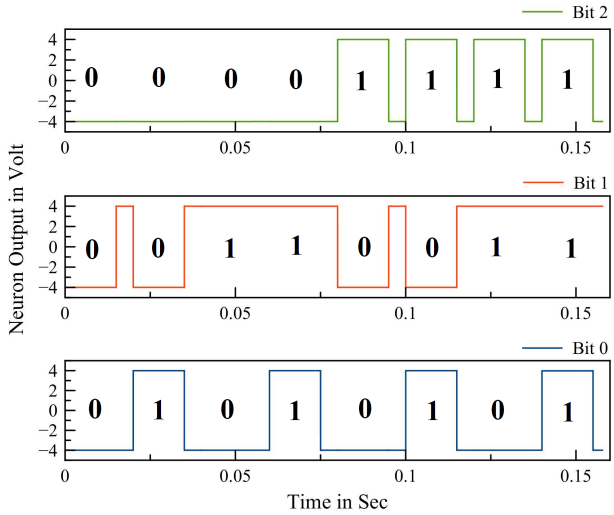


Fig. 7. Output from the output neurons

pattern "1" is given as input and its corresponding output "000" which is -4V in all three bits can be seen in the output neuron. The second pattern "2" is given as input after 5ms delay and its corresponding output "001" can be seen in the output. There is always a 5ms gap between two consecutive patterns. Thus, all the patterns are given as input sequentially and from the Fig. 7 it can be shown that the overall system response perfectly. For performance analysis, the system has been trained with noise free images and again with noisy input images. Input image with noise pixels is shown in Fig. 8. Error in the image has been calculated as the ratio of the number of noise pixel to the total number of pixels in the image. And the performance is calculated as the ratio between the no. of correctly recognized images to the total no. of test images.

$$\%of\ noise\ in\ the\ image = \frac{no.\ of\ noise\ pixels}{no.\ of\ total\ pixels\ in\ the\ image} \times 100\%$$

$$\%performance\ index = \frac{no.\ of\ correctly\ recognized\ images}{no.\ of\ total\ test\ images} \times 100\%$$

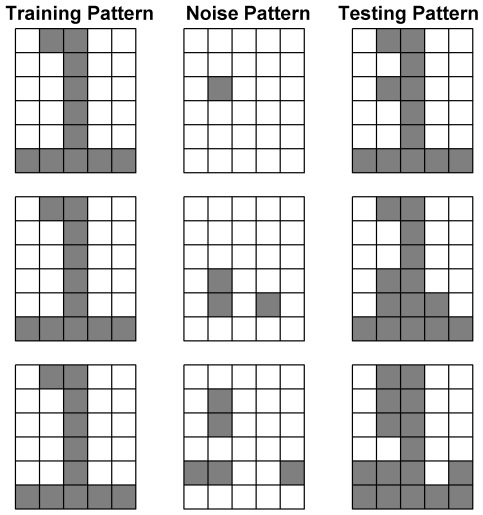


Fig. 8. Noisy test image

Overall performance of the system is given in Table I. The recognition rate drops rapidly as noise level increases.

TABLE I
PERFORMANCE OF THE SYSTEM

% of noise in the image	performance index (proposed method)	performance index (unsupervised method) [8]
0%	100%	100%
3.33%	100%	95%
10%	95%	85%
16%	62.5%	55%

For example, the recognition rate is reduced to 62.5% with 16% of the noise level. The main reason for incorrect image recognition is the variation of total resistance of memristance values that are not trained for input images. As a result, current going through the neuron changes, resulting in erroneous recognition.

V. CONCLUSIONS

The process of building an ANN circuit with memristors which can be used for pattern recognition has been presented. The proposed method is very simple and can be easily implemented in the circuit for a smaller number of patterns and pixels. Performance analysis result is satisfactory which is more immune to noise pixels compared to the method using unsupervised learning [8]. However, with increasing number of patterns and pixels implementation of this idea would be more difficult as it will require a device with high computational capability to train a lot memristors.

REFERENCES

[1] M. Turank, "Packet filtering by artificial neural network," in *Proceedings of the International Conference on Military Technologies (ICMT)*, pp. 1-4, 2015.

[2] M. lkuar, A. H. Iika and A. ifci, "Classification of breast cancer data with harmony search and back propagation based artificial neural network," in *Proceedings of the 22nd Signal Processing and Communications Applications Conference (SIU)*, pp. 762-765, 2014.

[3] T. W. Cenggoro, A. H. Kridalaksana, E. Arriyanti and M. I. Ukkas, "Recognition of a human behavior pattern in paper rock scissor game using backpropagation artificial neural network method," in *Proceedings of the 2nd International Conference on Information and Communication Technology (ICoICT)*, pp. 238-243, 2014.

[4] H. Zhang, Z. Wang and D. Liu, "A Comprehensive Review of Stability Analysis of Continuous-Time Recurrent Neural Networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, pp. 1229-1262, 2014.

[5] B. Hao and X. Dai, "A combined neural network design and application," in *Proceedings of the Second International Conference on Communication Systems, Networks and Applications (ICCSNA)*, vol. 1, pp. 211-214, 2010.

[6] X. Wu, V. Saxena and K. Zhu, "A CMOS spiking neuron for dense memristor-synapse connectivity for brain-inspired computing," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, pp. 1-6, 2015.

[7] S. P. Adhikari, H. Kim, R. K. Budhathoki, C. Yang and L. O. Chua, "A Circuit-Based Learning Architecture for Multilayer Neural Networks With Memristor Bridge Synapses," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 62, pp. 215-223, 2015.

[8] M. Chu, B, Kim, S. Park, H. Hwang, M. Jeon, B. H. Lee and B.-G. Lee, "Neuromorphic Hardware System for Visual Pattern Recognition With Memristor Array and CMOS Neuron," *IEEE Transactions on Industrial Electronics*, vol. 62, pp. 2410-2419, 2015.

[9] L. Xia, B. Li, T. Tang, P. Gu, X. Yin, W. Huangfu, P.-Y. Chen, S. Yu, Y. Cao, Y. Wang, Y. Xie and H. Yang, "MNSIM: Simulation platform for memristor-based neuromorphic computing system," *Proceedings of the Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pp. 469 - 474, 2016.

[10] P. Lorenzi, V. Sucre, G. Romano, R. Rao and F. Irrera, "Memristor based neuromorphic circuit for visual pattern recognition," in *Proceedings of the International Conference on Memristive Systems (MEMRISYS)*, pp. 1-2, 2015.

[11] Y. Yang, J. Mathew and D. K. Pradhan, "Matching in memristor based auto-associative memory with application to pattern recognition," in *Proceedings of the 12th International Conference on Signal Processing (ICSP)*, pp. 1463-1468, 2014.

[12] Y. Kaneko, Y. Nishitani, M, Ueda and A. Tsujimura, "Neural network based on a three-terminal ferroelectric memristor to enable on-chip pattern recognition," in *Proceedings of the Symposium on VLSI Technology (VLSIT)*, pp. 238-239, 2013.

[13] R. Danilo, H. Jarollahi, V. Gripon. P. Coussy, L. C.-Canencia and W. J. Gross, "Algorithm and implementation of an associative memory for oriented edge detection using improved clustered neural networks," in *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 2501-2504, 2015.

[14] H. Jarollahi, N. Onizawa, V. Gripon, T. Hanyu and W. J. Gross, "Algorithm and architecture for a multiple-field context-driven search engine using fully-parallel clustered associative memories," in *Proceedings of the IEEE Workshop on Signal Processing Systems (SiPS)*, pp. 1-6, 2014.

[15] B. Boguslawski, F. Heitzmann, B. Larras and F. Seguin, "Energy-Efficient Associative Memory Based on Neural Cliques," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 63, pp. 376-380, 2016.

[16] A. Rothenbuhler, "A memristor-based neuromorphic computing application," M. Eng. thesis, Boise State University, Boise, Idaho, May 2013.

[17] T. T. T. Tran, "Simulations of artificial neural network with memristive devices," M. Eng. thesis, Boise State University, Boise, Idaho, Dec. 2012.

[18] S. Rajasekaran and G. V. Pai, *Neural Networks, Fuzzy Logic, and Genetic Algorithms Synthesis and applications*, Prentice-Hall of India Private Limited, New Delhi - 110001, 2003.

[19] L. O. Chua and S. M. Kang, "Memristive devices and systems," *Proceedings of the IEEE*, vol. 64, pp. 209-223, 1976.

[20] F. Corinto, P. P. Civalleri and L. O. Chua, "Classification of mem-devices," in *Proceedings of the IEEE 15th International Conference on Nanotechnology (IEEE-NANO)*, pp. 629-932, 2015.

[21] X. Wu, V. Saxena and K. Zhu, "A CMOS spiking neuron for dense memristor-synapse connectivity for brain-inspired computing," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, pp. 1-6, 2015.

[22] C. Yakopcic, T. M. Taha, and R. E. P. Guru Subramanyam, *Advances in Neuromorphic Memristor Science and Applications*, Springer, 2015.

[23] F. Carbognani, F. Buergin, N. Felber, H. Kaeslin, and W. Fichtner, "Transmission gates combined with level-restoring cmos gates reduce glitches in low-power low-frequency multipliers," *EEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 16, pp. 830-836, 2008.

[24] S. B. Patil and R. D. Kanphade, "Design and analysis of a 2.4 GHz fully integrated 1.8V power amplifier in TSMC 180nm CMOS RF process for wireless communication," in *Proceedings of the International Conference on VLSI Systems, Architecture, Technology and Applications (VLSI-SATA)*, pp. 1-5, 2015.

Grid Current Regulation of a Three-Phase CSI Using a Backstepping Controller

T. K. Roy*, M. F. Pervej, and F. K. Tumpa

Dept. of Electronics & Telecommunication Engineering
Rajshahi University of Engineering & Technology, Rajshahi-6204, Bangladesh
*Email: roy.kanti03@gmail.com

Abstract—**This paper presents a nonlinear control strategy for a current source inverter (CSI) based three-phase grid-connected photovoltaic (PV) system to control the active and reactive power injection into the grid. The proposed control structure for the CSI based three-phase grid-connected PV system is composed of an outer current loop which is responsible for controlling the DC-side inductor current and the inner current control loop is responsible for controlling injected current into the grid. The proposed controller is designed recursively and the overall stability of the whole system is analyzed based on the formulation of control Lyapunov functions (CLFs) at different stages of the design process. Moreover, to extract the maximum power from PV arrays, a maximum power point tracker (MPPT) is used to ensure that PV arrays are operating at the maximum power point (MPP). Finally, the effectiveness of the proposed controller is tested on a CSI based three-phase grid-connected PV system under different atmospheric conditions. Simulation results show that the proposed controller can effectively meet the desired control objectives.**

Index Terms*—*Active power, backstepping controller, control Lyapunov function, current source inverter, reactive power

I. INTRODUCTION

Recently, photovoltaic (PV) systems have received a great deal of attention because of their distinctive advantages such as simplicity of allocation, no fuel cost, less maintenance cost, and widely distributed throughout the earth [1]-[3]. Another advantage is that it could be used for both modes such as stand-alone and grid-connected. However, due to the encouragements policies, most of the PV systems are installed for the grid connected plants [3]. The main purpose of a grid-connected PV system is to transfer maximum power into grid with a power factor as close to unity despite changing in atmospheric conditions. To achieve such objective, each grid-connected PV system has to perform two essential functions - one is to ensure that PV modules are operated at the maximum power point (MPP) tracking to extract the maximum power from PV arrays and another one is to inject a harmonic free sinusoidal current into the grid to improve the active and reactive power quality.

The PV array module shows a nonlinear voltage-current characteristic due to variations of solar irradiance and ambient temperature and consequently the PV output power is unpredictable. Thus, the MMP tracker must resist with a nonlinear and time-varying system to extract maximum power from the PV cell. To extract maximum power, many MPP tracking control algorithms have been proposed in the literature [5], [6] to operate the PV array at the MMP point. A fuzzy logic controller is proposed in [7] to control the MPP of a PV array system under changing atmospheric conditions. The MPPT control strategies based on perturb and observe (P&O) method is proposed in [8], [9]. However, P&O method has some limitations such as slow response speed, oscillations of the PV array output power and fails to quickly track the MPP under a rapidly change in atmospheric conditions. To overcome these problems, an incremental conductance (IC) method is proposed in [10] due to its conciseness and high tracking efficiency which is also used in this paper.

Due to the switching functions of DC-DC converters, CSI, and diode current of the PV cell, a grid-connected PV system is highly nonlinear. The CSI based PV system has significant advantages over VSI as discussed in [11]. Thus, to maintain the stable operation of a CSI based grid-connected PV system under the changes in atmospheric conditions a proper controller must be installed with the system. Generally, due to the simplicity and ease of installation linear controllers are used [12]. However, this linear controller is not suitable when the operating points of the system varied due to variations in atmospheric conditions. Thus, to ensure the stable operation of a grid-connected PV system over a wide range of operating point, the design of the nonlinear controller is very important. In the literature, several advanced control techniques are proposed to design the nonlinear controllers such as sliding mode controller (SMC) [13], model predictive controller (MPC) [14], feedback linearization controller (FBLC) etc. Among these, FBLC has been widely used to design a nonlinear controller for a grid connected PV system. To analyze the dynamical stability of a three-phase grid-connected PV system, a nonlinear FBLC is proposed in [16] to regulate the inverter output current injected into the grid. Though FBLC can maintain the system stability over a wide range of operating point, for successful implementation the system parameters and model should be known precisely. Practically, it is quite difficult to know the exact values of the system parameters. Moreover, FBLC cancels useful nonlinearities during the transformation from the nonlinear system to a fully or partially linearized system.

To avoid the cancellation of useful nonlinearities, a nonlinear backstepping controller is proposed in [4], [15], [18]-[20], which stabilizes the nonlinear terms instead of complete eliminations. To control the injected current into the grid, an SMC is proposed in [14], [17], to ensure the stable operation of a three-phase grid-connected PV system, where the controller is designed based on the time-varying sliding surface. Though SMC has low sensitivity to parameter variations, the selection of the time-varying surface is a difficult task. Moreover, the controllers as proposed in the aforementioned literature are designed only for the active power controlled, whereas reactive power control is not taken into account. Moreover, the controller is designed for a voltage source inverter (VSI) based three-phase grid-connected PV system.

This paper aims to control the CSI output current to inject the active and reactive power into the grid. To achieve these control objectives, the controller is designed using the nonlinear backstepping technique. The performance of the designed controller is evaluated on a CSI based three-phase grid-connected PV system under the changes in atmospheric conditions.

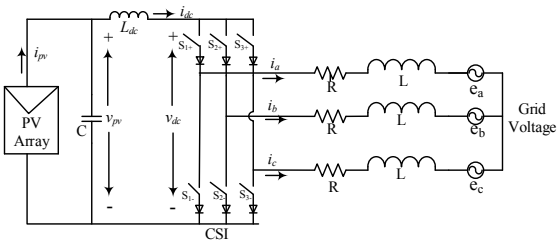


Fig.1. CSI based three-phase grid-connected PV system [11]

II. CSI Based PV System Modeling

The schematic configuration of a three-phase grid-connected PV system is shown in Fig. 1. In this configuration, a PWM based CSI is used whose output is connected to a three-phase grid via an output L filter to inject a purely sinusoidal current into the grid. From Fig. 1, it can also be seen that the CSI is a six-pulse inverter utilizing IGBT switches and to increase the reverse voltage capacity diodes are connected in series. Now by applying the Kirchhoff's law, the whole dynamical model of the system in abc frame can be represented by the following differential equations:

$$L\frac{di_a}{dt} = -Ri_a - e_a + \frac{v_{dc}}{3}(2S_a - S_b - S_c) \quad (1)$$

$$L\frac{di_b}{dt} = -Ri_b - e_b + \frac{v_{dc}}{3}(-S_a + 2S_b - S_c) \quad (2)$$

$$L\frac{di_c}{dt} = -Ri_c - e_c + \frac{v_{dc}}{3}(-S_a - S_b + 2S_c) \quad (3)$$

where S_a, S_b, and S_c are the input switching signals of PWM. The dynamical model as represented by equations (1)–(3) is nonlinear and time variant due to the input switching signals and diode current i_{pv}. In order to design the proposed controller, this model needs to be transformed into a direct-quadrature (dq) frame rotating at the angular frequency of the grid line ω.Thus, using the dq transformation; a nonlinear and time-invariant model can be written as:

$$\dot{I}_d = -\frac{R}{L}I_d + \omega I_q - \frac{1}{L}E_d + \frac{v_{dc}}{L}S_d \quad (4)$$

$$\dot{I}_q = -\frac{R}{L}I_q - \omega I_d - \frac{1}{L}E_q + \frac{v_{dc}}{L}S_q \quad (5)$$

where the symbols have their usual meaning which can be found in [18]. Now the DC-side dynamics of the system can be described by the following differential equations:

$$C\frac{dv_{pv}}{dt} = i_{pv} - i_{dc} \quad (6)$$

$$L_{dc}\frac{di_{dc}}{dt} = v_{pv} - v_{dc} - i_{dc}R_{dc} \quad (7)$$

where R_{dc} is the inductor interior resistance. In the synchronous rotating dq frame, the active and reactive power can be expressed in terms of I_d and I_q as follows:

$$\begin{cases} P_g = \frac{3}{2}\left(E_dI_d + E_qI_q\right) \\ Q_g = \frac{3}{2}\left(E_dI_q - E_qI_d\right) \end{cases} \quad (8)$$

During the steady-state condition, the average value of E_q is equal to zero [11]. As a result, equation (8) can be rewritten as:

$$\begin{cases} P_g = \frac{3}{2}E_dI_d \\ Q_g = \frac{3}{2}E_dI_q \end{cases} \quad (9)$$

From equation (9), it is clear that if the amplitude of E_d is constant, the maximum active and reactive powers delivered into the grid can be maintained by controlling the d-and q-axis currents of the grid. If it is assuming that the losses due to switching actions of the inverter are negligible then the following relation can be obtained based on power balance between the DC side and AC side of the inverter $P_{dc} = v_{dc}i_{dc} = P_g$. Using this power balance equation, equation (7) can be written as:

$$\frac{di_{dc}}{dt} = \frac{v_{pv}}{L_{dc}} - \frac{3}{2L_{dc}i_{dc}}E_dI_d - i_{dc}R_{dc} \quad (10)$$

Now the complete dynamical model of a CSI based three-phase grid-connected PV system can be described by the following set of differential equations:

$$\frac{dv_{pv}}{dt} = \frac{1}{C}(i_{pv} - i_{dc}) \quad (11)$$

$$\frac{di_{dc}}{dt} = \frac{v_{pv}}{L_{dc}} - \frac{3}{2L_{dc}i_{dc}}E_dI_d - i_{dc}R_{dc} \quad (12)$$

$$\frac{dI_d}{dt} = -\frac{R}{L}I_d + \omega I_q - \frac{1}{L}E_d + \frac{v_{dc}}{L}S_d \quad (13)$$

$$\frac{dI_q}{dt} = -\frac{R}{L}I_q - \omega I_d - \frac{1}{L}E_q + \frac{v_{dc}}{L}S_q \quad (14)$$

Here, equations (11)-(14) represent the complete dynamical model and the proposed controller is designed based on this model in the following section.

III. Proposed Controller Design

The control structure of a three-phase grid-connected PV system is composed of two parts. The first one is the MPPT control, which is used to extract the maximum power from PV arrays and the second one is the inverter switching control, which is used to achieve the following objectives:

- Inject a pure sinusoidal current into the grid with lower values of total harmonic distortion (THD), and
- Control the active and reactive power injection into the grid.

Thus, in order to resolve the aforementioned concerns, the proposed nonlinear backstepping controller design procedure is elaborately discussed as follows.

Step 1: According to the design purposes, the first tracking error between v_{pv} and v_{pvref} can be defined as:

$$e_1 = v_{pv} - v_{pvref} \quad (15)$$

The time derivative of e_1 using equation (11) can be written as:

$$\dot{e}_1 = \frac{1}{C}\left(i_{pv} - i_{dc}\right) - \dot{v}_{pvref} \quad (16)$$

In this case, the first CLF can be written as:

$$W_1 = \frac{1}{2}e_1^2 \quad (17)$$

Inserting the value of $\dot{e}_1$ from equation (16), the derivative of W_1 can be written as:

$$\dot{W}_1 = e_1\{\frac{1}{C}\left(i_{pv} - i_{dc}\right) - \dot{v}_{pvref}\} \quad (18)$$

At this point, the stabilizing function of i_{dc} needs to be selected in such a way that $\dot{W}_1$ would be negative semi-definite, i.e., $\dot{W}_1 \leq 0$. To do that, the synthetic value for i_{dc} can be chosen as:

$$\alpha = C(i_{pv} - \dot{v}_{dcref} + k_1e_1) \quad (19)$$

where $k_1 > 0$ is a positive constant which is used to tune the output responses. After selecting this, equation (18) can be written as:

$$\dot{W}_1 = -k_1e_1^2 \leq 0 \quad (20)$$

which is negative semi-definite. The remaining errors dynamics are discussed in the following steps.

Step 2: Since α is not the actual control input, the DC-link inductor current tracking error can be defined as follows:

$$e_2 = I_{dc} - \alpha \quad (21)$$

The time derivative of e_1 using equation (12) can be written as:

$$\dot{e}_2 = \frac{v_{pv}}{L_{dc}} - \frac{3}{2\,L_{dc}I_{dc}} E_d I_d - \frac{R_{dc}}{L_{dc}} I_{dc} - \dot{\alpha} \quad (22)$$

In order to stabilize the DC-link inductor current tracking error to zero, the CLF can be chosen as:

$$W_2 = W_1 + \frac{1}{2} e_2^2 \quad (23)$$

Inserting the value of $\dot{e}_1$ from equation (22), $\dot{W}_2$ can be written as:

$$\dot{W}_2 = -k_1 e_1^2 + e_1\left(\frac{v_{pv}}{L_{dc}} - \frac{3}{2\,L_{dc}I_{dc}} E_d I_d - \frac{R_{dc}}{L_{dc}} I_{dc} - \dot{\alpha}\right) \quad (24)$$

At this point, the stabilizing function of I_d needs to be selected in such a way that $\dot{W}_2$ would be negative semi-definite, i.e., $\dot{W}_2 \leq 0$. In such cases, the synthetic value for I_d can be chosen as:

$$\alpha_1 = \frac{2I_{dc}}{3E_d} L_{dc}\left(\frac{R_{dc}}{L_{dc}} I_{dc} + \dot{\alpha} - \frac{v_{pv}}{L_{dc}} + k_2 e_2\right) \quad (25)$$

After selecting this, equation (24) can be written as:

$$\dot{W}_2 = -k_1 e_1^2 - k_2 e_2^2 \leq 0 \quad (26)$$

which is negative semi-definite. Hence, the dynamics of inductor current tracking error is stabilized.

Step 3: In this step, the *d*-axis (active) current tracking error can be defined as:

$$e_3 = I_d - \alpha_1 \quad (27)$$

Using equation (13), $\dot{e}_3$ can be written as follows:

$$\dot{e}_3 = -\frac{R}{L} I_d + \omega I_q - \frac{1}{L} E_d + \frac{v_{dc}}{L} S_d - \dot{\alpha}_1 \quad (28)$$

Define the *q*-axis (reactive) current tracking error as follows:

$$e_4 = I_q - I_{qref} \quad (29)$$

Inserting the value of $\dot{I}_q$, the dynamics of e_4 can be written as:

$$\dot{e}_4 = -\frac{R}{L} I_q - \omega I_d - \frac{1}{L} E_q + \frac{v_{dc}}{L} S_q - \dot{I}_{qref} \quad (30)$$

In equations (28) and (30) the actual control inputs are appeared. At this stage, the objective is to design the control input signals S_d and S_q, in such a way that e_1, e_2, e_3, and e_4 converge to zero as $t \to \infty$. To achieve such objective, we consider the following CLF

$$W_3 = W_2 + \frac{1}{2} e_3^2 + \frac{1}{2} e_4^2 \quad (31)$$

The time derivative of W_2 is

$$\dot{W}_3 = \dot{W}_2 + e_3 \dot{e}_3 + e_4 \dot{e}_4 \quad (32)$$

Inserting the values of $\dot{e}_2$ from equation (28) and $\dot{e}_3$ from equation (30) into equation (32) yields

$$\dot{W}_3 = -k_1 e_1^2 - k_2 e_2^2 + e_3\left(-\frac{R}{L} I_d + \omega I_q - \frac{1}{L} E_d + \frac{v_{dc}}{L} S_d - \dot{\alpha}_1\right) + e_4\left(-\frac{R}{L} I_q - \omega I_d - \frac{1}{L} E_q + \frac{v_{dc}}{L} S_q - \dot{I}_{qref}\right) \quad (33)$$

From equation (33), the switching control laws S_d and S_q need to be selected in such a way that $\dot{W}_3$ would be negative semi-definite, i.e., $\dot{W}_3 \leq 0$, which asymptotically stabilize the error dynamics. To do that, the following control laws are chosen:

$$S_d = \frac{R}{v_{dc}} I_d - \frac{\omega L}{v_{dc}} I_q + \frac{E_d}{v_{dc}} + \frac{L}{v_{dc}} \dot{\alpha}_1 - \frac{L}{v_{dc}} k_3 e_3 \quad (34)$$

$$S_q = \frac{R}{v_{dc}} I_q + \frac{\omega L}{v_{dc}} I_d + \frac{E_q}{v_{dc}} + \frac{L}{v_{dc}} \dot{I}_{qref} - \frac{L}{v_{dc}} k_4 e_4 \quad (35)$$

As a result, the time derivative of W_3 can be simplified as:

$$\dot{W}_3 = -k_1 e_1^2 - k_2 e_2^2 - k_3 e_3^2 - k_4 e_4^2 \leq 0 \quad (36)$$

Equation (36) clearly indicates that the $\dot{W}_3$ is negative semi-definite and thus the system is asymptotically stable. The performance of the designed controller is shown in the following section.

IV. Controller Performance Evaluation

The aim of this section is to verify the effectiveness of the designed controller on a similar system as shown in Fig.1. To do that, an SPR 305-WHT based 100 kW PV module is considered in this paper. Table I represents the corresponding array electrical characteristics of the module SPR 305-WHT. The CSI is connected to the power grid through an *L* filter whose inductance is 10 mH. The parameters of the designed backstepping controller to tune the output responses are chosen as follows: $k_1 = 10$, $k_2 = 8$, and $k_3 = 12$. The performance of the controller is evaluated by considering the following two different operating conditions:

- Controller performance at standard atmospheric conditions, and
- Controller performance with changing solar irradiation.

Case I: *Controller performance at standard atmospheric conditions*

In this case study, the simulation is carried out at a standard atmospheric condition. The grid voltage and current responses of phase *A* are shown in Fig. 2. From Fig. 2, it can be seen that the grid current is not in phase with grid voltage as it is operating at a power factor of 0.80 rather than unity.

TABLE 1 Electrical Characteristics for one module of SPR 305-WHT

Parameter	Symbol	Value
Maximum power	P_m	100 kW
Open circuit voltage	V_{oc}	64.2 V
Short circuit current	I_{sc}	5.96 A
Maximum power voltage	V_{max}	54.7 V
Maximum power current	I_{max}	5.58 A
Number of series connected module per string	N_{ser}	5
Number of parallel strings	N_{par}	66

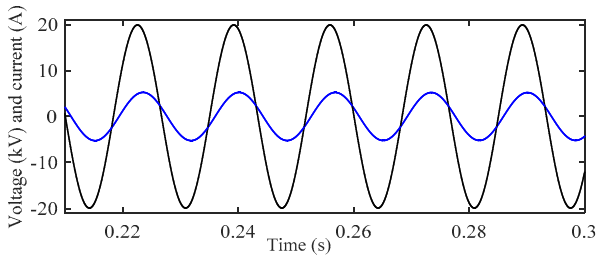


Fig. 2. Grid voltage (black line) and current (blue line) of phase *A* at standard operating condition

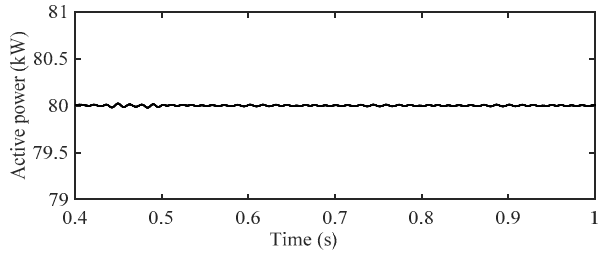


Fig. 3. Active power of the grid at standard operating condition

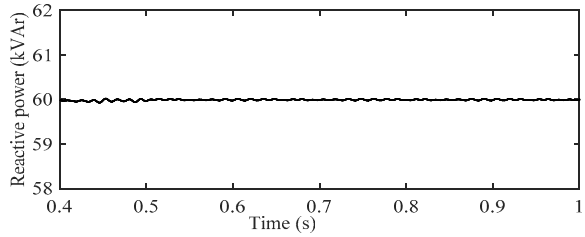


Fig. 4. Reactive power of the grid at standard operating condition

Thus, both active and reactive powers will be delivered to the grid. The corresponding active and reactive powers responses are shown in Fig. 3 and Fig. 4, respectively, which are injected into the grid. From Fig. 3 and Fig. 4, it can be seen that the designed controller can inject stable power into the power grid.

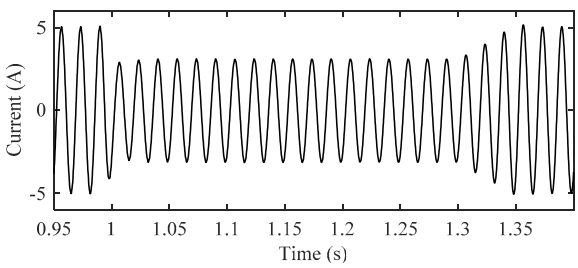


Fig. 5. Grid current with the change of solar irradiation

Case II: *Controller performance with changing solar irradiation*

In this case study, the simulation is carried out with changing solar irradiation while the temperature is kept constant at 25^0C. Initially, the solar irradiation level is set to 1 kW/m^2 and suddenly at t = 1 s, the solar irradiation level is step changed to 0.5 kW/m^2 and maintained this level till t=1.3 s and at last, back to 1 kW/m^2 after t=1.3 s. The resulting grid current is shown in Fig. 5, from where it can be seen that the grid current varies considerably in accordance with the variation of solar irradiation. Fig. 6 and Fig. 7 represent the corresponding active and reactive power of grid. The quality of the injected power into the grid for both cases is measured in term of the THD of the injected grid current. The THD response for this case is shown in Fig. 8 from where it is clear that the designed controller can maintain the IEEE-1547 THD standard as it is less than 5%.

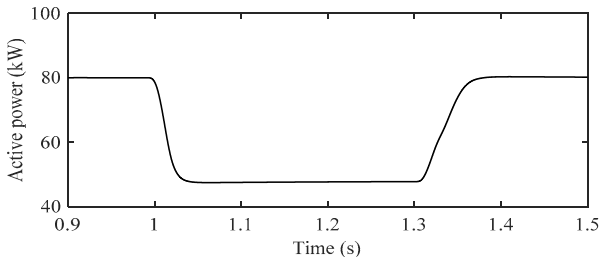


Fig. 6. Active power with changing solar irradiation

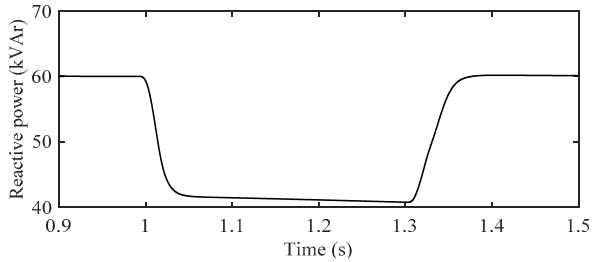


Fig. 7. Reactive power with changing solar irradiation

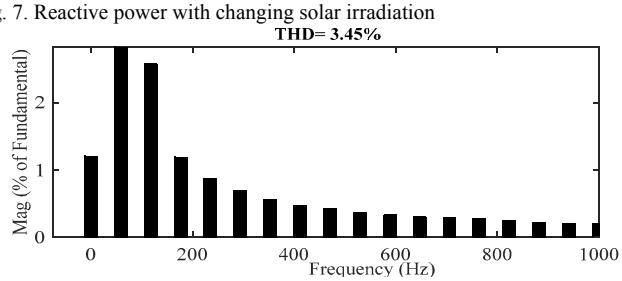


Fig. 8. THD of grid current with changing solar irradiation

V. Conclusion

In this paper, a nonlinear backstepping controller is designed to enhance the power quality of a CSI based three-phase grid-connected PV system. Simulation results, under different operating situations, are included to support the theoretical frameworks of the designed controller. From simulation results, it is clear that the designed controller provides a very satisfactory performance in terms of maintaining steady-state operation under various operating conditions. Furthermore, the designed controller has shown improvements in the power quality of the system by reducing the THD of the injected current into the grid. Future work will be devoted to the design of an adaptive backstepping controller by considering the parametric uncertainties of the system.

References

[1] A. E. Fadili, F. Giri, and A. E. Magri, "Reference voltage optimizer for maximum power point tracking in triphase grid-connected photovoltaic systems," *Int. Journal of Elect. Power & Energy Systems,* vol. 60, pp. 293-301, 2014.

[2] K. M. Tsang and W. L. Chan, "Three-level grid-connected photovoltaic inverter with maximum power point tracking*," Int. Journal of Energy Conversion and Management,* vol. 65, pp. 221-227, 2013.

[3] A. B. R.-Boué , R. G.-Valverde, F. de A. Ruz-Vila, José M. Torrelo-Ponce, "An integrative approach to the design methodology for 3-phase power conditioners in photovoltaic grid-connected systems," *Int. Journal of Energy Conver. and Management*, vol. 56, pp. 80-95, 2012.

[4] T. K. Roy, M. Morshed, F. K. Tumpa, M. F. Pervej, "Robust adaptive backstepping speed controller design for a series DC motor," in *Proc. of IEEE Int. WECECE*, December 2015.

[5] U. Zimmermann and M. Edoff, "A maximum power point tracker for long-term logging of PV module performance," *IEEE Journal of Photovoltaics*, vol. 2, no. 1, pp. 47-55, 2012.

[6] E. Koutroulis and F. Blaabjerg, "A new technique for tracking the global maximum power point of PV arrays operating under partial-shading conditions," *IEEE Journal of Photovoltaics*, vol. 2, no. 2, pp. 184–190, April 2012.

[7] F. Bouchafaa, D. Beriber, and M. S. Boucherit, "Modeling and control of a gird connected PV generation system," *18th Mediterranean Conference on Control & Automation*, pp.315-320, 23-25 June 2010.

[8] E. Bianconi, J. Calvente, R. Giral, E. Mamarelis, G. Petrone, C. A. R.-Paja, G. Spagnudo, and M. Vitelli, "Perturb and Observe MPPT algorithm with a current controller based on the sliding mode," *Int. Journal of Elect. Power & Ener. Systems,* vol. 44, pp. 346-356, 2013.

[9] Y. Yang and F. P. Zhao, "Adaptive perturb and observe MPPT technique for grid-connected photovoltaic inverters," *International Conference on Power Electronic and Engineering Application*, pp. 468-473, 2011.

[10] B. Liu, S. Duan, F. Liu, and P. Xu, "Analysis and improvement of maximum power point tracking algorithm based on incremental conductance method for photovoltaic array," in *Proc. IEEE Int. Conf. Power Electron. Drive Syst.*, pp. 637–641, 2007.

[11] P. P. Dash, and M. Kazerani, "Dynamic modeling and performance analysis of a grid-connected current-source inverter-based photovoltaic system," *IEEE Trans. On Sustainable Energy,* vol. 2, no. 4, Oct. 2011.

[12] R. Kadri, J.-P. Gaubert, and G. Champenois, "An improved maximum power point tracking for photovoltaic grid-connected inverter based on voltage-oriented control," *Int. Trans. on Industrial Electronics*, vol. 58, no. 1, January, 2011.

[13] I.-S. Kim, "Sliding mode controller for the single-phase grid-connected photovoltaic system," *Applied Energy*, vol. 83, pp. 1101-1115, 2006.

[14] A. Kotsopoulos, J. L. Darte, and M. A. M. Hendrix, "Predictive DC voltage control of single-phase PV inverters with small dc link capacitance," *in Proc. Int. Symp. Ind. Electron.*, pp.793–797, Jun. 2003.

[15] T. K. Roy, M. F. Pervej, S. Rahman, and F. K. Tumpa, "Nonlinear adaptive controller design for PWM based DC-DC boost converters using backstepping control scheme," *in Proc. of the 5th ICIEV,* 2016.

[16] D. Lalili, A. Mellit, N. Lourci, B. Medjahed, and E. M. Berkouk, "Input output feedback linearization control and variable step size MPPT algorithm of a grid-connected photovoltaic inverter," *Renewable Energy*, vol. 36, no. 12, pp. 3282–3291, Dec. 2011.

[17] I.-S. Kim, "Robust maximum power point tracking using sliding mode controller for the three-phase grid-connected photovoltaic system," *Solar Energy*, vol. 81, pp. 405-414, 2007.

[18] T. K. Roy, "Position control of a small helicopter using robust backstepping," in *Proc. of the 7th ICECE*, pp. 787-790, 2012.

[19] T. K. Roy, and M. Garrat, "Altitude control of an unmanned autonomous helicopter via robust backstepping controller under horizontal wind gusts," in *Proc. of 7th ICECE*, pp. 771-774, 2012.

[20] T. K. Roy, M. Garrat, H. R. Pota, and H. Teimoori, "Hover flight control of a small helicopter using robust backstepping and PID," in *Proc. of the 10th WCICA*, 6-8 July 2012.

A Wide Tuning Range and Constant Gain VCO for Radio Frequency Application

Abdullah Ibn Abbas,[1,*] and A.B.M. Harun-ur Rashid[2]

[1]Department of ECE, Concordia University, Montreal, Canada

[2]Department of EEE, BUET, Dhaka, Bangladesh

[*]abdullahbuet12@gmail.com

***Abstract*— A wide-tuning-range and constant gain ring voltage controlled oscillator (VCO) is presented for radio frequency application. The circuit is implemented in 90nm RF CMOS 1.2 V process technology. The biasing circuit utilizes the concept of voltage controlled voltage source to increase the tuning range as well as to obtain a constant gain (K_{VCO}) within the operating range. This results in a rail-to-rail control voltage (0 V-1.2 V). The oscillation frequency of the designed VCO ranges from 2.42 GHz to 12.12 GHz (S band to X band). The 1 MHz offset phase noise of the VCO for an output frequency of 12.12 GHz is found to be -78.87 dBc/Hz without applying any phase noise reduction technique. The circuit consumes a very low power of 2.47 mW at the maximum operating frequency of 12.12 GHz from a power supply of 1.2 V. The layout of the whole VCO circuit occupies an area of 0.001 mm^2.**

***Index Terms*— VCO, CMOS, phase-noise, ring oscillator, PLL, radio frequency (RF), voltage controlled voltage source, gain of a VCO.**

I. Introduction

As the technology is advancing, the need for high speed circuit is becoming inevitable. The continuous shrinking in the size of MOSFETs has increasingly attracted the research and development of low power, radio frequency (RF) CMOS integrated circuits. In this era of technology, the circuit designed should be on chip and occupy a small area. Phase Locked Loop (PLL) based frequency synthesizer are widely used in radio frequency synthesis. PLL is one of the key building blocks of an RF front-end transceiver. The integral part of a PLL is the Voltage Controlled Oscillator (VCO) which works at the highest frequency in the system. The frequency synthesizer requires their frequency to be tunable over a range of frequency bands. This requires the use of a VCO. VCOs may be of ring oscillator type or LC oscillator type.

The general trend in the design of high frequency PLL based frequency synthesizer is to use LC oscillators [1]-[4] because of its low phase noise characteristics. However, they suffer from narrow tuning range and also occupy a large die area due to the presence of on-chip spiral inductors. Ring oscillators on the other hand can be built in any standard CMOS process which require less die area than LC designs and have a wide tuning range. Furthermore, they have the particular desirable property that the inductors can be dispensed altogether which makes ring VCOs highly scalable. Also the wastefulness of ring oscillator is generally much smaller [5] and is more attractive in terms of phase noise trade-off. The literature works that use submicron process technology hardly present ring oscillators that operate above 9 GHz and with a wide tuning range. Some literature works show [6], [7] operating frequency above 9 GHz have been achieved in recent technology process (180nm, 90nm etc.) but without a constant gain and frequency tuning range not above 7 GHz. A voltage-controlled oscillator with a constant gain within full range of operating frequency upto 12 GHz in Si CMOS has not been reported yet.

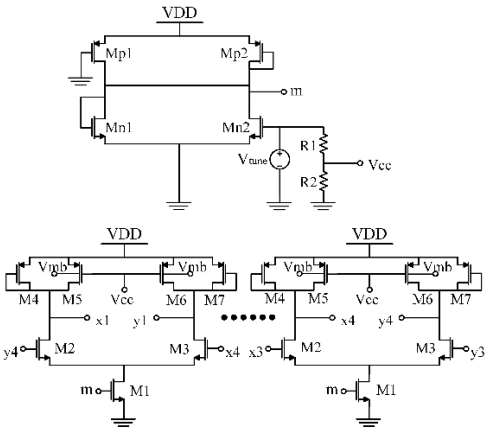


Fig. 1. Proposed VCO with biasing circuit.

In this paper the design of a ring VCO topology is presented which utilizes the concept of voltage controlled voltage source for dually controlling the frequency to achieve constant gain with a much wider frequency tuning range for RF application. The paper is organized as follows. In Section II, the design of the proposed VCO with improved topology is presented. Section III describes the schematic and layout of the proposed VCO whereas section IV demonstrates the simulation results with discussion. Section V features a comparison with other published works and the work is concluded in Section VI.

II. Design of The Proposed Differential Ring VCO

The topology of the proposed delay cell is shown in Fig. 1. It contains a source coupled differential pair and symmetric loads [8] which provide good control over delay and high dynamic power supply noise rejection. The differential delay cell is used to improve the performance of the ring oscillator. The frequency of oscillation of a ring oscillator depends on the delay of each cell while delay of each cell is inversely proportional to the current flowing through it. The higher the current through the cell the smaller is the delay and hence the higher is the oscillation frequency. The figure also shows the biasing circuit for the proposed VCO. It is based on a four stage differential ring oscillator [9]. To study the behavior of the current through each delay cell as the control voltage is varied, the gates of M5 and M6 are connected to various voltage points so as to achieve a wide tuning range. If the gates of M5 and M6 are connected to the ground with the gate

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

of the tail MOSFET connected to point m for the entire value (0-1.2 V) of the tuning voltage, V_{tune}, then the behaviour of the current is displayed by plot 1 in Fig. 2. Here the current has a wide variation but the variation is limited between 0.75 V to 1.2 V of the tuning voltage which will limit the desired linear variation of the oscillation frequency for the entire tuning voltage. When the gates of M5 and M6 are also connected to the point m as shown in Fig. 1, then the behavior of the current is represented by plot 2 in Fig. 2 which has a very narrow range of current variation. Again, if the gates of M5 and M6 are connected directly to the tuning voltage, the variation of current observed is indicated by plot 3 in Fig. 2. This curve has a wide current variation with a large tuning voltage range. However, it is not linear and as a result K_{VCO} will not be constant. Finally, plot 4 is obtained for the condition when the gates are connected to Vcc which has a wide range of current variation. Eventually, we select plot 4 and corresponding gates connection of M5 and M6 because of its highly linear behavior and wide current variation. The relationship between Vcc and the tuning voltage should be such that a good control over the delay is possible. This compelled us to incorporate a voltage controlled voltage source such that there exists a relationship between Vcc and the tuning voltage so as to obtain the desired wide frequency tuning range. The plot for the range of variation of m and Vcc as the tuning voltage is varied is presented in Fig. 3. If the oscillation frequency is controlled only through Vcc keeping the gate of tail MOSFET fixed at certain voltage, the variation of current through the delay cell would not have been linear. The desired linear variation in current is obtained through combined effect of Vcc and m as shown in Fig. 3. To achieve the load linearity [10] and to achieve higher operating frequency, the bulk of M5 and M6 are connected to Vmb where Vmb=1 V. M4 and M7 are connected as diode. The two resistors (R1 and R2 in Fig. 1) used in the biasing circuit act as a voltage controlled voltage source. The purpose of using such a circuit is to increase the tuning range while maintaining a constant K_{VCO}. Both Vcc and m are used as the control voltage to vary the delay of the proposed cell.

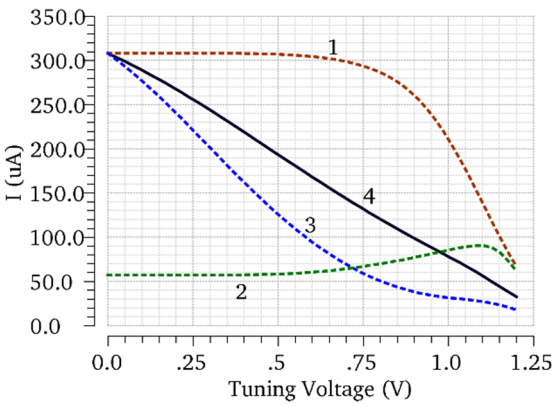


Fig. 2. Variation of current through each delay cell for different connection of Vcc as a function of tuning voltage.

III. Schematic and Layout of the Proposed VCO

The proposed VCO is designed using Cadence Electronic Design Automation tools in 90nm CMOS process technology. Virtuoso Schematic Editor is used for schematic design and all the simulation results are plotted in Cadence Spectre. Virtuoso Layout Editor is used for physical layout design. Fig. 4 shows the layout of the proposed VCO and it occupies a chip area of 0.001 mm^2. Assura QRC is used for parasitic extraction. The addition of parasitic elements degrades the performance of the circuit. We present the post-layout performance parameters with all the parasitic elements added to more accurately predict the behavior of the VCO circuit.

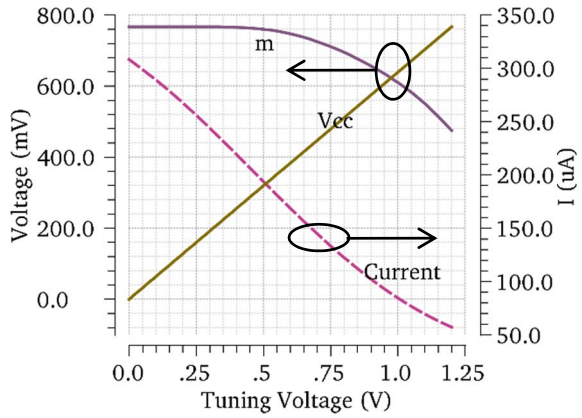


Fig. 3. Variation of Vcc and 'm' over tuning voltage and the variation of current through a delay cell for a fixed value of 'm'.

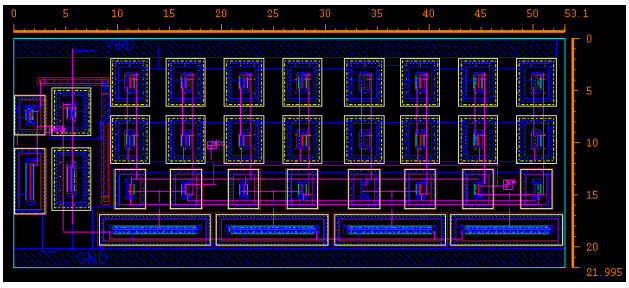


Fig. 4. Layout of the core VCO.

IV. Simulation and Discussion

We are interested in designing a differential ring VCO in 90nm CMOS process which can operate at a frequency as high as possible with a wide tuning range and constant K_{VCO}. In our design, we have not explicitly used any phase noise reduction technique and hence we cannot expect excellent phase noise performance, rather we will compare and justify the phase noise performance with the theoretical value obtained by well-established formulas.

A. Voltage Swing vs. Frequency Trade-off

For a differential ring oscillator, the frequency of oscillation can be approximated by the following expression [11]

$$f_0 = \frac{I_{tail}}{2NC_L V_{osc}} \tag{1}$$

where I_{tail}, C_L, and V_{osc} are the tail bias current of the differential pair, total load capacitance and the maximum single-ended voltage swing for f_0 at the output of each stage. From (1) it is clear that for a given number of stages, the oscillation frequency can be increased by increasing the I_{tail}. Furthermore, as f_0 increases, V_{osc} decreases. Therefore, we have adjusted the I_{tail} by adjusting the width of the differential pair so as to obtain not only an appreciable V_{osc} but also high f_0 and constant K_{VCO} at the same time. The blue dotted line in Fig. 5 shows the plot for the tuning range whereas the solid purple line shows the gain of the VCO, i.e., K_{VCO}. As the tuning voltage (V_{tune}) varies from 0 V to 1.2 V, the operating

frequency varies from 12.12 GHz to 2.42 GHz which corresponds to a percentage tuning range of 133.4%[1]. The plot of the K_{VCO} which is almost constant (-7.0 GHz/V) within the operating range is better than the previous published work.

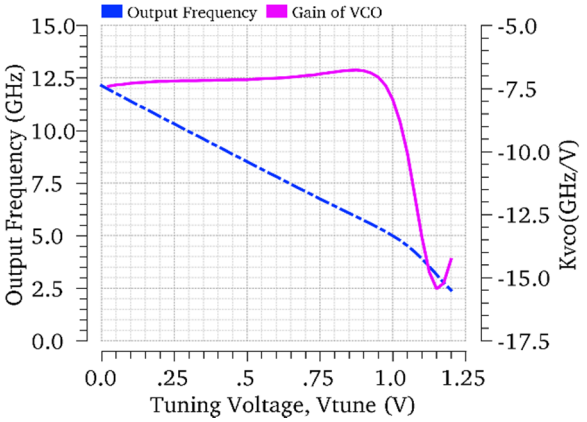


Fig. 5. Frequency tuning range and gain (K_{VCO}) versus tuning voltage.

B. Power vs. Phase Noise Trade-off

For a differential ring oscillator, the total power dissipation is given by [11]

$$P = NV_{DD}I_{tail} \tag{2}$$

where N is the number of stages and V_{DD} is the supply voltage. The sideband noise power is predicted by the Hajimiri phase noise model in [12]. The single-sideband phase noise for an N -stage ring oscillator is given by [13]

$$L(\Delta\omega) = \frac{2NFkT}{P}\left(\frac{\omega_0}{2Q\Delta\omega}\right)^2 \tag{3}$$

where $Q = \frac{N}{2}\sin\left(\frac{\pi}{N}\right)$ and is approximately 1.4 for a four stage ring oscillator. For a differential ring oscillator using short-channel devices, the lower bound on the single-sideband phase noise in the $1/f^2$ region is derived in [11] as

$$L\{\Delta f\} \geq \frac{18kTV_{DD}}{\pi^2 P}\left(\frac{2.5}{E_c L_{eff}}+1\right)\left(\frac{f_0}{\Delta f}\right)^2 N \tag{4}$$

where P is the power dissipation given by (2), E_c is the critical field in silicon, and L_{eff} is the gate length of the differential-pair devices. From (3) it is seen that there is a trade-off between phase noise and power dissipation.

C. Phase Noise Analysis of the Proposed Circuit

The single-sideband phase noise given by (3) can also be determined in terms of single-sideband power spectral density of the noise current source, $\overline{i_n^2}/\Delta f$, and is given by [12] and [14] as

$$L\{\Delta f\} = 10\log\left(\frac{\Gamma_{rms}^2}{q_{\max}^2}\cdot\frac{\overline{i_n^2}/\Delta f}{8\pi^2\Delta f^2}\right) \tag{5}$$

where Γ_{rms} is the rms value of the impulse sensitivity function (ISF) and Δf is the frequency offset from the carrier, and $\overline{i_n^2}/\Delta f$ represents the total current noise due to all the sources. Surprisingly, tail-current source noise in the vicinity of f_0 does not affect the phase noise. Therefore, only the noise of the differential transistors and the loads are taken into account. Assuming symmetrical rise and fall times, Γ_{rms} can be expressed as [14]

$$\Gamma_{rms} = \sqrt{\frac{2\pi^2}{3\eta^3}}\frac{1}{N^{1.5}} \tag{6}$$

where η is the proportionality constant and is equal to 0.9 for a differential oscillator. The dc value of ISF is given by [14]

$$\Gamma_{dc} = \frac{2\pi}{\eta^2}\frac{1}{N^2} \tag{7}$$

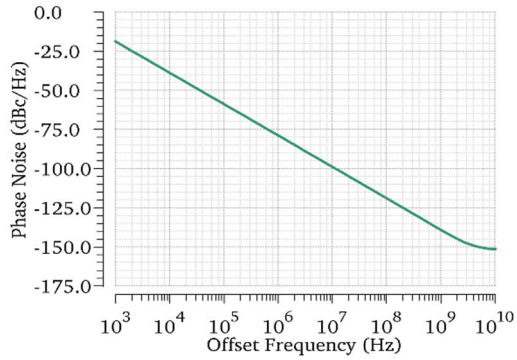


Fig. 6. Post-layout phase noise characteristics at 12.12 GHz operating frequency.

Now we predict the phase noise in the $1/f^2$ region of our designed four-stage differential ring. We observe from (4) that phase noise is a strong function of the number of stages which justifies the selection of a four-stage.

For this process we have a gate oxide thickness of $t_{ox,\mathrm{n}} = 2.8$ nm and $t_{ox,\mathrm{p}} = 2.85$ nm, threshold voltages of $V_{TN} = 0.5$ V and $V_{TP} = 0.45$ V, $L = 100$ nm. Using the process and geometry information, the total capacitance on each node is calculated to be $C_{total} = 3.432$ fF. Therefore, $q_{\max} = C_{total}V_{swing} = 2.4$ fC. The total noise power due to NMOS and PMOS devices is $\overline{i_n^2}/\Delta f = 4kT\gamma\mu C_{ox}(W/L)(V_{DD}/2 - V_T) = 2.2585\times10^{-23}\,\mathrm{A}^2/Hz$ where μ and V_T are replaced by μ_n, V_{TN} and μ_p, V_{TP} for NMOS and PMOS respectively. For a four-stage differential ring oscillator, $\Gamma_{rms}^2 \approx 9/N^3$. Therefore, according to (5) the predicted phase noise for our designed circuit is $L\{\Delta f\} = 10\log\left(\frac{6967.85}{\Delta f^2}\right) = -81.57$ dBc/Hz. The simulated phase noise plot of the VCO for f_0=12.12 GHz is shown in Fig. 6. From the figure it is noted that the phase noise at 1 MHz is -78.87 dBc/Hz which is consistent with the above predicted value of -81.57dBc/Hz. Fig. 7 shows the simulated phase noise plot at 1 MHz offset frequency for the entire operating range.

The power consumption of the core VCO together with the biasing circuit is 2.47 mW at the maximum oscillation frequency f_0 =12.12 GHz and 1.2 V supply voltage. The widely used figure-of-merit (FOM) of the designed ring VCO is expressed as [15]

[1]Frequency range/Centre frequency

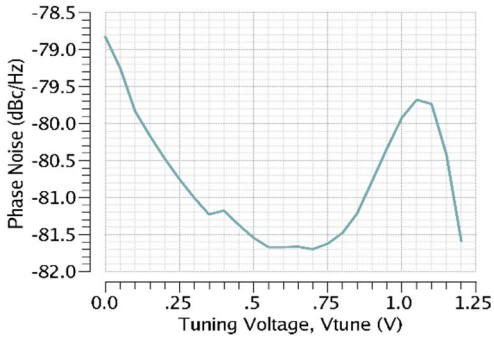


Fig. 7. Phase noise characteristics @ 1 MHz offset for the entire tuning range

TABLE 1
PERFORMANCE COMPARISON OF VCOS

Ref.	This work	[15]	[16]	[17]	[18]	[19]	[20]
Tech. (nm)	90	65	130	130	130	90	90
Tuning Range (GHz)	2.42-12.12	2-8	7.3-7.86	1-9	1.82-10.18	1.1-3.2	0.4-2
f_0 (GHz)	12.12	4.2	7.64	9	10.18	2.4	2
FTR (%)	133.4	120	7.3	160	139.4	97.67	70
Phase Noise (dBc/Hz)	-78.8	-101	-103	-105	-88.4	-94	-120
Offset (MHz)	1	1	1	10	1	1	0.2
Power (mW)	2.47	6.8	60	5.92	5	0.9	13
Area (mm^2)	0.001	0.038	0.009	0.013	N/A	0.00016	0.008
FOM	-157	-165	-163	-176	-96	-162	-175

$$FOM = L(\Delta f) - 20\log\left(\frac{f_o}{\Delta f}\right) + 10\log\left(\frac{Pmw}{1mw}\right) \quad (8)$$

where $L(\Delta f)$ is the phase noise at an offset frequency of Δf, f_o is the oscillation frequency, Pmw is the dc power consumption in milliwatts. Therefore, we calculate the final value of the *FOM* from the post-layout simulation parameters using (9) and it is -156.62.

V. COMPARISON WITH THE STATE-OF-THE-ART

The proposed ring VCO is compared to other reported ring oscillators in literature. A tabular comparison is summarized in Table 1. Compared with [15] and [17], this work exhibits a much wider frequency tuning range of 9.7 GHz with highly linear tuning characteristics. Reference [15] claims for linear and wide tuning range of 8 GHz but K_{VCO} is not as constant as reported by the present work. The phase noise reported by this work is comparable to [17] without applying any phase noise reduction technique. The power consumption at the maximum operating frequency of 12.12 GHz is also a minimum as compared to [15]-[19]. This work is also area efficient as compared to [15] and [16]. The FOM is also appreciable and is better than [17] and comparable to [16],[18],[20]. FOM of [17] and [18] is re-calculated here using (2) for the purpose of comparison. FOM can further be increased by improving the phase noise performance by incorporating some phase noise reduction technique.

VI. CONCLUSIONS

We proposed a four stage differential ring VCO topology for wide tuning range and constant K_{VCO}. The proposed VCO uses rail-to-rail voltage for its frequency control. The VCO designed can be used for 12 GHz applications and is implemented in 90nm RF CMOS 1.2 V process technology. Post-layout simulation results of the proposed VCO exhibit tuning range of 9.7 GHz from 2.42 GHz to 12.12 GHz, power consumption of 2.47 mW at a supply voltage of 1.2 V, phase noise of -78.87 dBc/Hz for f_0=12.12 GHz at 1 MHz offset frequency, a chip area of 0.001 mm^2 and an FOM of -156.62.

REFERENCES

[1] L. M. Franca-Neto, R. E. Bishop, and B. A. Bloechel, "64 GHz and 100 GHz VCO in 90 nm CMOS using optimum pumped method," in *Proc. IEEE Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers,* Feb. 2004, pp. 444-538.

[2] G. Ritzberger, J. Boch, and A. L. Scholtz, "45 GHz highly integrated phase-locked loop frequency synthesizer in SiGe bipolar technology," in *Proc. IEEE Int. Microw. Symp. Dig.,* Jun.2002, pp. 831-834.

[3] J. Craninckx, and M. S. J. Steyaert, "A 1.8-GHz low-phase-noise CMOS VCO using optimized hollow spiral inductors," *IEEE J. Solid-State Circuits,* vol. 32, no. 5, pp. 736-744, May 1997.

[4] C. Changhua, D. Yanping, and K.O. Kenneth, "A 50-GHz phase-locked loop in 130-nm CMOS," in *Proc. IEEE Custom Integrated Circuits Conf. (CICC),*Sep. 2006, pp. 21-24.

[5] R. Navid, T. H. Lee, and R. W. Dutton, "Minimum achievable phase noise of RC Oscillators," *IEEE J. Solid-State Circuits,* vol. 40, no. 3, pp. 630-637, Mar. 2005.

[6] V. De Heyn, G. V. Plas, J. Ryckaert, and J. Craninckx, "A fast start-up 3 GHz-10 GHz digitally controlled oscillator for UWB impulse radio in 90 nm CMOS," in *Proc. IEEE Eur. Solid-State Circuits Conf. (ESSCIRC),*Sep. 2007, pp. 484-487.

[7] H. Kodama, H. Okada, H. Ishikawa, and A. Tanaka, "Wide lock-range, low phase-noise PLL using interpolative ring-VCO with coarse tuning and frequency linearization," in *Proc. IEEE Custom Integrated Circuits Conf. (CICC),* Sep. 2007, pp. 349-352.

[8] D. Liang, and R. Harjani, "Design of low-phase-noise CMOS ring oscillators," *IEEE Trans.Circuits Syst. II,* vol. 49, no. 5, pp. 328-338, May 2002.

[9] L. Sang-yeop, S. Amakawa, N. Ishihara, and K. Masu, "Low-phase-noise wide-frequency-range differential ring-VCO with non-integral subharmonic locking in 0.18 µm CMOS," in *Proc. IEEE Eur.Microw. Integr. Circuits Conf. (EuMIC),*Sep. 2010, pp. 464-467.

[10] T. Li, J. Jiang, B. Ye, and X. Han, "Ultra low voltage, wide tuning range voltage controlled ring oscillator," in *Proc. IEEE ASIC Conf. (ASICON),* Oct. 2011, pp. 824-827.

[11] R. J. Betancourt-Zamora, and T. H. Lee, "Low phase noise CMOS ring oscillator VCOs for frequency synthesis," in *Proc. of the Second International Workshop on Design of Mixed-mode Integrated Circuits*, Jul. 1998, pp. 37–40.

[12] A. Hajimiri, and T. H. Lee, "A general theory of phase noise in electrical oscillators," *IEEE J. Solid-State Circuits,* vol. 33, no. 2, pp. 179-194, Feb. 1998.

[13] L. Dai, R. Harjani, *Design of High-Performance CMOS voltage-controlled oscillators*, Kluwer Academic Publishers, 2003.

[14] A. Hajimiri, S. Limotyrakis, and T. H. Lee, "Jitter and phase noise in ring oscillators," *IEEE J. Solid-State Circuits,* vol. 34, no. 6, pp. 790-804, Jun. 1999.

[15] S. Yoo *et al.*, "A 2-8 GHz wideband dually frequency-tuned ring-VCO with a scalable Kvco," *IEEE Microw. Wireless Compon.Lett.,* vol. 23, no. 11, pp. 602-604, Nov. 2013.

[16] H. Liu *et al.*, "A low-noise multi-GHz CMOS multi-loop ring oscillator with coarse and fine frequency tuning," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.,* vol. 17, no. 4, pp. 571-577, Apr. 2009.

[17] C. Li, and J. Lin, "A 1-9 GHz linear-wide-tuning-range quadrature ring oscillator in 130 nm CMOS for non-contact vital sign radar application," *IEEE Microw. Compon. Lett.,* vol. 20, no. 1, pp. 34-36, Jan. 2010.

[18] B. Fahs, W. Y. Ali-Ahmad, and P. Gomand, "A two-stage ring oscillator in 0.13-µm CMOS for UWB impulse ratio," *IEEE Trans. Microw. Theory Techn.,* vol. 57, no. 5, pp. 1074-1082, May 2009.

[19] C. Zhang, and J. Zhou, "A 2.4 GHz low power low phase noise ring VCO," in *Proc. IEEE Anti-Counterfeiting, Security and Identification (ASID) Conf.,* Aug. 2012, pp. 1-4.

[20] B. Jonathan *et al.,* "A low-complexity, low-phase-noise, low-voltage phase-aligned ring oscillator in 90 nm digital CMOS," *IEEE J. Solid-State Circuits,* vol. 44, no. 7, pp. 1942-1949, Jun. 2009.

Realization of Low-Voltage Ride Through Requirements for PMSGs in Wind Turbines Systems Using Generator-Rotor Inertia

Mohamed Abdelrahem, Muhammad Hosnee Mobarak, Ralph Kennel
Institute for Electrical Drive Systems and Power Electronics, Technical University of Munich (TUM)
E-mail: mohamed.abdelrahem@tum.de, hosnee.mobarak@tum.de, ralph.kennel@tum.de

***Abstract*—Recently, the electric power generation by wind energy conversion systems (WECSs) has increased significantly. Therefore, wind power generation systems are required to achieve the grid codes (GCs) requirements stated by the power system operators. In case of grid voltage dips, a mismatch between the generated active power from the generator and the active power delivered to the grid is produced. The traditional solution is using a braking chopper (BC) in the DC-link to dissipate this surplus active power. This paper proposes a low-voltage ride through (LVRT) strategy for variable-speed WECSs based on permanent magnet synchronous generators (PMSGs). The proposed method uses the PMSG-rotor inertia to store the surplus active power during the grid voltage dips. Therefore, no additionally hardware components is required. Moreover, a finite control set-model predictive control (FCS-MPC) system for the PMSG is proposed in order to improve the transient performance of the system. The performance of the proposed LVRT strategy is compared with the traditional BC solution for all the operation conditions via simulation results.**

***Index Terms*—Low-voltage ride through (LVRT), permanent magnet synchronous generator (PMSG), model predictive control.**

I. Introduction

Wind energy is rapidly growing in the recent years [1]. Among different types of wind turbines generators, due to several advantages (high efficiency, high power density, low maintenance costs, and good grid compatibility), permanent magnet synchronous generator (PMSG) is attractive concept. Due to the rapid integration of wind power generation systems into the grid, the grid codes (GCs) have been updated in many countries by the transmission and distribution system operators [2]. Among all the grid codes, the low-voltage ride through (LVRT) is the main concern for the wind turbine manufacturers. According to the LVRT requirements, during grid faults, the wind energy conversion system (WECS) must remain connected to the grid and must supply active and reactive power to help the grid to recover from the voltage dip [2].

In order to improve the LVRT capability of PMSG-based variable-speed wind energy conversion systems (WECSs), various approaches have been proposed in the last years. A braking chopper (BC) has been applied for the LVRT in [3]. However, the surplus active power is dissipated in the BC resistance . In [4], the WECS with PMSG is connected to the grid during grid faults using static synchronous compensator. This solution required extra components, which increases the cost. During grid faults, the power extracted from the wind turbine (WT) can be reduced using blade pitch angle control to limit the generated mechanical power [5]. However, the response of the mechanical system is very slow in comparison with the electrical system. An energy storage system in the DC-link is proposed in [6], which needs additional circuitry and control. LVRT requirements can be achieved by storing the surplus active power in the inertia of the PMSG-rotor as proposed in [7]. However, the DC-link voltage control is achieved by the machine side converter control system instead of the grid side converter control system during the voltage dip. Exchanging the references may be cause instability in the system due to the tuning of the PI controller.

In this paper, a low-voltage ride through (LVRT) strategy for variable-speed WECSs based on permanent magnet synchronous generators (PMSGs) is proposed. The proposed method uses the PMSG-rotor inertia to store the surprise active power during the grid voltage dips. The performance of the proposed LVRT strategy is compared with the traditional BC solution via simulation results.

II. Modeling of the WECS with PMSG

The block diagram of the WECS with PMSG is shown in Fig. 1. It consists of a PMSG mechanically coupled to the wind turbine directly via a stiff shaft. The stator windings of the PMSG are connected via a back-to-back full-scale voltage source converter (VSC) and a filter to the grid. The machine side converter (MSC) and the grid side converter (GSC) share a common DC-link with capacitance C_{dc} [As/V] with DC-link voltage u_{dc}[V].

A. Permanent magnet synchronous generator (PMSG)

The discrete model of the PMSG in rotor flux orientation is given by (see [8]):

$$\left.\begin{aligned} i_s^d[k+1] &= (1-\tfrac{T_s R_s}{L_s})i_s^d[k] + \omega_r T_s i_s^q[k] + \tfrac{T_s}{L_s}u_s^d[k] \\ i_s^q[k+1] &= (1-\tfrac{T_s R_s}{L_s})i_s^q[k] - \omega_r T_s i_s^d[k] - \tfrac{\omega_r T_s}{L_s}\psi_{pm} \\ &\quad + \tfrac{T_s}{L_s}u_s^q[k] \end{aligned}\right\} \quad (1)$$

$$T_e[k] = \frac{3}{2}n_p\psi_{pm}i_s^q[k] \quad (2)$$

where k is the current sampling instant; T_s [s] is the sampling time; i_s^d and i_s^q [A] are the stator currents of the PMSG; u_s^d and u_s^q [V] are the stator voltages of the PMSG; R_s [V/A] and L_s [Vs/A] are the stator resistance and inductance of the PMSG, ψ_{pm} [Vs] is the permanent-magnet flux linkage, and $\omega_r = n_p\omega_m$ [rad/s] is the electrical frequency of the rotor (with pole pair number n_p and mechanical angular speed ω_m)

978-1-5090-2963-1/16/$31.00 ©2016 European Union

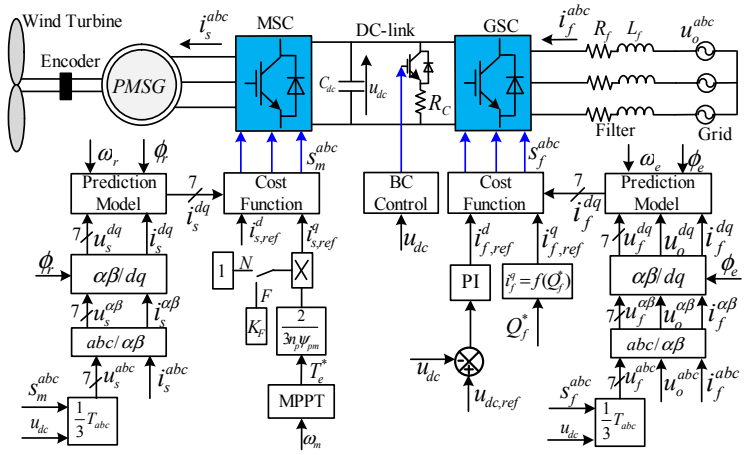


Figure 1: PMSG topology and control structure for variable-speed WECSs.

and T_e [Nm] is the electro-magnetic torque.
The output voltage of the MSC can be calculated as follows [8]:

$$\boldsymbol{u}_s^{abc} = \frac{1}{3} u_{dc} \boldsymbol{T}^{abc} \boldsymbol{s}_m^{abc} \tag{3}$$

where $\boldsymbol{s}_m^{abc} = (s_m^a, s_m^b, s_m^c)^\top \in \{0,1\}$ is the switching state vector of the MSC; u_{dc} is the DC-link voltage; and $\boldsymbol{T}^{abc}$ is the transformation matrix

$$\boldsymbol{T}^{abc} = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \tag{4}$$

describing the relation between switching state vector and output phase voltage vector of the converter. Considering all the switching combinations for 2-level converter eight switching states $(\boldsymbol{s}_m^{abc})^\top := \{000, 100, ..., 111\}$, and consequently, seven voltage vectors are obtained $\boldsymbol{u}_{s,0}^{\alpha\beta}, \boldsymbol{u}_{s,1}^{\alpha\beta}, ..., \boldsymbol{u}_{s,7}^{\alpha\beta}$ (note: $\boldsymbol{u}_{s,0}^{\alpha\beta} = \boldsymbol{u}_{s,7}^{\alpha\beta}$). Equation (3) can be expressed in the stationary/rotating reference frame by using the Clarke $\boldsymbol{T}_c$ and Park $\boldsymbol{T}_p(\phi)^{-1}$ transformation, respectively, given by (neglecting the zero sequence)

$$\boldsymbol{u}_s^{dq} = \underbrace{\begin{bmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{bmatrix}}_{=:\boldsymbol{T}_p(\phi)^{-1}} \underbrace{\frac{2}{3} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix}}_{=:\boldsymbol{T}_c} \boldsymbol{u}_s^{abc} \tag{5}$$

B. Filter and grid

In Fig. 1, a grid-connected voltage source converter is shown. It is connected to the grid via a RL-filter with resistance R_f [Ω] and inductance L_f [Vs/A]. The discrete model of the grid side filter is given by (see [8]):

$$\left.\begin{aligned} i_f^d[k+1] &= (1 - \tfrac{T_s R_t}{L_t}) i_f^d[k] + \omega_e T_s i_f^q[k] \\ &\quad + \tfrac{T_s}{L_t}(u_o^d[k] - u_f^d[k]) \\ i_f^q[k+1] &= (1 - \tfrac{T_s R_t}{L_t}) i_f^q[k] - \omega_e T_s i_f^d[k] \\ &\quad + \tfrac{T_s}{L_t}(u_o^q[k] - u_f^q[k]) \end{aligned}\right\} \tag{6}$$

where i_f^d and i_f^q [A] are currents of the grid side filter; u_o^d and u_o^q [V] are voltages of the grid; u_f^d and u_f^q [V] are the output voltages of the GSC.
The output voltages of the GSC can be expressed as

$$\boldsymbol{u}_f^{abc} = \frac{1}{3} u_{dc} \boldsymbol{T}^{abc} \boldsymbol{s}_f^{abc} \tag{7}$$

where $\boldsymbol{s}_f^{abc} = (s_f^a, s_f^b, s_f^c)^\top \in \{0,1\}$ is the switching state vector of the GSC. Again, considering all the switching combinations for 2-level converter eight switching states $(\boldsymbol{s}_f^{abc})^\top := \{000, 100, ..., 111\}$, and consequently, seven voltage vectors are obtained $\boldsymbol{u}_{f,0}^{\alpha\beta}, \boldsymbol{u}_{f,1}^{\alpha\beta}, ..., \boldsymbol{u}_{f,7}^{\alpha\beta}$. Equation (7) can be expressed in the stationary/rotating reference frame by using the Clarke $\boldsymbol{T}_c$ and Park $\boldsymbol{T}_p(\phi)^{-1}$ transformation as explained before.

III. Finite Control Set-Model Predictive Control

The proposed FCS-MPC algorithm takes the discrete states of the converter into account and the future converter performance is predicted for each sampling period. Then the optimal switching action that minimizes a predefined cost function is selected to be applied in the next sampling instant [9].

A. FCS-MPC for MSC

For the MSC, the chosen cost function is defined by

$$g_{MSC} = |i_{s,ref}^d[k+1] - i_s^d[k+1]| + |i_{s,ref}^q[k+1] - i_s^q[k+1]| \tag{8}$$

where $i_{s,ref}^d[k+1]$ and $i_{s,ref}^q[k+1]$ are the reference values of the $d-$ & q-axis currents. The reference d-axis current $i_{s,ref}^d[k+1]$ is always set to be zero and the reference q-axis current $i_{s,ref}^q[k+1]$ is calculated from the optimal torque $i_{s,ref}^q = \frac{2T_e^*}{3n_p\psi_{pm}}$. This optimal torque T_e^* is calculated from the maximum power point tracking (MPPT) algorithm [8], see Fig 1. For the prediction algorithm, (1) is evaluated for each of the possible seven voltage vectors, giving seven different current predictions. The voltage vector whose current prediction is minimizing the cost function (8) is applied at the next sampling instant.

B. FCS-MPC for GSC

For the GSC, the chosen cost function is defined by

$$g_{GSC} = |i_{f,ref}^d[k+1] - i_f^d[k+1]| + |i_{f,ref}^q[k+1] - i_f^q[k+1]| \tag{9}$$

where $i_{f,ref}^d[k+1]$ and $i_{f,ref}^q[k+1]$ are the reference values of the d- & q-axis currents. The reference d-axis current $i_{f,ref}^d[k+1]$ is calculated from the outer DC-link voltage control loop and the reference q-axis current $i_{f,ref}^q[k+1]$ is set to zero [8], see Fig 1. Again, (6) is evaluated for each of the possible seven voltage vectors, giving seven different current predictions. The voltage vector whose current prediction is minimizing the cost function (9) is applied at the next sampling instant.

Table I: Parameters of PMSG-based WECS.

Name	Nomenclature	Value
Wind turbine radius	r_t	1.6 m
Rated wind speed	v_{wrated}	20 m/sec
PMSG rated power	p_{nom}	20 kW
PMSG rated voltage (line-line]	u_s^{rms}	400 V
Number of pair poles	n_p	3
Stator resistance	R_s	0.2 Ω
Stator inductance	L_s	15 mH
Permanent magnet flux	ψ_{pm}	0.85
PMSG moment of inertia	Θ	0.8 $\frac{\text{kg}}{\text{m}^2}$
DC capacitor	C_{dc}	3mF
DC-link voltage	u_{dc}	700V
Grid line-line voltage	u_o	400V
Grid normal frequency	f_e	50Hz
Filter resistance	R_f	0.16Ω
Filter inductance	L_f	10mH

IV. LVRT Strategies

A. Braking Chopper (BC) strategy

The braking chopper (BC) or active crowbar system is proposed to be placed in the DC-link to dissipate the active power during the grid faults [3]. Braking chopper consists of a high power resistor R_C [Ω] with a series switch as shown in Fig. 1. The advantages of this structure are low cost and simple control structure. However, BC can only dissipate power and cannot improve the reactive power injection to the grid. The BC switch is turn on when the DC-link voltage reaches a certain threshold value (i. e. $1.1u_{dc}$).

B. Proposed LVRT Strategy

The proposed control system, Fig. 1, has the ability to work during the normal operation conditions or during the fault operation conditions. During the normal operation conditions the MPPT is achieved by the MSC control system via controlling the q-axis current of the PMSG i_s^q. Therefore, the switch in Fig. 1 will connected to position N which denote normal operation conditions. The DC-link voltage regulation is achieved by the GSC control system by controlling the d-axis current of the filter i_f^d.

During the fault operation condition, in order to store the surplus power in the PMSG-rotor inertia, the q-axis reference current of the PMSG is multiplied by the factor $k_F = \frac{\|\boldsymbol{u}_{o,f}\|}{\|\boldsymbol{u}_o\|}$, where $\|\boldsymbol{u}_{o,f}\|$ and $\|\boldsymbol{u}_o\|$ are the magnitude of the grid voltage during and before the fault, respectively. Therefore, the switch in Fig. 1 will connected to position F which denote fault operation conditions. Thus, the generated active power decreases and the difference between the WT output mechanical power and the generated power will be stored in the PMSG-rotor inertia. Therefore, the mechanical speed of the PMSG rotor increases. After the fault clearance, The MSC returns to the normal operations and the energy stored in the inertia will be released back to the grid. Consequently, the speed of the PMSG decreases.

V. Simulation Results

A simulation model of a 20 [kW] WECS with PMSG is implemented in Matlab/Simulink. The system parameters are listed in Table I. The implementation is shown in Fig. 1.

Fig. 2 shows the performance of the WECS during a three phase fault without applying any LVRT strategy. At $t = 0.2$ [s]

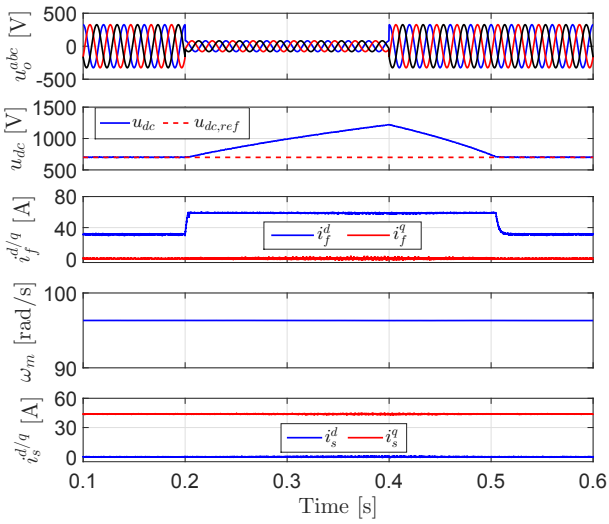


Figure 2: Performance of the WECS during a three-phase fault without LVRT strategy (from top): grid voltages $\boldsymbol{u}_o^{abc}$, Dc-link voltage (u_{dc}, $u_{dc,ref}$), currents of the filter (i_f^d, i_f^q), mechanical speed of the PMSG rotor ω_m, currents of the PMSG (i_s^d, i_s^q).

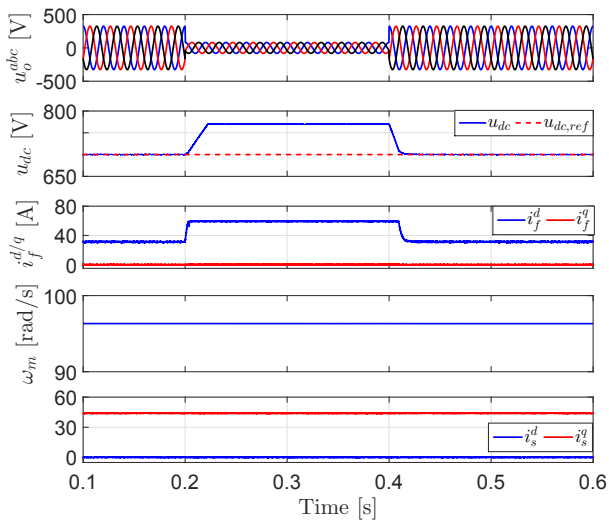


Figure 3: Performance of the WECS during a three-phase fault with BC (from top): grid voltages $\boldsymbol{u}_o^{abc}$, Dc-link voltage (u_{dc}, $u_{dc,ref}$), currents of the filter (i_f^d, i_f^q), mechanical speed of the PMSG rotor ω_m, currents of the PMSG (i_s^d, i_s^q).

a 80% voltage dip in the grid voltage for 200 [ms] has been occurred, see Fig. 2 (top). According to Fig. 2, the DC-link voltage increased to a very high value (approximately double the normal value), which is enough to destroy the DC-link capacitor. Also, the output current of the GSC increases to the maximum allowable value (i. e. $\boldsymbol{i}_{f,max} = 60$ [A]). However, the voltage dip did not have any effect on the MSC as the mechanical speed ω_m and the d- & q-axis current of the PMSG are constant. Thanks to the full decoupling between the grid and the PMSG.

The performance of the WECS under the same three phase fault with BC is illustrated in Fig. 3. When the DC-link voltage reached the threshold limit (i. e. $1.1u_{dc}$), the BC resistance is connected in parallel to the DC-link capacitor. Therefore, the surplus power is dissipated in the BC resistance and the DC-link voltage u_{dc} is kept constant at the threshold value. Again,

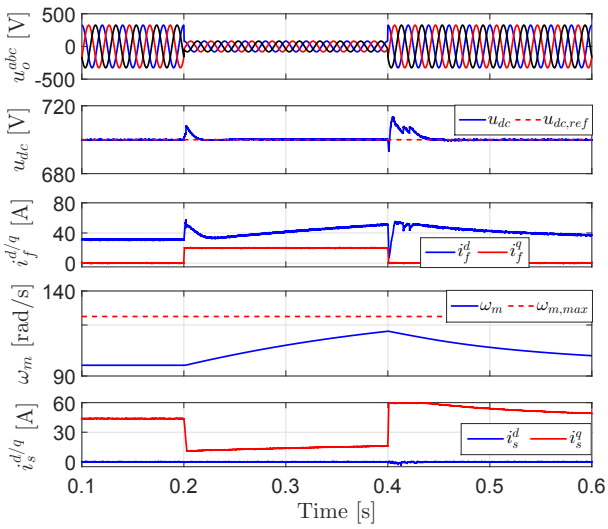


Figure 4: Performance of the WECS during a three-phase fault with the proposed LVRT strategy (from top): grid voltages $\boldsymbol{u}_o^{abc}$, Dc-link voltage (u_{dc}, $u_{dc,ref}$), currents of the filter (i_f^d, i_f^q), mechanical speed of the PMSG rotor ω_m, currents of the PMSG (i_s^d, i_s^q).

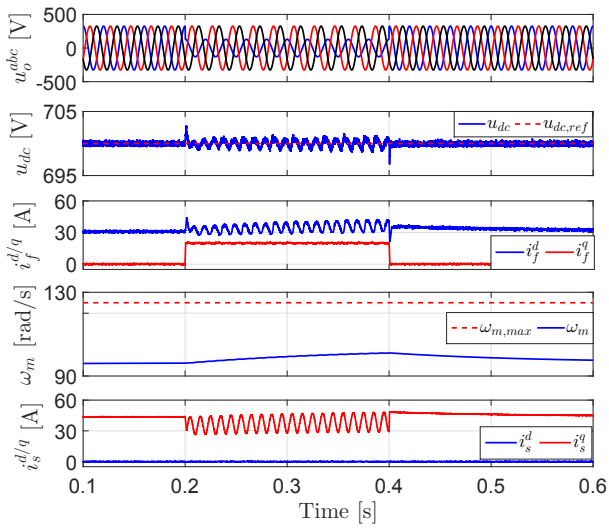


Figure 5: Performance of the WECS during a single-phase fault with the proposed LVRT strategy (from top): grid voltages $\boldsymbol{u}_o^{abc}$, Dc-link voltage (u_{dc}, $u_{dc,ref}$), currents of the filter (i_f^d, i_f^q), mechanical speed of the PMSG rotor ω_m, currents of the PMSG (i_s^d, i_s^q).

the mechanical speed ω_m and the d- & q-axis current of the PMSG are not affected by the fault (see Fig. 3). However, the main disadvantages of this BC solution is that the surplus power is simply dissipated in the BC resistance and it cannot realize the new grid code requirements for reactive power injection to the grid.

The performance of the WECS under the same three phase fault with the proposed LVRT strategy (i.e. using the generator inertia) is illustrated in Fig. 4. The DC-link voltage is kept almost constant at its reference value. However, a small deviation in the DC-link voltage appeared after the voltage dip appearance and clearance due to the swap of the MSC controller between the normal state and fault state. Moreover, a reactive current is injected to the grid during the fault using the q-axis current of the output filter i_f^q. Thus, the LVRT requirements are realized using the proposed LVRT strategy. The proposed LVRT strategy forces the generator to reduce the active power extracted during the fault, see Fig. 4. Hence, the generator current is reduced and also the generator electromagnetic torque. As a consequence, there is a torque mismatch in the mechanical system between the wind turbine and the PMSG, which causes the speed to increase, see Fig. 4. However, the speed of the PMSG still lower than the maximum allowable speed. Therefore, no stress on the mechanical system is produced. After the dip clearance, The MSC returns to the normal operations and the energy stored in the inertia is delivered to the grid. Therefore, the speed of the PMSG starts to decrease.

Fig. 5 shows performance of the WECS with the proposed LVRT strategy under a single-phase fault in phase a with 60% voltage dip. From the simulation results, it can be seen the results are similar to those of the case of symmetrical faults. The DC-link voltage is kept constant at its reference value. The oscillations of the DC-link voltage, output filter d-axis current i_f^d, and PMSG q-axis current i_s^q are due to the negative-sequence component.

VI. Conclusion

A FCS-MPC approach has been presented to fulfill the LVRT requirements in a variable-speed WECS based on a PMSG connected to the grid with a full scale back-to-back converter and filter. Faults in the grid cause a transient active power surplus in the WECS. The proposed LVRT strategy allows to store the energy surplus in the inertia of the PMSG-rotor, when either balanced or unbalanced grid faults appear. The performance of the proposed LVRT strategy is compared with the traditional braking chopper solution via simulation results. The results have shown that the proposed LVRT succeeded in keeping the DC-link voltage constant at its reference value and injecting active/reactive power to the grid during symmetrical or unsymmetrical faults.

References

[1] M. Liserre, R. Cardenas, M. Molinas, and J. Rodriguez, “Overview of Multi-MW Wind Turbines and Wind Parks”, *IEEE Transactions on Industrial Electronics*, Vol. 58, No. 4, pp. 1081–1095, April 2011.

[2] M. Tsili, S. Papathanassiou, “A review of grid code technical requirements for wind farms”, *IET Renewable Power Generation*, Vol. 3, No. 3, pp. 308–332, 2009.

[3] J. Conroy, R. Watson, “Low-voltage ride-through of a full converter wind turbine with permanent magnet generator”, *IET Renewable Power Generation*, Vol. 1, No. 3, pp. 182–189, 2007.

[4] B. Singh, et. al., “Static synchronous compensators (STATCOM): a review”, *IET Power Electronics*, Vol. 2, No. 4, pp. 297–324, 2009.

[5] Z. Liu, C. Liu, G. Li, “Power coordinated control of wind turbines with permanent magnet synchronous generator for low voltage ride through”, *IEEE PES General Meeting (Conference & Exposition)*, pp. 1–5, 2014.

[6] X. Guoyi, X. Lie, J. Morrow, “Power oscillation damping using wind turbines with energy storage systems”, *Renewable Power Generation*, Vol. 7, No. 5, pp. 449–457, 2013.

[7] S. Alepuz, A. Calle, S. Busquets-Monge, S. Kouro, B. Wu, “Use of Stored Energy in PMSG Rotor Inertia for Low-Voltage Ride-Through in Back-to-Back NPC Converter-Based Wind Power Systems”, *IEEE Transactions on Industrial Electronics*, Vol. 60, No. 5, pp. 1787–1796, 2013.

[8] M. Abdelrahem, C. Hackl, and R. Kennel, “Model Predictive Control of Permanent Magnet Synchronous Generators in Variable-Speed Wind Turbine Systems”, *in Proceedings of Power and Energy Student Summit (PESS 2016)*, Aachen, Germany, 19-20 January 2016.

[9] P. Cortes, M. Kazmierkowski, R. Kennel, D. Quevedo, J. Rodriguez, “Predictive Control in Power Electronics and Drives”, *IEEE Transactions on Industrial Electronics*, Vol. 55, No. 12 ,pp. 4312–4324, Dec. 2008.

Power Frequency Transformer-less Grid Connected Photovoltaic Arrays with Reduced DC Sources and Electronic Components

A. M. Mahfuz- Ur- Rahman,[1,*] and Md. Rabiul Islam[1]

[1]Department of Electrical & Electronic Engineering, Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh

[*]mahfuzbappi68@gmail.com

Abstract—**Photovoltaic system is one of the most promising renewable energy to meet the energy demand and reduce the global warming. Making the photovoltaic system efficient and compact is a challenge for the researchers. Multilevel inverters with high frequency transformer link eliminate the use of power frequency transformer for direct grid connected photovoltaic arrays. In this paper using high frequency transformer link a fifteen level inverter for 11 KV system is proposed with reduced dc sources and electronic components.**

Index Terms—**Multilevel inverter, high frequency transformer, power frequency transformer-less.**

I. Introduction

Renewable energy reduces global carbon emissions and the pace of investment has greatly increased as the cost of technologies fall and efficiency continues to rise. Solar energy one of the most auspicious form of renewable energy resists the climate change. Photovoltaic (PV) arrays with high frequency transformer link and multilevel inverters (MLI) [4]-[6] eliminate the use of power frequency transformer [1], [2]. Increase of the number of levels in MLI provides better output voltage waveform and lower total harmonic distortion (THD). But number of dc sources and switches increases with the increase of level of MLI. In this paper a fifteen level inverter consists of 3 dc sources and 10 switches per phase is proposed for 11 KV grid connected PV arrays. The voltage imbalance problem has solved using high frequency transformer link. The performance of existing topologies and proposed topology is also compared using MATLAB/SIMULINK.

II. Conventional 15-Level Cascaded MLI

A three phase 15 level cascaded H bridge inverter is shown in fig. 1. It consists of 7 cells and 7 dc sources per phase. A cell consists of 4 switches which is capable of producing three voltage levels +Vdc, -Vdc and 0. When another cell is added like fig. 1 then number of level increases by two. Thus 7 cells can produce 15 voltage levels and total 84 switches and 21 dc sources are required for 15 level three phase cascaded MLI.

III. Existing 7-Source 17-Switch Topology

This topology shown in fig. 2 has a series connected set of inverters which can produce a stepped output [7]. Each voltage sources connected with two switches except the first one. One switch is used to include the corresponding voltage source and another switch is used to omit that voltage source. The gate pulses are made such that the two switches do not conduct same time to prevent short circuit of the corresponding voltage source. Here 7 different voltages can be obtained as shown in Table. I by including or omitting the dc sources with the help of gate pulses. S1, S3 switches produces positive voltages and S2, S4 produces negative voltages. Thus it serves fifteen different voltages positive 7 voltages, negative 7 voltages and 0. 0 voltage can be obtained by turning off the S5 switch.

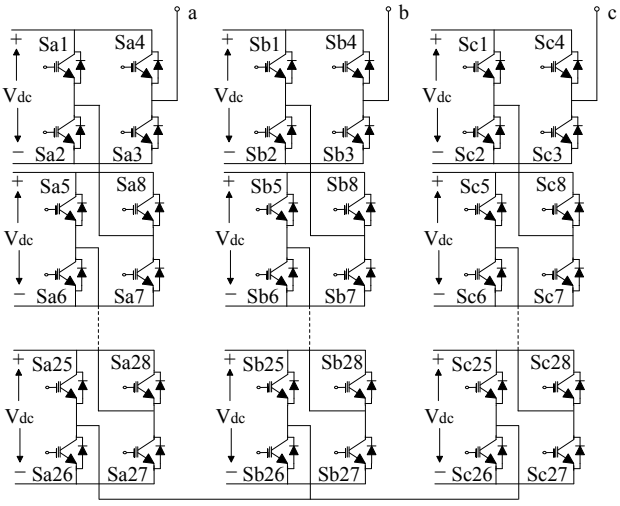


Fig. 1. Conventional 15 level three phase H bridge inverter.

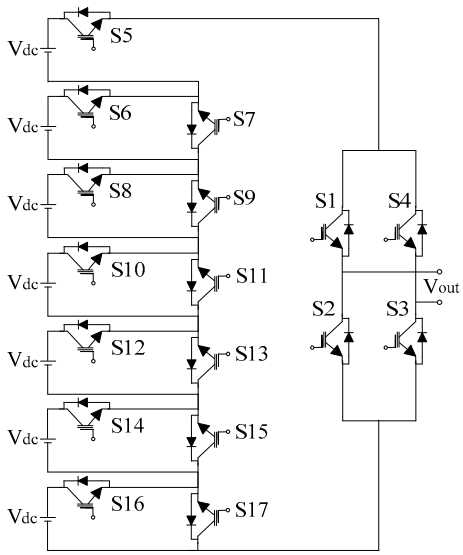


Fig. 2. 15 level 7-source 17-switch single phase inverter topology.

TABLE I
SWITCHING STATES OF 7-SOURCE 17-SWITCH TOPOLOGY

State	On Switches	Vout
1	S5, S7, S9, S11, S13, S15, S17	Vdc
2	S5, S6, S9, S11, S13, S15, S17	2Vdc
3	S5, S6, S8, S11, S13, S15, S17	3Vdc
4	S5, S6, S8, S10, S13, S15, S17	4Vdc
5	S5, S6, S8, S10, S12, S15, S17	5Vdc
6	S5, S6, S8, S10, S12, S14, S17	6Vdc
7	S5, S6, S8, S10, S13, S15, S16	7Vdc

So, for three phase 7-source 17-switch topology 21 voltage sources and 51 switches are required.

IV. EXISTING 7-SOURCE 11-SWITCH TOPOLOGY

A 7-source 11-switch topology for 15 level inverter is proposed in [8]. Only one switch will be conducting for a particular state as shown in fig. 3. When S11 is on then output voltage is Vdc. Output voltage is maintained 7Vdc when S5 is on. S1, S2, S3 & S4 are used to obtain positive and negative voltages. 0 Voltage is obtained by turning off all the switches. The main problem of this topology is switch voltage withstanding capability is too high.

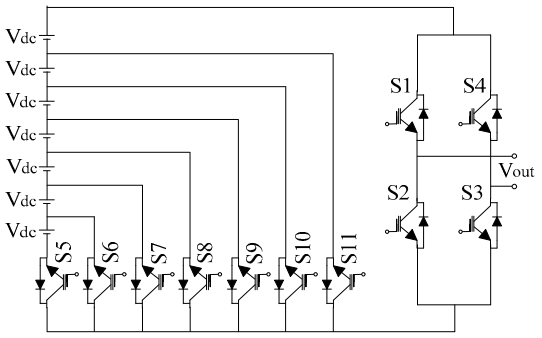


Fig. 3. 15 level 7-source 11-switch single phase inverter topology

TABLE II
SWITCHING STATES OF 7-SOURCE 11-SWITCH TOPOLOGY

State	On Switches	Vout
1	S11	Vdc
2	S10	2Vdc
3	S9	3Vdc
4	S8	4Vdc
5	S7	5Vdc
6	S6	6Vdc
7	S5	7Vdc

Table. II shows different voltage states of 7-source 11-switch topology. So, in this topology 21 sources and 33 switches are required for three phase system.

V. PROPOSED 3-SOURCE 10-SWITCH TOPOLOGY

In this proposed topology 3 dc sources Vdc, 2Vdc, 4Vdc and 10 switches are required. This topology is quite similar as the 7-source 17-switch topology. The difference is, here all the dc sources have different values and are connected with 2 switches. The 2 switches are used to include or omit the corresponding dc sources. The main difficulty of this proposed topology is to balance the voltage sources as required. This problem is solved by controlling the number of turns of secondary side of a multi-input multi-output high frequency transformer as shown in fig. 4 which also eliminates the use of power frequency transformer for a grid connected PV array system. The maximum power point tracker (MPPT) enhances the output power by 20% [3] connected with PV array as shown in the fig. 4. The switching states of phase A are shown in Table. III for the proposed topology.

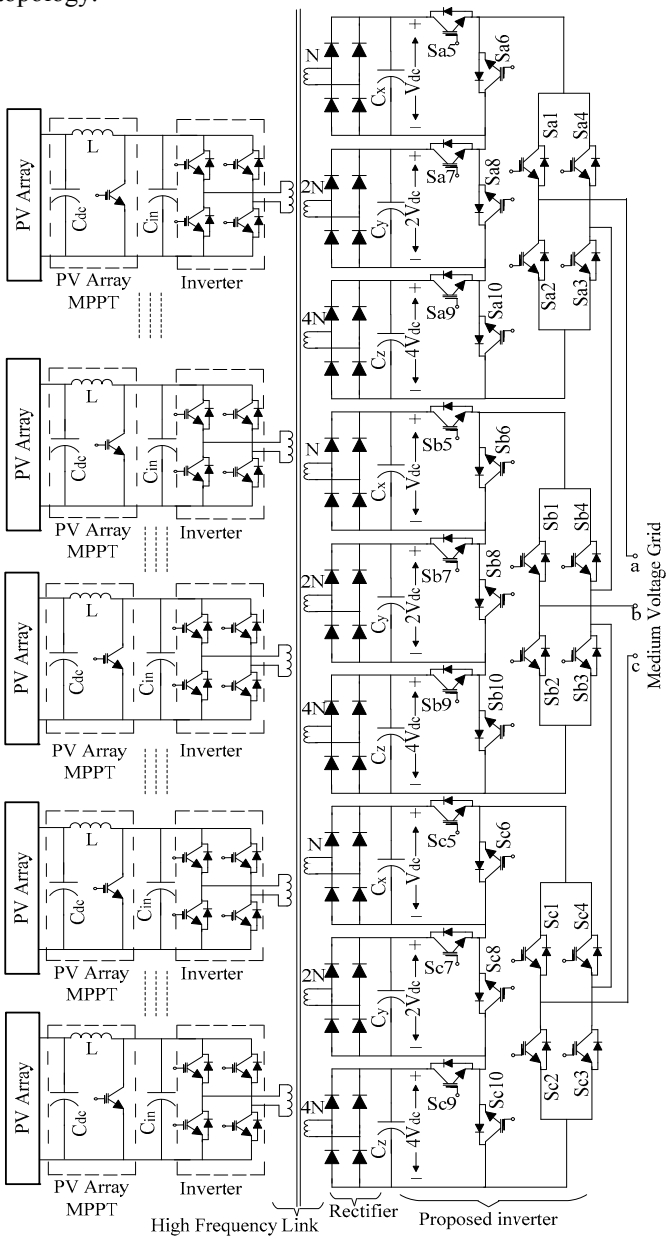


Fig. 4. Power frequency transformer-less proposed grid connected PV system with reduced dc sources and electronic components.

The proposed topology not only applicable for 15 level inverter but also a variety of levels can be made with this technique. In this topology the magnitudes of dc sources are related with the series of 2^x where x=0 to n-1. Here n is the total number of dc sources. If there are 5 sources then the magnitudes of the sources will be 2^0Vdc, 2^1Vdc, 2^2Vdc, 2^3Vdc, 2^4Vdc. Coefficients can be expressed as equation. 1

$$V_{coff} = 1,2,4,8.....2^{n-1} \quad (1)$$

$$S_n = 2n + 4 \quad (2)$$

$$V_{Lev} = 2\sum_{x=0}^{n-1} 2^x + 1 \quad (3)$$

S_n is the number of switches which is multiplied by 2 as each source needs 2 switches and then added with 4 because there needs 4 switches for positive and negative states of voltages.

In case of no. of voltage levels $V_{Lev,}$ summation of all the coefficients of dc sources are multiplied with 2 as positive and negative states of voltage and 1 is added for 0 Voltage. Table. IV shows up to 5 no. of sources the required no. of switches, voltage levels including magnitudes of the sources for the proposed topology.

TABLE III
SWITCHING STATES OF 3-SOURCE 10-SWITCH TOPOLOGY

State	On Switches	Vout
1	Sa6, Sa8, Sa10	0
2	Sa5, Sa8, Sa10	Vdc
3	Sa6, Sa7, Sa10	2Vdc
4	Sa5, Sa7, Sa10	3Vdc
5	Sa6, Sa8, Sa9	4Vdc
6	Sa5, Sa8, Sa9	5Vdc
7	Sa6, Sa7, Sa9	6Vdc
8	Sa5, Sa7, Sa9	7Vdc

From fig. 4 it can be seen that for 15 level inverter the proposed topology requires 30 switches and 9 voltage sources.

TABLE IV
PROPOSED TOPOLOGY FOR DIFFERENT MLI

No. of sources per phase	Magnitudes of Sources	No. of switches per phase	No. of Voltage levels
2	Vdc,2Vdc	8	7
3	Vdc,2Vdc,4Vdc	10	15
4	Vdc,2Vdc,4Vdc,8Vdc	12	31
5	Vdc,2Vdc,4Vdc,8Vdc,16Vdc	14	63

VI. GATE PULSE GENERATION & SIMULATION RESULTS

By comparing the modulating signal and carrier signal gate pulses are produced for the corresponding switches. In this simulation third harmonic injected pulse width modulation technique is used as it produces 15% higher fundamental voltage with the same dc bus voltage when the amplitude of the common mode signal is 1/6 times fundamental amplitude [9]. Between two schemes level shifted and phase shifted carrier based modulation level shifted is taken into consideration as it shows better harmonic spectra property. For 15 level inverter 14 in-phase disposition carriers (where all the carriers are in phase) are used for the pulse generation. For simulation following specifications are taken into consideration

Output line voltage= 11KV Carrier frequency= 5 KHz
Output frequency= 50Hz Modulating index=1
Converter level= 15 Number of phases=3

The gate pulse generation for Sa10 switch is shown in fig 5. From Table. III it can be seen that Sa10 switch remains in switching mode between state 4 and state 5. So, when the modulating signal remains between level 11 to 12 and level 4 to 5 then the comparison of modulating signal and carrier signal will cause for switching of Sa10 switch. Here level 11, 5 and level 12, 4 are the representation of state 4 and state 5 respectively in both positive and negative direction. Switch Sa10 is in conduction mode between state 1-2, state 2-3 and state 3-4. So, considering positive and negative the Sa10 will be in conduction when the modulating signal remains in 6 regions. The 6 regions are level 8-9, level 9-10, level 10-11, level 5-6, level 6-7 and level 7-8 as shown in fig. 5. By adding switching and conduction pulses the net pulse is produced.

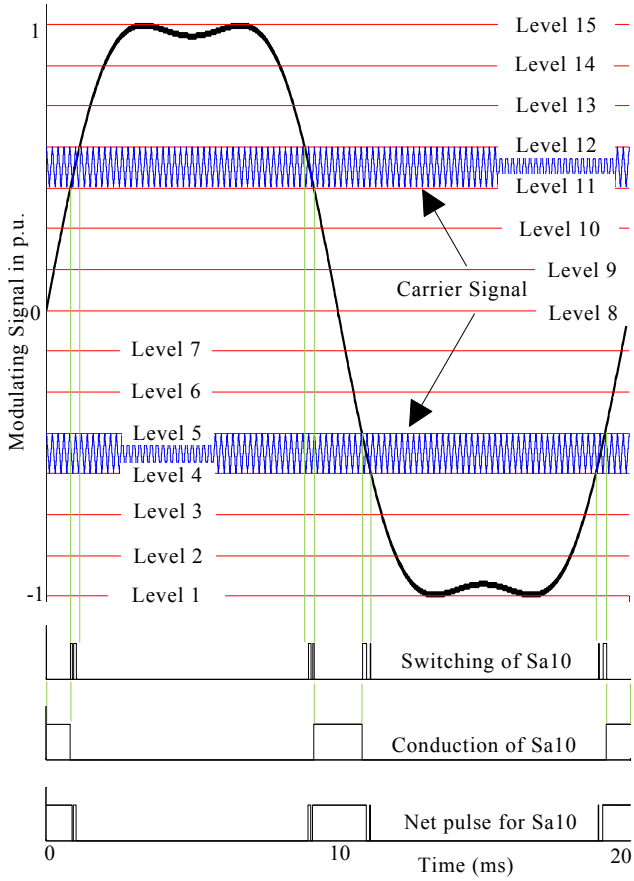


Fig. 5. Gate pulse generation for Sa10 switch

The gate pulses for other switches are also produced in the similar way. The gate pulses for other switches are shown in fig. 6.

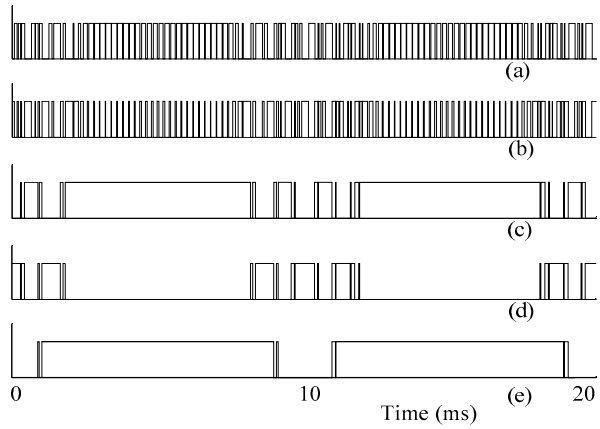


Fig. 6. Gate pulses corresponds to (a) Sa5 (b) Sa6 (c) Sa7 (d) Sa8 (e) Sa9.

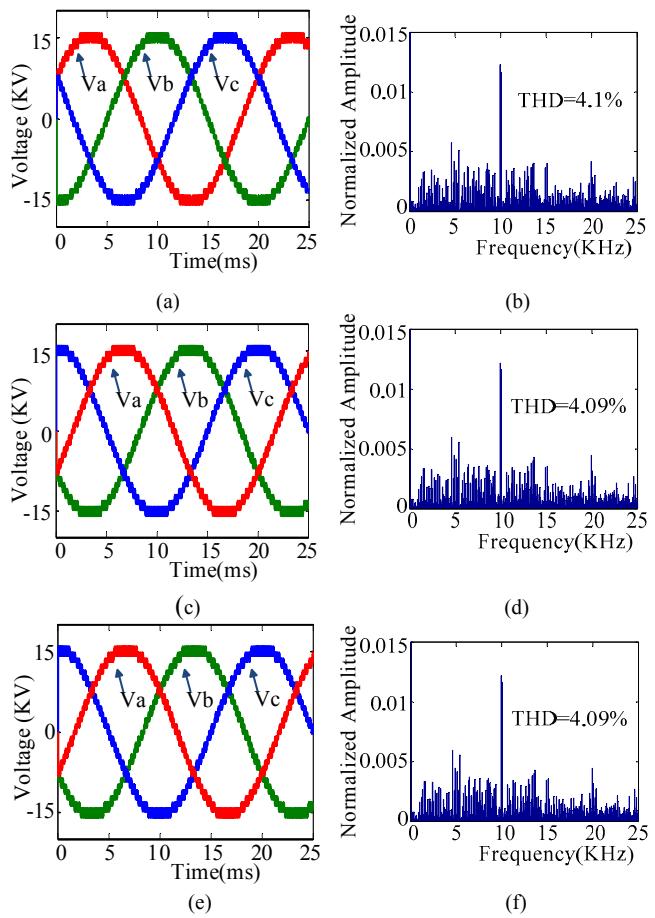


Fig 7. Output line voltage and THD corresponds to (a), (b) H bridge topology (c), (d) 7-source 17-switch topology (e), (f) 7-source 11-switch topology.

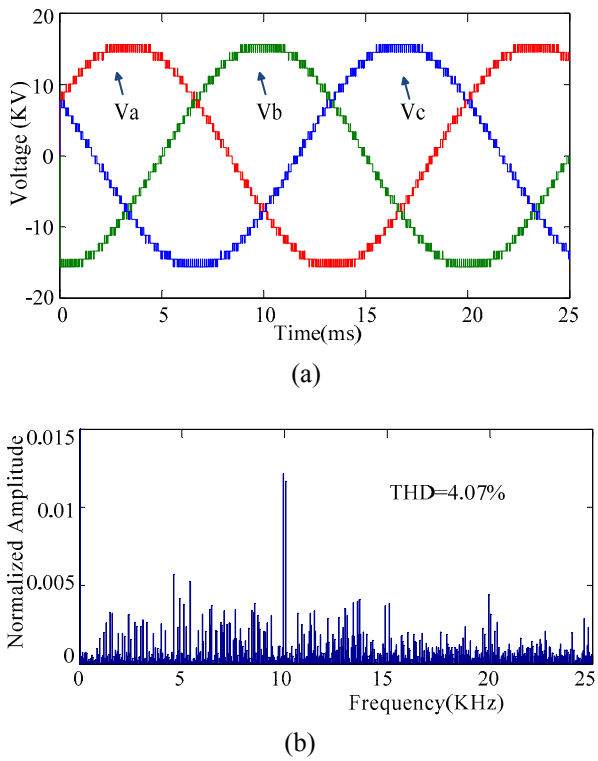


Fig. 8. (a) Output line voltage for proposed topology (b) THD of line voltage of proposed topology.

Fig. 7 shows the output line voltage and frequency spectrum of existing topologies. In fig. 8 the output line voltage and frequency spectrum of proposed topology is shown where THD= 4.07%.

Table. V & Table VI shows the comparison of different topologies with the proposed topology. Here conventional 1, conventional 2 & conventional 3 refer to H bridge topology, 7-sourse 17-switch topology and 7-source 11-switch topology respectively.

TABLE V
COMPARISON OF DIFFERENT TOPOLOGIES WITH PROPOSED TOPOLOGY FOR 15 LEVEL SINGLE PHASE INVERTER

	Conven-tional 1	Conven-tional 2	Conven-tional 3	Proposed
No. of sources	7	7	7	3
No. of switches	28	17	11	10

TABLE VI
COMPARISON OF DIFFERENT TOPOLOGIES WITH PROPOSED TOPOLOGY FOR 31 LEVEL SINGLE PHASE INVERTER

	Conven-tional 1	Conven-tional 2	Conven-tional 3	Proposed
No. of sources	15	15	15	4
No. of switches	60	33	19	12

VII. CONCLUSION

The proposed topology shows better performance with the increase of number of levels. This proposed topology not only reduces number of dc sources and electronic components but also enhances the output waveform than the existing topologies. So, using this topology a power frequency transformer-less, cost effective, efficient PV system can be developed.

REFERENCES

[1] M. R. Islam, Y. G. Guo, and J. G. Zhu, "A High-Frequency Link Multilevel Cascaded Medium Voltage Converter for Direct Grid Integration of Renewable Energy Systems," *IEEE Trans. on Power electronics*, Vol. 29, No. 8, pp. 4167-4182, August 2014.

[2] M. R. Islam, Y. G. Guo, and J. G. Zhu, "Multiple-Input Multiple-Output Medium Frequency-Link Based Medium Voltage Inverter for Direct Grid Connection of Photovoltaic Arrays," *International Conference on Electrical Machines and Systems*, Busan, Korea,Oct. 26-29, 2013, pp-202-207.

[3] M. R. Islam, Y. G. Guo, J. G. Zhu and M. G. Rabbani, "Simulation of PV Array Characteristics and Fabrication of Microcontroller Based MPPT," *Conference on Electrical and Computer Engineering ICECE 2010,* Dhaka, Bangladesh , 18-20 December 2010, pp-155-158.

[4] J. S. Lai and F. Z. Peng, "*Multilevel converters – a new breed of converters," IEEE Trans. on Industry Applications*, Vol. 32, No. 3 pp. 509–517, 1996.

[5] K. Corzine and Y. Familiant, "A new cascaded multi-level H-bridge drive," *IEEE Trans. on Power electronics*, Vol. 17, pp. 125 – 131, 2002.

[6] T. A. Meynard, M. Fadel and N. Aouda, "Modelling of multilevel converters," *IEEE Trans. on Industrial Electronics*, Vol. 44, No. 3, pp. 356–364, 1997.

[7] Ebrahim Babaei, "A Cascade Multilevel Converter Topology With Reduced Number of Switches", *IEEE Trans. on Power electronics*, Vol. 23, No. 6, pp. 2657-2664, 2008.

[8] T V V S Lakshmi, Noby George, S. Umashankar and D P Kothari, "Cascaded seven level inverter with reduced number of switches using level shifting PWM technique," *International Conference on Power, Energy and Control (ICPEC),* 2013, pp. 676-680.

[9] Joby Jose, G. N. Goyal and M. V, Aware, "Improved inverter utilization using third harmonic injection," *Power Electronics, Drives and Energy Systems (PEDES) & Power India*, 2010, pp. 1-6.

Precise Loss Evaluation Method of CHBMC with Reduced Switching Strategy for Direct Integration with Medium Voltage Smart Microgrids

Md. Mazharul Islam,[1,*] A. M. Mahfuz- Ur- Rahman,[1] and Md. Rabiul Islam,[1]

[1]Department of Electrical & Electronic Engineering, Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh
[*]mazharul.ruet11@gmail.com

Abstract—**Due to the accelerating crisis of the worldwide established fossil fuel and atomic energy system and the impact of green house gas emission led to the development of renewable generation units with medium voltage smart micro grids (e.g., 6-36 KV). The cascaded H-bridge multilevel (CHBM) converters are considered as strong competitors for step-up transformer less direct grid integration. But due to the limited space of installation, high cost of failure and for reliable operation of converters, the power losses in semiconductor devices are required to be designed, analyzed and minimized. In this paper an improved precise loss calculation method is proposed and this method is utilised to calculate the power losses in a 28.6 MVA, 11KV, optimum level with reduced switching modes using bus-clamping pulse width modulation (BCPWM) for CHBM converter. All the analysis are performed in MATLAB environment with reduced harmonic distortion as well as switching losses.**

Index Terms—**Cascaded H-bridge multilevel converter, bus clamping PWM, device voltage utilization factor, switching loss, conduction loss.**

I. Introduction

Currently there are over 432.9 GW of wind power generation and 227.1 GW solar photovoltaic (PV) power generations have been installed globally by 2015 and this amount is increasing day by day to mitigate accumulative crisis [12], [13]. With the arrival of new insulated gate bipolar transistor (IGBT) power modules, an enormously developed advanced magnetic material-based common high frequency link multilevel converter was proposed to meet the needs of future medium- or high-voltage converter system which is based on cascaded H-bridge configuration [1], [3], [6]. Fig. 1 represents basic block diagram of cascaded H-bridge converter for medium- and large-scale PV power plants.

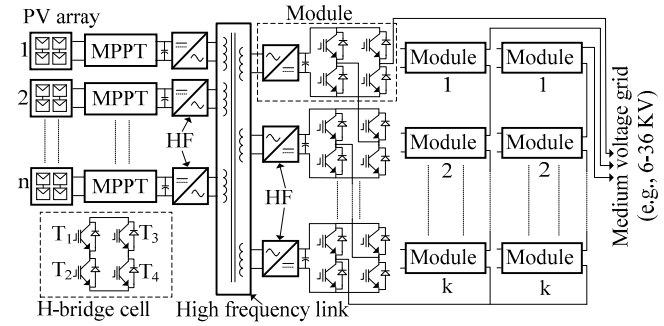


Fig. 1 High frequency link cascaded H-bridge medium voltage converter

Because of the special features, the CHBMC topology is considered as the best possible candidate for medium to high-voltage high-power application [2]. But a small percentage of loss among the total power supplied has a considerable effect on system operating cost, efficiency and reliability.

The overall losses in multilevel converter depend on modulation schemes, operating level and intrinsic behaviour of corresponding IGBTs. BCPWM shows a better spectra property with an extra capability of reducing switching losses [4]. Selection of optimum level is based on control complexity, power loss and device voltage utilization factor (DVUF) to achieve best performance/cost ratio [1], [2]. A higher DVUF is essential (e.g., 90% or above) to design converters with reduced power loss, where DVUF can be computed as:

$$\mathrm{DVUF} = \frac{V_{COM}}{V_{COM}@100\mathrm{FIT}} \tag{1}$$

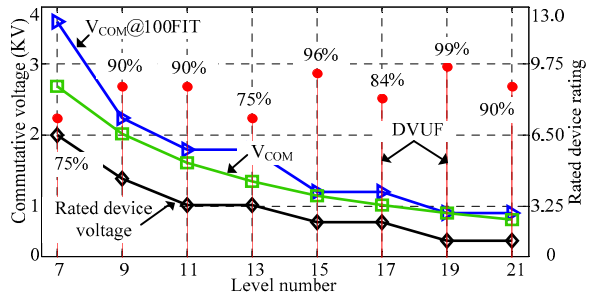
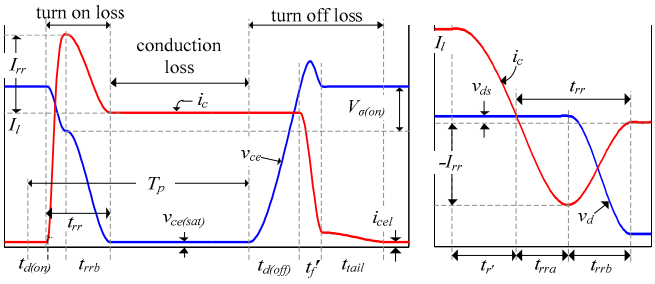


Fig. 2 Selection of level for precise loss calculation of 11KV CHBMC

Fig. 2 summarizes that, compared to 108 IGBTs in total in 3-phase for 19 level converter where 15 level with 84 IGBTs has a considerable lower switching & conduction loss though it has 4% lower DVUF value. That's why to design and analysis of losses in 15-level 11-KV CHBMC a suitable and commercially available IGBT module 5SNA 1500E250300 from ABB (rated voltage & current: 2.5 KV & 1,500 A) is considered in this paper for improved precise loss estimation method using curve fitting and interpolation technique based on measured voltage and current waveforms.

II. Dynamic Behaviours

Generally, the dynamic behaviours of IGBT modules are split based on the losses during switching and conduction mode operation. In modern fast recovery diodes used with IGBTs, the turn-on losses are less than 1% compared to the turn-off losses, so they are neglected for the analysis [5]. The estimation methods which summarize the total power loss per IGBT are switch turn-on, switch turn-off and diode turn-off losses. Fig. 3 indicates the approximated dynamic behaviour of IGBT module 5SNA 1500E250300 [5], [7]. Due to the variety of current in each device differs within the inverter,



(a) (b)

Fig. 3 Approximated behaviour of a) IGBT turn-on, IGBT turn-off and b) diode turn off at DUT condition

the loss evaluation in CHBMC is not an easy task compared to the loss estimation at device under test (DUT) condition.

III. Precise Loss Evaluation Method

A. Conduction Power Loss calculation

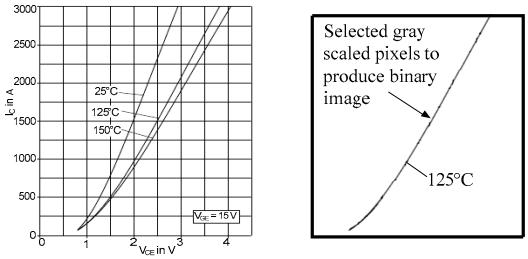
Analytical equations explicating the conduction power losses are very advantageous in the design procedure of CHBMC [8]. During the conduction mode IGBT collector-emitter voltage drop (v_{ce}) is approximated in

$$v_{ce} = v_{ce0} + R_c \cdot i_c \tag{2}$$

Where v_{ce0} represents IGBT on-state zero-current collector-emitter forward voltage drop and R_c is the collector emitter on-state resistance. For anti-parallel diodes

$$v_F = v_{F0} + R_F \cdot i_F \tag{3}$$

Where v_{F0} represents diode on-state zero-current forward voltage drop and R_F is on-state resistance. These important parameters can be obtained from IGBT datasheet [11].



(a) (b)

Fig. 4 (a) Typical IGBT on-state characteristics from data sheet, (b) Original gray-scale image

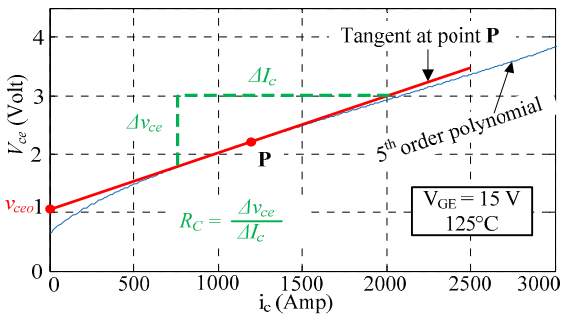


Fig. 5 Obtaining v_{ce0} and R_c from polynomial equation

Using pixel wise gray scale image processing and then using the curve-fitting tool for approximately 1000 points for the fig. 4(b), the 5th order polynomial equation of v_{ce} can be obtained using MATLAB and can be represented as follows

$$v_{ce} = S_1 i_c^5 + S_2 i_c^4 + S_3 i_c^3 + S_4 i_c^2 + S_5 i_c + S_6 \tag{4}$$

Coefficients are:

$S_1 = 2.235\times10^{-17}$
$S_2 = -1.996\times10^{-13}$
$S_3 = 7.118\times10^{-10}$
$S_4 = -1.294\times10^{-06}$
$S_5 = 0.002105$
$S_6 = 0.6739$

The slope of the tangent at point **P** $(\boldsymbol{i_{c1}}, \boldsymbol{v_{ce1}})$ can be computed as

$$\frac{dv_{ce}}{di_c} = 5S_1 i_c^4 + 4S_2 i_c^3 + 3S_3 i_c^2 + 2S_4 i_c^1 + S_5 = m_1 \tag{5}$$

At $i_c = i_{c1}$, $m_1 = R_c$

And the zero-current collector-emitter forward voltage is

$$v_{ce0} = v_{ce1} - (m_1 \,.\, i_{c1}) \tag{6}$$

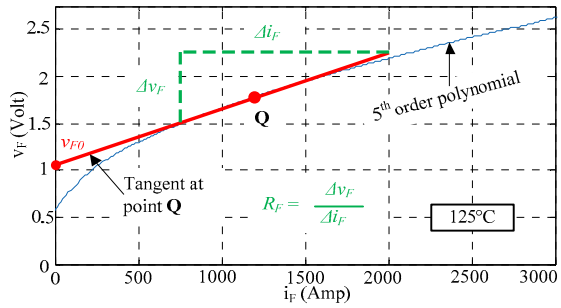


Fig. 6 Obtaining v_{F0} and R_F from polynomial equation

Similarly fig.7 represents diode forward characteristics with the following polynomial equation

$$v_F = D_1 i_F^5 + D_2 i_F^4 + D_3 i_F^3 + D_4 i_F^2 + D_5 i_F + D_6 \tag{7}$$

Coefficients are:

$D_1 = 3.298\times10^{-17}$
$D_2 = -2.907\times10^{-13}$
$D_3 = 9.866\times10^{-10}$
$D_4 = -1.669\times10^{-06}$
$D_5 = 0.001975$
$D_6 = 0.6265$

At point **Q** $(\boldsymbol{i_{F1}}, \boldsymbol{v_{F1}})$

$$R_F = \frac{dv_{ce}}{di_c}_{(i_c = i_{c1})} = m_2 \tag{8}$$

And

$$v_{F0} = v_{F1} - (m_1 \cdot i_{F1}) \tag{9}$$

So the instantaneous IGBT conduction losses can be found as

$$\begin{aligned} p_{cs}(t) &= v_{ce}(t) \cdot i_c(t) \\ &= v_{ceo}(t) \cdot i_c(t) + R_c \cdot i_c^2 \end{aligned} \tag{10}$$

And the average conduction power loss is

$$\begin{aligned} P_{ct} &= \frac{1}{2\pi}\int_0^{2\pi} [p_{cs}(t)] d(\omega t) \\ &= \frac{1}{2\pi}\int_0^{2\pi} [v_{ceo}(t) \cdot i_c(t) + R_c \cdot i_c^2]\, d(\omega t) \end{aligned} \tag{11}$$

$$\therefore\ P_{ct} = v_{ceo} \cdot I_{c,avg} + R_c \cdot I_{c,rms}^2 \tag{12}$$

Where $I_{c,avg}$ and $I_{c,rms}^2$ are the average and rms currents of IGBT. Similarly, the average diode conduction loss (P_{cd}) is

$$\therefore\ P_{cd} = v_{F0} \cdot I_{F,avg} + R_F \cdot I_{F,rms}^2 \tag{13}$$

The total conduction loss per phase with N chips IGBT switch can be expressed as

$$P_{cond/phase} = \sum_{n=1}^{N} [\, P_{ct}(n) + P_{cd}(n) \,] \tag{14}$$

B. Switching Power Loss

Approximation of switching energy using data sheet, based on switching function is the simplest method with further errors [9], [10]. The most accurate method for switching losses evaluation is to precise the collector current wave forms

according to switching transition and conduction mode operation of IGBTs and corresponding anti-parallel diodes.

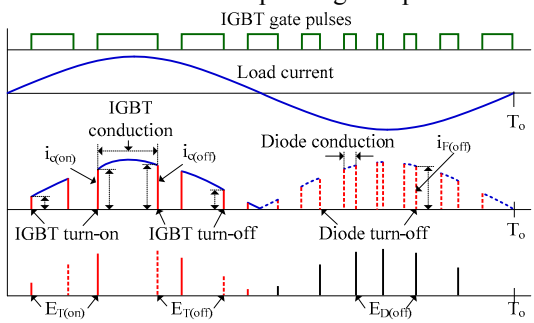


Fig. 7 Switching energy of IGBT and diode against their current at on/off switching moments.

Both switching losses (E_{on} & E_{off}) are proportional to the switching frequency and blocking voltage across the IGBTs.

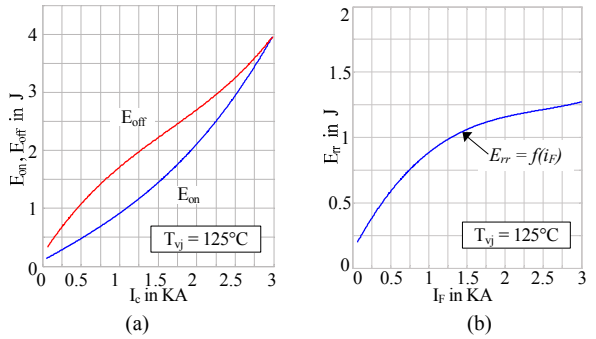


Fig. 8 Switching energies per pulse vs. collector current b) reverse recovery characteristics vs. forward current

The image processing based 5th order polynomial equation of switching energy (E_{on} , E_{off}) can be represented as follows:

$$E_{on} = a_1 i_c^5 + a_2 i_c^4 + a_3 i_c^3 + a_4 i_c^2 + a_5 i_c + a_6 \quad (15)$$

$$E_{off} = b_1 i_c^5 + b_2 i_c^4 + b_3 i_c^3 + b_4 i_c^2 + b_5 i_c + b_6 \quad (16)$$

Coefficients are:

$a_1 = -1.217\times10^{-25}$ $b_1 = 4.309\times10^{-19}$
$a_2 = 8.32\times10^{-22}$ $b_2 = -3.189\times10^{-15}$
$a_3 = 5.391\times10^{-11}$ $b_3 = 1.445\times10^{-10}$
$a_4 = 2.552\times10^{-8}$ $b_4 = -6.775\times10^{-7}$
$a_5 = 0.000738$ $b_5 = 0.002042$
$a_6 = 0.09619$ $b_6 = 0.2036$

Polynomial equation of reverse recovery characteristics of diode at the junction temperature 125°C, can be expressed as:

$$E_{rr} = b_1 i_F^5 + b_2 i_F^4 + b_3 i_F^3 + b_4 i_F^2 + b_5 i_F + b_6 \quad (17)$$

Where the coefficients are:

$c_1 = -2.34\times10^{-24}$
$c_2 = 3.211\times10^{-20}$
$c_3 = 5.448\times10^{-8}$
$c_4 = -0.0004056$
$c_5 = 1.108$
$c_6 = 127.3$

IGBT & diode switching power loss (P_{swt} & P_{rrd}) for fundamental period T_o can be expressed as follows:

$$P_{swt} = \frac{1}{T_o}\frac{V_{dc(min)}}{V_{dc(nom)}}\sum_{j=1}^{N}\left(E_{on_j}(i_c) + E_{off_j}(i_c)\right) \quad (18)$$

$$P_{rrd} = \frac{1}{T_o}\frac{V_{dc(min)}}{V_{dc(nom)}}\sum_{j=1}^{N}\left(E_{rrr_j}(i_F)\right) \quad (19)$$

Where $V_{dc(min)}$ and $V_{dc(nom)}$ are the minimum and nominal DC-link voltage of each H-bridge inverter cell.

IV. Simulation Results

Fig. 9 depicts the switching energy of IGBT and diode (E_{on} , E_{off} & E_{rr}) against their current at on/off switching moments for phase shifted pulse width modulation (PSPWM).

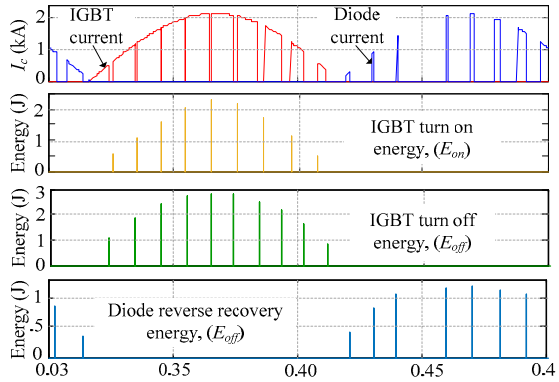


Fig. 9 Turn on/off energy loss of each IGBT and its anti-parallel diode reverse recovery energy loss in a 28.6 MVA, 11 kV PSPWM

Due to higher power loss and harmonic distortion in phase shifted carrier scheme, only level shifted configuration is considered in this paper for different modulation techniques.

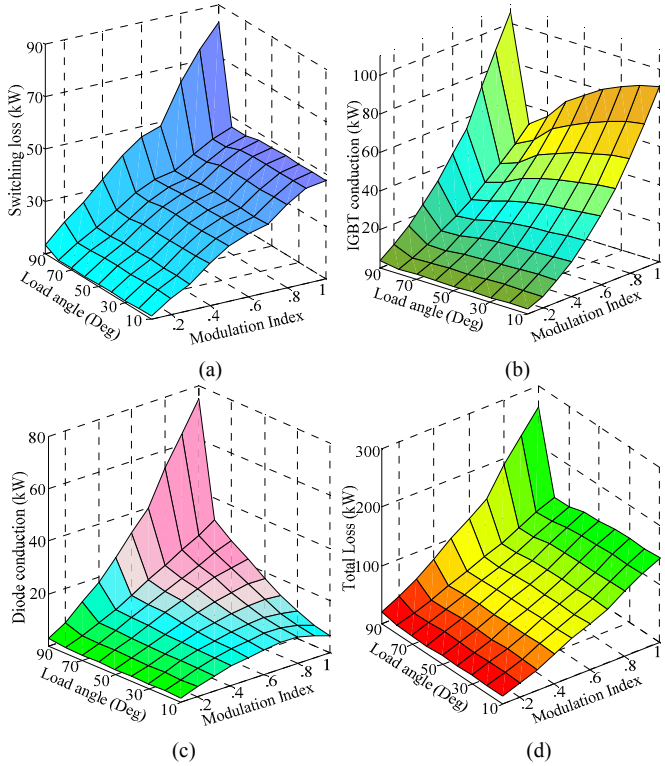


Fig. 10 a) Switching loss(IGBT & diode) b) IGBT conduction loss c) diode conduction loss and d) total loss, against modulation index and power factor for level-shifted SPWM (LSPWM) in a three-phase 11 level CHBMC.

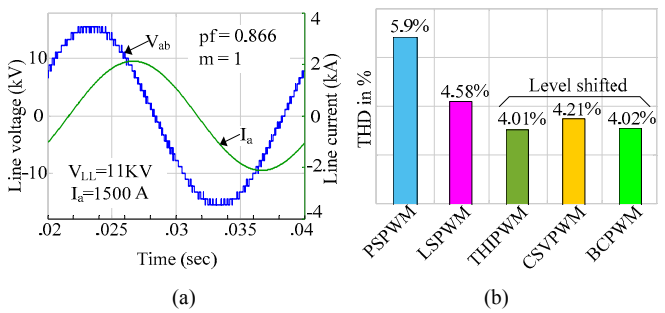


Fig. 11 a) Line voltage and current for 60° BCPWM, b) comparison of THDs in output line to line voltage for different modulation schemes

As depicted in Fig. 10(a) switching losses are almost independent of the power factor in CHBMC. No matter how much

the power factor is, the positive output current always flows through the IGBTs (T_1, T_3, T_2, T_4) and the anti-parallel diodes.

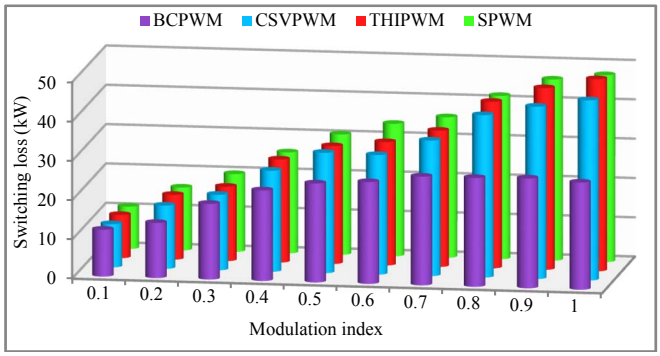


Fig. 11 Comparison of switching losses for different modulation schemes (e.g., SPWM, third harmonic injected PWM (THIPWM), conventional space vector PWM (CSVPWM), 60° bus clamping PWM (60° BCPWM).

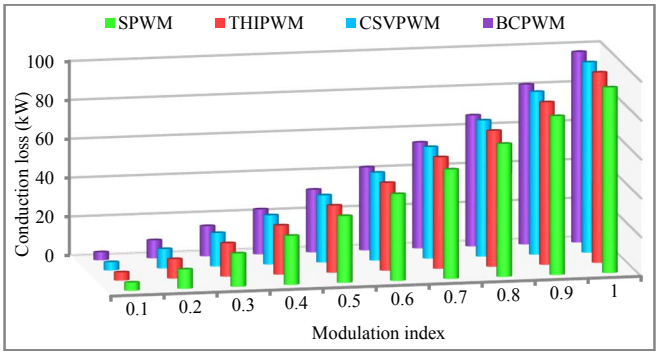


Fig. 12 Comparison of conduction losses for different modulation schemes

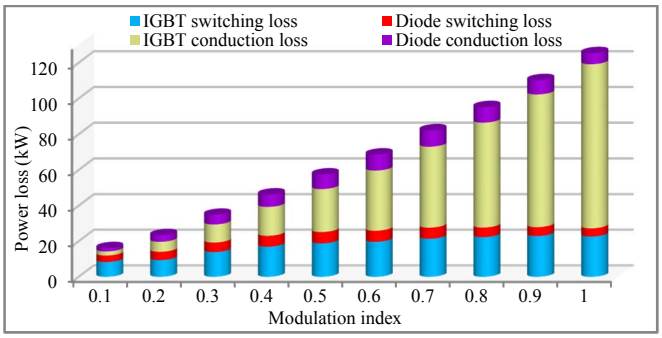


Fig. 13 Overall power losses against different modulation indexes for 11KV CHBM converter with 60° BCPWM

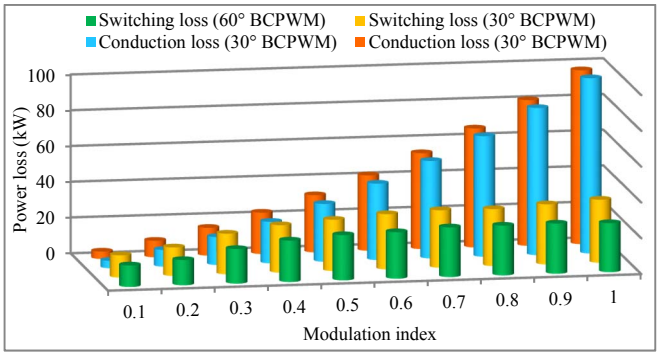


Fig. 14 Switching and conduction power losses for 60° and 30° BCPWM

Fig. 11 & 12 depict that the reduction of inverter total switching loss for 60° BCPWM is almost 42.6%, where the rise of conduction losses is only 3% compared to SPWM for same power factor and modulation index (e.g., pf =.866, m=1). 60° BCPWM has 22.5% lower switching loss and 0.7% higher conduction loss compared to 30° BCPWM as depicted in fig. 14. 60° BCPWM shows better harmonic spectra (e.g., 4.02%)

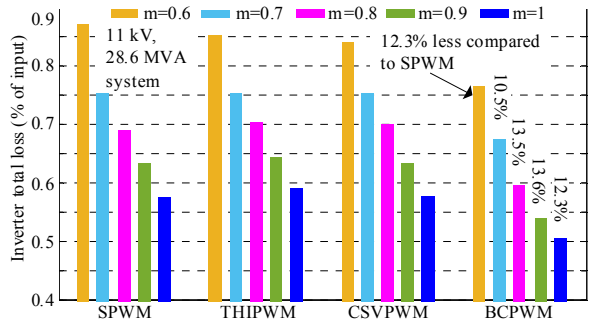


Fig. 15 Comparison of inverter total loss as a percentage of input power for different modulation schemes.

and only 0.5% power loss from input power which is almost 12.3% less compared to SPWM.

V. Conclusions

The presented method for switching and conduction power loss analysis was based on the image processing, simulation and numeric approach. The losses are calculated for different operating points with the approximation of higher order polynomial for the semiconductors characteristics instead of linear ones. Precise loss evaluation method is obligate for accurate power loss modelling and economical design of advanced CHBM converter.

References

[1] M. R. Islam, Y. G. Guo, and J. G. Zhu, *Power Converters for Medium Voltage Network,* Springer-Verlag GmbH, Heidelberg, Germany, 2014.

[2] M. R. Islam, Y. G. Guo, and J. G. Zhu, *Large Scale Renewable Power Generation: Advances in Technologies for Generation, Transmission and Storage*, 1st ed., chapter 5, pages 127–149. Springer-Verlag, Berlin Heidelberg, Germany.

[3] M. R. Islam, Y. G. Guo, and J. G. Zhu, "A High-Frequency Link Multilevel Cascaded Medium-Voltage Converter for Direct Grid Integration of Renewable Energy Systems," *IEEE Trans. Power Electron.,* vol. 29, no. 8, pp. 4167-4182, Aug. 2014

[4] G. Narayanan, H. K. Krishnamurthy, D. Zhao, and R. Ayyanar, "Advanced bus-clamping PWM techniques based on space vector approach," *IEEE Trans. Power Electron.*, vol. 21, no. 4, pp. 974–984, Jul. 2006.

[5] A.D. Rajapakse, A.M. Gole and P.L. Wilson, "Electromagnetic transient simulation models for accurate representation of switching losses and thermal performance in power electronic systems," *IEEE Trans. Power Delivery*, vol. 20, No.1, pp. 319-327, Jan. 2005.

[6] M. R. Islam, Y. G. Guo, and J. G. Zhu, "A Multilevel Medium-Voltage Inverter for Step-Up-Transformer-Less Grid Connection of Photovoltaic Power Plants," *IEEE J. Photovoltaics*, vol. 4, no. 3, pp. 881-889 May 2014

[7] A. D. Rajapakse, A. M. Gole, P. L. Wilson, "Approximation Loss Formula for Estimation of IGBT Switching Losses through EMTP-type Simulation", *International Conference on Power System Transients*, IPST05, Montreal, Canada, Paper No.:IPST05-184, June 2005.

[8] V. Dargahi, A.K. Sadigh, G. K. Venayagamoorthy, K. Corzine, "Analytic determination of conduction power losses in flying capacitor multicell power converter," *2015 IEEE Applied Power Electronics Conference and Exposition (APEC),* Charlotte, NC, 15-19 March 2015

[9] Alamri, Basem, and Mohamed Darwish. "Precise Modelling of Switching and Conduction Losses in Cascaded H-bridge Multilevel Inverters", *In Power Engineering Conference (UPEC)*, 2014 49th International Universities, pp. 1-6. IEEE, 2014.

[10] M. G. H. Aghdam, S. H. Fathi, A. Ghasemi, "The Analysis of Conduction and Switching Losses in Three-Phase OHSW Multilevel Inverter Using Switching Functions", *6th International Conference on Power Electronics and Drive Systems*, September 2005, Malaysia.

[11] Data sheet: ABB 5SNA 1500E250300 HiPak IGBT Module. Available: http://search.abb.com/library/Download.aspx?DocumentID=5SYA%201437-00&LanguageCode=en&DocumentPartId=&Action=Launch

[12] Global wind report, Annual market update:2015, Global Wind Energy Council, http://www.gwec.net

[13] Snapshot of global photovoltaic markets:2015, International Energy Agency-PVPS, http://www.iea-pvps.org

Performance Analysis of Modulation Techniques in Multilevel Inverters for Direct Grid Connected Photovoltaic Arrays

A. M. Mahfuz- Ur- Rahman,[1,*] Md. Mazharul Islam,[1] and Md. Rabiul Islam[1]

[1]Department of Electrical & Electronic Engineering, Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh
[*]mahfuzbappi68@gmail.com

Abstract—**Development of renewable energy has become crucial for the global warming and shortage of fossil fuel. Solar energy is one of the promising renewable energy to meet the energy crisis. One of the big disadvantage of solar array is the efficiency level remains between 14%-25%. So maximum power extraction and reduction of associate losses is a challenge for the researchers. Multilevel inverters are being used in grid connected photovoltaic arrays to enhance the output voltage. This paper analyses the performance of different existing and proposed modulation techniques for an 11 KV system applied in 15 level H bridge inverter in Matlab environment in a view to reduce the harmonics as well as switching loss.**

***Index Terms*—Cascaded multilevel inverters, pulse width modulation, total harmonic distortion, switching loss, PV arrays.**

I. Introduction

Burning fossil fuels creates carbon dioxide, the most dominating greenhouse gas contributing to global warming. On the contrary energy independence and environmental compatibility are two attractive features of PV systems. PV panels require additional equipment like inverter, power frequency step-up transformer for direct grid integration. A multiple-input multiple-output medium frequency-link based medium voltage inverters with maximum power point tracker (MPPT) for step-up transformer-less direct grid connection of PV arrays is proposed in [1]. The MPPT enhances the output power by 20% [2] and the medium frequency link provides electrical isolation between PV array and grid which overcomes common mode and voltage imbalance problems. Further improvement is done in this paper by analysing and comparing proposed pulse width modulation techniques with different types of existing PWM techniques like sinusoidal pulse width modulation (SPWM), Conventional Space vector pulse width modulation (CSVPWM), Third harmonic injected pulse width modulation (THPWM), Trapezoidal pulse width modulation (TRPWM), Sixty degree bus clamping pulse width modulation (SDBCPWM), Thirty degree bus clamping pulse width modulation (TDBCPWM). Here two modulation techniques Third harmonic injected sixty degree bus clamping pulse width modulation (THSDBCPWM) and Third harmonic injected thirty degree bus clamping pulse width modulation (THTDBCPWM) are proposed which show better frequency spectra as well as reduced switching loss.

II. Grid Connected Photovoltaic Array

A three phase PV inverter system for direct medium voltage grid integration is shown in Fig. 1. The inverter is connected to a primary winding of the medium frequency-link and each secondary winding works as an isolated source connected to H-bridge cell through a bridge rectifier. The total harmonic distortion (THD) reduces with the increase of number of levels of the H-bridge inverter. For a particular level of inverter the THD and switching loss can also be reduced by using proper modulation techniques. The performance of these modulation techniques will be analysed in next sections.

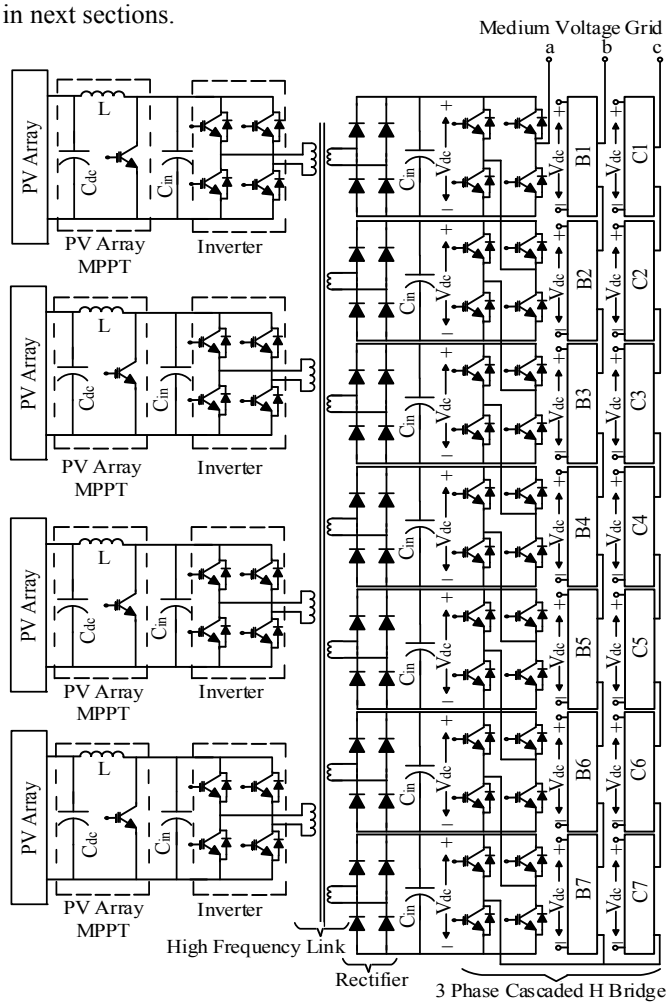


Fig. 1. 15 level three phase PV inverter for direct medium voltage grid integration.

III. Existing Modulation Techniques

A. Modulating Signals

The commonly used conventional modulation techniques are CSVPWM, SPWM, THPWM, SDBCPWM, TDBCPWM and TRPWM. The modulating signals are shown in Fig. 2.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

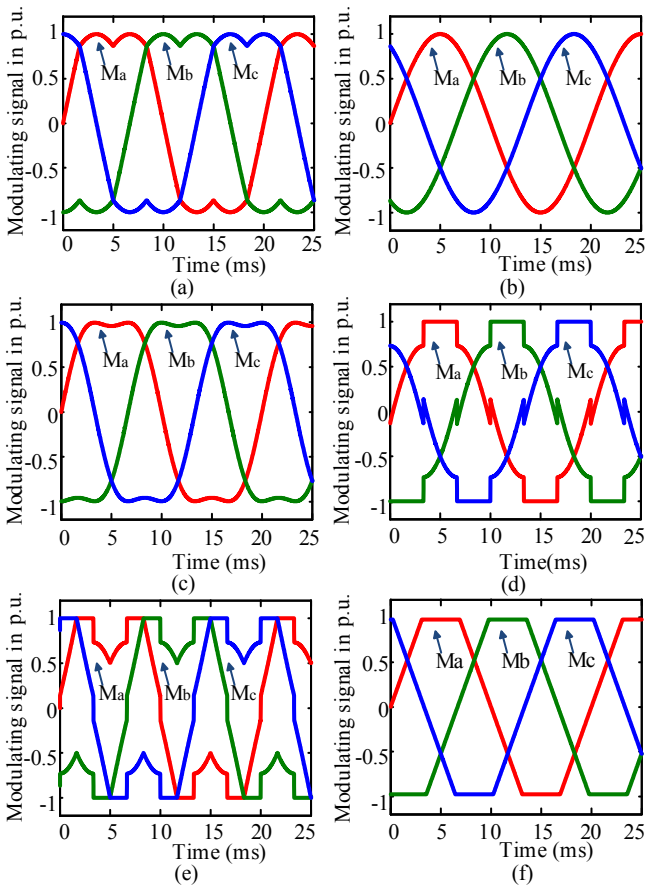


Fig. 2. Modulating Signal corresponding to (a) CSVPWM (b) SPWM (c) THPWM (d) SDBCPWM (e) TDBCPWM (f) TRPWM

The performance comparison of SPWM, THPWM and TRPWM is shown in [3]. The THPWM produces 15% higher fundamental voltage with the same dc bus voltage when the amplitude of common mode signal is1/6 times fundamental amplitude [4]. But if the amplitude of the third harmonic wave is 1/4 times the fundamental signal it shows a very good spectra property which is very close to CSVPWM.

The CSVPWM improves DC bus utilization, reduces current ripple and torque ripple compared to SPWM [5], [6]. CSVPWM is considered a benchmark for pulse width modulation techniques [5]–[8].

Bus clamping PWM (BCPWM) methods are used to reduce the switching loss of the inverter [5], [6]. BCPWM methods reduce harmonic distortion and pulsating torque in motor drives at high speeds [5], [6]. These techniques are also employed in multi-level converters [9].

The SDBCPWM and TDBCPWM methods are considered because these two are the highest and lowest values of line current THD respectively, among BCPWM methods [10]. So these two techniques represent the entire class of BCPWM methods. These BCPWM methods are described in [5], [6].

B. Output Voltage waveform and frequency spectrum

These six modulation techniques are simulated in Matlab environment where 14 level shifted in-phase disposition carriers (where all the carriers are in phase) are compared to these modulating signals and corresponding gate pulses are produced for the IGBTs in Fig. 1. Level shifted carrier scheme shows better harmonic spectra than phase shifted scheme. So level shifted scheme is taken into consideration. For simulation following specifications are taken into consideration.

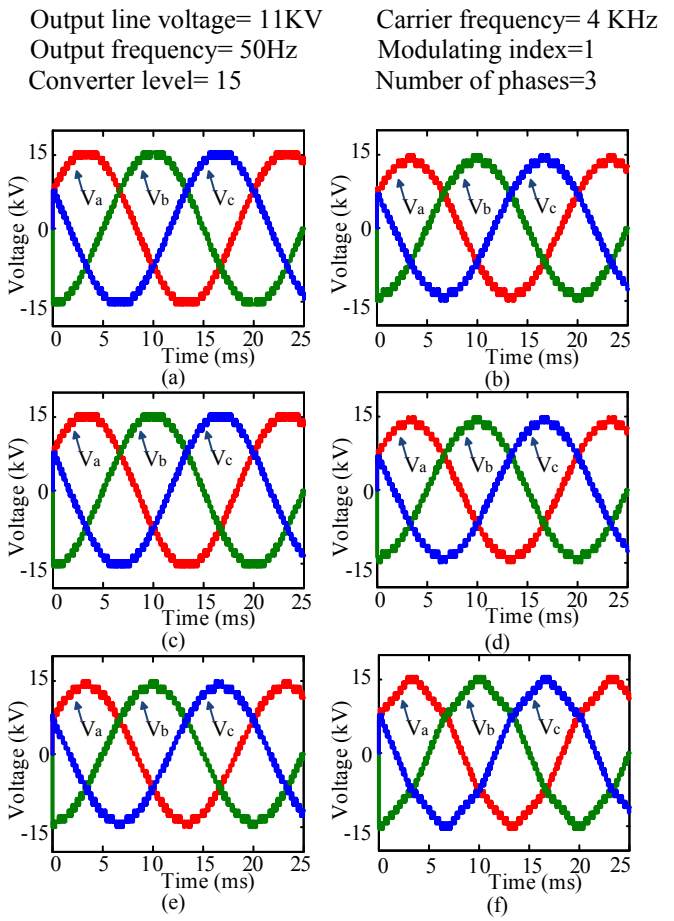


Fig. 3. Output line voltage corresponds to (a) CSVPWM (b) SPWM (c) THPWM (d) SDBCPWM (e) TDBCPWM (f) TRPWM.

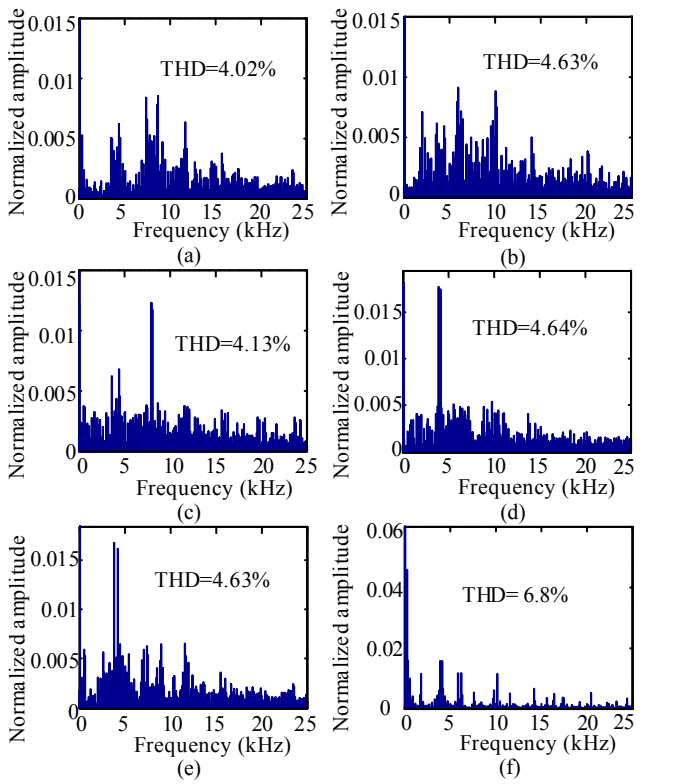


Fig. 4. Total harmonic distortion corresponding to (a) CSVPWM (b) SPWM (c) THPWM (d) SDBCPWM (e) TDBCPWM (f) TRPWM.

Fig. 3 and fig. 4 shows the output voltage waveform and frequency spectrum respectively. The CSVPWM shows the lowest and TRPWM shows the highest harmonic distortion among the analysed modulating signals.

IV. Third Harmonic Injected Bus Clamping PWM

A. Modulating Signals

The proposed modulating signals THSDBCPWM and THTDBCPWM are shown in Fig. 5.

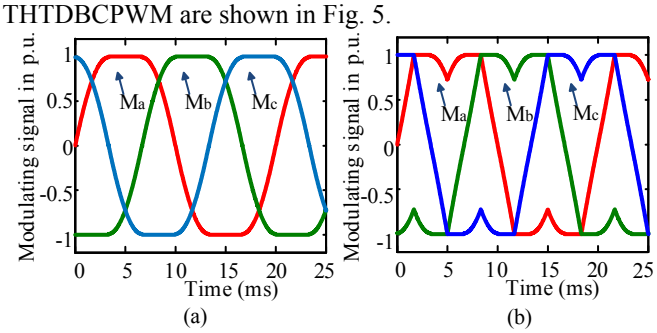


Fig. 5. Modulating signal corresponds to (a) THSDBCPWM (b) THTDBCPWM

From the modulating signals it is seen that for THSDBCPWM the signal is clamped for sixty degree per half cycle. In case of THTDBCPWM the modulating signal is clamped for thirty degree per quarter cycle. During this time no switching signals are made for IGBTs and the switching devices remains continuously on or off. Thus in this period there is no switching loss, only conduction loss occurs.

B. Formation of modulating signals

In case of third harmonic injected bus clamping PWM (THBCPWM) a common mode signal (CMS) is constructed from the third harmonic injected modulating signal.

For three phase inverters the three modulating signals are

$$M_{a1} = A_m \sin(wt) \quad (1)$$

$$M_{b1} = A_m \sin(wt - 120^0) \quad (2)$$

$$M_{c1} = A_m \sin(wt + 120^0) \quad (3)$$

The common mode signal for third harmonic injected pulse width modulation having tripled the fundamental frequency and k times the fundamental amplitude is added to (1)-(3) to form the modulating signals for THPWM.

$$M_{a2} = A_m \sin(wt) + k \sin(3wt) \quad (4)$$

$$M_{b2} = A_m \sin(wt - 120^0) + k \sin(3wt) \quad (5)$$

$$M_{c2} = A_m \sin(wt + 120^0) + k \sin(3wt) \quad (6)$$

For bus clamping of the third harmonic injected signal a common mode signal is needed.

Common mode signal V_{cm1}= Vc-max(Ma2,Mb2,Mc2)

Common mode signal V_{cm2}= -Vc-min(Ma2,Mb2,Mc2)

Where Vc is the peak of the carrier. By taking the combination of V_{cm1} & V_{cm2} different types of bus clamping PWM is possible.

For THSDBCPWM the common mode signal is formed by taking first 60 degree of V_{cm2} and next 60 degree of V_{cm1} and doing this in a periodic manner. But for THTDBCPWM the common mode signal is formed in a reverse manner of THSDBCPWM. That means at first V_{cm1} is taken for 60 degree then V_{cm2} is taken for the next 60 degree in a periodic manner.

This common mode signal can easily be formed by multiplying a periodic function of each common mode signal V_{cm1} and V_{cm2} and finally adding together.

Here, f1(α) and f2(α) are periodic function of α and defined as

f1(α)=0 when $0^0<\alpha<60^0$

=1 when $60^0<\alpha<120^0$

f2(α)=1 when $0^0<\alpha<60^0$

=0 when $60^0<\alpha<120^0$

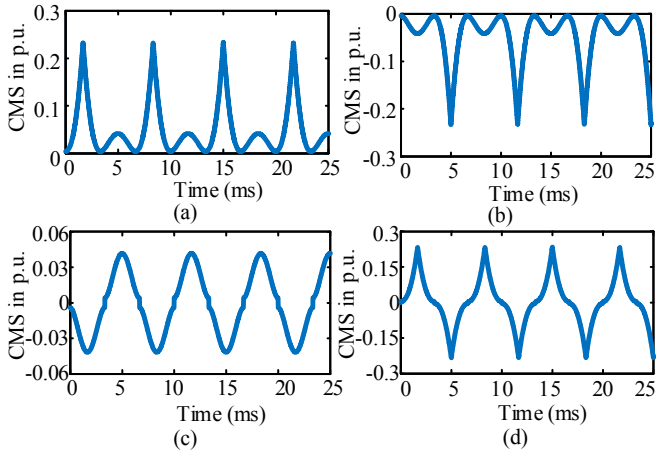


Figure 6. Common mode signal corresponding to (a) common mode signal 1 (b) common mode signal 2 (c) common mode signal for THSDBCPWM (d) Common mode signal for THSDBCPWM.

C. Output voltage waveform and frequency spectrum

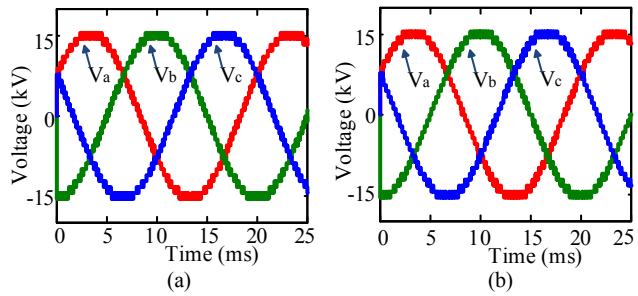


Fig. 7. Output line voltage corresponds to (a) THSDBCPWM (b) THTDBCPWM

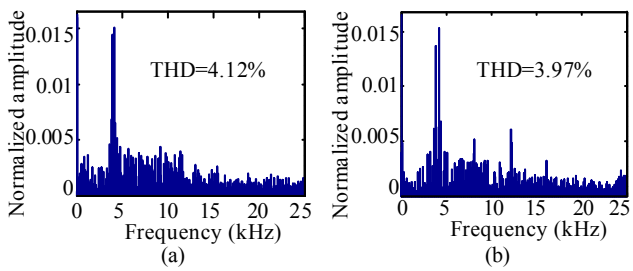


Fig. 8. Total Harmonic distortion corresponds to (a) THSDBCPWM (b) THTDBCPW

Fig. 9 shows the computer flow chart for the proposed modulating signal. The THD for proposed THSDBCPWM and THTDBCPWM are 4.12% and 3.97% as shown in Fig. 8. In Fig. 10 THD versus modulating index bar graph is shown for different modulating signals.

Inverter losses are shown in fig. 11 and fig. 12. Inverter loss includes switch turn on loss, switch turn off loss, diode turn off loss, switch conduction loss, diode conduction loss. Diode turn on loss is very negligible due to fast conducting during forward biased. The methods of calculating these losses are described in [11]. Using these methods and considering IGBT module 5SNA 1500E250300 from ABB (rated voltage & current: 2.5 KV & 1,500 A) different losses are calculated for the proposed modulation techniques. In fig.11 different losses for proposed techniques and in fig. 12 switching losses for different modulation techniques are shown with respect to different modulating index.

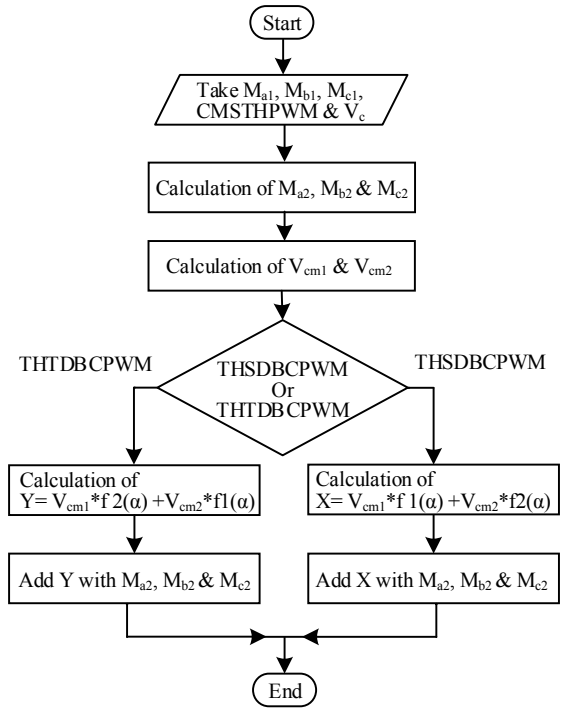


Fig. 9. Computer flow chart for obtaining proposed THBCPWM modulating signals.

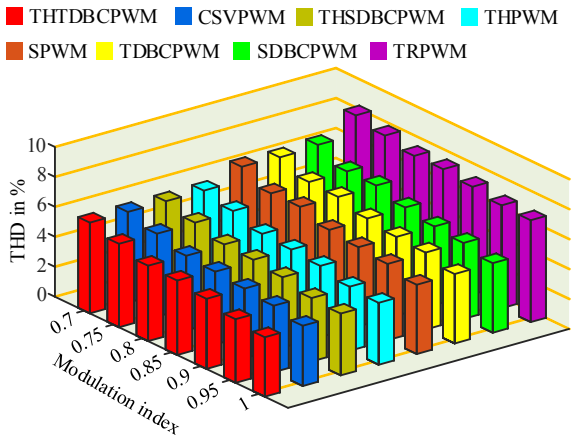


Fig 10. THD versus modulating index for different modulating signals.

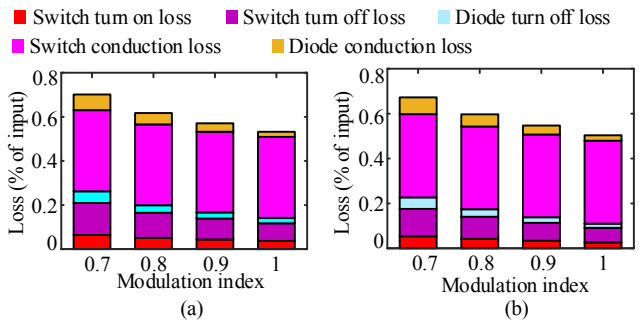


Fig. 11. Different losses corresponding to (a) THTDBCPWM (b) THSDBCPWM

From fig. 11 it can be seen that losses expressed in percentage of input reduces with the increase of modulating index. At modulating index=1, The THTDBCPWM and THSDBCPWM shows 0.53% and 0.50% total loss of input energy respectively. The THSDBCPWM causes the lowest switching loss 0.11% (when modulating index=1) among the analysed modulating signals as shown in fig. 12.

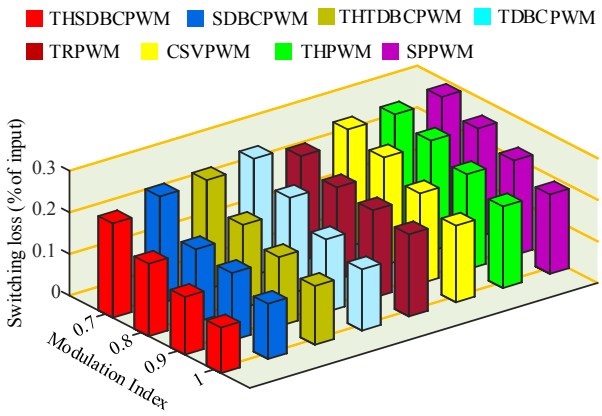


Fig. 12. Switching loss versus modulating index for different modulating signals.

V. Conclusion

The proposed THTDBCPWM & THSDBCPWM shows the better spectra property as well as low switching loss among all of the analysed modulating signals. Thus using these techniques in direct grid connected PV array inverter as in fig.1, a power frequency transformer-less, reduced THD & switching loss, cost effective PV system can be developed. This modulation techniques are also applicable for other power converter circuits.

References

[1] M. R. Islam, Y. G. Guo, and J. G. Zhu, "Multiple-Input Multiple-Output Medium Frequency-Link Based Medium Voltage Inverter for Direct Grid Connection of Photovoltaic Arrays," *International Conference on Electrical Machines and Systems*, Busan, Korea,Oct. 26-29, 2013, pp-202-207.

[2] M. R. Islam, Y. G. Guo, J. G. Zhu and M. G. Rabbani, "Simulation of PV Array Characteristics and Fabrication of Microcontroller Based MPPT," *Conference on Electrical and Computer Engineering ICECE 2010,* Dhaka, Bangladesh , 18-20 December 2010, pp-155-158.

[3] M. R. Islam, Y. G. Guo, and J. G. Zhu, "A High-Frequency Link Multilevel Cascaded Medium Voltage Converter for Direct Grid Integration of Renewable Energy Systems," *IEEE Trans. on Power electronics*, Vol. 29, No. 8, pp. 4167-4182, August 2014.

[4] Joby Jose, G. N. Goyal and M. V, Aware, "Improved inverter utilization using third harmonic injection," *Power Electronics, Drives and Energy Systems (PEDES) & Power India*, 2010, pp. 1-6.

[5] V. S. S. P. K. Hari and G. Narayanan, "Space-vector based hybrid pulse width modulation technique to reduce line current distortion in induction motor drives," *IET power electronics*, Vol. 5, No. 8, pp. 1463–1471, 2012.

[6] K. Basu, J. S. S. Prasad, and G. Narayanan, "Minimization of torque ripple in PWM AC drives," *IEEE Trans. on Industrial Electronics*, Vol. 56, No. 2, pp. 553–558, Feb. 2009.

[7] A. Ruiz-Gonzalez, M. J. Meco-Gutierrez, F. Perez-Hidalgo, F. Vargas- Merino, and J. R. Heredia-Larrubia, "Reducing acoustic noise radiated by inverter-fed induction motors controlled by a new PWM strategy," *IEEE Trans. on Industrial Electronics*, Vol. 57, No. 1, pp. 228–236, Jan. 2010.

[8] Z. Zhang, O. C. Thomsen and M. A. E. Anderson, "Discontinuous PWM modulation strategy with circuit-level decoupling concept of three-level neutral-point-clamped inverter," *IEEE Trans. on Industrial Electronics*, Vol. 60, No. 5, pp. 1897–1906, May 2013.

[9] Z. Zhang, O. C. Thomsen and M. A. E. Anderson, "Discontinuous PWM modulation strategy with circuit-level decoupling concept of three-level neutral-point-clamped inverter," *IEEE Trans. on Industrial Electronics*, Vol. 60, No. 5, pp. 1897–1906, May 2013.

[10] T. D. Nguyen, J. Hobraiche, N. Patin, G. Friedrich, and J. Vilain, "A direct digital technique implementation of general discontinuous pulse width modulation strategy," *IEEE Trans. on Industrial Electronics,* Vol. 58, No. 9, pp. 4445–4454, Sep 2011.

[11] Alamri, Basem, and Mohamed Darwish. "Precise Modelling of Switching and Conduction Losses in Cascaded H-bridge Multilevel Inverters", In Power Engineering Conference (UPEC), 2014 49th International Universities, pp. 1-6. IEEE, 2014.

Model Predictive Control for Low-Voltage Ride Through Capability Enhancement of DFIGs in Variable-Speed Wind Turbine Systems

Mohamed Abdelrahem, Muhammad Hosnee Mobarak, Ralph Kennel
Institute for Electrical Drive Systems and Power Electronics, Technical University of Munich (TUM)
E-mail: mohamed.abdelrahem@tum.de, hosnee.mobarak@tum.de, ralph.kennel@tum.de

***Abstract*—This paper proposes a finite control set-model predictive control (FCS-MPC) system for low-voltage ride through (LVRT) capability enhancement of doubly-fed induction generators (DFIGs) in variable-speed wind turbine systems (WTSs). The proposed FCS-MPC system takes the discrete states of the converter into account and the future converter performance is predicted for each sampling period. Then the optimal switching action that minimizes a predefined cost function is selected to be applied in the next sampling instant. The proposed LVRT strategy uses the DFIG-rotor inertia to store the surplus energy during the grid voltage dips. Moreover, the proposed LVRT strategy enhances the ability of the DFIG to inject active and reactive power to the grid during serious voltage dips. Simulation results are presented to validate the proposed LVRT strategy.**

***Index Terms*—Low-voltage ride through (LVRT), doubly-fed induction generator (DFIG), model predictive control.**

I. Introduction

The electrical power generation by wind turbines (WTs) has increased significantly during the last years [1]. Consequently, several countries have updated their grid codes (GCs) to involve specific requirements for wind power generation plants. One of the most challenging requirements is the low-voltage ride through (LVRT) capability, which requires that wind energy conversion systems (WECSs) remain connected to the grid during faults and inject active and reactive power to the grid [2]. Among various types of WECSs, doubly-fed induction generator (DFIG) is the most utilized one due to the advantages of lower converter rating (around 30% of the machine rating) and higher efficiency. To realize LVRT for DFIGs, several technical concerns appeared due to the grid faults must be properly addressed including the over-current in the stator and rotor circuits, and the over-voltage of the DC-link connecting the rotor side converter (RSC) and the grid side converter (GSC). To achieve the challenging demand of reactive power support during faults, the reactive power output from DFIGs needs to be maximized.

In order to enhance the LVRT capability of DFIG-based variable-speed wind energy conversion systems (WECSs), various approaches have been proposed in the literature. A crowbar circuit (CB) has been applied for the LVRT in [3]. However, once the crowbar is activated, DFIG becomes a grid-connected induction motor absorbing reactive power from the grid and loses its controllability. For better performance, some solutions install dynamic voltage compensator or dynamic resistor between the stator and grid to support stator voltage during grid faults [4]-[6]. However, they require additional hardware with its control, thus increase hardware cost and decrease the reliability. Other strategies design special control schemes to improve the utilization of the existing hardware. In [7], a LVRT strategy based on injecting demagnetizing currents from the RSC is proposed. Full DFIG control is kept, however a large rotor current capacity is needed. A robust controller in the $\alpha-\beta$ stationary frame is presented in [8], claiming full control in all LVRT cases. However, the results have been obtained with an oversized converter. In [9], the LVRT requirements has bee achieved by storing the surplus active power in the inertia of the DFIG-rotor. However, the traditional field oriented control (FOC) and voltage oriented control (VOC) have been employed to control the RSC and GSC, respectively. These control schemes suffers from poor transient response due to the limited bandwidth of the PI controllers.

In this paper, a finite control set-model predictive control (FCS-MPC) algorithm for LVRT capability enhancement of DFIGs in variable-speed WECSs is proposed. The proposed LVRT method uses the DFIG-rotor inertia to store the surplus energy during the grid voltage dips. Simulation results are presented to validate the proposed LVRT strategy.

II. Modeling of the WECS with DFIG

The block diagram of the WECS with DFIG is shown in Fig. 1. It consists of a wound rotor induction machine mechanically coupled to the wind turbine via a shaft and gear box with ratio $g_r \geq 1$. The stator windings of the DFIG are directly connected to the grid, whereas the rotor winding is connected via a back-to-back partial-scale voltage source converter (VSC), and a filter to the grid. The rotor side converter (RSC) and the grid side converter (GSC) share a common DC-link with capacitance C_{dc} [As/V] with DC-link voltage u_{dc}[V].

A. Doubly-Fed Induction Generator (DFIG)

The discrete model of the DFIG in stator/grid voltage orientation is given by (see [10]):

$$\left.\begin{aligned} i_r^d[k+1] &= i_r^d[k] + \tfrac{T_s}{\sigma L_s L_r}\big[-R_r L_s i_r^d[k] + (\omega_{sl}[k] L_r L_s - \\ &\quad \omega_s[k] L_m^2) i_r^q[k] + R_s L_m i_s^d[k] - \omega_r[k] L_m L_s i_s^q[k] \\ &\quad + L_s u_r^d[k] - L_m u_s^d[k]\big] \\ i_r^q[k+1] &= i_r^q[k] + \tfrac{T_s}{\sigma L_s L_r}\big[-R_r L_s i_r^q[k] - (\omega_{sl}[k] L_r L_s - \\ &\quad \omega_s[k] L_m^2) i_r^d[k] + R_s L_m i_s^q[k] + \omega_r[k] L_m L_s i_s^d[k] \\ &\quad + L_s u_r^q[k] - L_m u_s^q[k]\big] \end{aligned}\right\} \tag{1}$$

where k is the current sampling instant and T_s [s] is the sampling time. $\boldsymbol{u}_s^k = (u_s^d, u_s^q)^\top$ [V], $\boldsymbol{u}_r^k = (u_r^d, u_r^q)^\top$ [V],

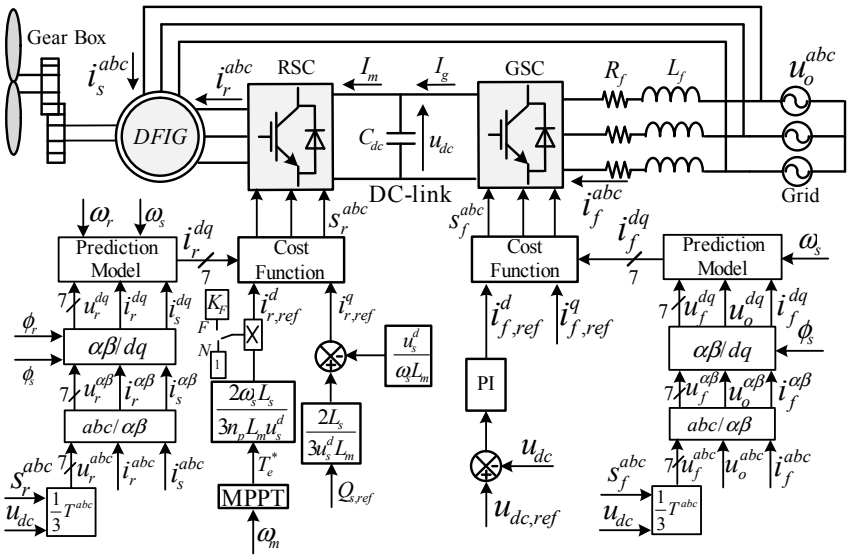


Figure 1: DFIG topology and control structure for variable-speed WECSs.

$\boldsymbol{i}_s^k = (i_s^d, i_s^q)^\top$ [A], and $\boldsymbol{i}_r^k = (i_r^d, i_r^q)^\top$ [A] are the stator and rotor voltages and currents in the rotating reference frame (k-coordinate system with axes d and q), respectively. ω_s, ω_r, and $\omega_{sl} := \omega_s - \omega_r$ [rad/s] are the stator, rotor, and slip electrical angular frequency ($\omega_r = n_p\omega_m$, where ω_m is the mechanical angular speed of the DFIG rotor and n_p is the number of pair poles). L_s, L_r, and L_m [Vs/A] are the stator, rotor and mutual inductances. R_s [Ω] and R_r [Ω] are stator and rotor winding resistances. $\sigma = 1 - \frac{L_m^2}{L_s L_r}$.
The rotor voltage of the DFIG (output voltage of the RSC) can be calculated as follows [10]:

$$\boldsymbol{u}_r^{abc} = \frac{1}{3} u_{dc} \boldsymbol{T}^{abc} \boldsymbol{s}_r^{abc} \tag{2}$$

where $\boldsymbol{s}_r^{abc} = (s_r^a, s_r^b, s_r^c)^\top \in \{0,1\}$ is the switching state vector of the RSC and $\boldsymbol{T}^{abc}$ is the transformation matrix

$$\boldsymbol{T}^{abc} = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \tag{3}$$

describing the relation between switching state vector and output phase voltage vector of the converter. Considering all the switching combinations for 2-level converter eight switching states $(\boldsymbol{s}_r^{abc})^\top := \{000, 100, ..., 111\}$, and consequently, eight voltage vectors are obtained $\boldsymbol{u}_{r,0}^{\alpha\beta}, \boldsymbol{u}_{r,1}^{\alpha\beta}, ..., \boldsymbol{u}_{r,7}^{\alpha\beta}$ (note: $\boldsymbol{u}_{r,0}^{\alpha\beta} = \boldsymbol{u}_{r,7}^{\alpha\beta}$). Equation (2) can be expressed in the stationary/rotating reference frame by using the Clarke $\boldsymbol{T}_c$ and Park $\boldsymbol{T}_p(\phi)^{-1}$ transformation, respectively, given by (neglecting the zero sequence)

$$\boldsymbol{u}_r^{dq} = \underbrace{\begin{bmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{bmatrix}}_{=:\boldsymbol{T}_p(\phi)^{-1}} \underbrace{\frac{2}{3}\begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix}}_{=:\boldsymbol{T}_c} \boldsymbol{u}_r^{abc} \tag{4}$$

B. Filter and grid

In Fig. 1, a grid-connected voltage source converter is shown. It is connected to the grid via a RL-filter with resistance R_f [Ω] and inductance L_f [Vs/A]. The discrete model of the grid side filter is given by (see [10]):

$$\left.\begin{aligned} i_f^d(k+1) &= (1 - \tfrac{T_s R_f}{L_f}) i_f^d(k) + \omega_e T_s i_f^q(k) \\ &\quad + \tfrac{T_s}{L_f}(u_o^d(k) - u_f^d(k)) \\ i_f^q(k+1) &= (1 - \tfrac{T_s R_f}{L_f}) i_f^q(k) - \omega_e T_s i_f^d(k) \\ &\quad + \tfrac{T_s}{L_f}(u_o^q(k) - u_f^q(k)) \end{aligned}\right\} \tag{5}$$

where i_f^d and i_f^q [A] are currents of the grid side filter; u_o^d and u_o^q [V] are voltages of the grid; u_f^d and u_f^q [V] are the output voltages of the GSC.
The output voltages of the GSC can be expressed as

$$\boldsymbol{u}_f^{abc} = \frac{1}{3} u_{dc} \boldsymbol{T}^{abc} \boldsymbol{s}_f^{abc} \tag{6}$$

where $\boldsymbol{s}_f^{abc} = (s_f^a, s_f^b, s_f^c)^\top \in \{0,1\}$ is the switching state vector of the GSC. Again, considering all the switching combinations for 2-level converter eight switching states $(\boldsymbol{s}_f^{abc})^\top := \{000, 100, ..., 111\}$, and consequently, eight voltage vectors are obtained $\boldsymbol{u}_{f,0}^{\alpha\beta}, \boldsymbol{u}_{f,1}^{\alpha\beta}, ..., \boldsymbol{u}_{f,7}^{\alpha\beta}$. Equation (6) can be expressed in the stationary/rotating reference frame by using the Clarke $\boldsymbol{T}_c$ and Park $\boldsymbol{T}_p(\phi)^{-1}$ transformation as explained before.

III. Maximum power point tracking (MPPT)

Wind turbines (WTs) convert wind energy into mechanical energy and, via a generator, into electrical energy. The mechanical (turbine) power of a WT is given by [11]:

$$p_t = c_p(\lambda, \beta) \underbrace{\frac{1}{2}\rho\pi r_t^2 v_w^3}_{\text{wind power}} \tag{7}$$

where $\rho > 0\,[\mathrm{kg/m^3}]$ is the air density, $r_t > 0\,[\mathrm{m}]$ is the radius of the wind turbine rotor (πr_t^2 is the turbine swept area), $c_p \geq 0\,[1]$ is the power coefficient, and $v_w \geq 0\,[\mathrm{m/s}]$ is the wind speed. The power coefficient c_p is a measure for the "efficiency" of the WT. It is a nonlinear function of the tip speed ratio

$$\lambda = \frac{\omega_m r_t}{g_r v_w} \geq 0 \qquad [1] \tag{8}$$

and the pitch angle $\beta \geq 0\,[°]$ of the rotor blades. The Betz limit $c_{p,\mathrm{Betz}} = 16/27 \approx 0.59$ is an upper (theoretical) limit of the power coefficient, i.e. $c_p(\lambda, \beta) \leq c_{p,\mathrm{Betz}}$. For typical WT,

the power coefficient ranges from 0.4 to 0.48 [11].
For wind speeds below the nominal wind speed of the WT, maximum power tracking is the desired control objective. Here, the pitch angle is held constant at $\beta = 0$ and the WT must operate at its optimal tip speed ratio $\lambda^\star$ (a given constant) where the power coefficient has its maximum $c_p^\star := c_p(\lambda^\star, 0) = \max_\lambda c_p(\lambda, 0)$. Only then, the WT can extract the maximally available turbine power $p_t^\star := c_p^\star \frac{1}{2}\rho\pi r_t^2 v_w^3$. Maximum power point tracking is achieved by the nonlinear speed controller

$$T_e^\star = k_p^\star \omega_m^2, \qquad k_p^\star := \frac{\rho\pi r_t^5}{2g_r}\frac{c_p^\star}{(\lambda^\star)^3} \tag{9}$$

which assures that the generator mechanical speed ω_m is adjusted to the actual wind speed v_w such that $\frac{\omega_m r_t}{g_r v_w} \stackrel{!}{=} \lambda^\star$ holds. According to (9) the optimum torque $T_e^\star$ can be calculated from the shaft speed ω_m.

IV. Finite control set-model predictive control

The proposed FCS-MPC algorithm takes the discrete states of the converter into account and the future converter performance is predicted for each sampling period. Then the optimal switching action that minimizes a predefined cost function is selected to be applied in the next sampling instant [12].

A. FCS-MPC for RSC

For the RSC, the chosen cost function is defined by

$$g_{RSC} = |i_{r,ref}^d(k+1) - i_r^d(k+1)| + |i_{r,ref}^q(k+1) - i_r^q(k+1)| \tag{10}$$

where $i_{r,ref}^d(k+1)$ and $i_{r,ref}^q(k+1)$ are the reference values of the $d-$ & q-axis currents. The value of the reference current $i_{r,ref}^d[k]$ is calculated from the optimum torque T_e^* and the value of $i_{r,ref}^q[k]$ is calculated from the reference stator reactive power $Q_{s,ref}[k]$ as follows [10]

$$\left.\begin{aligned} i_{r,ref}^d[k] &= \frac{2\omega_s[k]L_s}{3n_p L_m u_s^d[k]} T_e^*[k] \\ i_{r,ref}^q[k] &= \frac{2L_s}{3L_m u_s^d[k]} Q_{s,ref}[k] - \frac{u_s^d[k]}{\omega_s[k]L_m}. \end{aligned}\right\} \tag{11}$$

This optimal torque T_e^* is calculated from the maximum power point tracking (MPPT) algorithm, see section III. For the prediction algorithm, (1) is evaluated for each of the possible seven voltage vectors, giving seven different current predictions. The voltage vector whose current prediction is minimizing the cost function (10) is applied at the next sampling instant.

B. FCS-MPC for GSC

For the GSC, the chosen cost function is defined by

$$g_{GSC} = |i_{f,ref}^d(k+1) - i_f^d(k+1)| + |i_{f,ref}^q(k+1) - i_f^q(k+1)| \tag{12}$$

where $i_{f,ref}^d(k+1)$ and $i_{f,ref}^q(k+1)$ are the reference values of the d- & q-axis currents. The reference d-axis current $i_{f,ref}^d(k+1)$ is calculated from the outer DC-link voltage control loop and the reference q-axis current $i_{f,ref}^q(k+1)$ is set to zero, see Fig 1. Again, (5) is evaluated for each of the possible seven voltage vectors, giving seven different current predictions. The voltage vector whose current prediction is minimizing the cost function (12) is applied at the next sampling instant.

Table I: Parameters of DFIG-based WTS.

Name	Nomenclature	Value
Wind turbine radius	r_t	4.5 m
Rated wind speed	v_{wrated}	14 m/sec
DFIG rated power	p_{nom}	50 kW
DFIG stator voltage (line-line]	u_s^{rms}	400 V
Number of pair poles	n_p	2
Stator resistance	R_s	0.2448 Ω
Rotor resistance	R_r	0.4847 Ω
Stator inductance	L_s	80.76mH
Rotor inductance	L_r	82.12mH
Mutual inductance	L_m	77.2mH
DFIG moment of inertia	Θ	2.8 $\frac{\text{kg}}{\text{m}^2}$
DC capacitor	C_{dc}	3mF
DC-link voltage	u_{dc}	700V
Grid line-line voltage	u_o	400V
Grid normal frequency	f_e	50Hz
Filter resistance	R_f	0.16Ω
Filter inductance	L_f	12mH

V. Proposed LVRT Strategy

The proposed control scheme, Fig. 1, has the ability to work during the normal operation conditions or during the fault operation conditions. During the normal operation conditions the MPPT is achieved by the RSC control system via controlling the rotor d-axis current of the DFIG i_r^d. Therefore, the switch in Fig. 1 will connected to position N which denote normal operation conditions. The DC-link voltage regulation is achieved by the GSC control system by controlling the d-axis current of the output filter i_f^d.
During the fault operation condition, in order to store the surplus power in the DFIG-rotor inertia, the rotor d-axis reference current of the DFIG is multiplied by a factor $k_F = \frac{\|u_{s,f}\|}{\|u_s\|}$, where $\|u_{s,f}\|$ and $\|u_s\|$ are the magnitude of the stator/grid voltage during and before the fault, respectively. Therefore, the switch in Fig. 1 will connected to position F which denote fault operation conditions. Thus, the generated active power decreases and the difference between the WT output mechanical power and the generated power will be stored in the DFIG-rotor inertia.

VI. Simulation Results

A simulation model of a 50 [kW] WECS with DFIG is implemented in Matlab/Simulink. The system parameters are listed in Table I. The implementation is shown in Fig. 1.
Fig. 2 shows the performance of the WECS during a three phase fault with the proposed LVRT strategy under the rated wind speed of the wind turbine (i.e. $v_w = 14$ [$\frac{\text{m}}{\text{s}}$]). At $t = 0.2$ [s] a 70% voltage dip in the grid voltage for 200 [ms] has been occurred, see Fig. 2 (top). The proposed LVRT strategy forces the generator to reduce the active power extracted during the fault, see Fig. 2. As a consequence, there is a torque mismatch in the mechanical system between the wind turbine and the DFIG, which causes the mechanical speed of the DFIG to increase, see Fig. 2. However, the speed of the DFIG still lower than the maximum allowable speed (i.e. 1.3 [pu]). Therefore, no stress on the mechanical system is produced. After the fault clearness, the stored active power in the DFIG-rotor inertia is released back to the grid. Therefore, the mechanical speed of the DFIG decreases. The DC-link voltage is kept almost constant at its reference value. However, a very small oscillation in the DC-link voltage

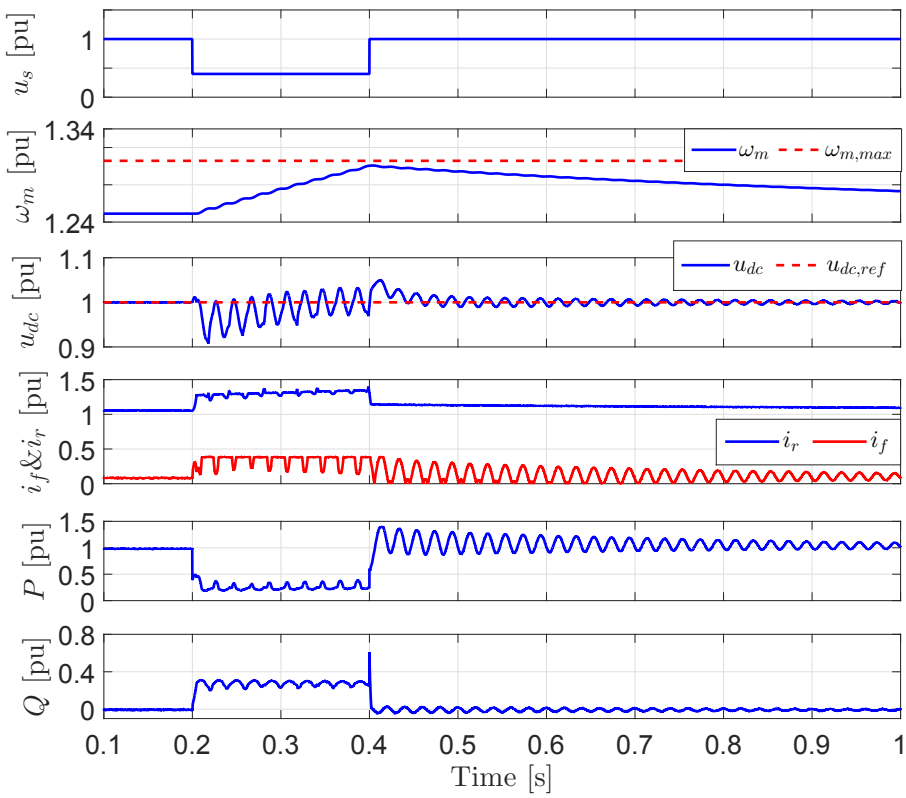


Figure 2: Performance of the WECS during a three-phase fault with the proposed LVRT strategy (from top): stator voltages $\boldsymbol{u}_s$, mechanical speed of the DFIG rotor ω_m, Dc-link voltage (u_{dc}, $u_{dc,ref}$), current of the rotor and output filter (i_r, i_f), active power injected to the grid P, reactive power injected to the grid Q.

appeared during the fault. The rotor current of the DFIG $\boldsymbol{i}_r$ and the output filter current $\boldsymbol{i}_f$ are lower than the maximum allowable value (i.e. 1.5 of the rated value). Moreover, using the proposed LVRT strategy, the DFIG has the ability to inject active and reactive power to the grid during the fault, see Fig. 2 (bottom). Thus, the proposed LVRT fulfill the grid codes requirements.

VII. Conclusion

A finite control set-model predictive control (FCS-MPC) approach has been presented to fulfill the LVRT requirements in a variable-speed WECS based on a doubly-fed induction generator (DFIG). The voltage dip in the grid causes a transient active power surplus in the system. The proposed LVRT strategy allows to store the energy surplus in the inertia of the DFIG-rotor. The performance of the proposed LVRT strategy is illustrated by simulation results. The simulation results have shown that the proposed control strategy can enable a safe LVRT of the DFIG based WECS as an operating generation unit of active and reactive power and fulfill the grid code requirements. Being a simple control strategy without additional hardware required, the proposed control is more superior to the existing methods including the traditional crowbar protection both in performance and in the improvement of grid reliability.

References

[1] M. Liserre, R. Cardenas, M. Molinas, and J. Rodriguez, “Overview of Multi-MW Wind Turbines and Wind Parks”, *IEEE Transactions on Industrial Electronics*, Vol. 58, No. 4, pp. 1081–1095, April 2011.

[2] M. Tsili, S. Papathanassiou, “A review of grid code technical requirements for wind farms”, *IET Renewable Power Generation*, Vol. 3, No. 3, pp. 308–332, 2009.

[3] J. Morren, S. Haan, “Ride through of wind turbines with doubly-fed induction generator during a voltage dip”, *IEEE Transaction on Energy Conversion*, vol. 20, no. 2, pp. 435-441, Jun. 2005.

[4] C. Wessels, F. Gebhardt, F. Fuchs, “Fault ride-through of a DFIG wind turbine using a dynamic voltage restorer during symmetrical and asymmetrical grid faults”, *IEEE Transaction Power Electrons*, vol. 26, no. 3, pp. 807-815, Mar. 2011.

[5] X. Yan, et. al., “Grid fault tolerant operation of a DFIG wind turbine generator using a passive resistance network”, *IEEE Transaction Power Electrons*, vol. 26, no. 10, pp. 2896-2905, Oct. 2011.

[6] P. Flannery, G. Venkataramanan, “A fault tolerant doubly fed induction generator wind turbine using a parallel grid side rectifier and series grid side converter”, *IEEE Transaction Power Electrons*, vol. 23, no. 3, pp. 1126-1135, May 2008.

[7] V. Mendes, C. de Sousa, S. Silva, J. Rabelo, W. Hofmann, “Modeling and ride-through control of doubly fed induction generators during symmetrical voltage sags”, *IEEE Transaction Energy Conversion*, vol. 26, no. 4, pp. 1161-1171, Dec. 2011.

[8] J. da Costa, H. Pinheiro, T. Degner, G. Arnold, “Robust controller for DFIGs of grid-connected wind turbines”, *IEEE Transaction on Industrial Electrons*, vol. 58, no. 9, pp. 4023-4038, Sep. 2011.

[9] D. Xie, et. al., “A Comprehensive LVRT Control Strategy for DFIG Wind Turbines With Enhanced Reactive Power Support”, *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 302-3310, Aug. 2013.

[10] M. Abdelrahem, C. Hackl, R. Kennel, “Encoderless Model Predictive Control of Doubly-Fed Induction Generators in Variable-Speed Wind Turbine Systems”, *to be published in Proceedings of The Science of Making Torque from Wind (TORQUE 2016) Conference*, Munich, 5–7 October 2016.

[11] C. Dirscherl, C. Hackl, and K. Schechner, “Modellierung und Regelung von modernen Windkraftanlagen: Eine Einführung”, *Chapter 24 in* Elektrische Antriebe – Regelung von Antriebssystemen, *D. Schröder*, 4th edition, Springer-Verlag Berlin Heidelberg, pp. 1540–1614, 2015.

[12] P. Cortes, M. Kazmierkowski, R. Kennel, D. Quevedo, J. Rodriguez, “Predictive Control in Power Electronics and Drives”, *IEEE Transactions on Industrial Electronics*, Vol. 55, No. 12 ,pp. 4312–4324, Dec. 2008.

Traffic Modelling of Low Dense Femtocellular Network

Jesmin Akhter[1,*] Md. Imdadul Islam,[1] and M. R. Amin[2]

[1]Department of Computer Science and Engineering, Jahangirnagar University, Saver, Dhaka 1342, Bangladesh

[2]Department of Electronics and Communications Engineering, Aftabnagar, Dhaka 1212, Bangladesh

[*]jesmin@juniv.edu

Abstract—**The prime objective of LTE (Long-term evolution) mobile cellular network is to ensure high speed data with good QoS. In LTE femtocellular network the tiny femtocell are integrated with macrocell to transfer overflow traffic (off loading) from macrocell to femtocell. A dense femtocellular network experiences frequent handover and huge interferences. Therefore in traffic modelling two parameters: the availability of traffic channel and threshold SNR (signal to noise ratio) at receiving end need to be considered together. In this paper a complete analytical model of traffic of low dense femtocellular network is developed using two dimensional Markov chain. Finally received SNR of UE (user equipment) is incorporated with the probability states of network traffic and the analytical results are verified with simulation.**

Index Terms—**Markov chain, probability of success, blocking probability and Engset traffic.**

I. Introduction

Femtocells are included inside a macro cell to off-load the traffic of macro cell and does not need central planning. Femto cells are actually low-cost and low-power cellular base station (radius of coverage within 10 to 20 meters) like the access point of a wireless LAN discussed in detail in [1-2]. The allocation of *time-frequency resource blocks* for users with new service (under macro cells and femto cells) and hanover are found in [3] for Macro-Femto overlay cellular networks. The paper measures performance of the network based on outage probability and throughput. The femtocell/macrocell integrated network resembles to underlay-overlay cellular network to improve capacity of the network.

In LTE a lot of femtocells exists in a macrocell hence possibility of huge number of handoff between femto to macro, macro to femto and femto to femto cells. In [4] authors provides as state transition chain of three nodes regarding mobility of users. The simulation results (profiler of number of transition) are shown for four 'transition probability matrix' and three base stations. A Survey of similar job is found in [5]. The extension of the analysis is found in [6]; where authors determine the 'connectivity probability' based on length of wireless link and the mobility of mobile users in femtocell/macrocell networks of LTE. The analytical model of probability of connectivity and outage probability is derived in details. The profile of probability of connectivity against communication range is shown taking number of femtocells, density of users and density of femtocells as parameters.

The rest of the paper is organized as: section II provides literature review of the work, section III deals with statistical analysis of traffic of femto cellular network with algorithm of traffic simulation, section IV provides both analytical and simulation work and finally section V concludes the entire analysis.

II. Related works

In [4] the arrival rate of femtocell consists of four parts: λ_{of} (arrival of newly originated call in femto cells), $\lambda_{h,mf}$ (handover arrival rate from macro to femto cells), $\alpha\lambda_{h,ff}$ (femto to femto handover provided that received SINR is greater than a threshold i.e. the condition of channel is satisfactory); where α is the probability that the received SINR is greater than threshold τ and $P_{D,m}\beta\lambda_{h,ff}$ (femto to femto handover provided that received SINR lies between two thresholds τ_1 and τ_2 i.e. the wireless link can be maintained marginally with macro eNB but all the channels of macro cell is occupied. Under this situation the call will be continued in femto cell); where β is call dropping probability of macro cell and is the probability that that received SINR lies between two thresholds τ_1 and τ_2. Here a network of N femto cells is considered for statistical analysis. One dimensional Markov chain is drawn taking call arrival rate of a femto cell: $(\lambda_{of}+\lambda_{h,mf}+\alpha\lambda_{h,ff}+P_{D,m}\beta\lambda_{h,ff})/N$ and call termination rate μ_f.

In [5] two dimensional Markov chain is drown taking aggregate arrival and termination rate of macro cell along x-direction and those of femtocell along y-direction. The closed form expression of probability states is determined from the Markov chain and the blocking probability of both macro and femto cell users are plotted agaist total arrival rate. Both the above paper considers high density femtocell model hence femto to femto handover is considered in aggregate traffic arrival rate.

III. Traffic model

The one dimensional Markov chain of [4] is shown in fig.1; where the total call arrival rate of a femto cell, $\lambda_T = (\lambda_{of}+\lambda_{h,mf}+\alpha\lambda_{h,ff}+P_{D,m}\beta\lambda_{h,ff})/N$ and call termination rate is μ_f.

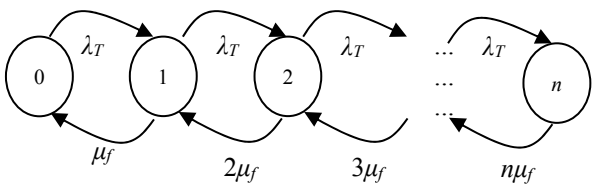


Fig.1 One dimensional Markov chain of [4]

Solving the chain the call blocking probability,

$$B_f = \frac{\left\{\frac{\lambda_T}{\mu_f}\right\}^n / n!}{\sum_{i=0}^{m}\left\{\frac{\lambda_T}{\mu_f}\right\}^i / i!}. \quad (1)$$

In this paper we first modify above chain into two dimensional Markov chain shown in fig.2; where the newly originated and handoff traffic are placed along x and y direction exclusively for a femto cell instead of macro-femto cell traffic of [5].

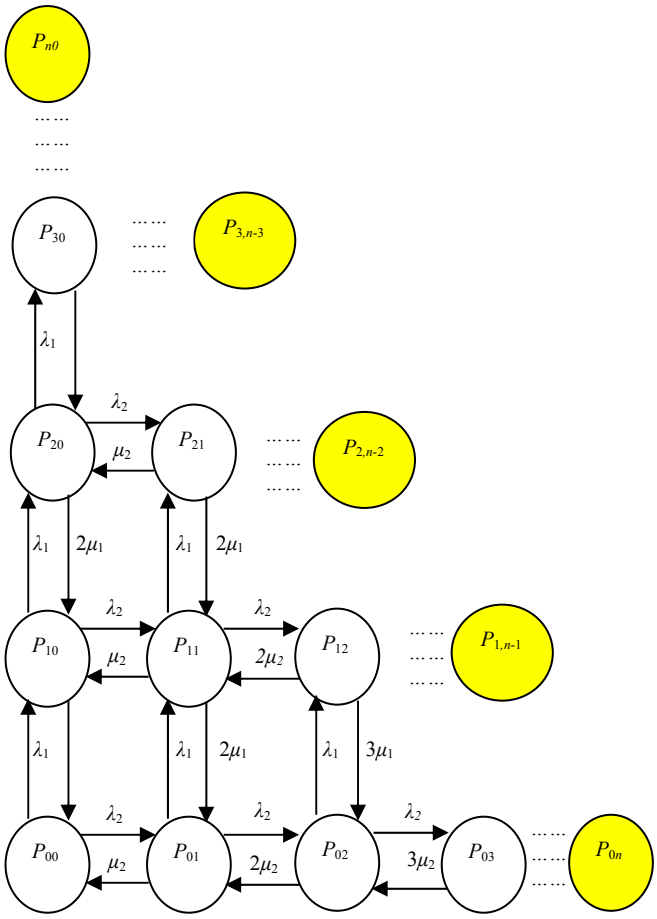


Fig.2 2D Markov chain of PCT-I traffic

Under Erlang's model (PCT-I traffic), the blocking probability is the sum of complete occupied state is derived as [6]:

$$B_f(A_1, A_2, n) = \frac{\sum_{y=0}^{n} \frac{A_1^y}{y!}\frac{A_2^{n-y}}{(n-y)!}}{\sum_{x_1=0}^{n}\sum_{x_2=0}^{n-x_1}\frac{A_1^{x_1}}{x_1!}\frac{A_2^{x_2}}{x_2!}}, \quad (2)$$

where, the offered traffic are, $A_1 = \lambda_{of} / \mu_f N$ and $A_2 = (\lambda_{h,mf} + \alpha\lambda_{h,ff} + P_{Dm}\beta\lambda_{h,ff}) / \mu_f N$ respectively.

Under Engset model (PCT-II traffic), the blocking probability is [7]:

$$B_n(A_1, A_2, n) = \frac{\sum_{y=0}^{n}\binom{M}{y}a_1^y\binom{M}{n-y}a_2^{n-y}}{\sum_{x_1=0}^{n}\sum_{x_2=0}^{n-x_1}\binom{M}{x_1}a_1^{x_1}\binom{M}{x_2}a_2^{x_2}}, \quad (3)$$

where, M is the number users/femtocell, the newly originated offered traffic per user is $a_1 = \lambda_{of} / \mu_f NM$ and that of handover is, $a_2 = (\lambda_{h,mf} + \alpha\lambda_{h,ff} + P_{Dm}\beta\lambda_{h,ff}) / \mu_f NM$.

Here two parameters α and β are related to received SNR under awgn. Let the test statistics at receiving end is the mean of R samples. Now, $T(n) = \frac{1}{R}\sum_{n=0}^{R-1} x(n)$; where any sample $x(n)$ has the pdf of $N(\mu', \sigma^2)$ then the test statistics $T(n)$ has the pdf of $N\left(\mu', \frac{\sigma^2}{R}\right) = N(\mu', \sigma_R^2)$

So the probability of getting the SNR above or equal τ is,

$$\begin{aligned} P_r\left\{\frac{T(n)}{\sigma_R} \geq \tau\right\} &= P_r\{T(n) \geq \tau\sigma_R\} \\ &= P_r\{T(n) \geq \tau'\} \\ &= Q\left(\frac{\tau - \mu'}{\sqrt{\sigma^2/R}}\right) = \alpha, \end{aligned} \quad (4)$$

where $Q(x) = \frac{1}{2} erfc\left(\frac{x}{\sqrt{2}}\right)$.

Again $P_r\{\tau_2 < T(n) < \tau_1\}$ can be derived as,

$$\beta = Q\left(\frac{\tau_2 - \mu'}{\sqrt{\sigma^2/N}}\right)\left\{1 - Q\left(\frac{\tau_1 - \mu'}{\sqrt{\sigma^2/N}}\right)\right\}. \quad (5)$$

For low dense femto cellular network the femto-femto cellular handoff is eliminated hence the Markov chain will take the form like Fig. 3.

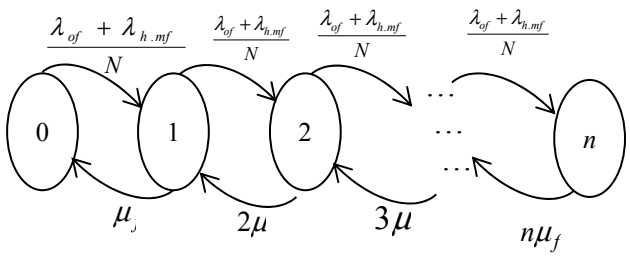


Fig. 3 Markov chain of a target femto cell of low dense network

Solving the chain, the call blocking probability is obtained as

$$B_f(n) = \frac{\left\{\frac{\lambda_{of} + \lambda_{h.mf}}{N\mu_f}\right\}^n / n!}{\sum_{i=1}^{m}\left\{\frac{\lambda_{of} + \lambda_{h.mf}}{N\mu_f}\right\}^i / i!}. \quad (6)$$

A. *Traffic Flow algorithm:*

Let us first introduce with the variables of simulation.

A = time of next call arrival
T_x = time when trunk number x is due for release (x = 1 or 2; $T_x = \infty$ if trunk x is already free)
L_x = 0 or 1 according to whether trunk x is free or busy
Q = number of waiting calls
Q_{max} = Maximum possible queue length
B = Number of blocked calls
C = number of calls completed since beginning of simulation
T = time of latest trunk seizure
M = total number of calls to be simulated
min(x, y, z) provide minimum value among *x, y* and *z*
I = Interarrival time
H = Service time

Step 1: At the beginning of simulation, trunk is free, set $L_1 = L_2 = C = A = Q = 0$, $T_1 = T_2 = \infty$. [Initialization]
Step 2: Take the Sample of interarrival time I
Step 3: Calculate the time of next call arrival $A = A + I$
Step 4: Take min (A, T_1, T_2).
the minimum value among time of next call arrival A and time T_x (when trunk number x is due for release. x =1 or 2; $T_x = \infty$ if trunk x is already free).
Step 5: If ($T_x < A$) then
Check whether number of waiting calls Q are present or not?

a) If ($Q > 0$) then
- Calculate $Q = Q$-1, $T = T_x$.
- Take the Sample of service time H and compute $T_x = T+H$

Else If ($Q = 0$) then
set $L_x = 0$, $T_x = \infty$

b) Set $C = C$+1.
c) Go to step 4.

Step 6: If ($A < T_x$) then
If (the number of calls completed since beginning of simulation(C) $\geq$ Total number of calls to be simulated (M)) then Exit.
Else if ($C < M$)
Check whether trunk x is free or busy that is whether $L_x = 0$ or 1?

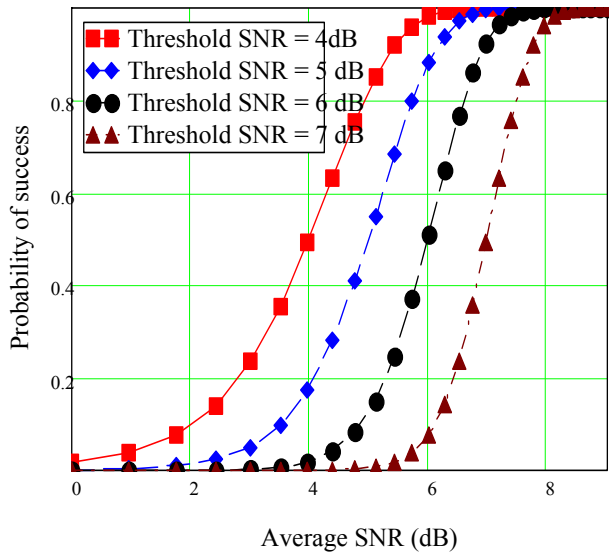
a) If ($L_x = 1$) then
- $Q = Q$+1.
- If ($Q > Q_{max}$) then $B = B$+1,$Q = Q$-1.
- Go to step 2.

b) Else If ($L_x = 0$) then
- $L_x = 1$, $T = A$
- Take the Sample of service time H
- Compute $T_x = T+H$
- Go to step 2.

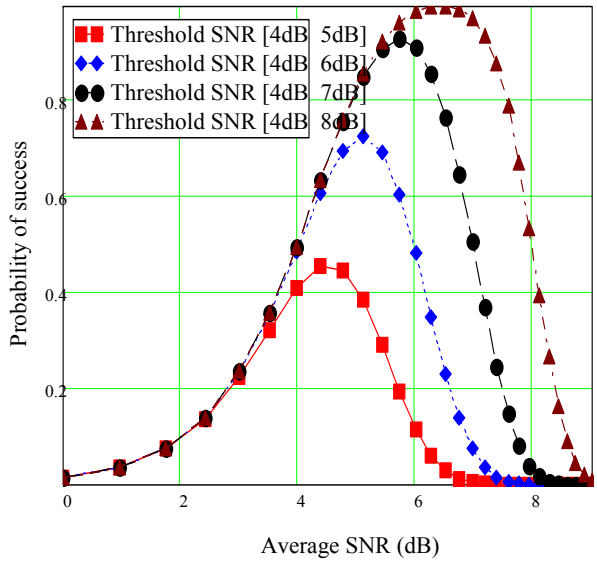
Step 7: Exit.

IV. RESULTS

In this paper we use two parameters α (the probability that the received SNR is greater than or equal to threshold a τ) and β (the probability that that received SNR lies between two thresholds τ_1 and τ_2) related to the fading and noise level of wireless channel. Since the length of wireless link in LTE much smaller than that of microcell of GSM hence the test statistics (average of received symbol) can be approximated by normal pdf (or lognormal pdf). The profile of α and β are shown in fig.1 (a) and (b) respectively against average SNR taking threshold SNR as a parameter. Both the SNR parameters are used with traffic parameter like thinning scheme of [8].

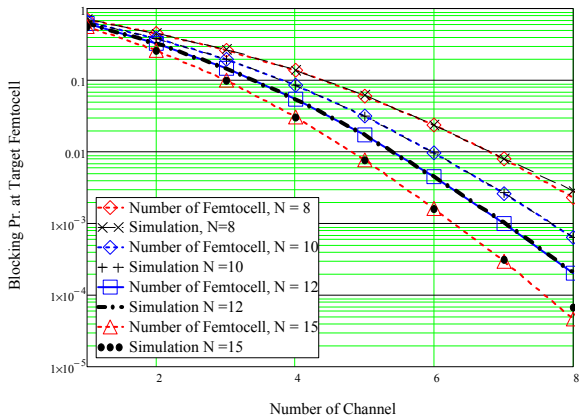


(a) Variation of parameter α against average SNR of wireless link

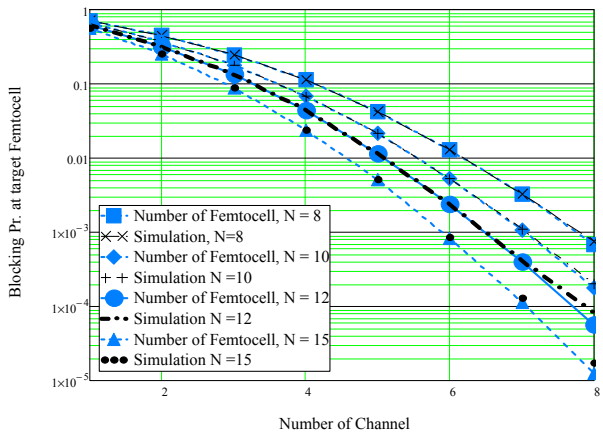


(b) Variation of parameter β against average SNR of wireless link
Fig.1 The profile of α and β

Taking the numerical values of traffic parameters, λ_{of} = 100 calls/min, $\alpha = 0.8$(threshold SNR 4dB and average SNR 5dB), $\lambda_{h,ff}$ = 58 calls/min, $\lambda_{h,mf}$ = 80 calls/min, $P_{D,m}$= 0.05, β=0.53(threshold SNR 4dB to 6 dB and average SNR 5.8dB)andμ_f=12 we plot the blocking probability at a femto cell against the number of traffic channels shown in fig.1(a) for the case of high dense femtocellular network of previous work. According to the algorithm of Monte Carlo simulation of previous chapter we also evaluate the simulation results

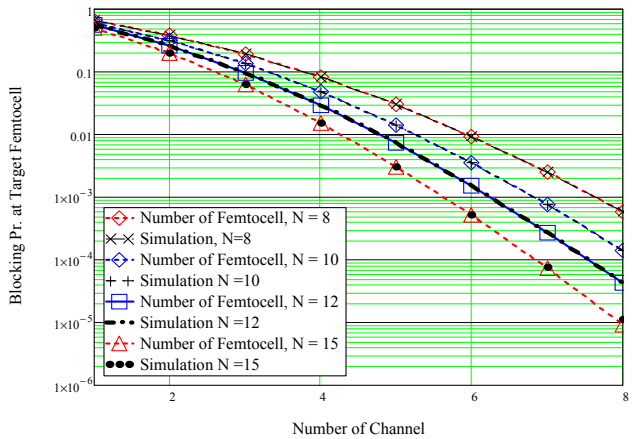


(a)

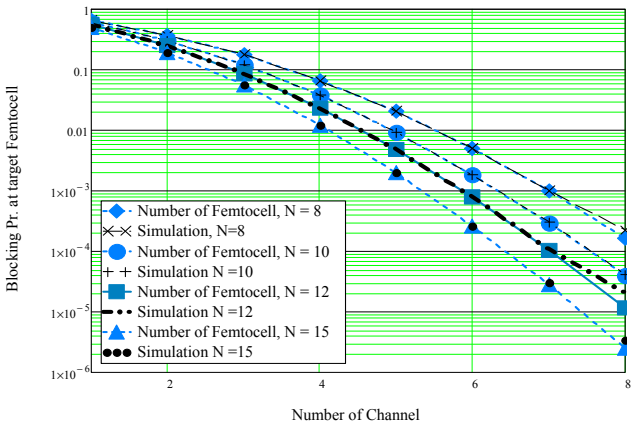


(b)

Fig.1 The 2D Erlang's and Engset traffic model (high dense)



(a)



(b)

Fig.2 The 2D Erlang's and Engset traffic model (low dense)

along with analytical values. The comparison of simulation and analytical results are shown in fig.1 (a) taking the number of femtocell N as a parameter. Similar results are shown in fig. 1(b) for Engset traffic taking the number users per femto cell as $M = 12$. The simulation results reveal 98% confidence level.

Taking the same traffic parameters we plot the blocking probability of low dense femto cellular network against the number of traffic channel for both PCT I and PCT II traffic shown in fig.1 (a) and (b) respectively. Here the blocking probability is found much lower than the case of high dense femtocells.

V. Conclusions

In this paper an analytical model of femtocellular traffic is developed changing the one dimensional state transition chain of previous literature to 2D chain. The paper incorporates the complementary of outage probability with traffic parameters to get more realistic scenario of blocking probability. The paper also consider awgn environment to derive complementary of outage probability for femtocell but still we have the scope to consider Rician fading model for such tiny cells when a cell is replete with objects of different size. The paper will be extended including call admission scheme to derive probability of forced termination under fading environment.

References

[1] Dakuri Chiranjeevi, B Rajakumar, M Devender, B Kiran, 'Performance Evaluation of LTE Femtocell in an Indoor Cellular Network,' International Journal of Electrical, Electronics and Data Communication, ISSN: 2320-2084 Volume-3, Issue-2, pp.20-23, Feb 2015

[2] Halidu Ibrahim Umar, Anas Adam Sulaiman, Musbahu Haruna Jibrin, 'Handoffs in Femtocell and Efficient Mobility Management in LTE Femtocell Network,' International Journal of Engineering Science and Computing, vol.6, no.4, pp. 4002-4006, April 2016

[3] Jongchan Lee and Moonho Lee, 'Resource Allocation with QoS Supporting in Macro-Femtocell Networks,' International Journal of Multimedia and Ubiquitous Engineering Vol. 10, No. 1, pp. 365-378 2015

[4] Mostafa Zaman Chowdhury and Yeong Min Jang, 'Handover management in high-dense femtocellular networks,'EURASIP Journal on Wireless Communications and Networking, pp.1-21, 2013.

[5] X. Ge, T. Han, Y. Zhang, G. Mao, C. X. Wang, J. Zhang, B. Yang, S. Pan, 'Spectrum and Energy Efficiency Evaluation of Two-Tier Femtocell Networks With Partially Open Channels,' IEEE Transactions on Vehicular Technology ,Volume:63, Issue: 3 , pp.1306 – 1319, November 2013.

[6] Md. Imdadul Islam, Md. Fazlay Rabbi, Risala Tasin Khan, and Jesmin Akhter, 'Performance Evaluation of Cognitive Radio Network Based on 2-D Markov Chain,' Journal of Telecommunications and Information Technology (JTIT), pp.39-44, vol.3, 2015.

[7] Md. Nurul Islam, Shamim, Ahmed, Ummy Habiba, Md.Imdadul Islam and M. R. Amin, 'Determination of Energy Efficiency of a Multi-user Wireless Network Based on Limited User Traffic Model' 8th International Conference on Electrical and Computer Engineering, pp.406 – 408, 2014, 20-22 Dec' 2014, Dhaka, Bangladesh.

[8] (Yuguang Fang, 'Thinning Scheme for Call Admission Control in Wireless Networks,' *IEEE Transactions on Computers*, vol.52, no.5, pp. 685-687, May 2003.

Traffic Modelling of WiMAX under BMAP with Length of Packet of General Distribution

Jesmin Akhter[1,*] Md.Imdadul Islam[2], and M. R. Amin[3]

[1,2]Department of Computer Science and Engineering, Jahangirnagar University, Dhaka, Bangladesh
[3]Department of Electronics and Communications Engineering, East West University, Aftabnagar, Dhaka 1212, Bangladesh
*jesmin@juniv.edu

Abstract—**The relation between offered traffic and the length of frame taking the batch size as a parameter are nonlinearly related under Batch Markovian Arrival Process (BMAP). Therefore the prediction of carried traffic and blocking probability against the length of frames or the batched size becomes more complicated compared to conventional packet or circuit switched traffic. The previous literature deals with M/1/G/K traffic for a network in generalized form but in this paper we bring the traffic model for WiMAX network. We compare the performance of subscriber station (SS) WiMAX network for M/M/1/K and M/1/G/K traffic cases considering BMAP. Finally we show the interrelationship among: offered traffic, the length of frame, carried traffic(related to throughput), the length of queue, the size of batch and blocking probability (related to packet loss) both analytically and graphically; where the previous literature show the results partially.**

Index Terms—**Subscriber station, MAP, offered traffic, frame length and blocking probability.**

I. Introduction

In store and forward switching the packets of variable length come at the router, which are stored for some instant because of FIFO (first in first out) queue of packets in the buffer. The traffic of a subnet depends on probability density function (pdf) of packet arrival rate and length of packet (service time). In circuit switched network for example in telephony, both the arrival of calls and theirs service time (or holding time) follow exponential pdf (or Passion's pdf) and expressed as M/M/n/K traffic known as Kendall's notation found in [1-2]. In packet switched network the pdf of service time is very complicated since a router may experiences both packets of length of exponential pdf (for example TCP/IP packet of Internet) and packet of fixed length (for example ATM cell has deterministic service time). The service time of packets in a router is modelled as the combination of exponential and deterministic pdf, which is known as general pdf. According to Kendall's notation such traffic is expressed as: M/G/n/K and its analysis is available in [3-5].

Markov Arrival Process (MAP) is a mathematical modelling technique of time between job arrivals (inter arrival) in a system. Similar work can be done by Markov chain also known as state transition chain which is solved by a set of linear equations called node or cut equations. One of the major drawback Markov chain lies in incorporation of large number of probability states which makes the model complicate. The MAP provides an equivalent state transition chain of few probability states with some assumption. The teletraffic engineering adopts three most cases of MAP are: PH Markov renewal process (PH-MRP), Markov Modulated Markov Process (MMPP) and Batch Markovian Arrival Process (BMAP). Mathematical modelling of MAP and performance of a network modelled by MAP is found in [6-10]. The objective of the paper is to analyze the performance of 4G network based on BMAP. The concept of BMAP and statistical analysis of BMAP is shown in next section. The BMAP is essential in 4G mobile cellular network since a user have the provision of communication of variable bit rate where he can negotiate with BS for required BW. According to the requirement of the mobile station (MS) the base station (BS) allocates the MS as much subcarriers of OFDMA as required. From the concept of statistical distribution of offered traffic and availability of traffic channels with respect to time, a group of users experience better service than an individual user. If several traffic channels are allocated to the group then the optimum utilization of channels is possible in a moderately loaded network. Such arrangement is called subscriber station (SS) prevalent in WiMAX. The traffic modelling of such SS based on combinations of 'Transition Matrix of Queue' and 'Transition Matrix of Connection' is expressed explicitly in [11-12]. In this paper we use BMAP to evaluate the performance of a SS of WiMAX.

The paper is organized as: section II provides statistical model of MAP specially Markov Modulated Poisson Process, Batch Markovian Arrival Process, concept of general service time, traffic of SS; section III provides the results based on theoretical analysis of section II and finally section IV concludes the entire analysis.

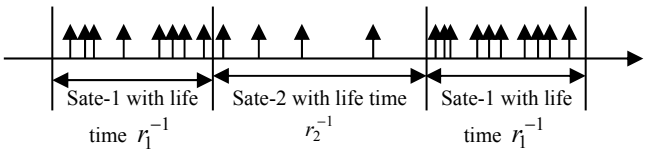
II. SYSTEM MODEL

A. Markov Modulated Poisson Process (MMPP)

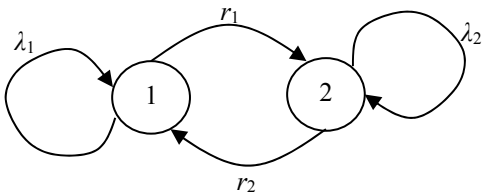
The Markov modulated Poisson process (MMPP) is a double stochastic Poisson process where arrival rate of any traffic depends on its probability state. The state transition chain of MMPP is characterized by two matrices: (**Q, D**) where **Q** is called generator and **D** is arrival matrix [13-14]. The elements of generator matrix are the inverse of sojourn time of states (or the transition rate among the states) where the elements of diagonal matrix are the arrival rate of states. For a 2 state MMPP the distribution of packet arrival rate on a state with sojourn time (inverse of transition rate) is shown in fig.1 (a) and the corresponding state transition chain is shown in fig. 1(b). The generator and diagonal matrix of two states MMPP is expressed as,

$$\mathbf{C}=\begin{bmatrix}-(r_1+\lambda_1) & r_1\\ r_2 & -(r_2+\lambda_2)\end{bmatrix},\ \mathbf{D}=\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix} \text{ and}$$

$$\mathbf{Q}=\mathbf{C}+\mathbf{D}=\begin{bmatrix}-r_1 & r_1\\ r_2 & -r_2\end{bmatrix}. \quad (1)$$



(a) Packet arrival and sojourn time



(b) State transition chain

Fig.1 Two states MMPP

B. Batch Markovian Arrival Process BMAP

In Batch Markovian Arrival Process (*BMAP*) several packets of a message arrives simultaneously similarly several packets are served in a batch service model. The size of the simultaneous packets or calls is called the size of the batch which is always greater than one [15-16].

If a BMAP of M+1 different possible batch size and it is in a transient state i then after elapsing the sojourn time of the state i, it will make a state transition to j with probability $P(m)_{i,j}$ where $1 \le m \le M$. The probability $P_{i,j}(m)$ satisfies,

$$\sum_{\substack{j\in S\\ j\neq i}} P_{i,j}(0) + \sum_{j\in S}\sum_{m\neq 1}^{\infty} P_{i,j}(m) = 1 \tag{2}$$

for all $j \in S$; $P_{i,j}(0)$ is called absorbing state of the BMAP.

In a BMAP, **D(0)** is the rate matrix of transition without arrival and **D(m)** is the rate matrices of transition with arrival of batch size m. The elements of the matrices: $D(0)_{i,j}=\lambda_i.P(0)_{i,j}$ for $i\neq j$, $D(0)_{i,i}=-\lambda_i$ and $D(m)_{i,j}=\lambda_i.P(m)_{i,j}$. The matrix **D**(0) has negative diagonal elements and positive off-diagonal element and matrices **D**(m) have non-negative elements. The summation of **D**(m) provides the infinitesimal generator matrix,

$$\mathbf{D} = \sum_{m=0}^{\infty} \mathbf{D}(m) \cdot \tag{3}$$

Let $\boldsymbol{\pi}$ be the stationary probability of the Markov process then $\boldsymbol{\pi}\mathbf{D} = 0$ and $\pi\mathbf{e} = 1$; where **e** is a column vector of 1's and π_j is the stationary probability that the arrival process in state j.

C. *Traffic with general service time*

In a circuit switched network both the arrival and service time of traffic follow exponential pdf is designated as: $M/M/n/n+k$ where M for Markovian arrival and service time, n is the number of server or trunk and k is the length of queue. In case of packets of deterministic service time (for example ATM cell of 53 bytes) the traffic is modeled as: $M/D/n/n+k$ where D for deterministic service time. In packet switching network the arrival of packet follows exponential pdf but the pdf of service time is actually the combination of exponential and deterministic called general pdf is represented as M/G/c/c+K where c is the number of server or channel and K is the size of buffer [17].

For single server case i.e. in M/G/1/K system the call blocking probability is found in closed form as [18-19]:

$$P_K(\rho) = \frac{\rho^{\left(\frac{\sqrt{\rho.e^{-s^2}}.s^2-\sqrt{\rho.e^{-s^2}}+2K}{2+\sqrt{\rho.e^{-s^2}}.s^2-\sqrt{\rho.e^{-s^2}}}\right)}.(\rho-1)}{\rho^{\left(\frac{\sqrt{\rho.e^{-s^2}}.s^2-\sqrt{\rho.e^{-s^2}}+K+1}{2+\sqrt{\rho.e^{-s^2}}.s^2-\sqrt{\rho.e^{-s^2}}}\right)}-1} \tag{4}$$

The throughput of the system, $Th(\rho) = \{1-P_0(\rho)\}\mu$ (5)

;Where $P_0(\rho)$ the is the probability of no arrival is found as,

$$P_0(\rho) = \frac{(\rho-1)}{\rho^{\left(\frac{2\sqrt{\rho.e^{-s^2}}.s^2-\sqrt{\rho.e^{-s^2}}+K+1}{2+\sqrt{\rho.e^{-s^2}}.s^2-\sqrt{\rho.e^{-s^2}}}\right)}-1} \tag{6}$$

Here s^2 is the square of co-efficient of variation of the service process and its typical value is ½.

D. Traffic of SS of WiMAX

The transition rate and arrival rate matrix of ith user is $\mathbf{Q}(i)$ and $\mathbf{D}(i)$ respectively. The transition probability matrix of ith user is obtained as, $\mathbf{U}(i) = (\mathbf{\Lambda}(i) - \mathbf{Q}(i))^{-1}\mathbf{\Lambda}(i)$. For batch arrival case the relation will be, $\mathbf{U}_b(i) = (\mathbf{\Lambda}_b(i) - \mathbf{Q}_b(i))^{-1}\mathbf{\Lambda}_b(i)$; where b is the size of batch and $b \in \{0, 1, \quad 2, .., L\}$. Each diagonal element of $\mathbf{\Lambda}(i)$ will be, $\frac{(\lambda(i)T_f)^k e^{-\lambda(i)T_f}}{k!}$; where $k = b$ i.e. the size of batch, T_f is the length of frame. For example, the diagonal probability matrix of two state MMPP case [20-21],

$$\mathbf{\Lambda}(k) = \begin{bmatrix} \frac{(\lambda_1 T_f)^k e^{-\lambda_1 T_f}}{k!} & 0 \\ 0 & \frac{(\lambda_2 T_f)^k e^{-\lambda_2 T_f}}{k!} \end{bmatrix} \tag{7}$$

;where T_f is the duration of a frame in seconds.

For a subscriber station of m users, probability matrix of the SS,

$$\mathbf{U} = \mathbf{U}(1) \otimes \mathbf{U}(2) \otimes \ldots \ldots \ldots \otimes \mathbf{U}(m) \tag{8}$$

The diagonal matrix under batch size of b [22],

$$\Lambda_a = \sum_{i+j+\ldots+k=a} \Lambda_i(1) \otimes \Lambda_j(2) \otimes \ldots \ldots \ldots \otimes \Lambda_k(m) \tag{9}$$

Where $i, j, \ldots \ldots \ldots, k \in \{0, 1, \quad 2, .. \ldots \ldots, L\}$

The steady probability sates,

$\mathbf{\Pi}.\mathbf{U} = \mathbf{\Pi}$ and $\mathbf{\Pi}.\mathbf{e} = 1$; where **e** is a vector of all 1s. (10)

The average offered traffic,

$$\rho(T_f) = \mathbf{\Pi}.\left[\sum_{a=0}^{N} a\mathbf{D}(a, T_f)\right].\mathbf{e}.T_f \tag{11}$$

;where N is the size of batch

E. Numerical example

Example-1: Let us consider two state MMPP traffic with parameters: $\lambda_1 = 0.656$, $\lambda_2 = 0.942$, $r_1 = 0.21$, $r_2 = 0.12$. Determine steady probability states.

$$\mathbf{Q}_1 = \begin{bmatrix} -0.21 & 0.21 \\ 0.12 & -0.12 \end{bmatrix}, \mathbf{\Lambda}_1 = \begin{bmatrix} 0.656 & 0 \\ 0 & 0.942 \end{bmatrix}, \mathbf{U} = (\mathbf{\Lambda}_1 - \mathbf{Q}_1)^{-1}\mathbf{\Lambda}_1$$

$$\Rightarrow \mathbf{U} = \begin{bmatrix} 0.779 & 0.221 \\ 0.088 & 0.912 \end{bmatrix}$$

The steady probability sates,

$\mathbf{\Pi}.\mathbf{U} = \mathbf{\Pi}$ and $\mathbf{\Pi}.\mathbf{e} = 1$; where $\mathbf{e}$ is a vector of all 1s.

Solving the steady probability sates,

$\Pi_1 = 0.285$ and $\Pi_2 = 0.715$.

Example-2:

Taking typical values of arrival and transition rates of three different connections as,

$\lambda_1 = 0.656$, $\lambda_2 = 0.942$, $\lambda_3 = 0.456$, $\lambda_4 = 0.442$, $\lambda_5 = 0.456$, $\lambda_3 = 0.456$, $\lambda_3 = 0.542$, $r_1 = 0.2$, $r_2 = 0.1$, $r_3 = 0.3$, $r_4 = 0.18$, $r_5 = 0.4$ and $r_6 = 0.2$. Determine transition probability matrix.

The $\mathbf{Q}_i$, $\mathbf{\Lambda}_i$ and $\mathbf{U}_i$ matrices are,

$$\mathbf{Q}_1 = \begin{bmatrix} -r_1 & r_1 \\ r_2 & -r_2 \end{bmatrix}, \mathbf{Q}_2 = \begin{bmatrix} -r_3 & r_3 \\ r_4 & -r_4 \end{bmatrix}, \mathbf{Q}_3 = \begin{bmatrix} -r_5 & r_5 \\ r_6 & -r_6 \end{bmatrix},$$

$$\mathbf{\Lambda}_1 = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \mathbf{\Lambda}_2 = \begin{bmatrix} \lambda_3 & 0 \\ 0 & \lambda_4 \end{bmatrix} \text{ and } \mathbf{\Lambda}_3 = \begin{bmatrix} \lambda_5 & 0 \\ 0 & \lambda_6 \end{bmatrix}$$

$$\mathbf{U}_1 = (\mathbf{\Lambda}_1 - \mathbf{Q}_1)^{-1}\mathbf{\Lambda}_1,\ \mathbf{U}_2 = (\mathbf{\Lambda}_2 - \mathbf{Q}_2)^{-1}\mathbf{\Lambda}_2 \text{ and}$$

$$\mathbf{U}_3 = (\mathbf{\Lambda}_3 - \mathbf{Q}_3)^{-1}\mathbf{\Lambda}_3$$

$$\mathbf{U}_1 = \begin{bmatrix} 0.779 & 0.221 \\ 0.088 & 0.912 \end{bmatrix}, \quad \mathbf{U}_2 = \begin{bmatrix} 0.647 & 0.353 \\ 0.187 & 0.813 \end{bmatrix},$$

$$\mathbf{U}_3 = \begin{bmatrix} 0.604 & 0.323 \\ 0.163 & 0.683 \end{bmatrix}$$

The Kronecker product of above 3 matrices (using Mathcad v.14) is:

$$\mathbf{U} = \mathbf{U}_1 \otimes \mathbf{U}_2 \otimes \mathbf{U}_3$$

$$U = \begin{pmatrix} 0.304 & 0.163 & 0.166 & 0.089 & 0.086 & 0.046 & 0.047 & 0.025 \\ 0.082 & 0.344 & 0.045 & 0.188 & 0.023 & 0.098 & 0.013 & 0.053 \\ 0.088 & 0.047 & 0.382 & 0.205 & 0.025 & 0.013 & 0.108 & 0.058 \\ 0.024 & 0.1 & 0.103 & 0.432 & 6.737\times 10^{-3} & 0.028 & 0.029 & 0.123 \\ 0.034 & 0.018 & 0.019 & 0.01 & 0.356 & 0.191 & 0.194 & 0.104 \\ 9.265\times 10^{-3} & 0.039 & 5.053\times 10^{-3} & 0.021 & 0.096 & 0.403 & 0.052 & 0.22 \\ 9.947\times 10^{-3} & 5.327\times 10^{-3} & 0.043 & 0.023 & 0.103 & 0.055 & 0.447 & 0.24 \\ 2.681\times 10^{-3} & 0.011 & 0.012 & 0.049 & 0.028 & 0.117 & 0.121 & 0.506 \end{pmatrix}$$

III. RESULTS

The variation of offered traffic against the length of frame is plotted in fig.2 for traffic of exponential service time, taking the batch size N as a parameter. Here the maximum length of T_f is taken as 1ms since WiMAX has frame of 5ms which is divided in to 8 sub-frame of length 5/8ms hence maximum value of T_f used in this paper is actually the length of a sub-frame. Offered traffic shows nonlinear relationship with length of frame T_f and converge to linear with increment of batch size N; whereas the above relation is found linear in circuit switch since offered traffic is the product of holding time and arrival rate.

Similar analysis is shown in fig.3 for SS of 3 user's case where the profile of offered traffic vs. length of frame is almost similar but offered traffic per channel becomes smaller because of statistical distribution of traffic among the channels.

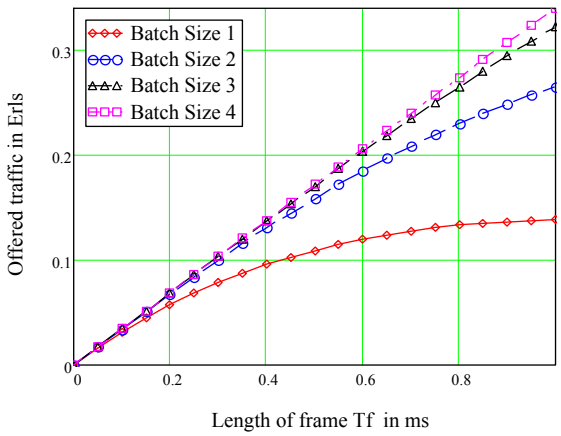


Fig.2 Profile of Offered traffic against the length of frame of single connection ($\lambda_1 = 0.692$, $\lambda_2 = 0.828$ $r_1 = 0.21$, $r_2 = 0.12$)

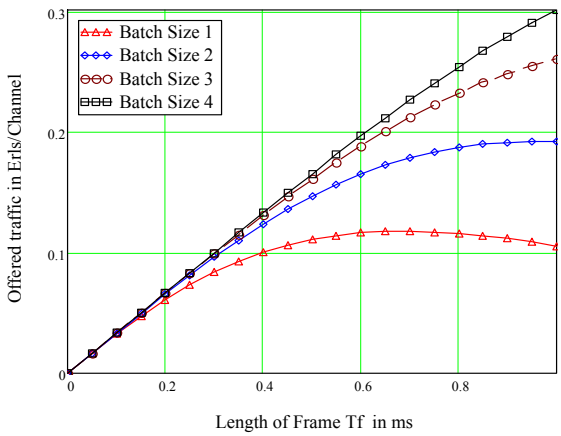


Fig.3 Profile of offered traffic against the length of frame (3 connections)

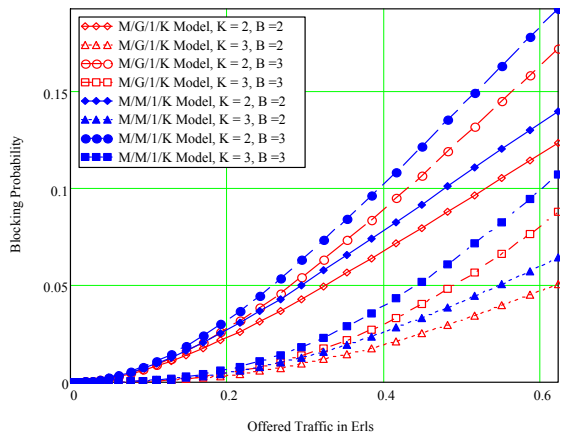


Fig.4 Profile of Blocking Probability against Offered traffic (single connection)

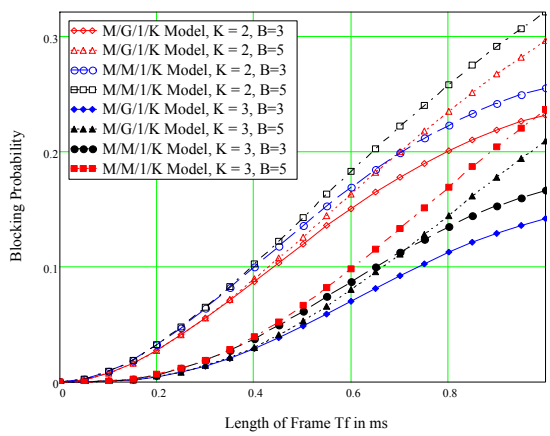


Fig.5 Profile of Blocking Probability against the length of frame (3 connections)

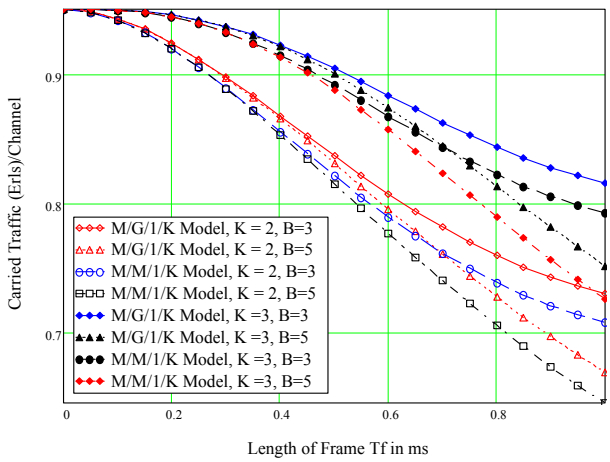


Fig.6 Profile of throughput (carried traffic) against the length of frame (3 connections)

Here transition rate and arrival rate matrix are formed taking the parameters of example-2 is like,

$$\mathbf{Q}_1 = \begin{bmatrix} -0.21 & 0.21 \\ 0.12 & -0.12 \end{bmatrix}, \quad \mathbf{Q}_2 = \begin{bmatrix} -0.35 & 0.35 \\ 0.18 & -0.18 \end{bmatrix}, \quad \mathbf{Q}_3 = \begin{bmatrix} -0.41 & 0.41 \\ 0.2 & -0.2 \end{bmatrix},$$

$$\mathbf{\Lambda}_1 = \begin{bmatrix} 0.656 & 0 \\ 0 & 0.942 \end{bmatrix}, \quad \mathbf{\Lambda}_2 = \begin{bmatrix} 0.456 & 0 \\ 0 & 0.442 \end{bmatrix}, \quad \mathbf{\Lambda}_3 = \begin{bmatrix} 0.456 & 0 \\ 0 & 0.542 \end{bmatrix}.$$

Taking the offered traffic $\rho(T_f)$ of *BMAP* and putting it in *M/G/1/K* model we get BMAP with service of general pdf hence we can model IP traffic of WiMAX of a SS. Fig.4 shows the profile of blocking probability vs. offered traffic of M/M/1/K and M/G/1/K traffic model taking batch size and the size of queue as parameters. The performance of M/G/1/K is found better than M/M/1/K at the same time performance deteriorates with increment of batch size but enhances with increment of size of buffer. Similar decision can be found from fig.5 and 6 where blocking probability increases but carried traffic/channel of decreases with increment of length of frame and batch size but with decrement of length of buffer. Above phenomena is consistent with conventional norm of teletraffic engineering although we got nonlinear relationship between frame length and offered traffic.

IV. Conclusions

In this paper we bring the M/M/1/K and M/G/1/K traffic model under BMAP and show the interrelationship among different traffic parameters for SS of WiMAX. The traffic model of the paper is equally applicable for another platform of 4G called LTE (Long Term Evolution) considering each sub-data string of a subcarrier (each parallel link of Orthogonal Frequency Division Multiplexing) as a separate channel. The entire work can be extended incorporating small scale fading of wireless link to evaluate SNIR (signal to noise and interference ratio) and SER (Symbol Error Rate). We also can model the pdf of aggregate interference of surrounding micro and femtocells to get the realistic results of SNIR and BER

References

[1] Md. Imdadul Islam, Md. Shahriar Karim, J. K. Dasand M. R. Amin, 'Evaluation of Traffic Parameters of Multidimensional Traffic of a Combined Link Using a Tabular Method,' International Conference on Computer and Information Technology,pp.161-165,Dec' 2008, KUET, Bangladesh

[2] Bin Li, Lizhong Li, Bo Li, Krishna M. Sivalingam and Xi-Ren, 'Call admission control for voice/data integrated cellular networks: performance analysis and comparative study,' *IEEE Journal on Selected Areas Communications*, vol.44, no.4, pp.718, May'2004

[3] Derek Doran, Lester Lipsky and Steve Thompson 'Cost-based Optimization of Buffer Size in M/G/1/N Systems Under Different Service-time Distributions,' pp.28-35, Ninth IEEE International Symposium on Network Computing and Applications, 2010

[4] J. Mac Gregor Smith, 'Properties and performance modeling of finite buffer M/G/c/K networks,' Computers and Operation Research, vol.52, Elsevier, pp.1-15, 2010

[5] Prince Bedell and J. MacGregor Smith, 'Topological Arrangements of M/G/c/K, M/G/c/c Queues in Transportation & Material Handling Systems,' Open Networks TND/MHS, pp.1-22, February 14, 2012

[6] Kang, S., Yong, H., Sung, D., Choi, B.: An application of Markovian arrival process (map) to modeling superposed atm cell streams. IEEE Transactions on Communications 50 (2002) 633–642

[7] HaruoAkimaru and Konosuke Kawashima, 'Teletraffic Theory and applications,' Springer-Verlag, London, Berlin, 1993

[8] Sang H. Kang, Dan K. Sung and Bong D. Choi, 'An Empirical real time approximation of waiting tine distribution in MMPP(2)/D/1,' IEEE Communications Letters, vol.2, no.1, pp.17-19, Jan1998

[9] Sang H. Kang, Yong Han Kim, Dan K. Sung and Bong D. Choi, 'An application of Markovian Arrival Process (MAP) to modeling superposed ATM cell stream,' IEEE Transaction on Communications, vol.50, no.4, pp.633-642, April 2002

[10] Sadrac K. Matendo, 'Some Performance Measures for Vacation Models With a Batch Markovian Arrival Process,' Journal of Applied Mathematics and Stochastic Analysis 7, no. 2, pp.111-124, Summer 1994

[11] Abdelali EL BOUCHTI Said EL KAFHALI And Abdelkrim HAQIQ, 'Performance Modeling and Analysis of Connection Admission Control in OFDMA based WiMAX System with MMPP Queuing', World of Computer Science and Information Technology Journal (WCSIT) ISSN: 2221-0741 Vol. 1, No. 4, 148 – 156, 2011

[12] Said El KafhaliAbdelali El Bouchti Mohamed Hanini And AbdelkrimHaqiq, 'Performance Analysis For Bandwidth Allocation In Ieee 802.16 Broadband Wireless Networks Using BMAP Queuing,' International Journal of Wireless & Mobile Networks (IJWMN) vol. 4, no. 1, pp.139-154, February 2012

[13] Hiroyuki Okamura, Tadashi Dohi, *Member,* and Kishor S. Trivedi, 'Markovian Arrival Process Parameter Estimation With Group Data,'IEEE/ACM Transactions on Networking, vol. 17, no. 4,pp.1326-1339, August 2009

[14] Chris Scheper and William J. J. Roberts, 'Anomaly Detection Using an MMPP-based GLRT,' International Journal of Network Security, Vol.17, No.6, PP.672-677, pp.672-677,Nov. 2015

[15] Jesus R. Artalejo, Antonio Gomez-Corral, 'Markovian arrivals in stochastic modelling: a survey and some new results,' SORT 34 (2), pp. 101-14, July-December 2010

[16] Joanna Virginia Rodríguez César, '(Batch) Markovian arrival processes: the identifiability issue and other applied aspects,' TESIS DOCTORAL, Department of Statistics UNIVERSIDAD CARLOS III DE MADRID, March 2015

[17] UmmyHabiba, Md. Imdadul Islam M. R. Amin, 'Performance Evaluation of VoIP Service of Cognitive Radio System Based on DTMC,' Journal of Information Processing Systems (Korea), J Inf Process Syst, Vol.10, No.1, pp.119~131, March 2014

[18] J. Mac Gregor Smith, 'M/G/c/K blocking probability models and system performance,' Performance Evaluation, vol.52, Elsevier, pp.237-267, 2003

[19] Prince Bedell and J. MacGregor Smith, 'Topological Arrangements of M/G/c/K, M/G/c/c Queues in Transportation & Material Handling Systems,' Open Networks TND/MHS, pp.1-22, February 14, 2012

[20] Jesmin Akhter, Md. Imdadul Islam, M.R. Amin, 'Performance Evaluation of Voice Traffic of IEEE 802.16e under Femto Cellular Network,' International Journal of Computer Applications (0975 – 8887), Volume 124, No.5, pp.37-41, August 2015

[21] Anupam Roy, Md. Imdadul Islam, And M. R. Amin, 'Performance Evaluation of Voice-Data Integrated Traffic in IEEE 802.11 and IEEE 802.16e WLAN,' WSEAS Transactions on Communications, Issue 7, Volume 12, pp.352-365, July 2013

[22] DusitNiyato and Ekram Hossain, 'Queue-Aware Uplink Bandwidth Allocation and Rate Control for Polling Service in IEEE 802.16 Broadband Wireless Networks,' IEEE Transactions on Mobile Computing, vo.5, no.6, pp.668-679, June 2006

Design of a Hummingbird Crypto Core Implementing BIST Technique

Md. Azizul Haque[1], and Md. Liakot Ali[2]

Institute of Information and Communication Technology, Bangladesh University of Engineering and Technology
Dhaka-1000, Bangladesh
[1]azhaque32@gmail.com
[2]liakot@iict.buet.ac.bd

Abstract— **In recent years, lightweight cryptography which focuses on designing new cryptographic primitives with small hardware footprint and with low average and peak power consumption, has received a lot of attention. Hummingbird, an ultra-lightweight cryptographic algorithm has been developed both with lightweight software and lightweight hardware implementations keeping resource-constrained devices in mind. In this paper, we describe FPGA implementation with both software simulation as well as HW implementation of Hummingbird algorithm on Altera Platform. Moreover, testability of a complex chip is of prime concern now a days. Built-In-Self-Test (BIST) is a widely accepted approach for this purpose. In this paper design of a Hummingbird Crypto Core implementing BIST technique is proposed and both software and hardware implementation is performed. Here LFSR is used as pseudorandom test pattern generator and Signature Analysis is used as a Data Compression Technique. The proposed Crypto Core is simulated in Quartus II simulation software for the Altera Cyclone II family device as well as hardware implementation done in Altera DE2 board and performance is analyzed and compared with the other research works on Hummingbird implementation.**

Index Terms*— *Lightweight Cryptography; Testability; BIST; FPGA.

I. Introduction

Information security is of prime concern for all the organizations in this era of Information and Communication Technology (ICT) [1]. All the business organizations are investing huge resources for this. The security threat encompasses in many applications such as internet e-commerce, internet banking, internet of things, mobile networks, automated teller machines (ATMs), copy protection against reverse engineering and software piracy, military and government to facilitate secret communication and many more. Cryptography plays an important role in security system to ensure confidentiality of the data. In cryptography the secret data is encrypted by a secret key and the encrypted data can only be deciphered if one has the key. The encryption algorithm is a set of well-defined steps to transform data from a readable format to an encoded format using the key so that the data become unreadable to the unauthorized users. This set of well-defined steps is called cipher. A number of algorithms on cryptography have been presented in the literatures [2-5]. At present AES has been proved as the strongest encryption algorithm declared by USA Govt. [5].

Depending on the different applications requirement speed, the physical area for the embedded hardware or its power budget are the things to consider while choosing a cryptographic algorithm. In today's world there is an increasing number of battery powered embedded systems like cell phones, networked sensors, smart cards, RFID which are used to store, access, manipulate or communicate sensitive data. It makes security an important issue for these devices. Low power and small area are the mandatory requirements for these resource constrained and battery powered devices. It has been found AES is not suitable for low cost embedded devices such as RFID tags, smart cards, wireless sensor nodes etc. due to their resource constraints and for this pervasive world new lightweight cryptographic algorithm is required [6-8]. Hummingbird has been proposed as ultra-lightweight cryptographic algorithm suitable for embedded system and shown that it is resistant to a number of attacks such as algebraic attacks, cube attacks, differential power attacks etc. The efficiency of the algorithm has been verified in software solution for a wide range of embedded applications [9-11].

Again testability of a complex chip is of prime concern in today's IC design. Testing a VLSI chip to guarantee its functionality is extremely complex, time consuming as well as expensive [12]. To deal with the testing problem at the chip level incorporating built-in self-test (BIST) capability inside a chip is a widely accepted approach [13]. When chip is complex then Built-In-Self-Test (BIST) is a norms of this day because external testing using ATE is not cost effective and less convenient in this case. BIST in Hummingbird crypto core is not reported yet. This paper focuses on implementing BIST in Hummingbird Crypto Core in Altera platform.

II. Overview of Hummingbird Cryptographic Algorithm

In Hummingbird algorithm both block cipher and stream cipher used combined with 16-bit block size, 256-bit key size, and 80-bit internal state. Figure 1, Figure 2 and Figure 3 illustrate the initialization, encryption and decryption processes of the Hummingbird cryptographic algorithm, respectively.

Both initialization and encryption consist of four 16-bit block encryption ciphers E_{ki} (i = 1, 2, 3, 4), four 16-bit internal state registers RSi (i = 1, 2, 3, 4), and a 16-stage Linear Shift Feedback Register (LFSR). In decryption block decryption ciphers D_{ki} (i = 1, 2, 3, 4) are used. The 256-bit secret key K is divided into four 64-bit subkeys k_1, k_2, k_3 and k_4 which are used in the four block ciphers respectively.

A. Initialization Process

Hummingbird initialization algorithm structure is shown in Figure 1. Four 16-bit random nonces $NONCE_i$ are first chosen

to initialize the four internal state registers RS_i (i = 1; 2; 3; 4), respectively, followed by four consecutive encryptions on the message RS1⊞RS3. The final 16-bit cipher text *TV* is used to initialize the LFSR. The 13th bit of the LFSR is always set to prevent a zero register.

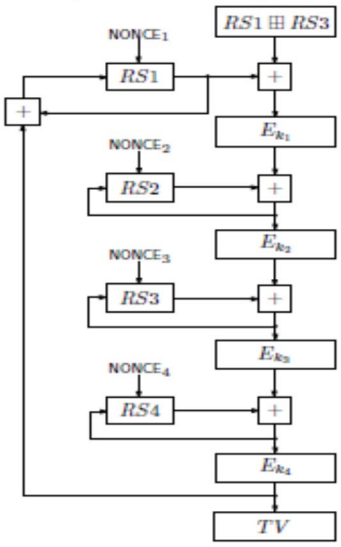


Fig. 1: Initialization Process

Algorithm: Hummingbird Initialization
Input: Four 16-bit random nonce $NONCE_i$ (i = 1; 2; 3; 4)
Output: Initialized four internal registers RSi_4 (i = 1; 2; 3; 4) and LFSR

1: $RS1_0 = NONCE_1$ [Nonce Initialization]
2: $RS2_0 = NONCE_2$
3: $RS3_0 = NONCE_3$
4: $RS4_0 = NONCE_4$
5: for t = 0 to 3 do
6: $V12_t = E_{k1}((RS1_t \boxplus RS3_t) \boxplus RS1_t)$
7: $V23_t = E_{k2}(V12_t \boxplus RS2_t)$
8: $V34_t = E_{k3}(V23_t \boxplus RS3_t)$
9: $TV_t = E_{k4}(V34_t \boxplus RS4_t)$
10: $RS1_{t+1} = RS1_t \boxplus TV_t$
11: $RS2_{t+1} = RS2_t \boxplus V12_t$
12: $RS3_{t+1} = RS3_t \boxplus V23_t$
13: $RS4_{t+1} = RS4_t \boxplus V34_t$
14: end for
15: $LFSR = TV_3 \mid 0x1000$ [LFSR Initialization]
16: return RSi_4 (i = 1; 2; 3; 4) and LFSR

B. Encryption Process

Hummingbird encryption algorithm is depicted in Figure 2. After a system initialization process, a 16-bit plaintext block PT_i is encrypted by first executing a modulo 2^{16} addition of PT_i and the content of the first internal state register RS1.The result of the addition is then encrypted by the first block cipher E_{k1}. This procedure is repeated in a similar manner for another three times and the output of E_{k4} is the corresponding ciphertext CT_i. Furthermore, the states of the four internal state registers will also be updated in an unpredictable way based on their current states, the outputs of the first three block ciphers, and the state of the LFSR.

Algorithm: Hummingbird Encryption
Input: A 16-bit plaintext PT_i and four internal registers RSi_t (i = 1; 2; 3; 4)
Output: A 16-bit ciphertext CT_i

1: $V12_t = E_{k1}(PT_i \boxplus RS1_t)$ [Block Encryption]
2: $V23_t = E_{k2}(V12_t \boxplus RS2_t)$
3: $V34_t = E_{k3}(V23_t \boxplus RS3_t)$
4: $CT_i = E_{k4}(V34_t \boxplus RS4t)$
5: $LFSR_{t+1} \leftarrow LFSR_t$ [Internal State Updating]
6: $RS1_{t+1} = RS1_t \boxplus V\ 34_t$
7: $RS3_{t+1} = RS3_t \boxplus V\ 23_t \boxplus LFSR_{t+1}$
8: $RS4_{t+1} = RS4_t \boxplus V\ 12_t \boxplus RS1_{t+1}$
9: $RS2_{t+1} = RS2_t \boxplus V\ 12_t \boxplus RS4_{t+1}$
10: return CT_i

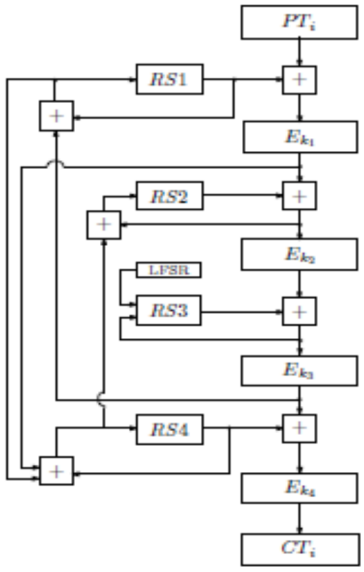


Fig. 2: Encryption Process

C. Decryption Process

Hummingbird decryption algorithm is illustrated in Figure 3. The decryption process follows the similar pattern as the encryption and a detailed description is shown in the following Algorithm 3.

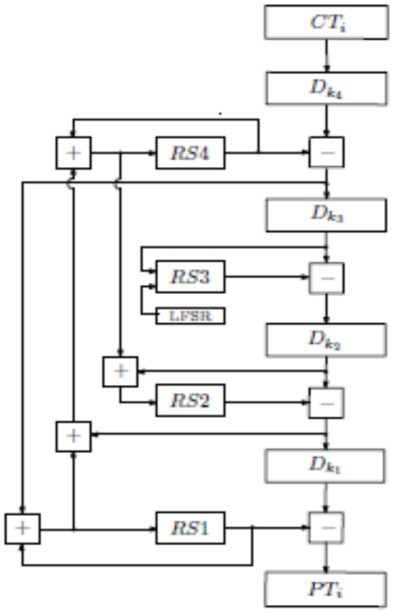


Fig. 3: Decryption Process

Algorithm: Hummingbird Decryption
Input: A 16-bit ciphertext CTi and four internal registers RSi_t (i = 1; 2; 3; 4)
Output: A 16-bit plaintext PTi

1: $V34_t = D_{k4}(CTi) \boxminus RS4_t$ [Block Decryption]

2: $V23_t = D_{k3}(V34_t) \boxminus RS3_t$
3: $V12_t = D_{k2}(V23_t) \boxminus RS2_t$
4: $PT_i = D_{k1}(V12_t) \boxminus RS1_t$
5: $LFSR_{t+1} \leftarrow LFSR_t$ [Internal State Updating]
6: $RS1_{t+1} = RS1_t \boxminus V34_t$
7: $RS3_{t+1} = RS3_t \boxminus V23_t \boxminus LFSR_{t+1}$
8: $RS4_{t+1} = RS4_t \boxminus V12_t \boxminus RS1_{t+1}$
9: $RS2_{t+1} = RS2_t \boxminus V12_t \boxminus RS4_{t+1}$
10: return PT_i

III. Proposed Work

In our proposed design we have incorporated BIST functionalities in the Hummingbird Crypto Core.

In our design, we have used the following blocks/modules:

1. Memory module
2. Initialization module
3. Input Selector
4. Test Pattern generator
5. Encryption
6. Decryption
7. Signal Analyser
8. Control Module

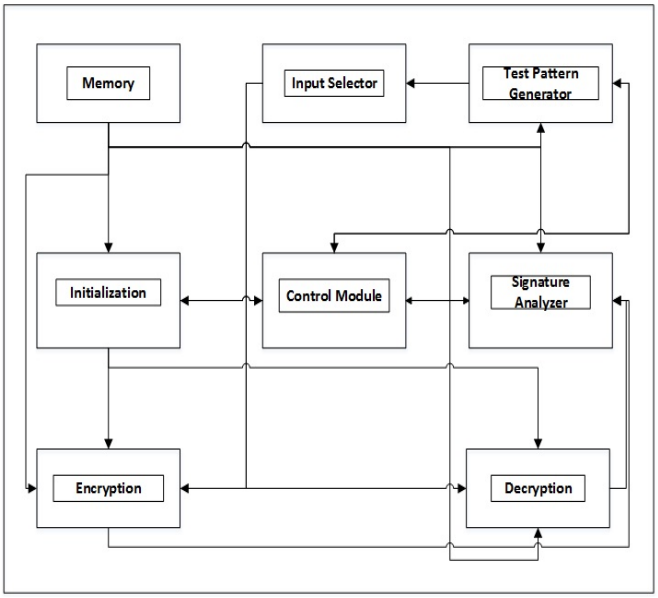


Fig. 4: Functional Block Diagram of the proposed design

The Hummingbird Algorithm operates on a block of 16 bits of input data and generates 16 bits of output. The length of the key used to encrypt/decrypt is 256 bits and there are 80 bits of internal states.

At the start, our desired Crypto Core is initialized. In this phase, internal registers and LFSR are initialized using NONCE and key values stored in the Memory which are preloaded. These initialized internal registers and LFSR are used in later phases in encryption/decryption modules. Depending on the mode of operation, the proposed crypto core can operate in Normal/Test mode. In Normal mode, both encryption and decryption modules works in parallel and provide output cipher text or plain text respectively for input of plain text or cipher text. In Test mode the proposed crypto core performs test for either encryption/decryption module, i.e. one at a time. In test mode, pseudorandom pattern generated by the test pattern generator is input to the desired test module (encryption/decryption) and then output is fed to the Signature Analyzer which produces a signature. This process is continued for the desired number of test count and the final signature is matched with the golden signature stored in the Memory. If it matches it indicates the functionality of the module under test is OK.

Figure 5 shows the flowchart of the design.

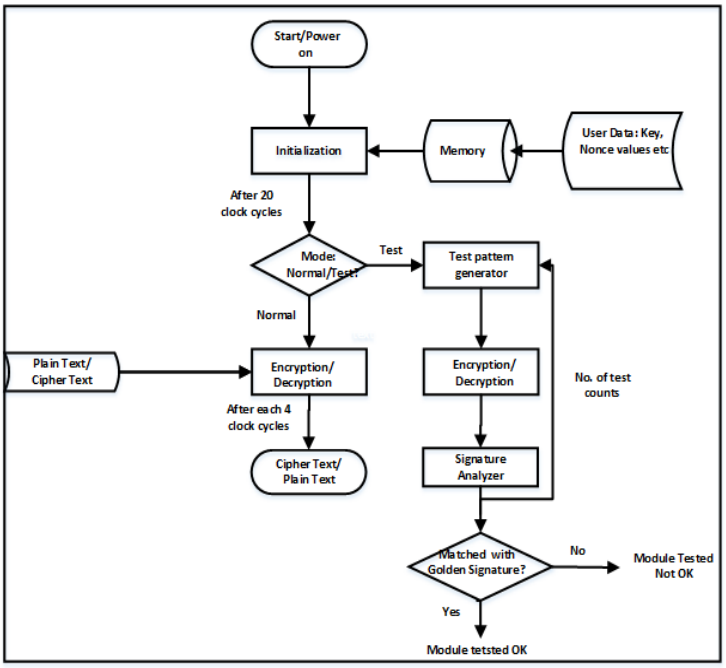


Fig. 5: Flowchart of the Hummingbird Crypto Core with BIST

IV. Experimental Results and Discussions

The proposed Hummingbird crypto core is implemented in FPGA device of family Altera, Cyclone II, EP2C5F256C7 device. The simulation is performed in Quartus II simulation software hardware implementation done in Altera DE2 Board.

A. HW Implementation

We have used Altera DE2 board for HW implementation. Below are the snapshots of HW implementation:

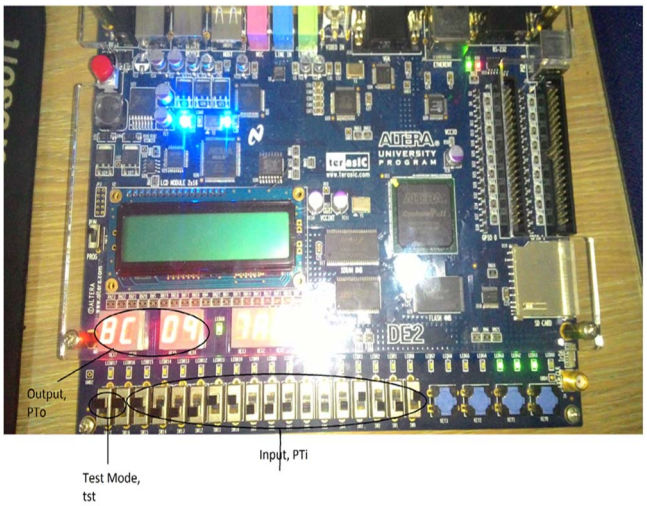


Fig. 6: Snapshot of HW Implementation

In the above setup, tst is set to 11 which indicate the module is in self-test. Here SW0 to SW15 is used as input PTi and HEX4 to HEX7 is used as output. Here PTo showed the

same value as PTi and thus proved the proper functionality in this mode.

B. Test Result

Encryption/Decryption functionality of the proposed Core is tested by using the output of the encryption block as the input to the decryption block. In this case we should get back the original plain text. In our design we have this arrangement to test the encryption followed by decryption.

TABLE I. INPUT AND OUTPUT IN TEST RESULT

Encryption Block input (PT_i)	Encryption Block output (CT_o)	Decryption Block input (CTx)	Decryption Block output (PTo)
A19C	6E5F	6E5F	A19C
49D1	7582	7582	49D1
3B79	05B0	05B0	3B79
6921	E054	E054	6921

In the above shown result obtained from simulation we have used plain text (PT_i) input stream as A19C, 49D1, 3B79, 6921 for Encryption module. After 4 clock cycles we got the encrypted output (CT_o) stream as 6E5F, 7582, 05B0, E054. In the next clock these cipher text used as input (CT_x) to the Decryption module. After another 4 clock cycles, we got the plain text (PT_o) output stream as A19C, 49D1, 3B79, 6921 which is found same as Plain Text (PT_i) input to the encryption module.

C. Comparison with Other Research Works

In this paper we have implemented Hummingbird Crypto Core with BIST. There are lots of works on Hummingbird but we don't find any other works which is implemented with BIST. In the below table a comparison is shown with other Hummingbird implementations. It is observed that the besides BIST functionality proposed design shows good throughput and efficiency comparing with the other research works.

TABLE II.PERFORMANCE COMPARISON WITH OTHER WORKS

Design/Cipher	Device	Total Occupied Slices/LAB	F_{Max} (MHz)	Throughput (Mbps)	Efficiency (Mbps/ # of Slices)
Hummingbird[14]	Spartan-3 XC3S200-5	273	40.1	160.4	0.59
Hummingbird[15]	Spartan-3 XC3S200-5	40	260.8	55.64	1.38
Hummingbird-2[16]	Spartan-3 XC 35200	273	40.1	160.4	0.59
Hummingbird[17]	Virtex-5 XC5V1X20T-2-FF-323	4242	152.905	NA	NA
This work	Cyclone II	282	59.3	238.4	0.85

V. CONCLUSION

Hummingbird is an ultra-lightweight cryptographic algorithm targeted for resource constrained devices. Also Built-In-Self-Test (BIST) is a widely accepted approach to ensure testability of an IC. This paper presented the design of Hummingbird Crypto Core implementing BIST. The design is implemented in FPGA device of Family Altera, Cyclone II. Simulation result of the design ensures that the design is functioning properly. The ciphered output from the design is verified by using it as input to the decryption and revert back the original plain text.

The work performed in this paper also exposes several new areas that can be explored. In future, rigorous performance analysis and optimization in terms of power, speed and hardware resources can be performed and also to improve fault coverage different approaches of BIST like recursive pseudo exhaustive two pattern generator can be explored.

REFERENCES

[1] W. Stallings, *Cryptography and Network Security: Principles and Practices*. Pearson Education, Inc. 2010.

[2] P. Mahajan and A. Sharma "A Study of Encryption Algorithms AES, DES and RSA for Security," Global Journal of Computer Science and Technology Network, Web & Security, vol. 13, Issue 15, Version 1.0 ,Year 2013

[3] A. Ramesh and A. Surliandi "Performance analysis of encryption algorithms for Information Security," published in 2013 International Conference on Circuits, Power and Computing Technologies (ICCPCT), paper 10.1109/ICCPCT.2013.6528957, pp. 840 – 844.

[4] G. Singh, and Supriya, "A Study of Encryption Algorithms (RSA, DES, 3DES and AES) for Information Security," International Journal of Computer Applications, vol. 67(19), pp. 33-38, April 2013.

[5] "Advanced encryption standard (AES)," Federal Information Processing Standards Publication (FIPS PUB) 197, National Institute of Standards and Technology (NIST), November, 2001.

[6] D. Engels, X. Fan, G. Gong, H. Hu, and E.M. Smith, "Hummingbird: Ultra-Lightweight Cryptography for Resource- Constrained Devices," 14th International Conference on Financial Cryptography and Data Security - FC 2010, Berlin, Germany, Springer-Verlag, 2010.

[7] A. Poschmann, "Lightweight Cryptography - Cryptographic Engineering for a Pervasive World," Ph.D. Thesis, Department of Electrical Engineering and Information Sciences, Ruhr-Universit aet Bochum, Bochum, Germany, 2009.

[8] T. C. Thanuja, "Review Paper on Lightweight Cryptography: Hummingbird on Different Platform," International Journal of Art & Humanity Science (IJAHS) , vol. 2 Issue 3, pp. 29-33, (May-June 2015).

[9] X. Fan, G. Gong, K. Lauffenburger, and T. Hicks, "FPGA Implementations of the Hummingbird Cryptographic Algorithm," IEEE International Symposium on Hardware -Oriented Security and Trust (HOST), 13-14 June, 2010.

[10] X. Fan, H. Hu, G. Gong, E.M. Smith, and D. Engels, "Lightweight Implementation of Hummingbird Cryptographic Algorithm on 4-Bit Microcontrollers," 1st International Workshop on RFID Security and Cryptography 2009 (RISC'09), pp. 838-844, 2009.

[11] R. Bhatia, "Study of Hummingbird Cryptographic Algorithm based on FPGA Implementation",2014 IJCSIT International Journal of Computer Science and Information Technologies, vol. 5(3), pp. 4426-4430, 2014.

[12] K. Nivita and A. Titus, "A BIST Circuit for Fault Detection Using Recursive PseudoExhaustive Two Pattern Generator," International Journal of Modern Engineering Research (IJMER), vol. 2, Issue.3, pp. 676-681, May-June 2012.

[13] Parag K. Lala, *An introduction to logic circuit testing*, Morgan&Claypool publishers, 2008.

[14] X. Fan, G. Gong, K. Lauffenburger, and T. Hicks "Design Space Exploration of Hummingbird Implementations on FPGAs". (2010) The citeseerx website. [Online] Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.233.1601.

[15] I. San and N. At, "Compact Hardware Architecture for Hummingbird Cryptographic Algorithm," 2011 21st International Conference on Field Programmable Logic and Applications, Chania, 2011, pp. 376-381.

[16] J. John "Performance Analysis of New Light Weight Cryptographic Algorithms," IOSR Journal of Computer Engineering (IOSRJCE) ISSN: 2278-0661, ISBN: 2278-8727, vol. 5, Issue 5, pp. 01-04 (Sep-Oct. 2012).

[17] N. Arora and Y. Gigras, "FPGA Implementation of Low Power and High Speed Hummingbird Cryptographic Algorithm," International Journal of Computer Applications(0975-8887), vol. 92- No.16, April 2014.

Simulation Study on the Effects of Changing Band Gap on Solar Cell Parameters

Mostofa Amir Rifat, Mohammad Imtiaz Morshed Khan, and Khairul Alam
Department of Electrical and Electronic Engineering, East West University
A/2, Jahurul Islam Avenue, Jahurul Islam City, Aftabnagar, Dhaka-1212, Bangladesh
mostofaamir044@gmail.com, imtiaz.morshed051@yahoo.com, kalam@ewubd.edu

Abstract—**We perform the numerical analysis of a CIGS solar cell parameters such as open circuit voltage, short circuit current, maximum power, fill factor, and external quantum efficiency as a function of absorber layer band gap. These parameters are known to be the key parameters of a solar cell to determine its performance. We change the band gap of the CIGS absorber layer by changing its alloy composition. ATLAS SILVACO is used to construct and simulate the CIGS solar cell structure with standard AM1.5 spectra. The open circuit voltage and the maximum power increase almost linearly with the band gap. However, the change in short circuit current and the fill factor with the CIGS bandgap does not show any formal relation. We found that the change in fill factor due to the band gap change is not significant compared to the change in open circuit voltage.**

Index Terms—**Solar cell; CIGS absorber layer;Open circuit voltage; Short circuit current density; Fill factor; Maximum power; External quantum efficiency**

I. Introduction

One of the major problems of this century is the limitation of fossil fuels. So, it is right time to search for new sources of renewable energy. Solar energy is the principle source of all energies. The technology through which it is possible to use solar energy to get electrical energy is called photovoltaic (PV) technology. The cell which is designed to convert photo energy into electrical energy is called solar cell. Photovoltaic technology can be categorized under different generations. The first generation PV technology is based on crystalline silicon. Second generation PV technology is based on thin film cells, which is relatively cheaper technology. The final PV technology is based on III-V semiconductor materials, which are known as multi-junction solar cells.

Toberer *et al*. [1] suggested that open circuit voltage should be an indicator of material quality. They presented the dependence of open circuit voltage on band gap. They used a step-edge absorption model to show that the photo voltage increases and photocurrent decreases systematically with increasing band gap. They also discussed the effect of changing thickness in increasing efficiency. They studied upon base thickness of 1μm and 3μm using conventional epitaxial growth and processing techniques, but without antireflection coating. Next, they used antireflection coating to study the behavior of External Quantum Efficiency (EQE). J. Krc and co-workers [2] have worked on optical and electrical modeling of thin film CIGS ($CuIn_xGa_{(1-x)}Se_2$) solar cells. For optical simulations of the structure, they use experimental values of complex refractive indices. The optical simulator Sun Shine [3, 4] and the electrical simulator Aspin [5, 6] were used to simulate optical and electrical properties and characteristics of the thin CIGS solar cell in their study. S. Sharbati and J. R. Sites [7] have studied the performance of solar cells for Zn(O, S)/ Cu(In, Ga)Se_2 heterojunction. They showed that with the increase in band gap of CIGS, open circuit voltage increases. Mustafa M. A. Hussein and co-workers [8] studied ITO/CdTe/Al_2O_3/Si/Au thin film solar cell and they proposed a maximum value of quantum efficiency and fill factor at a thickness of 1.7 nm Al_2O_3. The I-V characteristics are measured by KEITHLEY computerized system. The contact wires are soldered to each layer of samples by indium alloy soldering. This type of soldering ensures good contacts, and low temperature keeps the structural properties of films unchanged. The fill factor and efficiency are calculated from I-V characteristics for different thickness of Al_2O_3 layer. They declared Gold as back contact and ITO as front contact for their hetero junction structure.

In this work, we vary the band gap of CIGS layer by varying the *x*-composition of $CuIn_xGa_{(1-x)}Se_2$ to observe the resultant variations in open circuit voltage, short circuit current, maximum power, fill factor, and external quantum efficiency. We found that the open circuit voltage and the maximum power increase almost linearly with the *x*-composition. However, short circuit current and fill factor show very non-linear behavior with *x*-composition. Note that the band gap of $CuIn_xGa_{(1-x)}Se_2$ increases almost linearly with the *x*-composition.ATLAS is used to simulate solar cells. The simulation data from ALTAS is exported to MATLAB for post processing.

II. Simulation Approach Using ATLAS SILVACO

Semiconductor band gap is one of the key factors that influences losses and hence the conversion efficiency of a solar cell. To see the band gap influence on solar cell parameters such as open circuit voltage, short circuit current and fill factor, we simulate a solar cell structure using SILVACO simulation package with CIGS as the absorbing layer. The *x*-composition of CIGS alloy was varied to vary the band gap of CIGS.

The simulation structure is shown in Fig. 1. The structure consists of 0.4 micron TCO, 0.04 μm n type CdS, 3 μm p type CIGS absorber layer, 0.4 μm Mo front and back contacts. Using "`model`" statement of SILVACO, specified and required physical modelsare declared for simulating the structure. The physical models are divided into five classes

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

mobility, recombination, carrier statistics, impact ionization, and tunneling. We use Shockley-Read-Hall (SRH) recombination model as described below.

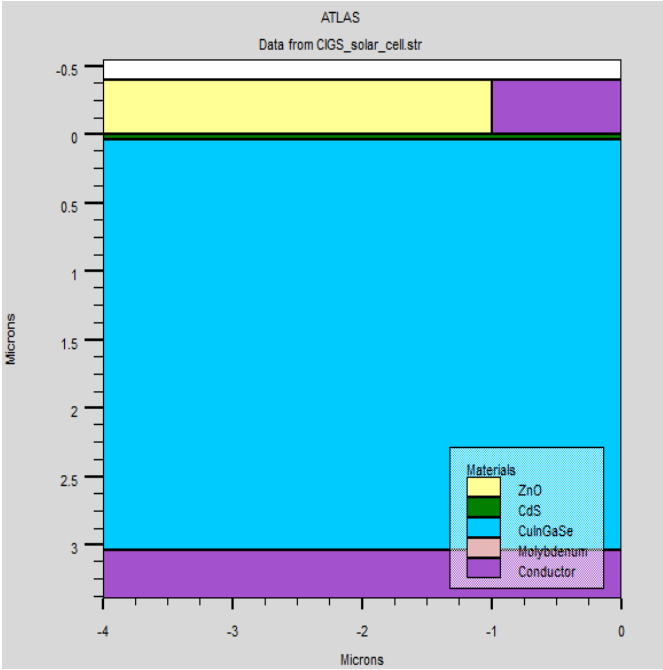


Figure 1: The CIGS solar cell structure used for simulation

$$R_{SRH} = \frac{pn - n_{ie}^2}{TAUP0[n + n_{ie}\exp(\frac{ETRAP}{kT_L})] + TAUN0[p + n_{ie}\exp(\frac{-ETRAP}{kT_L})]}$$

Here, ETRAP is the difference between the trap energy level and the intrinsic Fermi level, T_L is the lattice parameter in degree kelvin, and TAUN0 and TAUP0 are the electron and hole lifetimes. Using "srh" parameter of the model statement, this recombination model is declared. The materials parameters are listed below. Other parameters are default from ATLAS database.

TABLE I

MATERIAL PARAMETERS USED FOR SIMULATION

Parameters	**ZnO**	**CdS**	**CIGS**
Doping concentration	5×10^{17} cm^{-3}	1×10^{18} cm^{-3}	1×10^{16} cm^{-3}
Doping type	n type	n type	p type
Band gap	3.3 eV	2.4 eV	varied
Dielectric permittivity	9 F/m	10 F/m	13.6 F/m
Electric affinity	4 eV	3.75 eV	3.89 eV
Electron mobility	50 cm^2/Vs	10 cm^2/Vs	300 cm^2/Vs
Hole mobility	5 cm^2/Vs	1 cm^2/Vs	30 cm^2/Vs

III. Result and Analysis

The J-V characteristics for different *x*-compositions are shown in figure 2.

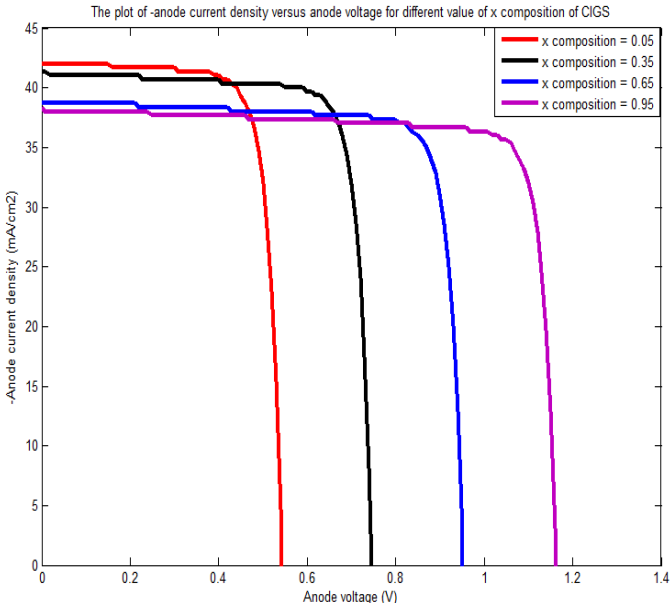


Figure 2: Anode current density versus anode voltage characteristics for different *x* composition of CIGS solar cell

The J-V curves look as they are expected to look, and the changes in open circuit voltage and short circuit current with band gap (*x*-composition) are clearly observed. Different parameters of solar cells such as open circuit voltage (V_{oc}), short circuit current density (J_{sc}), maximum power (P_m), fill factor (FF), external quantum efficiency (EQE) are studied.Those are presented below.

A. Open Circuit Voltage

From our simulation, the following profile of open circuit voltage was found with changing *x*-composition of CIGS. In figure 3, it is clear that open circuit voltage increases with the increase in *x*-composition of CIGS material.The variation of open circuit voltage is almost linear with *x*-composition.

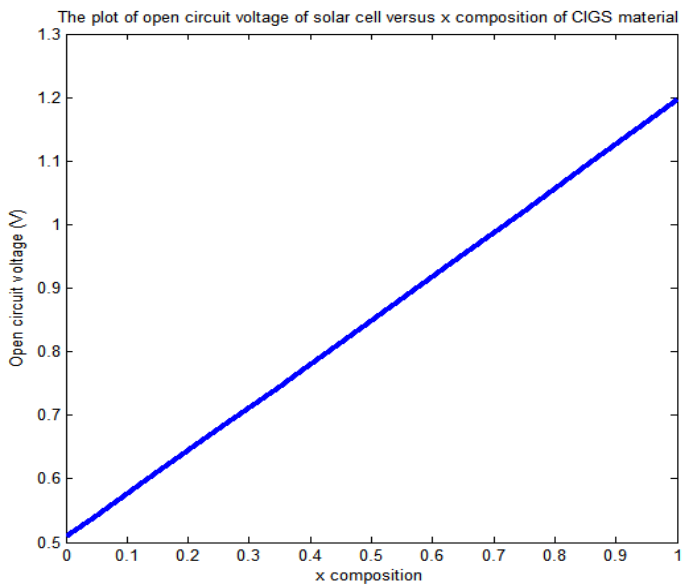


Figure 3: Open Circuit Voltage versus x composition for CIGS solar cell

B. Short Circuit Current

According to simulation, the profile of short circuit current density with respect to the *x*-composition of CIGS material is resented in the following figure 4. The short circuit current of the CIGS solar cell is decreased with increasing *x*-composition of the CIGS material, and the change does not follow any particular relationship. The current density changes from 42 mA/cm^2 to 37.6 mA/cm^2 when the *x*-composition is varied from 0 to 1.

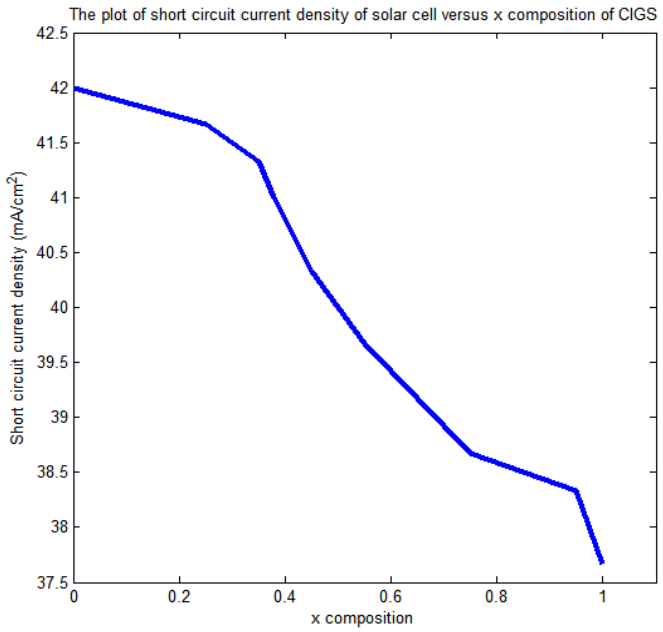


Figure 4: Short Circuit Current Density versus x composition for CIGS solar cell

C. Maximum Power

The maximum power is the multiplication of maximum voltage and maximum current. For different value of *x*-composition, simulation is first done. From the P-V curves for each *x*-composition, maximum power is found. Finally, the maximum power profile is found as shown in figure 5.

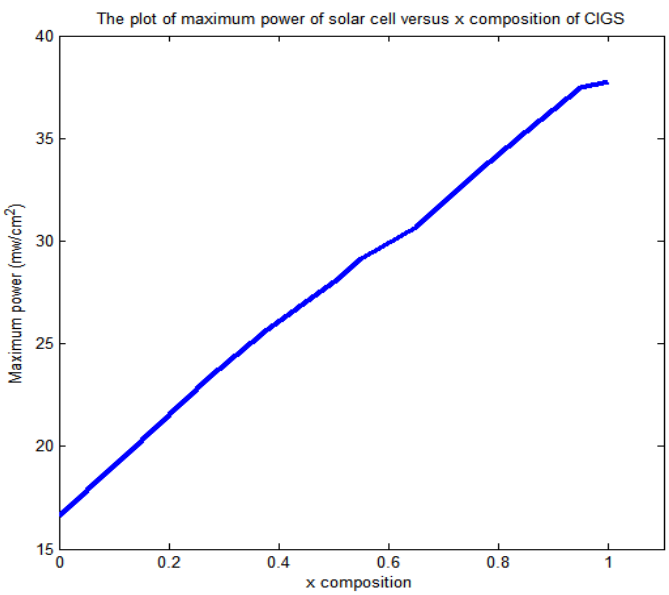


Figure 5: Maximum power versus *x* composition for CIGS solar cell

From the simulation result, it is clear that with the increase in *x*-composition of the CIGS material, the maximum power of CIGS solar cell increases. Theoretically, the maximum power increases with the band gap, peaks and then decreases. The peak occurs at a band gap value of 1.67 eV. With *x*-composition, the band gap of CIGS changes from 1.01 to 1.67 eV, that is, the band gap does not exceed the optimum value, and therefore, we do not see a decline nature in the maximum power curve.

D. Fill Factor

The fill factor profile is quite interesting for this CIGS solar cell. The profile is shown in figure 6.

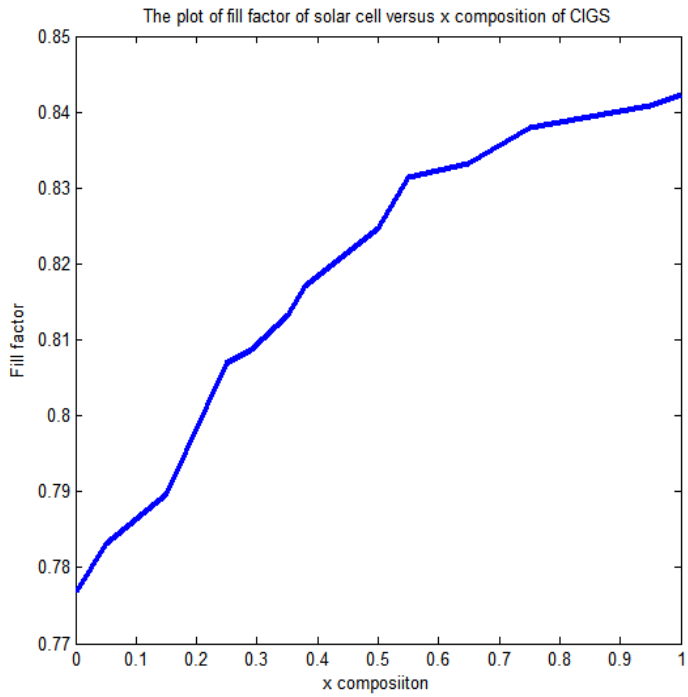


Figure 6: Fill factor versus x composition for CIGS solar cell

The fill factor is known to increase with open circuit voltage. As the V_{oc} increases with *x*-composition, so does the fill factor. However, when the V_{oc} changes from 0.51 V to 1.2 V, the fill factor change from 0.778 to 0.841. In other word, the change in fill factor is not significant compared to the change in V_{oc}.

E. External Quantum Efficiency

External Quantum Efficiency (EQE) was studied for a range of optical wavelength from 0.3 micron to 2.5 micron. Figure 7 represents different EQE versus optical wavelength characteristics for different *x*-composition values. From this characteristic, it is clear that the cell with lower *x*-composition can be used for longer range of optical wavelength. So, the design of an efficient solar cell over a long range of optical wavelength of incident photon on the surface should consist of CIGS ($CuIn_xGa_{(1-x)}Se_2$) layer having smaller x composition.

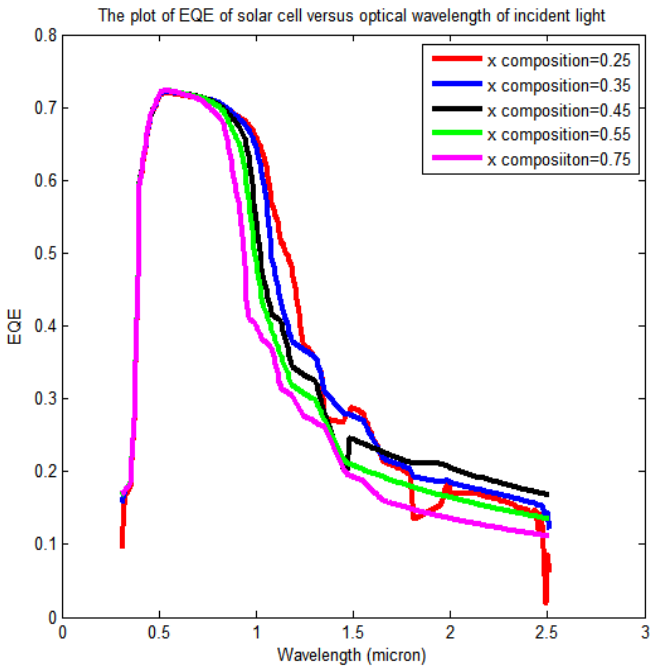


Figure 7: External Quantum Efficiency (EQE) versus optical wavelengthof CIGS solar cell for five different value of x-composition

The photon energy has an inverse relation with the optical wave length. With the increase in optical wavelength, the energy of photon decreases. Photon having energy less than the band gap of the absorption layer of solar cell will not be absorbed. So, the EQE at longer wavelength becomes zero as the CIGS absorber layer becomes transparent to the incident photons. At small wave length, the photo generated electrons go to excited energy level, and then come to the conduction band edge by thermalization. This thermal loss reduces EQE at short wave length.

IV. Conclusion

A numerical study on the effects of changing band gap on solar cell parameters is performed. ATLAS SILVACO is used to create a solar cell structure. This structure is simulated for different *x*-composition of the CIGS layer under AM1.5 spectrum illumination. Then open circuit voltage, short circuit current, maximum power, fill factor, and external quantum efficiency are plotted to see the variations of these parameters with variation of band gap. It is seen that open circuit voltage increases with increase in *x*-composition of CIGS and the short circuit current decreases with increase in *x*-composition of CIGS. The maximum power should increase, reach at a peak and then decrease. But in this study the range of band gap of CIGS varies from 1.01 eV to 1.67 eV when the *x*-composition is changed from 0 to 1; which is not enough to see that characteristics. Fill factor also increased with increase in *x*-composition; but this change is not as significant as open circuit voltage. The external quantum efficiency is found for different *x*-composition of CIGS layer, where flatter upper peak is found for smaller *x*-composition. It indicates that cell with lower *x*-composition can be used for longer range of optical wavelength.

Reference

[1]: S. Eric Toberer, C. Tamboli Adele, S. Myles, and K. Sarah, "Analysis of Solar Cell Quality Using Voltage Matrics", *Conference Paper*, NREL/CP-5200-54120, June 2012

[2]: J. Krc, G. Cernivec, A. Campa, J. Malmström, M. Edoff, F. Smole, and M. Topic, "Optical and electrical modeling of Cu(In,GA)Se_2 solar cells", *optical and quantum electronics*, vol. 38, PP.1115-1123, 2006

[3]: J. Krc, PhD Thesis, University of Ljubljana, 2002

[4]: J.Krc, F. Smole and M. Topic,"*Prog. Photovolt., Res. Appl.*"vol. 11, PP.15, 2003

[5]: M. Topic, F. Smole and J. Furlan, *J. Appl. Phys.* vol.79, PP.8537, 1996

[6]: G. Cernivec, J. Krc, F. Smole and M. Topic, *Thin solid films,* vol.60, PP.511–512, 2006

[7]: S. Sharbati and R. SitesJames, "Impact of the Band Offset for n-Zn(O,S)/p-Cu(In,Ga)Se2 Solar Cells", *IEEE Journal of photovoltaics*, vol.4, PP.2, 2014

[8]: M. A. Hussein Mustafa, M. Abdullah Manal, R. Saleh Wasan and A. Aadem Kadhim, "I-V Characteristics of ITO/CdTe/Al2O3/Si/Au Thin Film Solar Cell", *International Journal of Current Engineering and Technology*, E-ISSN 2277 – 4106, P-ISSN 2347 – 5161 ©2014 INPRESSCO® , All Rights Reserved, Available at http://inpressco.com/category/ijcet

[9]: Online available:http://www.pveducation.org/pvcdrom/solar-cell-operation/open-circuit-voltage

[10]: Online available:http://www.pveducation.org/pvcdrom/solar-cell-operation/short-circuit-current

[11]: S. Sharbati and R. Sites James, "Impact of the Band Offset for n-Zn(O,S)/p-Cu(In,Ga)Se2 Solar Cells", *IEEE Journal of photovoltaics*, vol.4, PP.2, 2014

Estimation of the Most Effective Rhythm for Human Identification using EEG Signal

Md. Mahmudul Hasan,[1,*] Md. Hanif Ali Sohag,[2] Md. Eakub Ali,[3] and Mohiuddin Ahmad[4]

[1,2,3,4]Department of Electrical and Electronic Engineering, Khulna University of Engineering & Technology, Khulna-9203, Bangladesh

[*]mahmudul.hasan.eee.kuet@gmail.com

Abstract— **Human identification using a special biological feature has become a promising field for the purpose of security system. Electroencephalogram (EEG) signals are the signature of human mind and can be used confidently as a strong biometric identifier. As EEG signals are consist of five different frequency bands, this paper represents a general methodology to determine the most effective rhythm for human identification. Using the different features from different rhythms in time and frequency domain, four neural networks are developed for the classification approach. Comparison of the designed neural networks shows that beta rhythm gives the best performance with a very low mean square error whereas delta rhythm gives the worst performance with comparative higher mean square error for identifying a person. It is concluded that beta rhythm is the most effective frequency band for human identification using EEG in resting and problem solving condition.**

Keywords— Electroencephalogram (EEG); Biometrics; Frequency band; Rhythm; Human Identification; Artificial Neural Network (ANN).

I. Introduction

Human identification is the first step of authentication and security systems. If the first is done, the latter two steps are almost done. The term 'identification' sometimes creates confusion with the other two security related terms named 'authentication' and 'authorization'. Actually identification means to detect or identify a specific person, authentication indicates to verify the claimed identity of a person, and authorization means the official approval [1]. Identification refers to identifying an individual from others or distinguishing a specific person from a group of people [2].

Human identification using a special physiological or behavioral feature of a person is known as biometrics. Electroencephalogram (EEG) based biometrics has become a promising field for research as it conveys a lot of information of human brain. It is distinct for a specific person under a specific condition and can be used as a unique biometric identifier [3-5]. Unlike other conventional human identification systems, EEG biometrics is more confident, very difficult to reproduce, and impossible to steal or get under threat. Many unexpected incidents occurred worldwide closely related with conventional biometrics like chopping of finger for desired fingerprint in Malaysia [6], use of fingerprint membrane to duplicate it [7] and just use of wood glue to hack the iPhone 5's touch id [8]. The unexpected incident of Boston marathon bombing provides strong evidence how facial detection systems can fail to identify a person from database [9].

EEG signals are composed of five different frequency bands which are known as delta, theta, alpha, beta, and gamma [10-12]. The properties of these rhythms are shown in Table I.

TABLE I. Typical Frequencies and Amplitudes of EEG Bands

Rhythm	Freq. (Hz)	Amp. (μV)	Location	Reason
Delta	1-4	20-200	Frontal lobe	Deep sleep
Theta	4-8	10	Middle, temporal	Drowsiness
Alpha	8-13	20-200	Entire lobe	Relaxation
Beta	13-30	5-10	Frontal, Occipital	Concentration and thinking
Gamma	36-44	10-12	Somatosensory cortex	Long term meditation

This paper represents a novel methodology to identify a person by using the most effective rhythm. The data acquisition was done using the BIOPAC data acquisition unit Acqknowledge® 4.1 Software [13], feature extraction was done using MATLAB 2012, and artificial neural network (ANN) was used for the classification and identification approach.

The latter part of the paper is presented as follows: Section II views the proposed methodology of determining the most effective rhythm. Section III gives the brief description of EEG signal processing and analysis steps. The feature extraction and comparison of different rhythms are described in section IV. Section V gives a focus on classification approach and experimental results. Finally the paper is ended with a conclusion and associated future works in Section VI.

II. Proposed Methodology

A. Proposed Model

The model proposed in this paper for individual identification using different frequency bands of EEG signal is shown in Fig. 1. After the EEG signal acquisition using BIOPAC and MP 36, the signal is preprocessed and separated into four different rhythms (δ, θ, α and β) by using band pass filter. Then the features are extracted using 'AcqKnowledge 4.1' software and MATLAB 2012. The extracted features for each rhythm during different trials are used to produce the input vector in the neural network classifier. The classification leads to individual identification. Comparing the accuracy and mean square error of the designed neural

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

networks, a decision is taken about what rhythm of EEG signal is best for human identification.

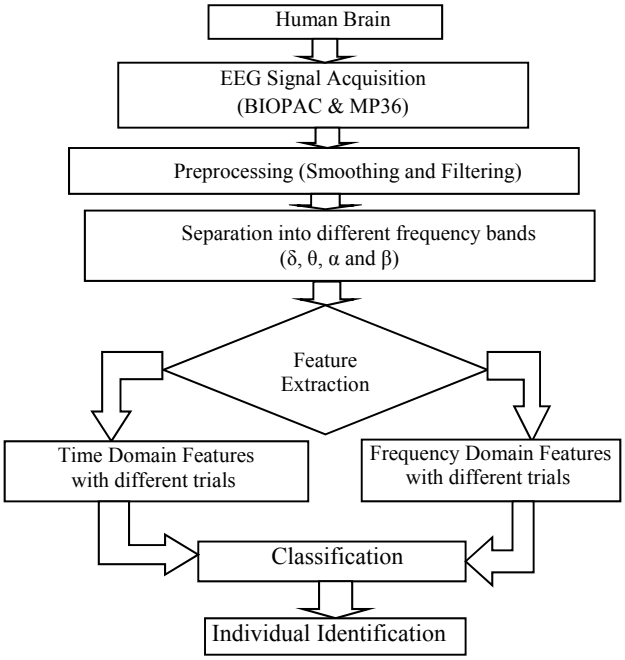


Fig 1. Block diagram of the proposed methodology

B. EEG Signal Acqusition and Experimental Conditions

The experiment for this research work was done in research was done in Biomedical Engineering Laboratory (BME), KUET, khulna-9203, Bangladesh. The experiment was done on three young healthy male subjects. All of them were healthy, controlled and were not affected by any mental disease or psychological problem. The experimental condition was eyes closed and solving some specific mathematical problem mentally. All the subjects were told to solve some specific problems by an interpreter who told the problems. Total 25 trials were taken in which each trials was independent of another and the specified problems were different. The first 20 trials were used to design the classifier and the rest 5 trials were used as test input data.

C. Experiment Setup

1) Hardware description: In this experiment we used a personal computer based data acquisition system named 'MP system'. The BIOPAC data acquisition unit MP36 and MP150 were connected to a personal computer. Fig. 2 shows the pictorial view of EEG measurement in BME Lab, KUET.

2) *Analytic software:* The software used for the analysis of the EEG signal were BIOPAC Student Lab Pro and AcqKnowledge 4.1. After the data acquisition system receives signal from elctrodes and transducers, Biopac Student Lab software displays the relative waveforms in the computer screen or monitor. 'AcqKnowledge' is a very user friendly software for biomedical signal analysis. The EEG signal can be measured, analyzed and transformed here from time to frequency domain.

Fig 2. Pictorial view of EEG measurement in BME Lab KUET

III. EEG Signal Processing And Analysis

A. Filtering, Artifact Removal and Smoothing

During the EEG signal acquisition, the raw EEG signal is collected and saved. But the raw signal contains a lot of artifact and noise. The artifacts may EMG noise due to hand movement and the line frequency which must be removed. The raw EEG signal was made free from those artifacts using different artifact removal process in MATLAB environment. As the line frequency was 50 Hz, the signal was need to pass through a low pass filter. A band pass Finite impulse Response (FIR) filter was used for filtering between 0.5 Hz to 44 Hz.

Smoothing is nothing but the moving averaging of the raw signal. The smoothing process is to calculate a mean value from some data point and to replace it with the calculated average value.

B. Separation of Frquency Band

After filtering and smoothing, the EEG signal was separated in different frequency range by AcqKnowledge software. The separated frequency bands can be distinctively selected and analyzed in time and frequency domain. The pictorial view of AcqKnowledge software for analysis of EEG frequency bands is shown in Fig. 3.

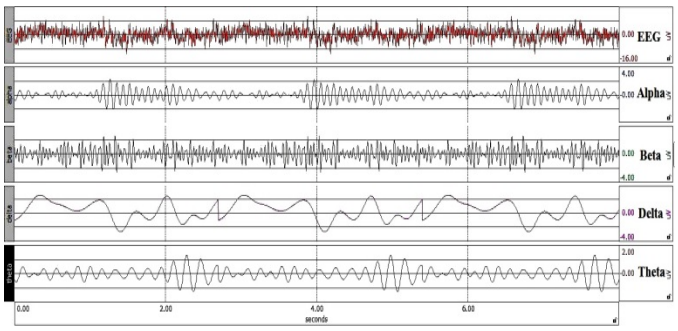


Fig 3. Separation of EEG bands in AcqKnowledge Software

C. Time Domain and Frequecy Domain Analysis

In the time domain, the features extracted from the different frequency bands are maximum value, standard deviation, skewness and kurtosis. These are statistical methods. In the frequency domain, we work in the Fourier space where the magnitude is plotted against frequency. In this research work the features extracted in the frequency domain were Fast Fourier transform (FFT) and Power Spectral Density (PSD). Power Spectral Density (PSD) shows how the average power is distributed as a function of frequency in spectral component. PSD of an EEG signal can expressed mathematically using the following Eq.(1).

$$P_{EEG}(f) = \frac{1}{T_s}\left|T_s \sum_{i=0}^{N} f_{output} e^{-i2\pi f i T_s}\right|^2 \quad (1)$$

Here, $P_{EEG}(f)$ possesses a unit of (μV2)/Hz. T_s signifies the sampling time.

IV. Experimental Data

A. Comparison of the Features of a Specific Rhythm for an Individual

In the time domain we considered four features i.e. maximum value, standard deviation, skewness and kurtosis. In the frequency domain, we considered the mean and maximum values of PSD and FFT. The features of beta rhythm which were extracted in time domain for subject 1 are shown in Table II.

TABLE II. Response of Beta Rhythm in Time Domain For Subject 1

Rhythms	TRIAL	Features in Time Domain			
		Max. value (μV)	S.D (μV)	Skew (μV)	Kurtosis (μV)
Alpha	1	7.84	1.095	-0.2129	5.535
	2	6.16	1.062	-0.0137	4.461
	3	4.72	1.048	0.0048	3.187
	4	5.47	1.049	0.0072	5.772
	5	5.15	1.054	0.0007	3.853
Beta	1	6.74	1.275	0.0575	3.635
	2	3.98	1.303	-0.0123	3.294
	3	3.59	1.332	-0.0110	2.570
	4	4.13	1.337	-0.0168	3.243
	5	4.10	1.658	-0.0123	3.261
Delta	1	10.25	3.319	-0.2583	2.264
	2	6.61	3.414	-0.4236	1.725
	3	6.62	2.555	-0.9913	3.816
	4	5.18	3.391	0.2101	2.727
	5	5.13	3.138	-0.1493	2.913
Theta	1	3.10	1.968	0.0029	3.141
	2	3.10	1.825	-0.0429	3.321
	3	2.62	1.920	-0.0345	3.153
	4	2.27	1.920	-0.0016	2.173
	5	2.86	1.720	-0.0354	2.719

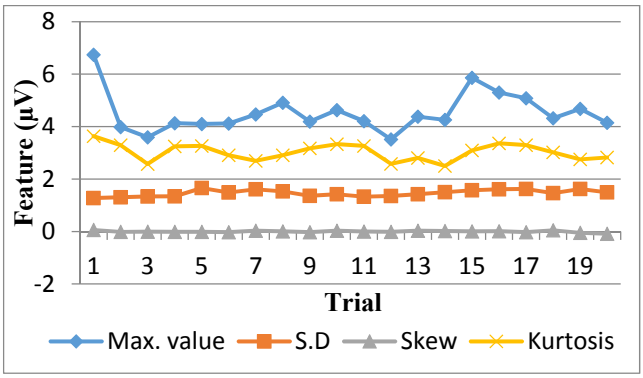


Fig 4. Intra-individual features of beta rhythm in different trials for subject 1 in time domain

Fig. 4 shows the small variation of the time domain features of beta rhythm with 20 independent trials. Similarly, the maximum and mean values of PSD and FFT also varied a little.

B. Comparison of the Features of a Specific Rhythm for Different Individuals

The experiment was done in three subjects in the same condition. If we plot the features found in time and frequency domain with respect to 20 different trials for different subjects, we find the figure like Fig. 5. Fig. 5 shows that the mean value of PSD varies within a small range for an individual but is distinct for different individuals. This large difference in changes of features with different trials can be used for identifying an individual from another.

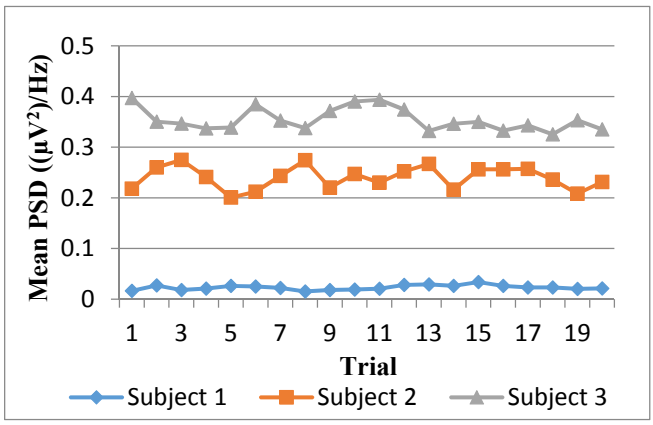


Fig 5. Comparison of inter-individual feature in different trials in frequency domain

V. Classification Results And Discussion

The classification approach is needed for the purpose of identifying an individual from others. As in this research work we dealt with four frequency bands, we found four scenarios for classification. Gamma rhythm was not considered for its less contribution in the given experimental condition. The networks were built using four features from time domain and four features from frequency domain. Feed forward back propagation algorithm was applied to build the neural networks. MATLAB build in 'nntool GUI' was used for the purpose of designing the neural networks.

A. Scenario 1:Delta Rhythm

The first network was built using the features of delta band of EEG signal. This network identified the three subjects separately with accuracy of 92.82%, 95.67% and 93.78%. Here, the mean square error was 0.1436 at epoch 2. The overall regression was 0.88067 which means that the targets and outputs are not fit enough to identify an individual.

B. Scenario 2:Theta Rhythm

The second network was designed using the features of theta band. This network was able to identify the three subjects separately with accuracy of 96.85%, 96.75% and 97.85% respectively. The mean square error was 0.1128 at epoch 4.

C. Scenario 3:Alpha Rhythm

The features extracted from time and frequency domain from Alpha band of EEG signal were used as input to this neural network. The three subjects were identified with the neural network with accuracy of 98.20%, 97.67% and 95.78% respectively. The overall regression was 0.9876 and the mean square error of this network was 0.1027 which indicate that the network is not good enough.

D. Scenario 4:Beta Rhythm

The fourth network used the features of beta band as input and here we got the accuracy of 99.87%, 100% and 100% respectively for subject 1, subject 2 and subject 3. The overall regression was 1 which indicates that the target and output values are fitted enough. The validation performance was 4.6816×10^{-7} at epoch 15 which is shown in Fig. 6. This network gives the best performance than the previous networks.

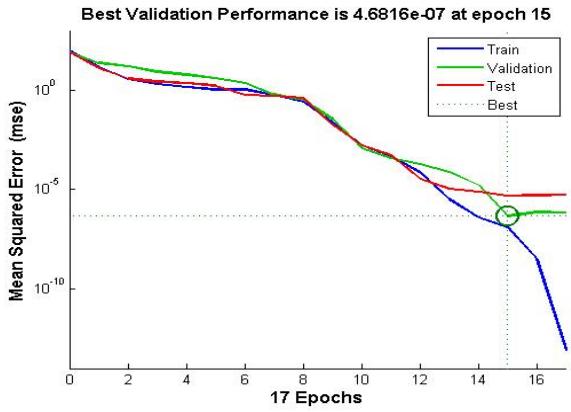


Fig 6. Mean square error for Neural Network designed with features of Beta rhythm

TABLE III. COMPARISON OF PERFORMANCE OF THE THREE DESIGNED NEURAL NETWORKS

Attributes		Delta	Theta	Alpha	Beta
Mean Square Error		0.1436	0.1128	0.1027	4.6816×10^{-7}
Regression (Overall)		0.88067	0.9076	0.9876	1
Accuracy of identification	Sub 1	92.82%	96.85%	98.20%	99.87%
	Sub 2	95.67%	96.75%	97.67%	100%
	Sub 3	93.78%	97.85%	95.78%	100%

The relative comparison of the four neural networks for four different scenarios is shown in Table III. From the comparison, it is crystal clear that the network designed using the features of beta rhythm gives best performance with least mean square error of 4.6816×10^{-7}. Here we got the overall accuracy of 100% using confusion matrix. As the research work was done with a specified condition of eyes closed and solving a specific problem, this condition is in favor of beta brainwave. Delta rhythm gives the worst performance having a mean square error of 0.1436, because in the given experimental condition Delta rhythm is not prominent.

VI. CONCLUSION

From this offline analysis of EEG signal, it is easily understood that EEG brainwave signals have special features which can be used for human identification. As EEG signal is composed of different frequency bands, any of these bands separately can be used for the identification approach and beta rhythm have a special significance in this case. Because when someone is in relaxed state and focusing on a specific problem, beta rhythm is prominent for this condition. So here we found the least mean square error of 4.6816×10^{-7} On the other hand, delta bands are found in deep sleep stage, so it contributes a little and we got a comparative high mean square error of 0.1436. With increasing the number of people, the accuracy of identification may decrease, which can be overcome by using EEG headset with more number of channels. To conclude, it can be said that beta brainwave signal can alone provide special significance for EEG biometrics which can be used for cyber security as well as security for upcoming future internet of things (IoT).

ACKNOWLEDGMENT

This work was partially supported by Higher Education Quality Enhancement Project (HEQEP), UGC, Bangladesh under Subproject “Postgraduate Research in BME”, CP#3472, KUET, Bangladesh.

REFERENCES

[1] J. Klonovs, C. Kjeldgaard Petersen, H. Olesen, A. Hammershø, “Development of a Mobile EEG-based Biometric Authentication System."IEEE Vahicular magazine, pp. 81-89, Volume: 8, Issue:1, February 2013.

[2] C. He, X. Lv, and J. Wang, “ Hashing the mar coefficients from EEG data for person authentication,” IEEE Int’l Conf. on Acoustics, Speech and Signal Processing (ICASSP), pp. 1445–1448, April 2009.

[3] A. Riera, A. Soria-Frisch, M. Caparrini, C. Grau, and G. Ruffini, “Unobtrusive biometric system based on electroencephalogram analysis,” EURASIP J. Adv. Sig. Proc., 2008.

[4] C. R. Hema, M. P. Paulraj, and H. Kaur, “Brain signatures: A modality for biometric authentication,” Int’l Conf. Electronic Design (ICED), pp. 1–4, December 2008.

[5] Q. Gui, Z. Jin, W. Xu, “Exploring EEG-based Biometrics for User Identification and Authentication,” IEEE conference on Signal Processing in Medicine and Biology Symposium (SPMB), pp. 1-6, Dec. 2014.

[6] J. Kent, “Malaysia car thieves steal finger,” BBC News, 31 March, 2005.

[7] T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino, “Impact of artificial gummy fingers on fingerprint systems,” Datenschutz und Datensicherheit, 26(8), 2002.

[8] S.Curtis, “iPhone 5s fingerprint sensor 'hacked' within days of launch,” The Telegraph, 23 Sep 2013.

[9] B. Schneier, "The Boston Marathon Bombing: Keep Calm and Carry On," The Atlantic 15 April, 2013.

[10] D. Cohen, The Secret Language of the Mind. London: Duncan Baird Publishers, 1996.

[11] M. Teplan, "Fundamentals of EEG Measurement," Measurement Science Review, vol. 2, pp. 1-11, 2002.

[12] S. Sanei, J.A. Chambers, EEG Signal Processing. John Wiley & Sons Ltd, West Sussex, 2007.

[13] Reference manuals of AcqKnowledge software guide and MP system hardware guide. (2008). BIOPAC Systems Inc., 42 Aero Camino, Goleta, CA – 0067, USA.

Millimeter Wave Generation Based on Optical Frequency Multiplication in Radio over Fiber Systems

Tahmid Aziz Chowdhury,[1,*] Md. Mahfuzul Azam Ghuri,[1] and Refat Kibria[2]
[1]Department of EEE, Shahjalal University of Science and Technology
Sylhet-3114,Bangladesh
[2]Department of CSE, Shahjalal University of Science and Technology
Sylhet-3114,Bangladesh
*tahmid.aziz@gmail.com

***Abstract*—In this paper an overview of the generation schemes for millimeter wave signals has been reported and a comparative analysis is made to decide their suitability to modern radio over fiber systems. Photonic millimeter-wave (mm-wave) generation using optical frequency multiplication (OFM) is considered as a cost-effective solution in modern radio over fiber (RoF) systems. This paper provides an insight to the advancement of this field in recent times by reviewing the pros and cons of a few of the latest OFM architectures. Simulation of an RoF link based on an optical frequency multiplication technique is performed and results show that the link exhibits satisfactory performance for a range upto 30 km single mode fiber.**

***Index Terms*—Radio over Fiber (RoF), Millimeter Wave (mm-wave) Generation, Optical Frequency Multiplication (OFM), Mach-Zehnder Modulator (MZM).**

I. Introduction

The merging of radio frequency and optical fiber technologies for the distribution of millimetre wave signal has been a topic of great interest and extensive efforts has been devoted in this matter in recent years [1]. Radio over fiber (RoF) technology has emerged, as a result of this endeavour as the key to the distribution of wireless communication signal in modern systems [2]. RoF provides significant advantages over existing cable systems because of its seemingly unlimited bandwidth and extremely low attenuation of optical fibers. Dynamic resource allocation, cost effective expandability and ability to provide independent infrastructure for multi-service operation gives RoF great potential to support secure, cost-effective, and high-capacity wireless access for the future provisioning of broadband, interactive, and multimedia wireless services.

Millimeter wave can be classified as electromagnetic spectrum that spans between 30 GHz to 300 GHz, which corresponds to wavelengths from 10 mm to 1 mm [3]. In today's world the emergence of newer and better services requiring more and more data throughput are causing the conventional microwave bands ranging from 900 MHz-20 GHz to rapidly saturate. Therefore, special attention has been given towards higher and higher segments of the frequency spectrum, in the quest for systems with increased capacity. The use of mm-wave frequencies offers large bandwidths in the wireless domain and also overcomes the problem of spectral congestion at lower frequency ranges [4].

Within the massive segment of frequency spectrum of which millimeter wave comprises, the 60 GHz frequency band has gained significant interest recently. This is mostly due to the availability of unlicensed spectrum worldwide around 60 GHz which is continuous and less restricted in terms of power usage limits [5]. Unfortunately though, the generation of millimeter wave signal at such high frequency is a huge challenge using traditional electronic devices, and hence researchers have resorted to the optical domain. A good number of techniques have been proposed over the last few years and a solution which supports the need of future RoF systems is being sought out. The recent development of optical frequency multiplication scheme has paved a way for cost efficient and simple answer to the generation of millimeter wave.

The paper is organized as follows: Section II gives a summary of the methods used for the generation of millimeter wave at 60 GHz. Section III reviews the concept of optical frequency multiplication, lists some notable schemes proposed in recent times and gives a comparative analysis among them. In section IV generation of 60 GHz millimeter wave using one of the methods is simulated and its transmission performance in an RoF link is discussed in section V.

II. Millimeter Wave Generation Techniques

In recent years, a considerable research effort has been carried out regarding the generation of low phase noise millimeter wave signal and their transport over fiber [6]. Table I lists the advantage and disadvantage of the techniques employed for

TABLE I
Comparison among various enabling techniques of generating mm wave signal

Technique	Advantage	Disadvantage
Direct Modulation	Low cost and simple construction	Frequency limitation and poor transmission integrity
External modulation	Higher frequency, greater flexibility and scalability	Larger insertion loss, requires optical filter, and high drive voltage
Optical heterodyning	No dispersion effect, high spectral purity, high freq. oscillator is not required	Complicated light source with narrow linewidth is required
Up- and Down conversion	Free from fiber dispersion effect	Requires high freq. modulator, oscillator
Optical transceiver	requires a single transciever	High RF loss, WDM incompatibility

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

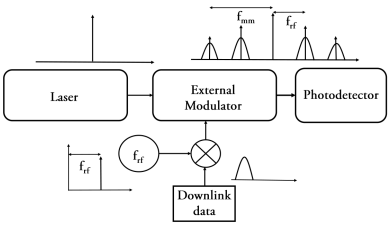


Fig. 1. Externally modulated laser source with OFM

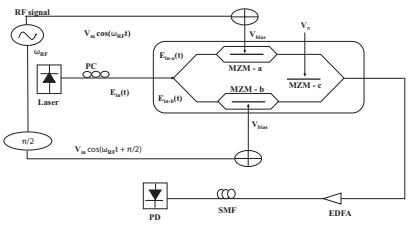


Fig. 2. Optical Frequency multiplication scheme using a single integrated MZM and no optical filtering.

generating mm-wave over the years. From the table it is clear that these methods suffers from serious impairments which limits their application in modern RoF systems. The scheme based on external intensity modulator is a relatively mature and low-cost technique because the external intensity modulator has shown high reliability and has been widely employed in the real optical transmission systems for high-speed signal modulation. Up to now, three different modulation schemes have been demonstrated, such as double sideband (DSB), single sideband (SSB), and optical carrier suppression (OCS), to generate optical mm-wave [7]. The simpler DSB modulation suffers from fading due to chromatic dispersion which makes it incompetent for long distance transmission whereas, SSB and OCS needs dual arm MZM. However, the frequency response of these modulators above 40 GHz is not satisfactory and hence it is not capable to facilitate 60 GHz RoF links. Using optical heterodyning, high mm-wave frequencies can be obtained avoiding high speed electronic and broad bandwidth optical devices. Although, it may involve high cost of setup and control. Remote up- and down- conversion at base station (BS) requires a high frequency RF oscillators and mixers at BS, which increases the system cost to a great extent. Although the simplest BS architecture can be achieved by electro absorption transceiver, the device suffers from chromatic dispersion in fiber and lack of WDM compatibility makes this ineligible for high data rate RoF systems. The method of optical frequency multiplication removes the need of high frequency RF oscillators and phase noise problems, reduces the system cost and enables generation of high frequency mm-wave with greater spectral purity.

III. Optical Frequency Multiplication

Generation of millimeter wave in RoF and its processing is done in optical domain for the sake of simplicity and reduction of cost of implementation. Optical frequency multiplication offers a cost effective solution by using low frequency RF oscillators and narrow-band electrical/optical (E/O) converters.

A. Principle of OFM

The main principle of optical frequency multiplication is to generate higher order optical harmonics in an E/O converter such as an MZM using low frequency RF signal. Many higher order harmonics are generated because of sub-carrier modulation which is caused by the non-linear response of E/O converter. By virtue of beating of any higher order harmonic pair or by beating a higher order harmonic with the carrier signal results in a photodetector results in a generated millimeter wave. The process is illustrated in Fig. 1. The downlink signal is up-converted to one of the sub harmonic (f_{rf}) of the desired millimeter wave frequency (f_{mm}) and after modulation two sidebands with spacing equal to f_{mm} are selected to beat at the photodetector. The two main approaches of OFM technique are dicussed in this paper.

B. Tehcniques Proposed

The basic approach is to use the higher order harmonics. By choosing suitable bias and modulation index, higher order sidebands are excited and multiple sidebands are generated with a spacing equal to a sub-harmonic f_{rf}. Then two harmonic spaced f_{mm} apart are optically filtered out and beating of the two sidebands results in the desired mm-wave. The multiplication factor then depends upon harmonic generation efficiency and response of the modulator. A method to generate millimeter-wave signals using an external optical modulation technique was proposed by O'Reilly et al. in 1992 [8]. A frequency-doubled electrical signal was optically generated by biasing the Mach-Zehnder modulator (MZM) to suppress even-order optical sidebands. A 36-GHz millimeter-wave signal was generated when the MZM was driven by an 18-GHz microwave signal. Such a system was employed for a remote delivery of video services. The multiplication factor can be increased by employing multiple modulators instead of a single modulator.

The second method is to use optical double sideband suppressed carrier (ODSB-SC) modulation where the carrier is suppressed by using either a notch filter or an interleaver at MZM output. In this way beating the first order sidebands results in a millimeter wave having frequency double to the local oscillator f_{rf}. In [9] a continuously tunable millimeter wave generation scheme was proposed using an external modulator and notch filter as shown in Fig. 2. The biasing of modulator was done in such way that the odd order sidebands are suppressed and only the even order sidebands are present at the notch filter output. Biasing the MZM at the null transmission point also suppress the carrier and hence can be used in this method. Again since only a fixed wavelength notch filter was needed, fiber bragg grating (FBG) could be used to filter out the carrier. In [10], a DWDM millimeter wave architecture has been proposed and experimentally demonstrated which uses phase modulator and optical interleaver. In this scheme, the bias drift problem of MZM has been avoided and subsequently stability is highly increased. Moreover, the interleaver is much less dependent on temperature compared to FBG, which improves the performance of this scheme.

However, ODSB modulation is prone to chromatic dispersion and hence suffers performance degradation after long distance transmission. A possible solution is to use optical single sideband modulation (OSSB) and incorporate OFM tecnique with it [11]. Techniques employing the OSSB modulation using dual parallel MZM has been reported. One particular scheme uses a single integrated MZM for the generation of

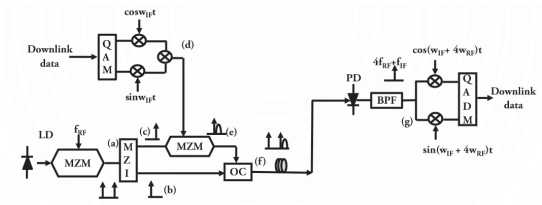


Fig. 3. ROF link for 60 GHz Millimeter-Wave system

millimeter wave and needs no narrowband filter to select the desired sideband [12]. This characteristics makes the system a suitable contender for WDM multiplexed RoF links. A conceptual diagram of the proposed system is presented in Fig. 2. An external integrated MZM that consists of three sub-MZMs where one sub-MZM (MZM-a or MZM-b) is embedded in each arm of the main modulator (MZM-c) is key to generating optical millimeter-wave signals. A two-tone light wave with frequency separation of four times the modulation frequency can be generated using only one external modulator. An example RoF link based on this scheme is discussed in the next section and analyzed to investigate the transmission performance.

TABLE II
GLOBAL PARAMETERS

Parameter	Value
Bit rate	10^9 bits/s
Sequence Length	512 bits
Samples per bit	64
Symbol rate	2.5×10^9symbols/s
Number of samples	32768
Dispersion	17 ps/nm/km
Attenuation	0.2 db/km

IV. SIMULATION EXPERIMENT SETUP

An RoF link based on the scheme proposed in [12], illustrated in Fig. 3 has been simulated in the Optisystem 12.0 software for analysis. The light source is a laser diode (LD) which has a linewidth of 10 MHz. The lightwave is emitted from the continuous laser diode with a central frequency of 193.1 THz and injected into an LN-MZM ($LiNbO_3$ Mach-Zehnder Modulator). All the global parameter associated with the measurement is listed in the Table II.

The LN-MZM is biased with 4 V and is driven by 15 GHz RF local oscillator with amplitude of 3.06. The output of the LN-MZM mainly contains two second order side bands with a frequency of 193.07 THz and 193.13 THz. The signal is passed through an MZI (Mach-Zehnder Interferometer) to separate the two sidebands. The delay between the two arms of the MZI is set at 25 ps and sidebands at 193.13 THz and 193.07 THz are separated. The sideband at 193.13 THz is SSB modulated by 16-QAM downlink signal. The downlink signal is generated by mapping a 10 GB/s pseudo-random bit sequence (PRBS) to a 16-ary vector and then modulating it by a 5 GHz signal. For SSB modulation another LN-MZM is used and a broadened signal at 193.135 THz is generated, it is 5 GHz away from the original 193.13 THz sideband. SSB modulation is responsible for generating this extra tune. The sideband at 193.07 THz remains unmodulated. The 193.13, 193.135 and 193.07 THz signals are combined by an optical coupler. Then the combined signal is transmitted to the Base

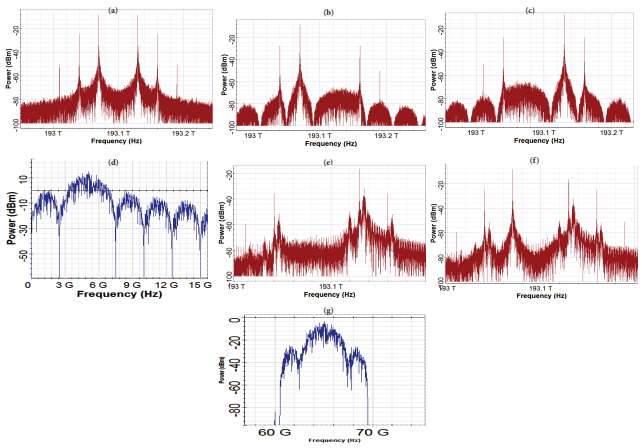


Fig. 4. Optical and RF spectra of various stages of the simulated ROF link.

Station via SMF having power attenuation of 0.2 dB/km and dispersion 17 ps/nm-km. To increase the signal strength EDFA (Erbium Doped Fiber Amplifier) with a gain of 20 dB is used.

V. RESULTS AND DISCUSSION

Simulation output of different stages of the architecture exhibited in Fig. 3 are shown in Fig. 4. From the output of the first MZM in Fig. 4(a), it is evident that the optical carrier and odd order sidebands are totally suppressed leaving the even order sidebands. Two second order sidebands are separated by an MZI shown in Fig. 4(b) and Fig. 4(c), and the difference of higher even order sideband with the second order is more than 17 db. The two second order sidebands are generated at 193.07 THz and 193.13 THz while the latter one is SSB modulated by the data signal in Fig. 4(d). Due to SSB modulation a broadened sideband is created and higher order sidebands are suppressed which is shown in Fig. 4(e). The unomodulated sideband is coupled with the modulated signal as shown in Fig. 4(f) and transmitted through single mode fiber (SMF). The transmitted optical signal is detected by a high speed photodetector, which gives the RF output. Using the fiter, the 65 GHz frequency band which is carrying the downlink data is separated as in Fig. 4(g).

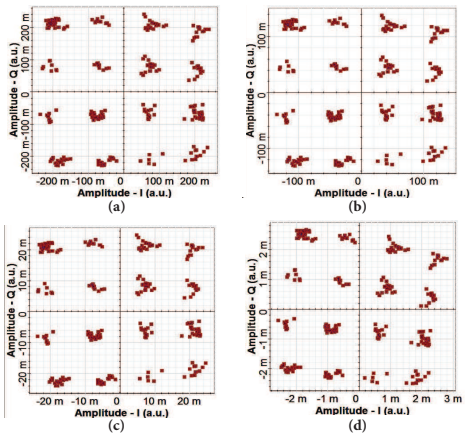


Fig. 5. Constellation diagram of downlink data for a) back to back, b) 10 km, c) 20 km, d) 30 km SMF.

The downlink data then undergoes performance analysis. The constellation diagrams for various lengths of fiber is shown in Fig. 5. For the back-to-back case in Fig. 5(a) the

points maintain sufficient distance although slight expansion occurs because of the noise in laser diode and photodetector. Its is seen that the distance between adjacent points become narrower as the length of the fiber is increased. Even after 30 km transmission each of the constellation points can be identified clearly which is shown in Fig. 5(d).

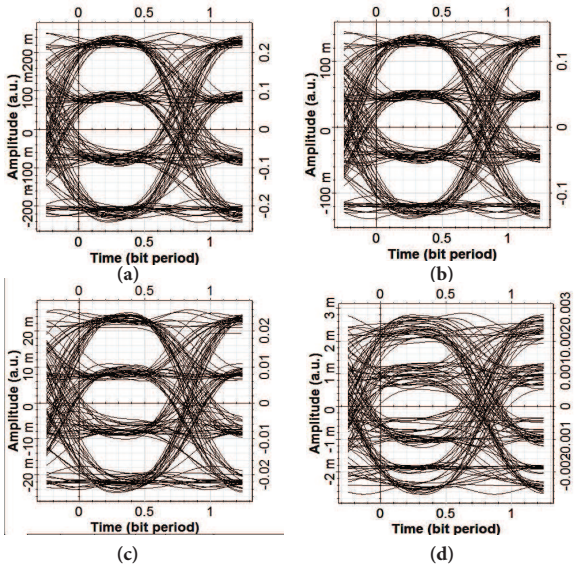


Fig. 6. Eye diagram of I-branch of downlink data for a) back to back, b) 10 km, c) 20 km, d) 30 km SMF.

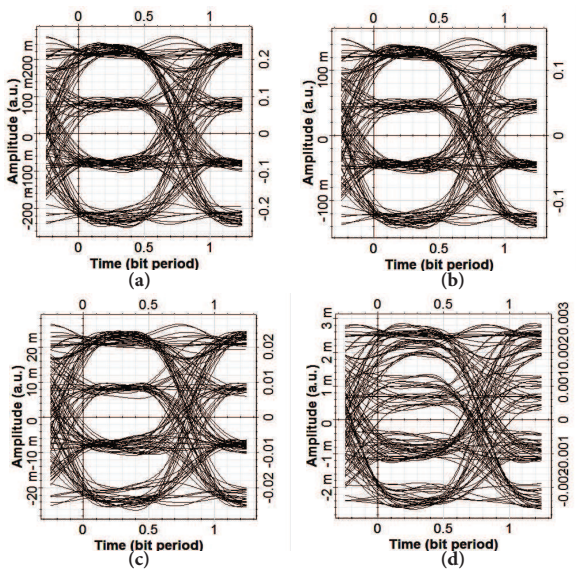


Fig. 7. Eye diagram of Q-branch of the downlink data for a) back to back, b) 10 km, c) 20 km, d) 30 km SMF.

Similarly, The eye diagrams for various lengths of fiber in Fig. 6 and Fig. 7 shows eye opening is gradually decreasing with increasing lengths of fiber. After propagating through 30 km fiber, eye is still narrowly open for both I- and Q-branch which is evident in Fig. 6(d) and Fig. 7(d) suggests that the system can be implemented for transmission of data without significant error for fiber length upto 30 km.

VI. Conclusions

Millimeter wane signal is considered as the precise solution for the upcoming RoF links and a generation technique which has high potential has been discussed in this paper. A system which can facilitate the needs of high bit rate link is simulated and analyzed. Our results show that, the system shows great integrity after being transmitted through a fiber of length up to 30 kilometers. This has brought us to a conclusion that, optical frequency multiplication combined with single sideband modulation can be used to generate mm waves of high spectral purity. However much deeper study to optimize such systems by changing various parameters must be done before considering them for practical applications. This work only simulated the downlink part whereas, for a full duplex system uplink is essential. Again, the link with a centralized source can reduce cost and complexity of the RoF systems. Hence, thorough analysis of optical frequency multiplication techniques for enabling future millimeter wave RoF architecture is strongly suggested.

Acknowledgment

This work is supported by the University Research Center of Shahjalal University of Science and Technology.

References

[1] M. Sauer, A. Kobyakov, and J. George, "Radio over fiber for picocellular network architectures," *J. Lightwave Technol.*, vol. 25, no. 11, pp. 3301–3320, Nov 2007.

[2] D. Wake, M. Webster, G. Wimpenny, K. Beacham, and L. Crawford, "Radio over fiber for mobile communications," in *Microwave Photonics, 2004. MWP'04. 2004 IEEE International Topical Meeting on*, Oct 2004, pp. 157–160.

[3] A. D. Olver, "Millimeterwave systems-past, present and future," *IEE Proceedings F - Radar and Signal Processing*, vol. 136, no. 1, pp. 35–52, Feb 1989.

[4] Z. Jia, J. Yu, G. Ellinas, and G. K. Chang, "Key enabling technologies for optical & wireless networks: Optical millimeter-wave generation, wavelength reuse, and architecture," *Journal of Lightwave Technology*, vol. 25, no. 11, pp. 3452–3471, Nov 2007.

[5] S. K. Yong and C.-C. Chong, "An overview of multigigabit wireless through millimeter wave technology: Potentials and technical challenges," *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, pp. 1–10, 2007.

[6] N. Mohamed, S. M. Idrus, and A. B. Mohammad, "Review on system architectures for the millimeter-wave generation techniques for RoF communication link," in *RF and Microwave Conference, 2008. RFM 2008. IEEE International*, 2008, Conference Proceedings, pp. 326–330.

[7] V. Sarup and A. Gupta, "A study of various trends and enabling technologies in radio over fiber (RoF) systems," *Optik - International Journal for Light and Electron Optics*, vol. 126, no. 20, pp. 2606–2611, 2015.

[8] J. O'Reilly and P. Lane, "Remote delivery of video services using mm-waves and optics," *Journal of Lightwave Technology*, vol. 12, no. 2, pp. 369–375, Feb 1994.

[9] G. Qi, J. Yao, J. Seregelyi, S. Paquet, and C. Belisle, "Generation and distribution of a wide-band continuously tunable millimeter-wave signal with an optical external modulation technique," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 10, pp. 3090–3097, Oct 2005.

[10] J. Yu, Z. Jia, L. Xu, L. Chen, T. Wang, and G.-K. Chang, "DWDM optical millimeter-wave generation for radio-over-fiber using an optical phase modulator and an optical interleaver," *IEEE Photonics Technology Letters*, vol. 18, no. 13, pp. 1418–1420, July 2006.

[11] J. Beas, G. Castanon, I. Aldaya, A. Aragon-Zavala, and G. Campuzano, "Millimeter-wave frequency radio over fiber systems: A survey," *Communications Surveys & Tutorials, IEEE*, vol. 15, no. 4, pp. 1593–1619, 2013.

[12] C. T. Lin, P. T. Shih, J. Chen, W. Q. Xue, P. C. Peng, and S. Chi, "Optical millimeter-wave signal generation using frequency quadrupling technique and no optical filtering," *IEEE Photonics Technology Letters*, vol. 20, no. 12, pp. 1027–1029, June 2008.

Detection of Torsional Oscillations in Line-Start IPM Motor Drives Using Motor Current Signature Analysis

S. F. Rabbi, Student Member, IEEE and M. A. Rahman, Life Fellow, IEEE
Department of Electrical and Computer Engineering
Memorial University of Newfoundland, St. John's, NL, Canada, A1B 3X5
sfr560@mun.ca, arahman@mun.ca

Abstract—**Line-start interior permanent magnet (LSIPM) motors are vulnerable to limit cycles and often experience severe torsional oscillations. This paper presents a technique for detection of torsional oscillations in LSIPM motors based on stator current signatures. Rotor speed variations associated with torsional oscillations in a LSIPM motor influences the electrical supply, and introduces variable amplitude lower and upper sidebands in the stator current. In this paper, frequency domain analysis of the nonstationary current signals is carried out for detection of the sideband frequency components superimposed on the fundamental. The proposed technique is validated by performing finite element analysis (FEA) of a 3-phase 4-pole 1HP LSIPM motor drive. Experimental investigations have been carried out for the same motor drive in order to validate the performance of the proposed method under various practical operating conditions. Based on FEA and experimental results, the proposed technique can successfully detect the onset of torsional oscillations in the drive system without any vibration sensor.**

Index Terms—**Limit cycle, permanent magnet motors, torsional oscillations, Fourier transform, simulation.**

I. Introduction

A line-start interior permanent magnet (LSIPM) motor is an induction start synchronous motor. It is a type of IPM motor that can be operated either from a fixed frequency direct power supply or by a variable speed drive (VSD) system. LSIPM motors are subjected to inherent torsional oscillations [1]-[3]. Also, applications like reciprocating pumps and compressors create periodic load torque variations, inducing torsional oscillations in the system. Detection of the level of oscillation is crucial in order to avoid failures in the mechanical transmission system, and is traditionally carried out using external vibration sensors such as accelerometers/encoders mounter on the shaft of the motor [3]. Vibration sensors have a short lifetime and the cost for the maintenance of the vibration monitoring system is generally high. Sustained torsional oscillation introduces severe stress on the mechanical drive system. The onset of such oscillations can be viewed as a mechanical failure of the motor drive and the drive system needs to be protected against any potential occurrence of sustained torsional oscillations.

Diagnosis of faults in ac motor drives has been an issue of interest for several decades. A popular method of fault diagnosis is the motor current signature analysis (MCSA) [3]-[10]. MCSA typically involves the analysis of stator current signals to detect various internal faults such as symmetrical and unsymmetrical faults, magnet demagnetization, insulation failures, overheating, winding turn-turn faults, etc. [3]-[5]. In recent years, MCSA has been used to detect mechanical failures in permanent magnet motor drive systems. These failures include bearing damage, gear-box faults, magnet defects, faults due to air-gap eccentricity and mechanical unbalance of the rotor [5-9]. Recent applications of MCSA based techniques involve time-frequency analysis of the stator current signal using various signal processing techniques [3-9]. The electrical signatures of different fault conditions are established through comparison of the stator current between healthy operation and the appropriate fault condition. The majority of MCSA-based fault diagnostics focuses on closed loop feedback controlled variable speed permanent magnet motor drives. MCSA-based fault diagnostics with a focus on the detection of torsional oscillations in LSIPM motor drives has not been conducted. As LSIPM motors are becoming more popular, there exists a need to develop a technique for the diagnosis of torsional oscillations in LSIPM motors for protection against unwarranted electromechanical failures [8].

This paper presents a procedure for detection of torsional oscillations in LSIPM motors based on stator current signature analysis. Torsional oscillations result in electromagnetic field variations inside the motor, and thus modulate the stator currents. Frequency-domain analysis of the current signal reveals the induced modulation, expressed as variable amplitude lower and upper sidebands superimposed on the fundamental supply frequency. In this paper, the short-time Fourier Transform (STFT) is applied for analysis of the nonstationary current signals in order to identify the induced sideband frequency components. The proposed method is validated by carrying out finite element simulation for a 3-phase 4-pole 208V, 1HP LSIPM motor drive. Experimental investigations have been carried out for the same motor drive in order to validate the performance of the proposed method under nonideal operating conditions. Both simulation and experimental results are presented and analysed in this paper.

II. Analysis of Torsional Oscillations in IPM Motors

The equation of motion of an IPM motor drive can be written as,

$$J\frac{d^2\delta}{dt^2} = T_s(\delta) - T_d\frac{d\delta}{dt} - T_L \tag{1}$$

where J is the inertia of the motor, T_s is the synchronous torque of the motor, T_d is the damping torque coefficient, T_L is the load torque for the motor and δ is load angle.

The synchronous torque of an IPM motor can be represented as the summation of the magnet torque and reluctance torque components. The equation of the synchronous torque is given below [1]:

$$T_s(\delta) = \frac{3pE_0V}{2\omega_s X_{sd}}sin\delta + \frac{3pV^2\left(X_{sd} - X_{sq}\right)}{4\omega_s X_{sq}X_{sd}}sin2\delta \tag{2}$$

where V is the supply voltage per phase, E_0 is the motor back-emf, p is the number of pole pairs, X_{sd}, X_{sq} are d-q axis inductances, respectively, ω_s is the synchronous angular frequency and δ is electrical load angle.

Fig. 1 shows the synchronous torque for an IPM motor. The maximum or pull out torque occurs at the load angle δ_{max} which is between an electrical angle of 90^0 and 180^0. The load torque is assumed constant, represented by a straight-line in the figure. When such a fixed load is applied to the motor, the points of intersection between the load line and the synchronous torque curve define the operating electrical load angle. The load angle δ_s' is the unstable load point where the net motoring torque and the load torque are balanced. The angle δ_s represents the stable load point sought during motor synchronization. If synchronization is successful following stabilization due to inertia and damping, the operating load angle will quickly converge to δ_s. Successive variations in load conditions or in the power supply results in a periodic change in the load angle. Minor oscillations in the rotor speed are eventually damped out by the cage along with the shaft stiffness and viscous damping. In this event, the IPM motor operates within a practically stable limit cycle with the load angle exhibiting low amplitude periodic variation. In the event of large oscillations in the load torque, the cage becomes unable to damp out the oscillations and the load angle δ changes successively over a wide range. This creates large pulsations in the developed electromagnetic torque. As a result, the whole electromechanical system starts swinging as the rotor speed periodically oscillates in order to remain synchronized with the field. At this point, the motor operates in an unstable limit cycle, exhibiting high amplitude torsional oscillations. Fig. 2a shows the trajectories of the load angle δ for stable, practically stable and unstable operations of the LSIPM motor. The sign of δ is neglected. Fig. 2b illustrates the stable, practically stable and unstable limit cycles for a LSIPM motor. For unstable limit cycles, a sustained torsional oscillation takes place which creates significant mechanical vibrations and may lead to serious damage of the electro-mechanical assembly.

III. Stator Current Signature for Detection of Torsional Oscillations

The rotor speed due to torsional oscillations can be expressed as:

$$\omega_r(t) = \omega_{r0} + \sum_{j=1}^{m} \Delta\omega_{rj} \sin(\omega_j t + \xi_j) \qquad (3)$$

where ω_r is the instantaneous rotor speed, ω_{r0} is the steady state synchronous speed, $\Delta\omega_{rj}$ is the amplitude of the jth oscillation, ξ_j is the phase shift from the torque oscillation, ω_j is the frequency of the jth oscillation and t is the time.

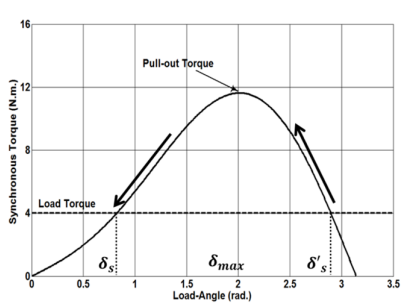


Fig. 1. Synchronous torque of an IPM motor.

The rotor angle during the oscillation can be determined from the following equation:

$$\theta_r(t) = p\int_0^t \omega_r(\tau)\, d\tau$$
$$= \theta_{r0} + \sum_{j=1}^{m} \Delta\theta_{rj} \sin(\omega_j t + \gamma_j) \qquad (4)$$

where θ_r is the instantaneous rotor electrical angle, θ_{r0} is the steady state rotor angle, $\Delta\theta_{rj}$ is the amplitude of the jth oscillation, γ_j is the phase shift and p is the no. of pole pairs. During normal behavior of the motor, the rotor lags behind the stator field and can be expressed as:

$$\theta_r(t) = \theta_{r0} = p\omega_{r0}t - \delta \qquad (5)$$

Fig. 3 illustrates a simplified magnetic equivalent circuit of an IPM motor. The total flux in the air-gap is a combination of the flux generated due to the magnets in the rotor and the flux generated by a rotating magnetic field due to 3-phase stator currents. The total air-gap flux can be stated by the following equation [6]:

$$\phi_g(t) = \sum_{h=1}^{n} \left[\phi_{sh}\cos(\omega_{sh}t + \lambda_{sh}) + \phi_{rh}cos(\omega_{sh}t + \psi(t))\right] \qquad (6)$$

where ϕ_g is the air-gap flux, ϕ_{sh} is the peak amplitude of the rotating magnetic flux, ϕ_{rh} is the peak amplitude of the flux due to the permanent magnets, λ_{sh} is the phase shifts, $\omega_{sh} =$

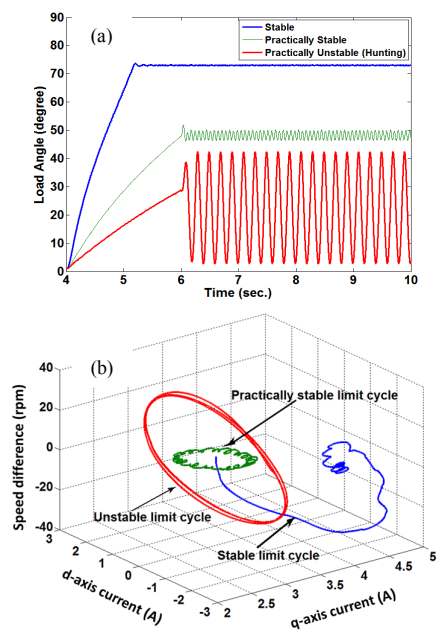


Fig. 2. (a) Load angle trajectories during limit cycles and (b) limit cycles in a line start IPM motor.

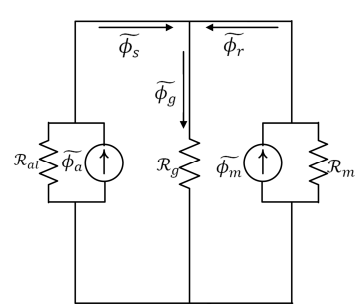


Fig. 3. Simplified magnetic equivalent circuit of an IPM motor

$p\omega_{r0} = 2\pi f_{sh}$, h indicates the number of harmonics (1, 3, 5, ...) and $\psi(t) = \sum_{j=1}^{m} \Delta\theta_{rj} \sin(\omega_j t + \gamma_j)$.

The back electromotive force (EMF) due to the air-gap flux induced in the stator coils is calculated from the following equations:

$$E(t) = \frac{d}{dt}\left(\phi_g(t)\right)$$
$$= -\sum_{h=1}^{n}[\omega_{sh}\phi_{sh}\sin(\omega_{sh}t + \lambda_{sh})]$$
$$-\sum_{h=1}^{n}\left[\omega_{sh}\phi_{rh}\sin\left(\omega_{sh}t + \psi(t)\right)\right]$$
$$-\sum_{h=1}^{n}\left[\phi_{rh}\sin\left(\omega_{sh}t + \psi(t)\right)\left\{\sum_{j=1}^{m}\Delta\theta_{rj}\omega_j\cos\left(\omega_j t + \gamma_j\right)\right\}\right] \quad (7)$$

Equation (7) can be simplified as:

$$E(t) = -\sum_{h=1}^{n}[V_{sh}\sin(\omega_{sh}t + \lambda_{sh})]$$
$$-\sum_{h=1}^{n}\left[V_{rh}\sin\left(\omega_{sh}t + \psi(t)\right)\right]$$
$$-\sum_{h=1}^{n}\frac{\phi_{rh}}{2}\left[\sum_{j=1}^{m}\beta_{rj}\left[\sin\{(\omega_{sh} \pm \omega_j)t + \psi(t) \pm \gamma_j\}\right]\right] \quad (8)$$

where,

$$\beta_{rj} = \Delta\theta_{rj}\omega_j \quad (9)$$

Neglecting the stator phase resistance, the phase current is directly proportional to the induced voltage in the stator windings. The stator phase current I_a can be written as:

$$I_a(t) = \sum_{h=1}^{n}[I_{sh}\sin(\omega_{sh}t + \sigma_{sh})]$$
$$+\sum_{h=1}^{n}\left[I_{rh}\sin\left(\omega_{sh}t + \psi(t)\right)\right]$$
$$+\sum_{h=1}^{n}\frac{\phi_{rh}}{2}\left[\sum_{j=1}^{m}\beta_{r1j}\left[\sin\{(\omega_{sh} \pm \omega_j)t + \psi(t) \pm \gamma_j\}\right]\right] \quad (10)$$

where σ_{sh} is the phase shift and β_{r1j} is a fraction of the amplitude of the modulation current depending on the magnitude and frequency of the rotor oscillation.

It is evident from (10) that the stator current is both amplitude and phase modulated during rotor oscillations, becoming a non-stationary signal with time varying statistical properties. The Short-time Fourier transform (STFT) is an extension of the Fourier Transform, allowing time-frequency analysis of nonstationary signals. The STFT of a discrete stator current signal I[n] can be expressed as follows:

$$STFT_I[n,k] = I_a[n,k] = \sum_{m=n}^{n+(N-1)} I_a[m]\, e^{-j\frac{2\pi k}{N}m} \quad (11)$$

where n is the time and k is the frequency of the discrete signal (k =0, 1, 2, …N-1), N is the size of the window and m is the index.

Frequency domain analysis of the stator current signal using the STFT shows that the modulated current signal during oscillations contains lower and upper sidebands around the fundamental frequency and its multiples. The frequencies of the sidebands can be expressed as:

$$[\{f_1^1, f_1^2, \dots\}, \{f_3^1, f_3^2, \dots\}, \dots] = \sum_{h=1}^{n}\left[f_{sh} \pm \sum_{j=1}^{m} f_j\right] \quad (12)$$

The sideband components around the fundamental frequency will have the following frequencies,

$$[\{f_s^l\}, \{f_s^2\}] = \left[\left\{f_s - \sum_{j=1}^{m} f_j\right\}, \left\{f_s + \sum_{j=1}^{m} f_j\right\}\right] \quad (13)$$

where f_s^1 and f_s^2 are the set of lower and upper sideband frequencies, respectively, f_j is the frequency of the j_{th} oscillation and f_s is the supply frequency.

IV. Results & Discussions

A 3-phase 4-pole 208V 1HP line start IPM motor model is studied using both finite element simulation and experimental investigation. ANSYS Maxwell software is used for time stepping finite element (FE) analysis of the motor. The motor is run directly from a Y-connected 3-phase, 208V balanced fixed frequency 60Hz ac supply. Torsional oscillation is introduced in the motor at t=3s by periodic variations of the load torque and is cleared after t=8s. The rotor speed and rotor angle of the motor before and during torsional oscillation are presented in Figs. 4a and 4b, respectively. The phase current waveform is depicted in Fig. 4c. The modulation in the current signal starts instantaneously when the rotor oscillates and the level of modulation depends on the frequency and amplitude of the oscillations. The d-q axis current waveforms are illustrated in Figs. 4d and 4e, respectively. The Fourier spectrum of the current signal during hunting is illustrated in Fig. 5a. The frequencies of the rotor oscillation are 5Hz, 20Hz and 28Hz. As a result, symmetric sidebands appear in the current spectrum along with the fundamental. The frequencies of the lower sidebands are 55Hz, 40Hz and 32Hz, respectively. The upper sidebands contain frequencies of 65Hz, 80Hz and 88Hz, respectively. The frequency spectra of d-q axis currents are shown in Fig. 5b and 5c, respectively. The frequencies of

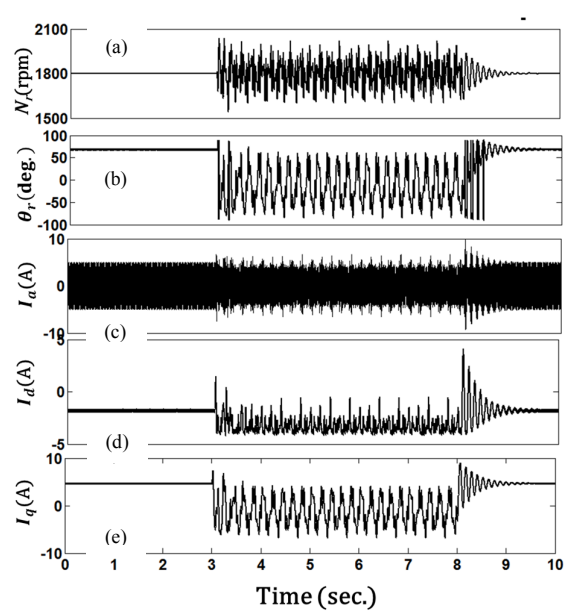


Fig. 4. Simulation results: (a) speed, (b) rotor angle, (c) phase current, (d) d-axis current and (e) q-axis current.

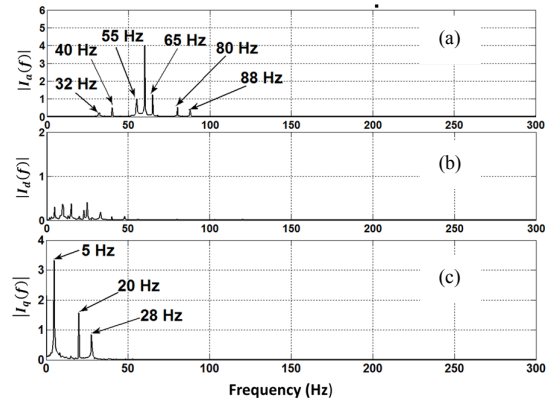


Fig. 5. Frequency spectra: (a) phase current (b) *d*-axis current and (c) *q*-axis current.

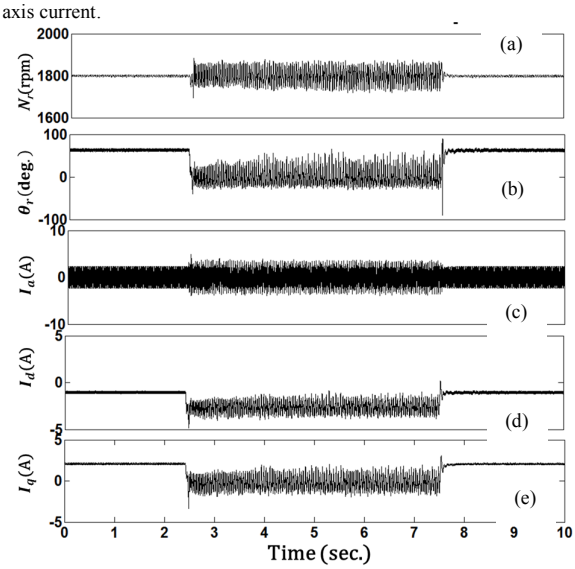


Fig. 6. Experimental results: (a) speed, (b) rotor angle, (c) phase current, (d) *d*-axis current and (e) *q*-axis current.

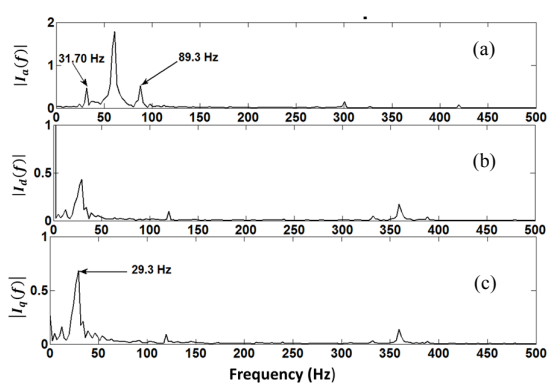


Fig. 7. Frequency spectra: (a) phase current (b) *d*-axis current and (c) *q*-axis current.

the rotor oscillation are present in the *q*-axis current spectrum. The *d*-axis current spectrum exhibits the phase modulation due to the torsional oscillations. Consequently, both the phase and amplitude modulation effects can be detected by analyzing the spectrum of *d*-*q* axis currents in a LSIPM motor.

The experimental speed, rotor angle and phase current waveforms during torsional oscillations for 60Hz supply frequency are illustrated in Figs. 6a-6c, respectively. An ac dynamometer is used as a loading machine, allowing the inducement of torsional oscillations in the system. The rotor suffers from sustained oscillations around the synchronous speed. This oscillatory behavior influences the stator current, with modulation depicted in the current signal. The *d*-*q* axis currents are presented in Figs. 6d and 6e, respectively. The modulation in the current signal is picked up by the *d*-*q* axis currents. However, the magnitude of the stator current remains below the motor rating. In addition, the phase currents remain balanced. As a result, a traditional overcurrent relay or a differential motor protection relay will not pick up the effects of torsional oscillations.The frequency spectra of the phase and *d*-*q* axis current waveforms are presented in Figs. 7a-7c, respectively. The primary frequency of the rotor oscillation is 29.3 Hz and it is detected by the *q*-axis current spectrum. The induced sidebands have frequencies of 31.70Hz and 89.3Hz as displayed in Fig. 7a.

V. Conclusions

This paper introduces a current signature analysis technique for detection of torsional oscillations in LSIPM motors. The efficacy of the proposed method is validated by performing finite element simulations and experimental investigations on a 1HP LSIPM motor drive. Based on the simulation and experimental results, the developed method can detect any rotor oscillation without the need of a vibration sensor. It is fast and non-intrusive. It is suitable to be a part of the protection system for LSIPM motors drive systems.

VI. Acknowledgement

This authors wish to thank Mr. M. Little for his assistance in this works.

References

[1] S. F. Rabbi and M. A. Rahman, "Critical criteria for successful synchronization of line start IPM motors", *IEEE Journal of Emerging and Selected Topics in Power Electronics*, Vol. 2, No.2, pp. 348-358, June 2014.

[2] Z. Li, J. B. Park, Y. H. Joo, B. Zhang, and G. Chen, "Bifurcations and chaos in a permanent-magnet synchronous motor", *IEEE Transactions on Circuits and Systems-I: Fundamental and Applications,* Vol. 49, No. 3, pp. 383-387, March 2002.

[3] W. Wang and O. A. Jianu, "A smart sensing unit for vibration measurement and monitoring", *IEEE/ASME Transactions on Mechatronics*, Vol. 15, No. 1, pp. 70-78, February 2010

[4] S. A. Saleh, R. Ahshan and M. A. Rahman, "Performance evaluation of an embedded *d–q* WPT-based Digital protection for IPMSM Drives", *IEEE Transactions on Industry Applications*, vol. 50, no.3, pp. 2277 - 2291, May 2014.

[5] M. A. S. K. Khan and M. A. Rahman, "Development and implementation of a novel fault diagnostic and protection technique for IPM motor drives", *IEEE Transactions on Industrial Electronics*, Vol. 56, No. 1, pp. 85–92, September 2008.

[6] P. Ostojic, A. Banerjee, D. C. Patel, W. Basu and S. Ali, "Advanced Motor Monitoring and Diagnostics", *IEEE Transactions on Industry Applications*, vol. 50, no.5, pp. 3120 -3127, September 2014.

[7] W. G. Zanardelli, E. G. Strangas and S. Aviyente, "Identification of Intermittent Electrical and Mechanical Faults in Permanent-Magnet AC Drives Based on Time–Frequency Analysis", *IEEE Transactions on Industry Applications*, Vol. 43, No. 4, pp. 971-980, July/August 2007.

[8] H. Kim, "On-line mechanical unbalance estimation for permanent magnet synchronous machine drives", *IET Electric Power Applications*, vol. 3, no. 3, pp. 178–186, May 2009.

[9] F. J. T. E. Ferreira, B. Leprettre and A. T. Almeida, "Comparison of protection requirements in IE2-, IE3-, and IE4-class motors" *IEEE Transactions on Industry Applications*, March 2016.

[10] H. Kim, S. B. Lee, S. Park, S. H. Kia and G. A. Capolino, "Reliable detection of rotor faults under the influence of low-frequency load torque oscillations for applications with speed reduction couplings", *IEEE Transactions on Industry Applications*, vol. 52, no.1, pp. 1460 - 1468, March 2016

A TDPSOLA Based Concatenation Technique for Bengali Text to Speech Synthesis System Subachan

Kamrunnahar Swarna,[1,*] Abu Naser[1]
[1]Department of Computer Science and Engineering, Shahjalal University of Science and Technology
Sylhet-3114, Bangladesh
[*]swarna.sustcse@gmail.com

Abstract—**Creating an intelligible as well as a natural text to speech synthesizer has been the ultimate goal of researchers for the past 30 years; and concatenative synthesis provides the most natural speech. It is usual to have distortions in the concatenation points in concatenative speech synthesis, and therefore generating audible clicks in the synthesized speech. To solve this problem, several signal processing concatenation algorithms exist, such as TDPSOLA, FDPSOLA, MBROLA etc. This paper addresses the problem of audible discontinuities at the concatenation points of diphones in Bengali speech synthesizer Subachan and solving it using TDPSOLA. In the process of doing this, we detected correct pitch mark locations of diphones, detected voiced and unvoiced speech frames of diphones and finally concatenated those diphones using TDPSOLA after rescaling them to remove energy mismatches. As a result, the audible clicks in the concatenation points are removed and speech with much better quality is generated.**

Index Terms **– Bengali text to speech, speech processing, concatenation point smoothing, pitch marking, voiced and unvoiced detection, TDPSOLA.**

I. Introduction

A usable natural text to speech synthesizer is necessary considering its various uses in our lives, e.g. reading ebooks, being used by speech impaired people to talk to others etc. Long term usage includes man machine communication i.e. when robots will be able to talk to humans, they will be able to speak in Bengali [1]. During the past decade several approaches were taken to create a usable and good quality speech synthesizer for Bengali Language. One of them is Bangla Vaani which is a Kolkata based Bengali text to speech synthesizer in which partnemes [2] were used as speech units. Katha is another Bengali speech synthesizer which is developed using festival [3]. SUST produced a more efficient diphone based synthesizer named Subachan [4]. The use of diphones as speech units and minimizing the number of diphones significantly reduced memory requirements.

In the current architecture of Subachan, diphones are selected based on phonetic algorithm and they are concatenated one after another without applying any signal processing concatenation algorithms. Therefore, frequent audible clicks generated in the concatenation points of the diphones remain in the synthesized speech. Fig. 1 shows concatenation of two diphones without applying any signal processing algorithms.

The speech resulted from this concatenation has clicks in the concatenation point.

This was solved by applying TDPSOLA [5][6][7]. TDPSOLA is a highly efficient diphone concatenation algorithm which requires only 7 computations per sample [1]. We chose TDPSOLA because of its simplicity and efficiency.

II. Signal processing among Bengali synthesizers

There were several approaches for making text to speech synthesizers in Bengali language, but most of the approaches were focused on natural language processing module.

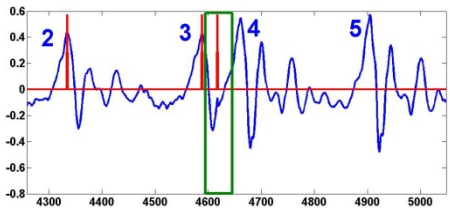


Fig. 1 Concatenation point of diphones 'রা' and 'আশ' in the current system. The area inside the rectangle shows the concatenation point of two diphones. We clearly see the phase mismatch in the concatenation point which produces audible click.

Since partnemes were used as speech units in Bangla Vaani synthesizer, a signal processing method named ESNOLA was used for concatenation.

The development of Katha was mainly focused on NLP module, and the concatenation was automatically handled by festival [3].

All of the previous works in Subachan were focused on NLP module; no signal processing methods were applied [4].

It is apparent that no previous approaches were made for applying TDPSOLA among Bengali speech synthesizers.

III. Methodology

Applying TDPSOLA in the concatenation point and rescaling diphones according to their energies remove audible clicks from the concatenation point. In doing so, first we find pitch marks in the diphones, then detect voiced unvoiced speech frames to decide whether to rescale those diphones or not. Finally, the diphones are rescaled if necessary and then concatenated by overlap adding.

A. *Finding Pitch Marks*

Pitch marking is the crucial part of a successful implementation of TDPSOLA [8]. Failure to mark pitch locations accurately produces distorted synthesized speech.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

The pitch marking method in [8] was chosen for this concatenation algorithm. The pitch marking algorithm places pitch marks in either positive or negative peak positions, the ones with the highest amplitude values. This property creates problem when the pitch marking algorithm places pitch marks in different types of peaks (positive/negative) in adjacent diphones. Concatenation in this case results in poor synthesis.
For example, wrong pitch mark locations of diphone 'ওম' are shown in fig. 2. Red vertical straight lines are used throughout the paper to indicate pitch mark locations.

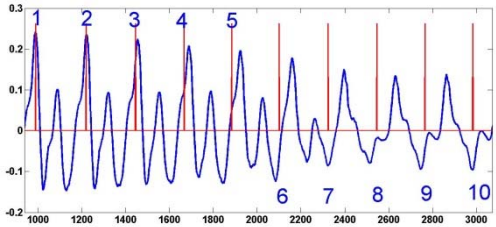


Fig. 2 Pitch marks in their wrong positions for diphone 'ওম'. Red vertical straight lines indicate pitch marks.

In the first block of this pitch marking algorithm, the signal is squared to compute energy curve for further calculations, thus converting negative peaks to positive. In the equation $E(n) = hanning\ window * s(n)^2$, we used original signal $s(n)$ instead of squaring it.
In the second block of the algorithm, absolute values of the speech signal were taken for processing; this also converts negative peaks to positive. In the equation $|s(c(k) + [-p_{max}, \ldots, p_{max}])| \times hanning\ window$, we used $s\big(c(k) + [-p_{max}, \ldots, p_{max}]\big)$ instead of absolute values.

After those modifications, the algorithm returns our intended result i.e. it places pitch marks only in the positive peak positions. The correct pitch mark locations are shown in Fig. 3 of the diphone 'ওম' resulting from the modifications.

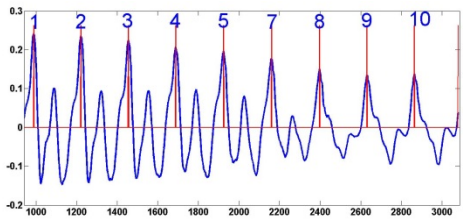


Fig. 3 Pitch marks in their correct positions in diphone 'ওম' after changing the algorithm. The pitch mark locations or peaks are numbered so that beginners can understand.

The pitch marking algorithm has been efficiently implemented in MATLAB and the paper along with the implementation were available from the authors via email.

B. *Voiced and Unvoiced Detection*

We need to detect if the frames in the boundaries of diphones in concatenation point are voiced or unvoiced for performing rescaling. First we used zero crossings and short time energy based method [9] to detect voiced and unvoiced frames in the diphone. It does not return expected results in certain diphones, especially in VV types or in the subtypes of VC, CV diphones where C is voiced. Finding short-time energy and zero crossing rates unreliable measures to decide voiced and unvoiced frames, we decided to use periodicity and short time energy as the features to differentiate between voiced and unvoiced speech frames. Periodicity was selected as a feature, exploiting the facts that voiced speeches are quasi periodic and unvoiced speeches are non-periodic. We found an autocorrelation pitch i.e. periodicity detector online [10] and used it in the concatenation algorithm. It is implemented in MATLAB and is shared by the author of [11].
In this pitch detector, if a signal frame is periodic i.e. if it has pitch; it is assigned a non zero positive pitch period value, otherwise zero. We mark the frames which have non zero positive pitch period values as voiced and others as unvoiced. Though it returns correct results for the most of the diphones, the ones which do not include any voiced frames are incorrectly assigned non zero positive pitch period values.
But our observation revealed that diphones with all unvoiced frames have negative log magnitude values for all of the frames. Therefore, we mark all of the frames of that diphone as unvoiced. After that, we use this voiced or unvoiced detection method in the concatenation algorithm.

C. *Concatenation*

For concatenation, at first pitch marks of both diphones are detected and stored in arrays, let's say $pm_1(m)$ and $pm_2(n)$ respectively. The pitch marking algorithm automatically inserts pitch marks in the first and last locations of each diphone, so we will ignore the first and last pitch mark of a diphone while processing. Then the diphone which is in the right side of a concatenation point is rescaled [7] if needed, followed by overlap add in the concatenation point.

1) Rescaling: Rescaling a speech signal means increasing or decreasing its energy i.e. adjusting its loudness. It is needed when two adjacent diphones have significant energy differences in their concatenation point. The energy mismatch creates click in the concatenation point.
When both diphones have voiced frames in the concatenation point, we take a sample; say s_1 from the left diphone's last peak position and another sample s_2 from the right diphone's first peak position. Let us denote the ratio of s_1 and s_2 as r.
When $0.2 \leq r \leq 1$, we multiply $d_2(n)$ i.e the right diphone by $0.9 \times r$, so that it has only 90% of energy than that of the left diphone, $d_1(n)$. Ratio is chosen to be greater than 0.2 because rescaling with a ratio lower than this leads to unintelligible speech.

$$d_2(n) = d_2(n) \times 0.9 \times r$$

The decrease of energy of the right diphone than the left one causes smooth transition between diphones and makes the speech output more pleasing to hear. After rescaling, pitch

marks are detected in the rescaled diphone and $pm_2(n)$ is replaced by the newly found pitch marks. Fig. 4 and fig. 5 show the concatenation point before and after rescaling.

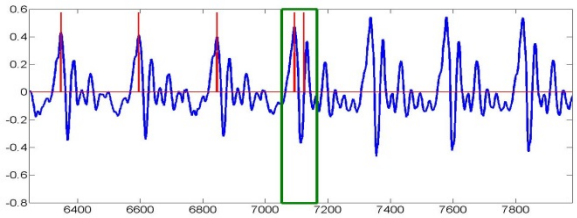


Fig. 4 Overlap added signal without rescaling. The concatenation point is shown using the rectangle. The increase in amplitudes is clearly noticeable as the signal progresses to the right from the concatenation point. It causes clicks in the concatenation point.

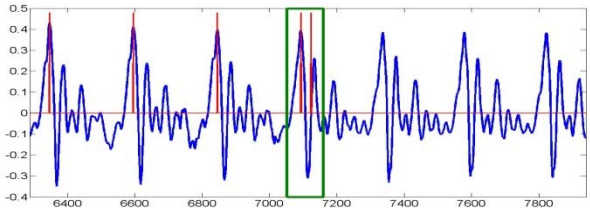


Fig. 5 After rescaling, we see the decreasing amplitudes from the concatenation point which removes the audible click in the concatenation point.

2) Overlap Add: Finally speech is synthesized by overlap adding the rescaled diphones. Frames in the diphone boundaries are overlap added so that the last peak of $d_1(m)$ and the first peak of $d_2(n)$ are aligned. The algorithm is described below:

$$output \rightarrow initially\ empty$$

Let's say that m and n indicate the lengths or the number of points of diphones $d_1(m)$ and $d_2(n)$ respectively. pt is the number of points starting from the last peak location of $d_1(m)$ to the end of $d_1(m)$.

The length of the frames which will be overlap added is indicated by j and is the number of points starting from the first point of $d_2(n)$ to the point $pm_2(2) + pt$ i. e. $j = (pm_2(2) + pt)$. Here $pm_2(2)$ is the first peak location of $d_2(n)$. Equation (1) shows the frame $mid(k)$ which is created by overlap adding the excised frames from both diphones.

$$mid(k) = (d_1(x) + d_2(y)) \div 2 \qquad (1)$$

where,

$$k = 1, 2, \dots, j$$
$$x = (m - j) + 1, (m - j) + 2, \dots, m$$

and

$$y = 1, 2, \dots, j$$

$d_1(x)$ is shown by the rectangle which includes peak 3 and $d_2(y)$ is shown by the rectangle which includes peak 4 in fig. 6.

Then the speech waveform is synthesized by taking the remaining parts of the diphones by removing the excised frames.

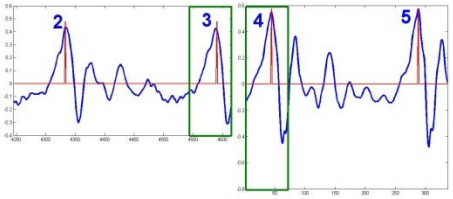


Fig. 6 Rectangles showing the overlap add frames excised from the diphones in the concatenation point. Pitch marks are numbered from left to right consecutively.

if $d_2(n)$ is the last diphone of the word,

$$temp(u) = begin(w); mid(j); end(z); silence$$

else

$$temp(u) = begin(w); mid(j)$$

where,

$$begin(w) = d_1(w), \qquad w = 1,2, \dots, (m - j)$$

and,

$$end(z) = d_2(x), \qquad x = j, j + 1, \dots, n$$
$$output = output(L); temp(u) \qquad (2)$$

Here, semi colon indicates that the signals are concatenated one after another. In (2), L is the length of the synthesized waveform so far. It is noticeable that the remaining part of $d_2(n)$ is not concatenated in the synthesized speech waveform unless it is the last diphone of the word. The reason is that, it will be treated as the left diphone in the next concatenation point. For this, all the arrays and variables of $d_1(n)$ are replaced by the corresponding arrays and variables of $d_2(n)$ in the next iteration of synthesis. In the following iterations, unlike the first time, $begin(w)$ is calculated by removing the frame for overlap adding excised from the right diphone (left diphone in this iteration) of the previous iteration. That is,

$$begin(w) = d_1(x),\ where\ x = k, k + 1, \dots, (m - j)$$
$$here, k\ is\ the\ length\ of\ previous\ frame\ for\ overlap\ adding$$

Fig. 7 shows the speech resulted after rescaling and overlap adding the diphones 'রা' and 'আশ'. Peak 3 and 4 from fig. 1 are overlap added and so now they are merged into one peak.

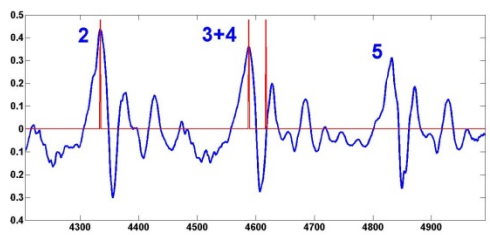


Fig. 7 This is the concatenated signal after rescaling and overlap adding the diphones 'রা' and 'আশ'. Peak 3 and 4 are merged into one peak. The audible click is totally gone from the concatenation point.

IV. Results and Discussions

We tested the algorithm in a full paragraph and 268 different words which covered most of the diphones in our system. Following chart in fig. 8 shows the success rate of removing click sounds in the concatenation point of diphones.

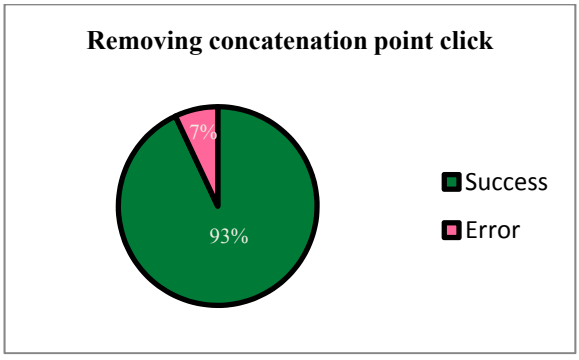


Fig. 8 Success rate of removing audible clicks by the concatenation algorithm in Subachan

Errors occur when the frames in the diphone boundaries of concatenation point are unvoiced and they have energy mismatches. The rescaling was avoided in these cases since unvoiced speeches are assigned random pitch mark locations and so they produce unpredictable ratios. Therefore, when those diphones have energy mismatches, the click cannot be removed. The pitch marking algorithm sometimes fails in the cases of voiced consonants which have lower energies than the voiced vowels in the same diphone but the click generated in this case is less noticeable.

We had 20 listeners to listen to the full paragraph and the words generated using the concatenation algorithm and the same paragraph and words without using it. Afterwards, they were asked which outputs sounded more natural and more intelligible. Their responses are shown in the following charts in fig. 9 and fig. 10.

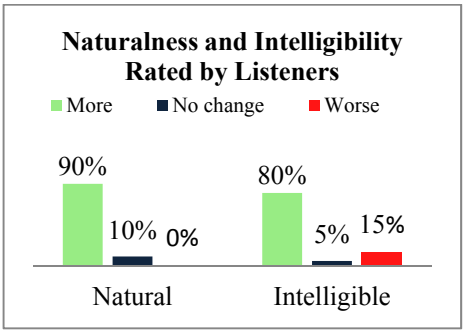


Fig. 9 Performance of the concatenation algorithm in sentence level

The results show that the output speech has significant improvements in the naturalness of speech after using the concatenation algorithm. In sentence level, 90% of the listeners said that the output was more natural, 10% sensed no change. 80% of the listeners said that the output speech was more intelligible, 5% sensed no change and 15% said that the intelligibility was worse. In word level, however, 95% of the listeners said that the output was more natural, 5% sensed no change. 75% of the listeners said that the output speech was more intelligible, 5% sensed no change and 20% said that the intelligibility was worse.

V. Conclusions

We have shown how TDPSOLA could be used for concatenating two diphones in a Bengali speech synthesizer. We modified the pitch marking algorithm to get correct results, and then used a pitch detection algorithm for finding voiced or unvoiced speech frames. Next we concatenated the diphones using the derived results. The output generated using this method is free from clicks in the concatenation points hence generated more natural output speech. Although, we have shown it for diphone synthesis, the same method can be used for syllable based text to speech synthesis system to remove clicks in the concatenation points of the syllable units.

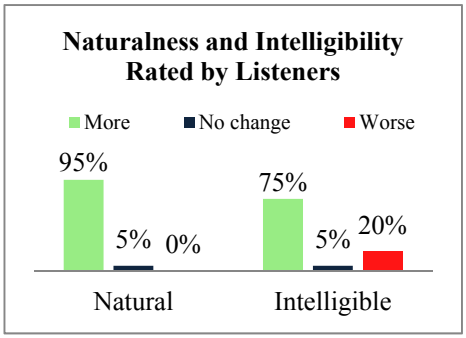


Fig. 10 Performance of the synthesis algorithm in word level

Acknowledgment

KS thanks Professor Dr. Md. Shahidur Rahman for his help throughout the project; his vast knowledge in speech processing showed lights during the darkest moments of this endeavour.

References

[1] T. Dutoit, "High-quality text-to-speech synthesis: An overview," *Journal of Electrical and Electronics Engineering Australia,* vol. 17, pp. 25-36, 1997.

[2] S. K. D. Mandal and A. K. Datta, "Epoch synchronous non-overlap-add (ESNOLA) method-based concatenative speech synthesis system for Bangla," *SSW,* 2007.

[3] F. Alam, P. K. Nath, and M. Khan, "Text to speech for Bangla language using festival," 2007.

[4] A. Naser, D. Aich, and M. R. Amin, "Implementation of subachan: Bengali text to speech synthesis software," *ICECE*, 2010.

[5] Prosodic modification of speech, research.cs.tamu.edu/prism/lectures/sp/l19.pdf.

[6] O. Türk, "New methods for voice conversion," Doctoral Dissertation, pp. 32-34, Boğaziçi University, 2003.

[7] C. Hamon, E. Mouline, and F. Charpentier, "A diphone synthesis system based on time-domain prosodic modifications of speech," in *Proc. ICASSP'89*, 1989, pp. 238-241.

[8] V. Goncharoff and P. Gries, "An algorithm for accurately marking pitch pulses in speech signals," in *Proc. SIP'98*, 1998, pp. 281–284.

[9] R. G. Bachu, S. Kopparthi, B. Adapa, and B. D. Barkana, "Separation of voiced and unvoiced using zero crossing rate and energy of the speech signal," in *Proc. ASEE'08*, 2008, pp. 1-7.

[10] L. R. Rabiner (2015), Autocorrelation Pitch Detector, http://www.mathworks.com/matlabcentral/fileexchange/52083-autocorrelation-pitch-detector/all_files.

[11] L. R. Rabiner, "On the use of autocorrelation analysis for pitch detection," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 25, no. 1, pp. 24-33, Feb. 1977.

Examining the Inaccuracies of Citations and Referencing: A Case of a Computer Science and Engineering Journal

Sujit Kumar Basak
Université du Québec à Montréal (UQÀM)
Montreal, Canada
sujitbasakmca@gmail.com

Abstract—**Citations and referencing accuracy play a crucial role for readers. The aim of this study was to analyse inaccuracies of citations and referencing in the international journal of computer science and engineering published in 2015 (volume no. 7, issue nos. 1-12). This study was an experiment-based and results revealed that referencing errors of volume 7 in issue no. 1 has the highest error rate of 5% and on the other hand, issue nos. 6 and 8 have the lowest error rate of 1.4%. Similarly, in the issue no. 2, error rate is 2.5%; issue nos. 3 and 4, error rate are 2.3%; issue no. 5, error rate is 2.5%; issue no. 7, error rate is 1.7%; issue no. 9, error rate is 1.5%; issue no. 10, error rate is 3.1%; issue no. 11 and 12, error rate are 3% and 1.7% respectively. However, the results also shown that citation errors in volume 7 for the issue nos. 3, 5, 7, and 12 do not have any errors, nevertheless, issue no. 4 has the highest error rate of 3.7%, followed by 3.6% (issue no. 9), 1.3% (issue nos. 1, 6, and 11), 1.2% (issue no. 8), and 1.1% (issue no. 2). This study recommends for academics and researchers to use reference management software tools, namely, EndNote, Mendeley, RefWorks, etc. to avoid errors and redundancy in order to make more accurate citations and referencing than manual systems.**

***Keywords*— Computer Science & Engineering, Citations, Inaccuracies, Referencing.**

I. Introduction

Citations are an essential for scientific manuscripts to confirm statements by authors to help readers in understanding the context of articles and it is important for the scientific communication to explain rationale of papers to evaluate metrics for academic performance such as impact factor, Hirsch factor, SCImago, and country rank [1]. According to [2], in scholarly publications, citations are more important in the age of electronic distribution to provide valuable functions in electronic publishing, namely, an actionable, clickable link to more information about citations.

Accuracy of references is important for the transmission of scientific knowledge to make easy search of the electronic databases [3] and [4]. In addition, [5] stated that accurate references allow interested readers to read additional publications that are related to a specific journal article. According to [3] and [4], accuracy of references increases a quality and credibility of study for authors and journals. Several researchers stated reference lists of journals, serves a several significant factors such as background information in a particular area of interests; assessing faculty production for contract negotiations; constructing citation indexes; evaluate a rigor of an article; development of journals, publishers, and author's ranking; establish a credibility for authors as researchers; and finally, identification and location for articles that readers wishes to consult [14], [17] and [18]. Furthermore, referencing serves as a resource for readers since readers of literature, in particular, students and those are preparing manuscripts for publications are very often retrieve cited documents. Nevertheless, citation accuracy allows readers to easily locate references [19].

II. Problem Statement

According to [15], five leading social work journals suffered referencing list errors. In addition, several researchers stated in medical journals, reference accuracy for many years have been a major problem for editors and readers [6-9]. Reference [10] examined a total 100 citations of volume 22 and concluded errors within citations. However, the inaccuracy of references are obstacles to retrieve the original information from a computer because it cannot retrieve or recognize any sort of inaccuracy [11]. According to [5], if citations are inaccurate, then it becomes time-consuming, frustrating, or even impossible to retrieve for readers. References [12] and [13] indicated as cited in [15] "a lack of clear education as one of the main factors contributing to reference errors and also posited that the pressure to publish may lead authors to neglect their reference lists when preparing to submit manuscripts (p. 173). Moreover, [14] also indicated "carelessness and misuse of language" led to inaccuracies in references (p. 42)". In spite of the reference accuracy of surgical journals blamed "authors do not check their references or may not even read them", finally, they also criticized that even reviewers are not checking their references [16].

III. Aim and Objectives

The aim of this study is to examine inaccuracies of citations and referencing of volume 7 in the international journal of computer science and engineering (IJCSE) published in 2015. In order to achieve the aim, following objectives were taken into consideration.

- To investigate inaccuracies of citations against referencing for the volume 7 in the IJCSE; and
- To investigate inaccuracies of referencing against the main source of articles for the volume 7 in the IJCSE.

IV. Research Questions

This study's purpose is to answer two research questions and these questions are following:

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

- To what extent does citations have inaccuracies against referencing for the volume 7 in the IJCSE?
- To what extent does referencing have inaccuracies against the main source of articles for the volume 7 in the IJCSE?

V. Research Hypotheses

This study has two research hypotheses and these hypotheses are:

- There are no inaccuracies of citations against referencing for the volume 7 in the IJCSE (H_1); and
- There are no inaccuracies of referencing against the main source of articles for the volume 7 in the IJCSE (H_2).

VI. Literature Reviews

A study conducted by [20] in five core entomology journals on bibliographical reference citations and results revealed that an average 30.1% of citations in journals contained 1 or more errors. In addition, the study further shown most of the errors occurs in titles of journals. A similar study conducted by [5] on five biomedical informatics journals with a total of 37 articles that was issued first in 2004 and a total of 656 references included. Results revealed that at least one error included 34.3% with 225 articles. Thus, out of 225 references, a total of 311 errors were identified. However, the study also indicated reference error rate by journals ranged from 22.1% to 40.7% and most of the errors occurs in the following fields, namely, names of journal (31.2%), title (17.7%), page (7.4%), year (3.5%), and volume (1.3%). In a study by [21] on comparison of referencing and citations for the inventive engineering and sciences journal (volume 2, issue 6 and 7) and a total of 5 articles were included in the study. The results of the study revealed that a total of 208 errors. On the other hand, another study by [22] examined referencing and citations for the international journal of recent technology and engineering and the study included a total of 13 articles (volume 3 and issue 6) and found a total of 304 errors.

A research by [23] on Archives of Dermatology and British Journal of Dermatology and the study randomly selected a total of 240 references (60 per journal) that was checked with original articles. Results revealed citation error rates were 41% and the quotation error rates were 35%. Nevertheless, the study furthermore, indicated that only 36% of the references were errors free. On the other hand, a similar study conducted by [24] on the accuracy of references in sport science journals and study randomly selected 100 articles that was published in 2009-2011 (volume 5). The results of the study revealed a total of 12% error rate were in the following fields' author's name (6.5%), journal title (3.5%), article title (2%), page number (1%), and year of publication (0.5%).

In [25], an experiment was conducted on the journal of information and management published in 2007 and results revealed errors in six fields, namely, author(s), article title, publication year, volume, and page numbers. However, the study further indicated that a higher percentage of errors in the journal of information and management were 49.1% and the average percentage were for authors 56%, page number 22%, the article title 15%, volume 3%, publication year 2%, and the journal title 2%.

A research by [26] to observe the accuracy of references in articles published in indexed journals of Nepal's between January 2007 and December 2008. The study randomly selected a total of 30 references from three journals such as Kathmandu University Medical Journal (KUMJ), Nepal Medical College Journal (NMCJ), and Journal of Nepal Medical Association (JNMA). However, all references were checked with the original source of articles in six fields and results revealed that errors in the KUMJ 23.3%, errors in the NMCJ 11.6%, and errors in the JNMA 11.1% respectively. In addition, the study also indicated most common errors were found in all three journals in the following fields, namely, author's name and journal name.

A study conducted by [17] on citation errors in Canadian Journal of Anesthesia between the volumes of 1990 and 1994. The study randomly selected one hundred references from each year publications in order to scrutinize fields such as authors' names, article title, journal title, volume number, page numbers, and year. Results revealed that references errors 48% in 1990 and 22% in 1994 respectively. In addition, the study also indicated errors in the field of title and author have most common errors about 70% either of the two years.

Reference [27] conducted a study on reference errors of four paediatric journals, namely, Acta Paediatrica (Acta Paediatr), Archives of Disease in Childhood (Arch Dis Child), Journal of Pediatrics (J Pediatr), and Pediatrics and results of the study revealed an average error was 36% in the Acta Paediatr followed by Arch Dis Child 22%, J Pediatr 29%, Pediatrics 32% respectively. However, the study also indicated that all journals had major errors with a percentage of 1% in the Acta Paediatr followed by 1% of the Arch Dis Child, 2% in the J Pediatr, and 4% in the Pediatrics 4% respectively.

A study conducted by [28] on the comparison in order to scrutinize citations and reference errors for the computer science and engineering journal (volume 6 and number 5) and inventive engineering and sciences journals (volume 2 and issue 7) published in 2014. Results revealed a considerable less error in the inventive engineering and sciences journals as compared to the computer science and engineering journal.

VII. Methodology

In this study, a total of 20 articles (315 references) was included from the IJCSE published in 2015 (volume 7 and issue numbers 1-12). Each text citation of all articles was checked against a specific reference and all the references were checked with original articles through Google, Google Scholar. However, scrutinisation was based on the six fields such as journal/conference name, author(s), article name, publication year, volume number, and the page number. Collected data were analysed using MS Excel. Many authors did not follow journal instructions in terms of citations and referencing and this study did not consider any errors for the inconsistency of 20 articles (315 references). However, this study considered errors in giving information (citations and referencing) in articles that did not match while comparing in the above mentioned six fields for citations versus referencing and referencing versus an original source of articles.

VIII. Results

Table I shown that *issue no. 1* has the highest error rate of 5%, but on the other hand, journal/conference name, publication year, and volume have the same number of errors

and similarly, article name and page number also have a same number of errors. *Issue no. 6 and 8* have the lowest error rate of 1.4%. In addition, article name, publication year, page number has the equal number of errors for the issue no. 6 and 8. However, for the issue no. 8, volume and page number have the equal number of errors. In the case of *issue no. 2* error rate is 2.5% and it has got equal number of errors with the publication year, volume, and page number. For the *issue nos. 3 and 4,* errors rate is 2.3% and it has got the equal number of errors with author(s) and article name (issue no. 3). On the other hand, publication year and volume (issue no. 3) have the same number of errors. Furthermore, for the issue no. 4, it has got the equal number of errors with the publication year, volume, and page number. For the *issue no. 5,* errors rate is 2.5% and it has the equal number of errors with journal/conference name and author(s) but, on the other hand, article name, publication year, volume and page number also have the same number of errors. For the *issue no. 7*, error rate is 1.7% and the equal number of errors with the article name and volume number, but, on the other hand, publication year and page number have the same number of errors. For the *issue no. 9*, error rate is 1.5% and it has got the equal number of errors with the publication year and page number. For the *issue no. 10*, error rate is 3.1% and the equal number of errors it has got with the journal/conference name and article name but, on the other hand, publication year, volume number, and page number have the same number of errors. For the *issue no. 11 and 12*, error rate are 3% and 1.7%, respectively but on the other hand, it also has an equal number of errors with the publication year, volume, and page number.

TABLE I
ERRORS IN DIFFERENT FIELDS FOR REFERENCING AGAINST THE MAIN SOURCE OF ARTICLES

I	ER	II	III	IV	V	VI	VII
I1	5%	1	2	0	1	1	0
I2	2.5%	10	5	1	0	0	0
I3	2.3%	2	3	2	0	0	1
I4	2.3%	8	2	1	0	0	0
I5	2.5%	2	2	0	0	0	0
I6	1.4%	12	6	2	2	1	2
I7	1.7%	5	2	1	0	1	0
I8	1.4%	2	8	4	1	0	0
I9	1.5%	3	7	2	1	0	1
I10	3.1%	1	6	1	0	0	0
I11	3%	5	12	2	0	0	0
I12	1.7%	4	5	1	0	0	0
I: Issue number				**ER**: Error			
II: Journal/Conference name				**III**: Author(s)			
IV: Article name				**V**: Publication year			
VI: Volume				**VII**: Page number			

Fig. 1 shown errors in the IJCSE published in 2015 (volume 7, issue nos. 1-12). In the case of *issue no. 1*, author(s) have highest errors and on the other hand, volume and publication year have the lowest errors. In addition, volume and publication errors are equal in numbers. In the case of *issue no. 2 and 4*, journal/conference name have highest errors and, nevertheless, article name has the lowest errors. In the case of *issue no. 3*, author(s) name have highest errors and in spite of lowest errors in the page number. In the case of *issue no. 5*, author(s) name and journal/conference name have equal errors. In the case of *issue no. 6*, journal/conference name has the highest errors and on the other hand, volume number has the lowest errors. In the case of *issue no. 7*, journal/conference name have highest errors and even if the lowest errors in the volume number and article name. In the case of *issue no. 8*, author(s) name have highest errors and in spite of lowest errors in the publication year. In the case of *issue no. 9*, author(s) name have highest errors and however, lowest errors in the publication year and page number. In the case of *issue no. 10*, author(s) name have highest errors and, nevertheless, lowest errors in the journal name and article name. Finally, for the *issue no. 11 and 12*, author(s) name have highest errors and on the other hand, lowest errors in the article name for the IJCSE.

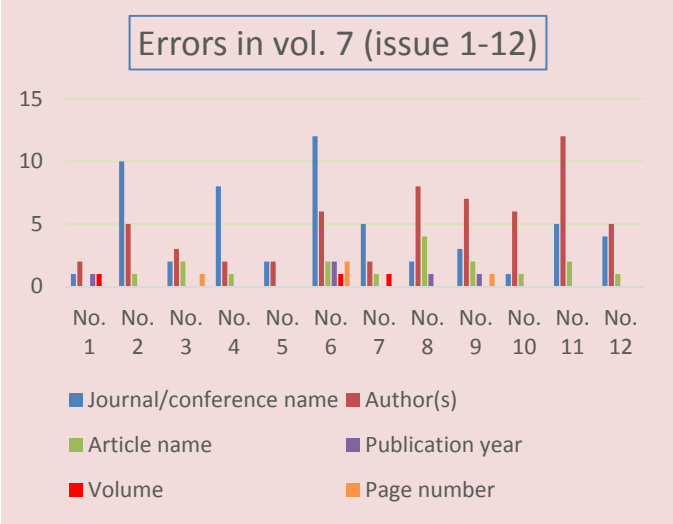


Fig. 1 Errors in different fields in volume 7.

TABLE II
ERRORS IN DIFFERENT ISSUE NOS. FOR TEXT CITATIONS AGAINST REFERENCES

Table II has shown that issue nos. 3, 5, 7, and 12 do not have any errors in terms of cross referencing (text citations versus referencing) but on the other hand, issue no. 4 has the highest errors 3.7%, followed by 3.6% (issue no. 9), 1.3% (issue no. 1, 6, and 11), 1.2% (issue no. 8), and finally 1.1 (issue no. 2).

Issue nos.	**A**	**B**	**Errors**
Issue no. 1	25	18	1.3%
Issue no. 2	41	35	1.1%
Issue no. 3	19	19	0%
Issue no. 4	26	7	3.7%
Issue no. 5	10	10	0%
Issue no. 6	35	26	1.3%
Issue no. 7	16	16	0%
Issue no. 8	22	18	1.2%
Issue no. 9	22	6	3.6%
Issue no. 10	25	25	0%
Issue no. 11	57	41	1.3%
Issue no. 12	17	17	0%
A: Total Number of References		B: Total Number of Text Citations Does Match with Referencing	

IX. CONCLUSION AND DISCUSSION

This section describes that have been revealed for each of the hypotheses and furthermore, it has also investigated a total

of 20 articles included 315 references (volume 7 and issues 1-12) and all the references were examined against the main source of articles. The reason being more errors and omissions in the citations and referencing is because it is hard for academics and researchers to be consistent while doing their citations and referencing using the manual system. They can easily avoid this type of errors and omissions if they use reference management software tools such as EndNote, Mendeley, RefWorks, etc. in order to maintain their citations and referencing. According to [29], Mendeley reference software can import more data from the Google Scholar among EndNote, Mendeley, RefWorks.

Table I shown that in the *first research question*, referencing errors 41% against the main source of articles and on the other hand, Table II shown that in the *second research question*, a total of 13.5% errors in the text citations in the cross referencing.

The *first hypothesis* was rejected because Table II has shown that citation errors against referencing in the IJCSE (volume no. 7, issue nos. 1-12). The *second hypothesis* was also rejected because Table I shown that referencing errors against the original source of articles in the IJCSE (volume no. 7, issue nos. 1-12).

The novelty of this study was designed to examine citations and references inaccuracies of the IJCSE (volume no. 7, issue nos. 1-12). The aim of this study was to analyse citations and referencing inaccuracies and findings of this study shown errors for citations and referencing in the IJCSE published in 2015 (volume no. 7, issue nos. 1 to 12). The limitation of this study is conducted only within the volume 7 but a larger number of volume might provide the basis for more detailed statistical analysis. However, this would be considered for the future investigation. Finally, this study recommends for researchers and academics to use reference management software to avoid errors and redundancy in order to make more accuracy for citations and referencing their articles.

REFERENCES

[1] H. Jergas and C. Baethge, "Quotation accuracy in medical journal articles-a systematic review and meta-analysis," *PeerJ,* pp. 1-20, 2015.

[2] C. A. Meyer, "Reference accuracy: Best practices for making the links," *Journal of Electronic Publishing,* vol. 11, pp. 1-14, 2008.

[3] V. J. Roach, T. K. Lau, and K. W. D. Ngan, "The quality of citations in major international obstetrics and gynecology journals," *American Journal of Obstetrics and Gynecology,* vol. 177, pp. 973-975, 1997.

[4] M. F. McLellan, L. D. Case, and M. C. Barnett, "Trust, but verify. The accuracy of references in four anesthesia journals," *Anesthesiology,* vol. 77, pp. 185-188, 1992.

[5] D. Aronsky, J. Ransom, and K. Robinson, "Accuracy of references in five biomedical informatics journals," *Journal of the American Medical informatics Association,* vol. 12, pp. 225-228, 2005.

[6] R. F. Browne, P. M. Logan, M. J. Lee, and W. C. Torreggiani, "The accuracy of references in manuscripts submitted for publication," *Canadian Association of Radiologists Journal,* vol. 55, pp. 170-173, 2004.

[7] K. W. Hinchcliff, "Accuracy of references," *Journal of the American Veterinary Medical Association,* vol. 198, pp. 11-12, 1991.

[8] K. W. Hinchcliff, N. J. Bruce, J. D. Powers, and M. L. Kipp, "Accuracy of references and quotations in veterinary journals," *Journal of the American Veterinary Medical Association,* vol. 202, pp. 397-400, 1993.

[9] S. Holt, R. Siebers, A. Suder, R. Loan, and O. Jeffery, "The accuracy of references in Australian and New Zealand Medical Journals," *The New Zealand Medical Journal,* vol. 113, pp. 416-417, 2000.

[10] A. E. O'Connor, W. Lukin, L. Eriksson, and C. O'Connor, "Improvement in the accuracy of references in the journal emergency medicine Australasia," *Emergency Medicine Australasia*, vol. 25, pp. 64-67, 2013.

[11] R. Siebers, "How accurate are references in clinical chemistry?" *Clinical Chemistry,* vol. 47, pp. 606-607, 2001.

[12] J. H. Sweetland, "Errors in bibliographic citations: A continuing problem," *The Library Quarterly,* vol. *59*, pp. 291-304, 1989.

[13] A. White, "Reference inaccuracies in two counseling journals," *Counselor Education and Supervision,* vol. 26, pp. 286-292, 1987.

[14] J. E. Fenton, H. Brazier, A. D. Souza, J. P. Hughes, and D. P. McShane, "The accuracy of citation and quotation in otolaryngology/head and neck surgery journals," *Clinical Otolaryngology,* vol. 25, pp. 40-44, 2000.

[15] S. E. Wilks, and C. A. Spivey, "Views of reference list accuracy from social work journal editors and published authors," *Advances in Social Work,* vol. 5, pp. 172-181, 2004.

[16] J. T. Evans, H. I. Nadjari, and S. A. Burchell, "Quotational and reference accuracy in surgical journal: A continuing peer review problem," *JAMA,* vol. 263, pp. 1353-1354, 1990.

[17] M. Asano, K. Mikawa, K. Nishina, N. Maekawa, and H. Obara, (1995). "Improvement of the accuracy of references in the Canadian Journal of Anesthesia," *Canadian Journal of Anesthesia,* vol. 42, pp. 370-372, 1995.

[18] M. K. Taylor, "The practical effects of errors in reference lists in nursing research journals," *Nursing Research,* vol. 47, pp. 300-303, 1998.

[19] L. Schulmeister, "Quotation and reference accuracy of three nursing journals," *Image: the Journal of Nursing Scholarship,* vol. 30, pp. 143-146, 1998.

[20] C. Kristof, "Accuracy of reference citations in five Entomology Journals," *American Entomologist,* vol. 43, pp. 246-251, 1997.

[21] S. K. Basak, "Examining the references and citations accuracy of the inventive engineering and sciences journal," *International Conference on Electrical Engineering and Information Communication Technology,* 2015, pp. 1-4.

[22] S. K. Basak, "Examining the consistency of citation, reference accuracy for the articles published in an international journal: Recent technology and engineering," *International Journal of Research in Information Technology,* vol. 3, pp. 52-56, 2015.

[23] P. M. George and K. Robbins, "Reference accuracy in the dermatologic literature," *Journal of the American Academy of Dermatology,* vol. 31, pp. 61-64, 1994.

[24] M. Zasa, "The accuracy of references in five sport science journals," *Science and Sports,* vol. 30, pp. e31-e33, 2015.

[25] K. Davies, "Reference accuracy in library and information science journals," in *Aslib Proceedings,* 2012, vol. 64, pp. 373-387.

[26] P. Adhikari, "Accuracy of references in indexed journals of Nepal," *Nepal Med Coll J,* vol. 11, pp. 130-132, 2009.

[27] A. Vargas-Origel, G. Gómez-Martinez, and M. A. Vargas-Nieto, "The accuracy of references in paediatric journal," *Arch Dis Child,* vol. 85, pp. 497-498, 2001.

[28] S. K. Basak, "A survey of accuracy of citations and references comparison in articles published in computer science and engineering journal and the inventive engineering and science journal," *International Journal of Research in Information Technology,* vol. 3, pp. 204-208, 2015.

[29] S. K. Basak, "A comparison of researcher's reference management software: RefWorks, Mendeley, and EndNote," *Journal of Economics and Behavioral Studies,* vol. 6, pp. 561-568, 2014.

Highly Birefringent Bisymmetric Photonic Crystal Fiber for Terahertz Wave Guidance

Md Nafiz Amin,[1,*] and Mohammad Faisal[1]

[1] Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology, Dhaka-1205, Bangladesh

[*] nafiz.amin.bd@ieee.org

Abstract— **A novel design for a highly birefringent photonic crystal fiber (PCF) with all circular holes for transmission in the terahertz (THz) regime is proposed. It shows a very high birefringence ~0.05 in the 0.7-1.5 THz frequency range. Using finite element method, we calculated the important optical properties like birefringence, material loss, confinement loss, and dispersion and investigated the effects of varying the fiber parameters on them. Design simplicity of the proposed PCF leads to great fabrication advantage over other complicated patterns. Therefore, the fiber is likely to find important uses in different THz polarization maintaining applications.**

Index Terms—**Birefringence, photonic crystal fiber, polymer, terahertz.**

I. Introduction

THz electromagnetic transmission (0.1-10 THz) has been an area of increasing interest and extensive research over the past few years because of its promising applications in sensing [1], imaging [2], spectroscopy [3], solid state detectors [4], security, and medical diagnostics etc. The selection of a suitable THz waveguide material except dry air has always been a difficulty due to large material absorption loss in this frequency range. Different research efforts in designing application specific waveguides involving metallic wires [5], metal coated dielectric tube [6], and polymer [7] have faced challenges due to high material loss, bending loss, inefficient mode coupling etc. Photonic crystal fibers (PCFs) have shown promising applications in THz transmission domain due to their design flexibility and simultaneous air-solid guiding mechanism, providing better control over optical properties like loss, dispersion, birefringence, single mode operation etc. In particular, high birefringence can be obtained in PCFs by simply introducing asymmetry in the core and cladding periodic structure. These highly birefringent fibers have important applications in polarization-sensitive THz time domain spectroscopy [8].

Since the first demonstration of highly birefringent PCFs using subwavelength air holes in the core by Atakaramians et al. [9], a number of designs have been proposed. Honeycomb cladding PCF has shown a birefringence of 0.022 at 1 THz [10]. PCF with rotated elliptical holes has shown a birefringence value of 0.0445 [11], while asymmetric slotted core has achieved as high as 0.075 [12]. The last two suffer from practical fabrication difficulties due to presence of complicated patterns. Very recently, Islam et al. have demonstrated low material loss all circular hole PCF with birefringence value of 0.045 [13], with moderate confinement in the core region. Wu et al. have used slightly enlarged circular air holes near the core and have obtained birefringence value up to 0.035 in the range 2-4 THz [14].

In this paper we propose a simple bisymmetric single mode PCF with all circular holes that shows higher birefringence than the previously proposed fibers over the 0.7-1.5 THz band. The birefringence is found to be as high as 0.0495 at 1 THz with optimum selection of parameters, which can reach up to 0.06 by design variation. The effective material absorption loss is found to be very low (~0.2 cm^{-1} around 1 THz) in the working range of 0.7-1.5 THz. Confinement loss is enhanced for the undesired leaky modes by squeezing the lattice of air holes, while a low value is ensured for the confined modes. Smooth dispersion profile is obtained with the considered design parameters. The main advantage of the all circular core design is fabrication simplicity in terms of existing commercial methods like drill and draw [15], [16] relative to slotted core or elliptical core designs.

II. Design Considerations and Analysis

A cross-section view of the proposed PCF is displayed in Fig. 1. The cladding part is formed by a triangular lattice distribution of subwavelength sized air holes on the background of cyclic olefin copolymer (namely, TOPAS). This material has an almost constant refractive index of 1.5258 in the 0.1-1.5 THz range [16]. Also, low bulk material loss, insensitivity to humidity [17], and high glass transition temperature due to different grades of constituents [18] etc. makes it a very suitable material for PCF fabrication.

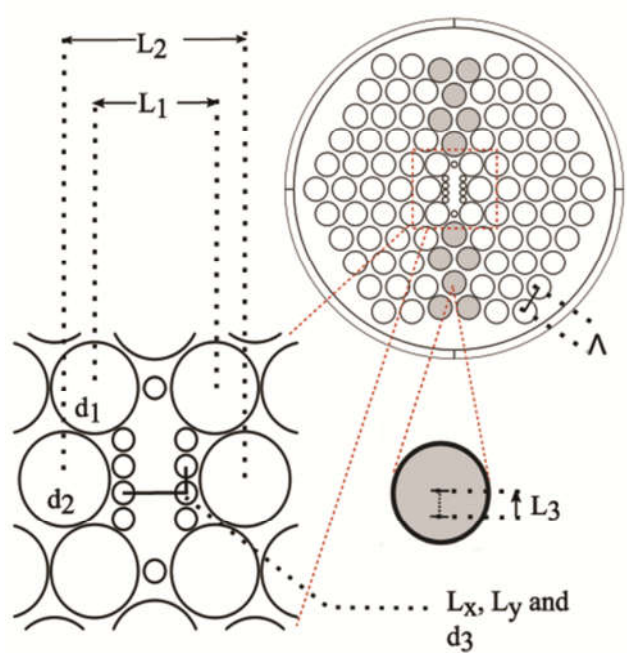


Fig. 1 Cross-section and the magnified core and cladding views of the bisymmetric photonic crystal fiber.

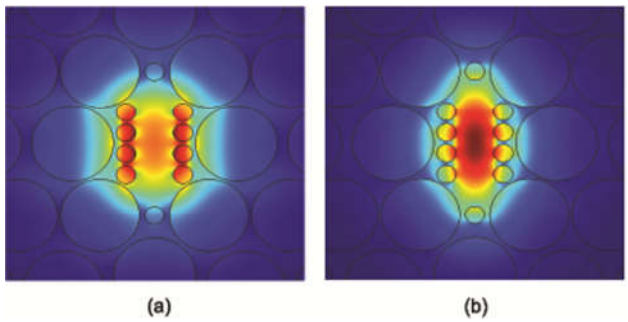


Fig. 2 Electric field patterns of (a) x-polarized and (b) y-polarized mode at 1 THz.

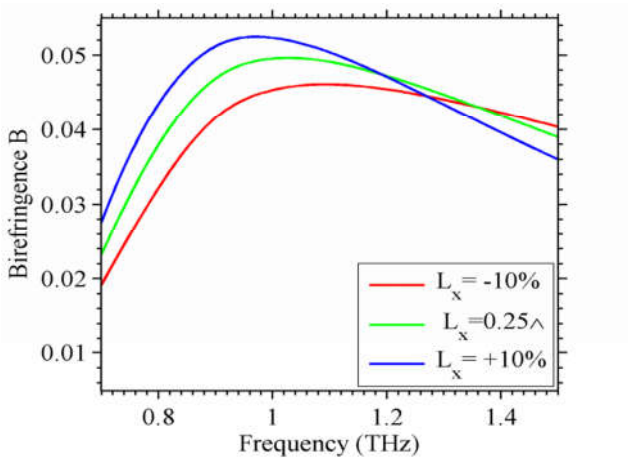


Fig. 3 Birefringence versus frequency for variations of horizontal distance of the small holes from the core L_x with L_2=1.8Λ, d_2=0.9Λ, and d_3=d_1/4.

The distance between two adjacent air holes is Λ. All the holes in the cladding have a diameter d_1, except two slightly larger holes in the first ring with diameter d_2, as denoted in the magnified image of the core. Moreover, the regular and larger holes in the first ring are displaced outward and inward respectively from their expected ring positions, and their gaps are denoted by L_1 and L_2. These modifications in the first ring effectively break the cladding symmetry and induce birefringence. Additional asymmetry is created in the core region by introducing rows of small sized holes with diameter d_3. The holes are spaced at a horizontal distance L_x from the center and with uniform spacing L_y between each other. Two other same sized holes are also placed in the first cladding ring. Aside from inducing birefringence, all these small sized holes transmit a significant portion of the confined mode energy and thus help in reducing the effective material loss caused by index guiding in the material. The shaded holes in the cladding region are slightly displaced from their lattice positions towards the core by a length L_3. This is done to squeeze the core region as much as possible, which keeps away the multiple higher order modes from confinement, as it will be apparent from the loss characteristics of these modes later on.

In the design investigation we have kept fixed values for the parameters Λ=200 µm, d_1=0.85Λ, L_1=1.2Λ, L_3=0.143Λ, and L_y=0.25Λ. Fig. 2 shows the tightly confined electric mode field patterns of the two orthogonally polarized modes at 1 THz. The key properties have been investigated using full vector finite element method (FEM). To prevent reflection and back-scattering, perfectly matched layer (PML) has been employed at the circular boundary.

III. Results and Discussion

The optical properties of the PCF have been analysed and the impacts of the key design parameters' variation on them have been investigated. The parameter variations are kept in the range as such to make the fiber physically realizable with non-overlapping air holes. The mode birefringence of a fiber is defined as

$$B = |\eta_x \quad \eta_y|, \tag{1}$$

where η_x and η_y are the effective mode indices of the two polarized modes. Fig. 3 shows the variation of birefringence as a function of frequency for L_x=0.25Λ and for a variation by 10% from this selected value. In each case, the birefringence profile is a downward concave function of frequency with peak around 1 THz. At low frequencies, pushing the small holes away from core, i.e. increasing L_x increases birefringence, while the case is opposite at higher frequencies. This phenomenon can be attributed to the dependence of asymmetry ratio (ratio of the lengths of a mode field along y-axis and x-axis) on both frequency and the core structure. At low frequencies the x-mode reaches well across the small holes, while y-mode is confined in a small region, as seen in Fig. 2. Pushing the small holes further from the core impacts the confinement of y-mode more, as it gets almost confined in the central region, while x-mode still passes through the small holes. In this case, increasing L_x enhances the contrast in effective indices. However, at high frequencies both the modes are tightly confined in a small area. Pushing the holes further

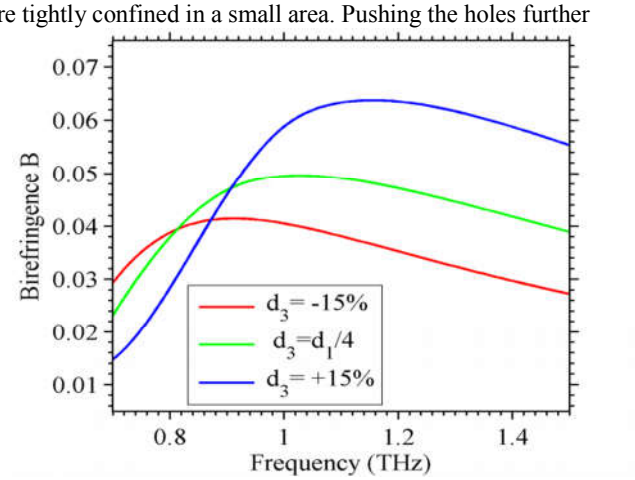


Fig. 4 Birefringence versus frequency for variations of size of the small holes d_3 with L_2=1.8Λ, d_2=0.9Λ, and L_x=0.25Λ.

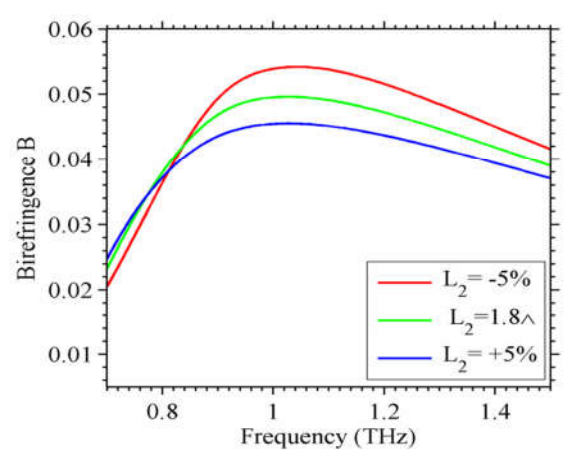


Fig. 5 Birefringence versus frequency for variations of distance between large holes L_2 with d_2=0.9Λ, L_x=0.25Λ, and d_3=d_1/4.

away (i.e. increasing L_x) reduces air hole contribution for both the modes, eventually reducing index contrast, and causing a drop in birefringence.

Fig. 4 shows the influence of varying the small hole size d_3 in the core from its selected value of $d_1/4$. As seen, increasing d_3 significantly increases birefringence value at a particular frequency, while the peak moves towards higher frequencies. At our selected hole size, birefringence at 1 THz is 0.0495 and can reach up to 0.059 by increasing d_3. However, in case of increase of d_3 by 15%, the solid bridge between two small holes becomes as narrow as 1.12 µm, which might introduce fabrication difficulties.

The cladding asymmetry notably depends on the position and size of the two large holes in the first cladding ring, and the effects of their variations are prominent in the value of birefringence. Fig. 5 shows the effect of altering the hole spacing L_2 from the selected value of 1.8Λ. Bringing the holes towards the core gradually increases birefringence over a large portion of the bandwidth, which is restrained by the presence of smaller holes. Fig. 6 shows the change of birefringence with variation of the large hole sizes from d_2=0.9Λ. Enlarging these air holes enhances cladding asymmetry, deforms the core, and leads to higher birefringence. Other fiber parameters have similar but less significant effects on birefringence.

Another important optical property in the THz domain is the effective material loss (EML) defined as [19]

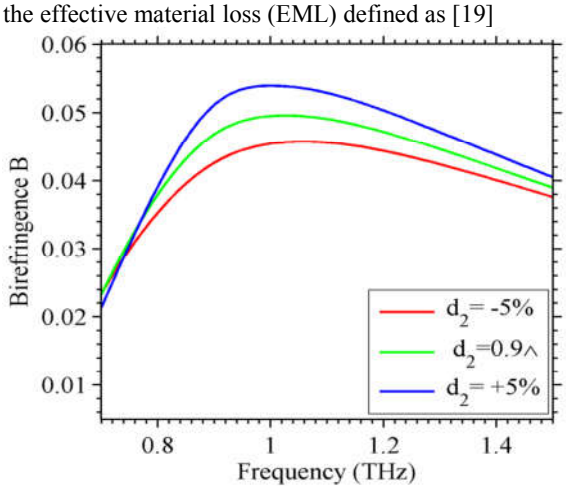


Fig. 6 Birefringence versus frequency for variations of size of the large holes d_2 with L_2=1.8Λ, L_x=0.25Λ, and d_3=d_1/4.

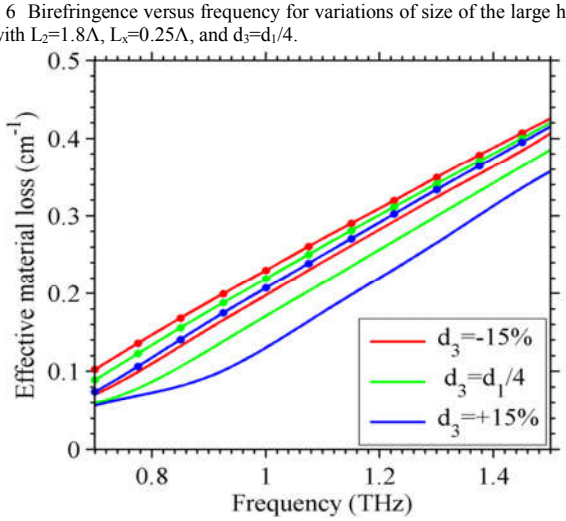


Fig. 7 Effective material loss versus frequency for variations of size of the small holes d_3 with L_2=1.8Λ, d_2=0.9Λ, and L_x=0.25Λ for both x-axis and y-axis (marked line) modes.

$$\alpha_{eff} = \sqrt{\frac{\varepsilon_0}{\mu_0}}\left(\frac{\int_{mat}\eta_{mat}|E|^2\alpha_{mat}\,dA}{2\left|\int_{all}S_z dA\right|}\right). \quad (2)$$

Here, ε_0 and μ_0 are the permittivity and permeability of vacuum, η_{mat} is the refractive index, α_{mat} is bulk material absorption loss, E is the modal electric field, and S_z is the z-component of the Poynting vector. The bulk material loss of TOPAS is 0.06 cm^{-1} at 0.4 THz and increases at a rate of 0.36/THz [13]. EML is a measure of the transverse modal loss in a waveguide with inhomogeneous loss coefficient profile. Fig. 7 shows the increasing trend of EML with frequency for variation of small hole size d_3 for both the polarized modes.

The loss is greater for the y-axis mode than its counterpart, which is due to the fact that more power propagates through the solid material for y-mode than the x-mode, incurring more absorption loss. Fig. 8 shows the EML variation for varying the horizontal displacement of the small holes L_x. At 1 THz the EML is 0.22 cm^{-1} for the x-mode and 0.17 cm^{-1} for the y-mode for the selected fiber parameters. As evident from the two EML figures, enlarging the small holes (increasing d_3) or bringing them closer (decreasing L_x) reduces the EML. This happens because more power propagates through the holes in the above two cases, reducing the material contribution to guiding.

Next we investigate the confinement loss of the fiber, given by

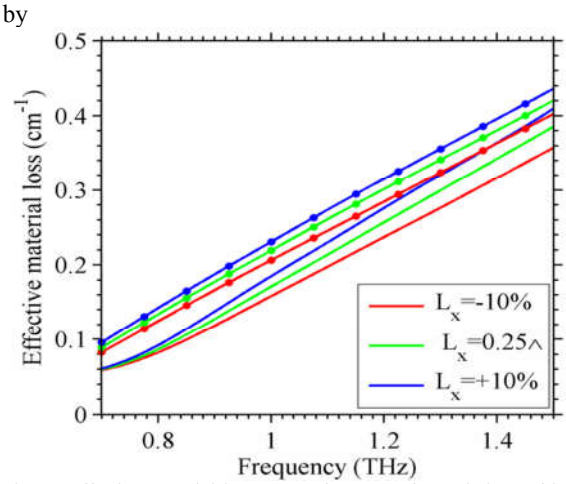


Fig. 8 Effective material loss versus frequency for variations of horizontal distance of the small holes from the core L_x with L_2=1.8Λ, d_2=0.9Λ, and d_3=d_1/4 for both x-axis and y-axis (marked line) modes.

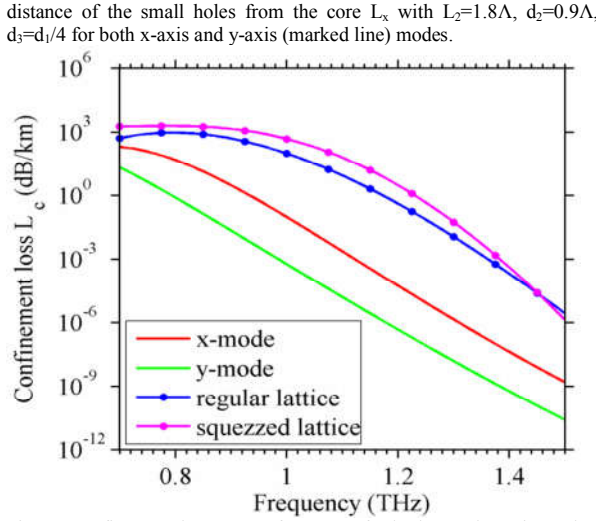


Fig. 9 Confinement loss versus frequency for both x and y-axis modes with five air hole rings. The marked lines represent confinement loss of the first leaky mode in regular and squeezed lattice conditions.

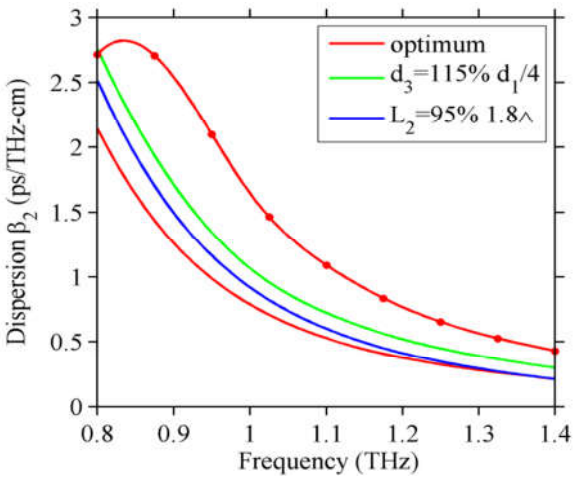


Fig. 10 Dispersion profile of the y-axis mode and the impacts of changing d_3 and L_2 on it. The marked line represents dispersion profile of the x-axis mode for optimum fiber parameters.

$$L_c = 40\pi.\frac{Im[\eta_{eff}]}{\lambda.\ln 10}. \quad (3)$$

Fig. 9 shows the confinement loss for both the polarized modes for the selected fiber parameters and five cladding air hole rings. At 1 THz, the losses are 9.5×10^{-2} dB/km and 5.8×10^{-4} dB/km for x-mode and y-mode respectively, in both cases decreasing with frequency. This value can be further reduced by increasing the fiber air-filling fraction, i.e. increasing the number of rings or the air hole size d_1. We also show the loss profiles of the first leaky mode for both regular cladding lattice as well as when squeezed by L_3. In regular condition, the loss is 97.3 dB/km at 1 THz, and in squeezed condition it increases to 476.8 dB/km. This case is generally true for all the leaky modes, justifying the squeezing of lattice without affecting the key fiber properties.

Since the material dispersion of TOPAS is nearly zero over the range 0.1-2 THz [20], the only significant kind of dispersion is the waveguide dispersion, which is expressed as

$$\beta_2 = \frac{2}{c}\frac{d\eta_{eff}}{d\omega} + \frac{\omega}{c}\frac{d^2\eta_{eff}}{d\omega^2}, \quad (4)$$

where $\omega=2\pi f$ is the angular frequency. Fig. 10 shows the dispersion profile in the range 0.8-1.4 THz, where the dashed line typifies the x-mode and the other lines represent y-mode. At the optimum fiber parameters the x-mode shows a dispersion of 1.64 ps/THz-cm, while it is 0.79 ps/THz-cm for the y-mode. The impacts of varying parameters like d_3 and L_2 on the y-mode are investigated to understand the nature of dispersion. It is observed that the case of increasing d_3 or decreasing L_2, both of which increase the birefringence value also increase the dispersion, particularly at lower frequencies. The dispersion is very flat at higher frequencies irrespective of the fiber parameters.

IV. Conclusions

We have proposed a novel design for highly birefringent photonic crystal fiber with all circular holes in the THz regime. By using three different sizes of circular air holes and adjusting their positions around the core, we have found very high birefringence of ~0.05 at 1 THz. Low values of effective material loss and confinement loss are observed in the relevant frequency range of 0.7-1.5 THz, and low dispersion ranging from 0.95 to 0.3 ps/THz-cm in the range of 0.95-1.4 THz is found. From fabrication point of view, the design simplicity with all circular holes puts a great advantage. Therefore, the proposed PCF is expected to find important uses in THz polarization maintaining applications.

References

[1] M. M. Awad and R. A. Cheville, "Transmission terahertz waveguide-based imaging below the diffraction limit," *Appl. Phys. Lett.*, vol. 86, no. 22, 2005.

[2] Z. D. Taylor, J. Garritano, S. Sung, N. Bajwa, D. B. Bennett, B. Nowroozi, P. Tewari, J. W. Sayre, J. P. Hubschman, S. X. Deng, E. R. Brown, and W. S. Grundfest, "THz and mm-wave sensing of corneal tissue water content: In vivo sensing and imaging results," *IEEE Trans. Terahertz Sci. Technol.*, vol. 5, no. 2, pp. 184–196, 2015.

[3] L. Ho, M. Pepper, and P. Taday, "Terahertz spectroscopy: Signatures and fingerprints," *Nat. Photonics*, vol. 2, pp. 541–543, 2008.

[4] X. He, N. Fujimura, J. M. Lloyd, K. J. Erickson, A. A. Talin, Q. Zhang, W. Gao, Q. Jiang, Y. Kawano, R. H. Hauge, F. Léonard, and J. Kono, "Carbon nanotube terahertz detector," *Nano Lett.*, vol. 14, no. 7, pp. 3953–3958, 2014.

[5] K. Wang and D. M. Mittleman, "Metal wires for terahertz wave guiding," *Nature*, vol. 432, no. 7015, pp. 376–379, 2004.

[6] B. Bowden, J. A. Harrington, and O. Mitrofanov, "Silver/polystyrene coated hollow glass waveguides for the transmission of Terahertz radiation," *Opt. Lett.*, vol. 32, no. 20, pp. 2945–2947, 2007.

[7] L.-J. Chen, H.-W. Chen, T.-F. Kao, J.-Y. Lu, and C.-K. Sun, "Low-loss subwavelength plastic fiber for terahertz waveguiding.," *Opt. Lett.*, vol. 31, no. 3, pp. 308–310, 2006.

[8] M. B. Byrne, M. U. Shaukat, J. E. Cunningham, E. H. Linfield, and A. G. Davies, "Simultaneous measurement of orthogonal components of polarization in a free-space propagating terahertz signal using electro-optic detection," *Appl. Phys. Lett.*, vol. 98, no. 15, p. 151104, 2011.

[9] S. Atakaramians, S. Afshar V., B. M. Fischer, D. Abbott, and T. M. Monro, "Low loss, low dispersion and highly birefringent terahertz porous fibers," *Opt. Commun.*, vol. 282, no. 1, pp. 36–38, 2009.

[10] Y. Hou, F. Fan, Z. W. Jiang, X. H. Wang, and S. J. Chang, "Highly birefringent polymer terahertz fiber with honeycomb cladding," *Opt. Int. J. Light Electron Opt.*, vol. 124, no. 17, pp. 3095–3098, 2013.

[11] N. Chen, J. Liang, and L. Ren, "High-birefringence, low-loss porous fiber for single-mode terahertz-wave guidance," *Appl. Opt.*, vol. 52, no. 21, pp. 5297–5302, 2013.

[12] R. Islam, S. Habib, G. K. M. Hasanuzzaman, R. Ahmad, S. Rana, and S. F. Kaijage, "Extremely High-Birefringent Asymmetric Slotted-Core Photonic Crystal Fiber in THz Regime," *IEEE Photonics Technol. Lett.*, vol. 27, no. 21, pp. 2222–2225, 2015.

[13] R. Islam, M. Selim Habib, G. K. M. Hasanuzzaman, S. Rana, and M. Anwar Sadath, "Novel porous fiber based on dual-asymmetry for low-loss polarization maintaining THz wave guidance," *Opt. Lett.*, vol. 41, no. 3, pp. 440–443, 2016.

[14] Z. Wu, X. Zhou, Z. Shi, H. Xia, J. Huang, X. Jian, and W. Wu, "Proposal for high-birefringent terahertz photonic crystal fiber with all circle air holes," *Opt. Eng.*, vol. 55, no. 3, p. 037105, 2016.

[15] M. A. van Eijkelenborg, M. C. J. Large, A. Argyros, J. Zagari, S. Manos, N. A. Issa, I. Bassett, S. Fleming, R. C. McPhedran, C. M. de Sterke, and N. A. P. Nicorovici, "Microstructured polymer optical fibre.," *Opt. Express*, vol. 9, no. 7, pp. 319–327, 2001.

[16] K. Nielsen, H. K. Rasmussen, A. J. Adam, P. C. Planken, O. Bang, and P. U. Jepsen, "Bendable, low-loss Topas fibers for the terahertz frequency range," *Opt. Express*, vol. 17, no. 10, pp. 8592–8601, 2009.

[17] W. Yuan, L. Khan, D. J. Webb, K. Kalli, H. K. Rasmussen, A. Stefani, and O. Bang, "Humidity insensitive TOPAS polymer fiber Bragg grating sensor.," *Opt. Express*, vol. 19, no. 20, pp. 19731–19739, 2011.

[18] C. Markos, A. Stefani, K. Nielsen, H. K. Rasmussen, W. Yuan, and O. Bang, "High-Tg TOPAS microstructured polymer optical fiber for fiber Bragg grating strain sensing at 110 degrees.," *Opt. Express*, vol. 21, no. 4, pp. 4758–4765, 2013.

[19] R. Islam, G. K. M. Hasanuzzaman, M. S. Habib, S. Rana, and M. A. G. Khan, "Low-loss rotated porous core hexagonal single-mode fiber in THz regime," *Opt. Fiber Technol.*, vol. 24, pp. 38–43, 2015.

[20] H. Bao, K. Nielsen, H. K. Rasmussen, P. U. Jepsen, and O. Bang, "Fabrication and characterization of porous-core honeycomb bandgap THz fibers.," *Opt. Express*, vol. 20, no. 28, pp. 29507–17, 2012.

Ultra-low Loss THz Waveguide with Flat EML and Near Zero Flat Dispersion Properties

Sharafat Ali,[1,*] Tamal Sarkar,[2] Anamika Day,[3] Md. Rabiul Hasan,[4] J. R. Mou,[4] Sohel Rana,[5] M. Islam,[1] N. Ahmed[1]

[1]Centre of Excellence, Advanced Communication Engineering Cluster, School of Computer & Communication Engineering, University Malaysia Perlis, Pauh, 02600 Arau, Perlis, Malaysia.

[2]Department of Electrical and Electronic Engineering, Bangladesh University of Engineering & Technology, Dhaka 1000, Bangladesh.

[3]University of Asia Pacific, Dhaka 1215, Bangladesh.

[4]Department of Electronics & Telecommunication Engineering, Rajshahi University of Engineering & Technology, Rajshahi 6204, Bangladesh.

[5]Department of Electrical & Electronic Engineering, Bangladesh University, Mohammadpur, Dhaka 1207, Bangladesh.

*sharafat.ali@ieee.org

Abstract—**This paper represents an ultra-low loss waveguide with flat effective material loss (EML) and near zero dispersion flattened properties suitable for THz wave transmission. A novel type of air hole position is proposed as the design of the photonic crystal fiber. Numerical analysis exhibits an EML of 0.562 dB/cm at the operating frequency of 1.0 THz and a flat EML of 0.57±0.0225 dB/cm from 1.0-2.0 THz band. The proposed porous core fiber shows a low EML of 0.473 dB/cm with negligible confinement and bending losses at 0.9 THz frequency. This waveguide also exhibits near zero flat dispersion of 0.7±0.3 ps/THz/cm in the frequency band from 1.2 to 2.0 THz.**

***Index Terms*—Porous core waveguide, THz waveguide, flat EML, flat dispersion, low loss waveguide.**

I. Introduction

Terahertz (THz) regime is generally defined in the frequency range from 0.1 to 10 THz [1]. It is also known as the 'THz gap' due to the difficulties to explore this frequency band in practical applications. This regime has already taken the focus of researchers for its many potential applications in the field of astronomy, spectroscopy, biomedical sensing, non-invasive imaging, security sensitive areas such as drugs, explosives or weapons monitoring in a non-destructive manner, hybridization of DNA and communications etc. [2]. One of the main reasons for THz technologies staying beyond the reach of industrialization and commercialization is the lack of low loss transmission waveguide for THz spectrum. Thus the existing THz systems mostly depend on the free space propagation and as a result the propagation suffers from transmitter receiver alignment related issues, uncertain absorption loss due to surrounding atmosphere etc. Numbers of possible waveguides for Terahertz (THz) propagation have been proposed in the past few years using different dielectric and metallic substances [3]. However, most of the reported waveguides in the past were unable to satisfy the primary criteria of low loss wave transmission in the THz regime [3].

Among the reported waveguides for THz wave propagation, sub-wavelength porous fiber has presented remarkable properties regarding the THz transmission for its advantages of low material absorption loss [4], tuning dispersion [5] and birefringence [6] in the THz frequency band. Initial researches on porous core photonic crystal fiber have shown that various guiding properties with lower material loss are achievable by changing the fiber structure i.e., the arrangement of the air holes [7-9]. Up to present date, many researchers have proposed different porous core PCF designs with remarkable guiding properties such as; low EML [10], near zero flat dispersion [1] and high birefringence [11]. In 2013, Liang *et al.* [1] reported a porous core fiber having a flat dispersion range of 0.17 THz with absolute dispersion variation less than 2.5 ps/THz/cm with the material absorption loss of 0.432 cm^{-1}. In 2014, Kaijage *et al.* [12] reported an octagonal porous core fiber with absorption loss of 0.05 cm^{-1} at 1 THz frequency. In 2014, Imran *et al.* [5] reported a hybrid porous core fiber design that exhibits a near zero f at dispersion of ±0.18 ps/THz/cm over 0.8 THz band with an EML of 0.056 cm^{-1}. Same year, Shaopeng *et al.* [13] investigated a triangular porous core using elliptical air holes with hexagonal cladding. Their designed porous fiber shows 0.1-5 THz broadband single-mode transmission. In 2015, Islam *et al.* reported different hybrid porous core fiber [6, 14, 15] showing low EML at 1 THz operating frequency with high birefringence [6] and near zero flat dispersion properties with negligible bending loss [14]. In 2016, Sharafat *et al.* reported a THz porous core waveguide with elliptical air holes distributed in triangular arrangement in the core region [16]. This waveguide showed a high birefringence of 7×10^{-2} for having high asymmetry in the design and also exhibited a low EML of 0.1 cm^{-1} with dispersion flattened properties. However, none of these previously reported fibers to this point have showed any flat EML response with the THz frequency. Few months ago in 2016, Sharafat *et al.* [17] reported a new concept of designing hybrid core in the porous fibers and reported a low and flat EML over 1.5-5.0 THz frequency range with large band of near zero flat dispersion. However, the proposed thickness of the porous fiber was too large to fabricate easily.

In this report, we have presented a porous fiber waveguide using a novel hybrid-core structure. Our designed fiber shows a low flat EML of 0.57±0.0225 dB/cm from 1.0-2.0 THz band. The proposed porous core fiber shows a low EML of 0.473 dB/cm with negligible confinement and bending losses at 0.9

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

THz frequency and a low EML of 0.562 dB/cm at 1.0 THz operating frequency. The proposed porous fiber exhibits a near zero flat dispersion of 0.7±0.3 ps/THz/cm in the frequency domain over 0.8 THz band. With such exclusive guiding properties, this reported hybrid core porous fiber can be strongly considered as a very good waveguide for low loss communication applications in the THz regime.

II. Novel Hybrid Core Porous PCF Design

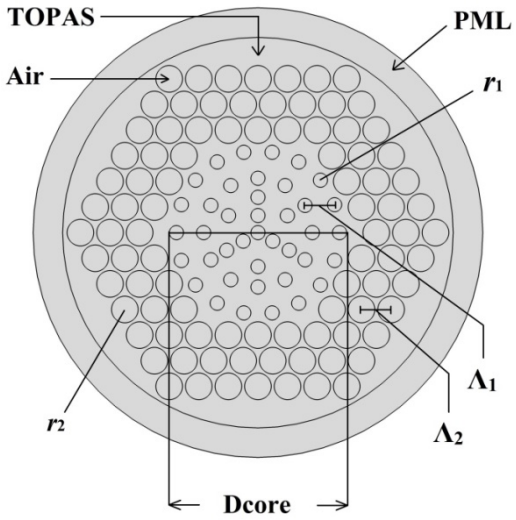


Fig. 1 Design of the proposed hybrid core porous fiber.

Fig. 1 exhibits the arrangement of air holes for the proposed hybrid core design. We have used grade 5013 TOPAS (chemically known as COC or cyclic-olefin copolymer) as the fiber host material. It has low bulk material loss of only 1.0 dB/cm at f = 1 THz frequency and a constant refractive index of n = 1.5258 in the THz band from 0.1 – 2.0 THz, which are beneficial for near zero dispersion. The hosting polymer material has chemical inertness, humidity insensitivity and also it does not absorb water.

The hybrid core design has 40 air holes with a combination of triangular and circular distributions. Two circular air hole rings are surrounding the outer region of the triangular region of the core. The radius of the core air holes is r_1=16.75 µm and pitch is denoted as Λ_1=Dcore/7, where the Dcore (450 µm) is the diameter of the core. The cladding has three hexagonal rings and here the air hole radius is r_2=40.3 µm and pitch is Λ_2=92.7 µm. The cladding of the fiber has 90 air holes in three hexagonal rings together. The air holes have been distributed closely in the cladding so that the light transmission gets a tight confinement inside the hybrid core. The designing parameters of the core and cladding region result in a high air filling ration and low refractive index around the core area. This reduces the material absorption loss or EML. The proposed hybrid core of the porous fiber introduces new type of geometrical sequence of air hole distribution in the core region and it shows a large near zero flat dispersion property with low losses.

III. Guiding Properties of the Proposed Waveguide

The guiding properties of wave transmission and different modes are investigated by using COMSOL Multiphysics 4.2. A full vector Finite Element Method (FEM) with Perfectly Matched Layer (PML) of 13.8% thickness of the fiber radius is used for the numerical analysis. Previous reports on the porous core waveguides have shown that a lower EML can be achieved with higher core porosity [14, 16, 18]. In this design the porosity is taken at 22.17% due to higher porosity than this value will result in larger air holes in the core and create challenges for the fabrication process. We have investigated the design for different core diameters to achieve the lowest material at 22.17% porosity for single mode propagations. The material loss or EML (α_{eff}) can be expressed by [16]:

$$\alpha_{eff} = \frac{\alpha_{mo}}{\alpha_{mat}} = \frac{1}{2}\left(\frac{\varepsilon_0}{\mu_0}\right)^{1/2} \frac{\int_{mat} n_{mat} |E|^2 dA}{\left|\int_{all} S_z dA\right|} \qquad (1)$$

Here in this equation (1), α_{mo} is fundamental mode loss, α_{mat} is material absorption loss, ε_0 is the permittivity and μ_0 is the permeability of free space, n_{mat} is the refractive index and s_z is pointing vector of z - component. Here $s_z = ½\ \vec{E}\times\vec{H}.z$ where E is the electric H is the magnetic field.

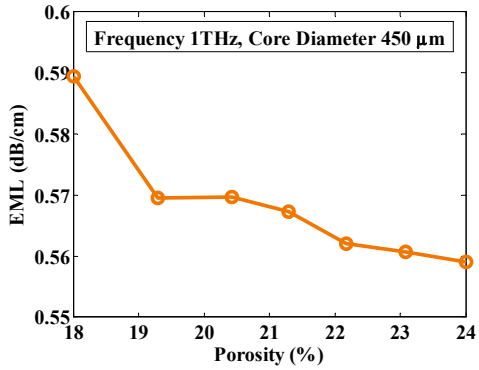


Fig. 2 EML at different core porosities for the proposed design.

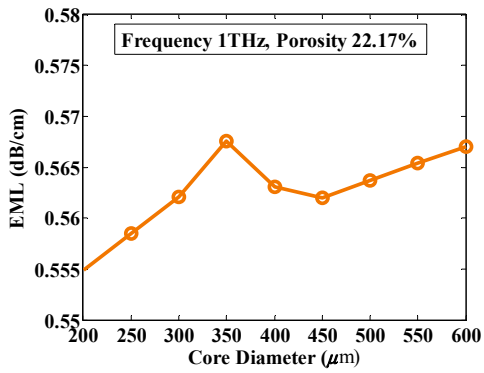


Fig. 3 EML at different core diameters for the proposed design.

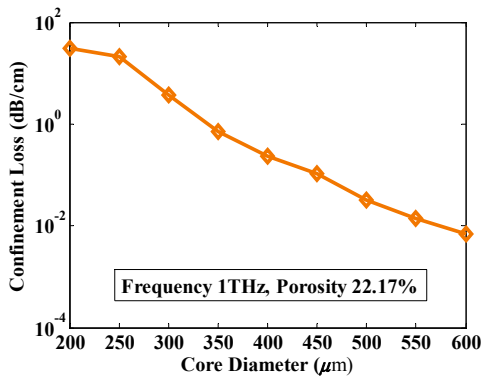


Fig. 4 Confinement losses at different core diameters for the proposed design.

For optimal design parameters 450 µm is chosen as core diameter (*D*core). Fig. 2 shows the EMLs at 450 µm *D*core and 1 THz frequency for different core porosity. As said earlier, it can be seen that lower EMLs are found at higher porosities. But for the ease of the fabrication process, the porosity is taken 22.17%. Fig. 3 shows the EML values for different core diameters at 22.17% core porosity for 1THz operating frequency. It can be seen from Fig. 3 that the lower EMLs are found at 450 µm and below 300 µm core diameters.

Usually porous core fibers have three losses that are mainly concerning, namely: EML, confinement loss and bending loss. The conf nement loss of the fiber, *Lc* can be expressed by [16]:

$$\alpha_{CL} = 8.686\frac{2\pi f}{c}\mathrm{Im}(n_{eff}) \qquad (2)$$

Where f is the frequency of the light, constant c is the speed of the light in vacuum and $\mathrm{Im}(n_{\mathrm{eff}})$ is the imaginary part of the refractive index of the guided mode. The equivalent refractive index for the bending of fiber is described as [16]:

$$n_{eq}(x,y) = n(x,y)(1+\frac{x}{R}) \qquad (3)$$

Here $n_{\mathrm{eq}}(x,y)$ is the modified equivalent index distribution, $n(x,y)$ *is* the original refractive index, R is the bending radii and x is the distance from the centre of the fiber to the bending point. At first the equivalent refractive index is determined for a particular bending radii from equation (3) and then the loss is calculated from equation (2) using the value of n_{eq}.

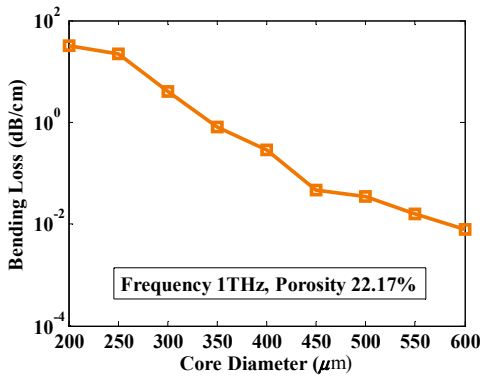


Fig. 5 Bending losses at different core diameters for the proposed design.

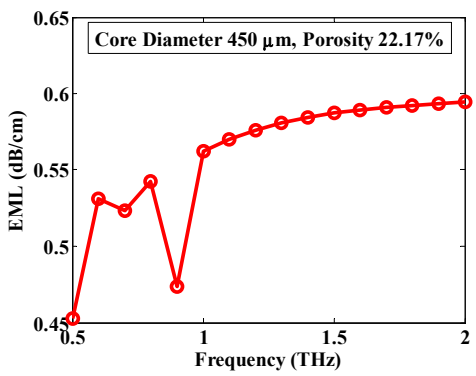


Fig. 6 EML as a function of frequency of the proposed design.

Fig. 4 and 5 show the confinement and bending losses of the proposed porous fiber at 22.17% core porosity. Judging from these results, it is very clear that, a negligible confinement and bending loss can be achieved for the core the core diameters higher than 350 µm. Thus, 450 µm core diameter and 22.17% core porosity are chosen as optimum design parameters.

The proposed hybrid core porous fiber shows an EML of 0.473 dB/cm at 0.9 THz frequency and an EML of 0.562 dB/cm at 1.0 THz operating frequency (Fig. 6). Fig. 7 shows that a 1 THz band of flat EML from 1.0-2.0 THz frequencies at 0.57±0.0225 dB/cm is resulted for the proposed porous design.

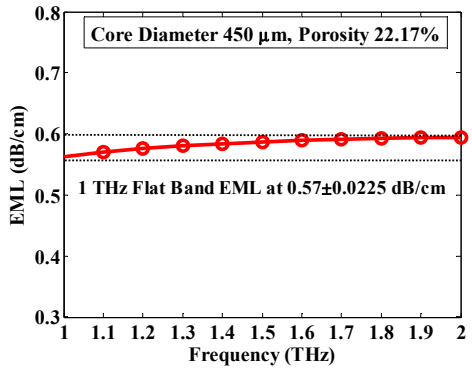


Fig. 7 Flat EML response of the proposed novel waveguide design.

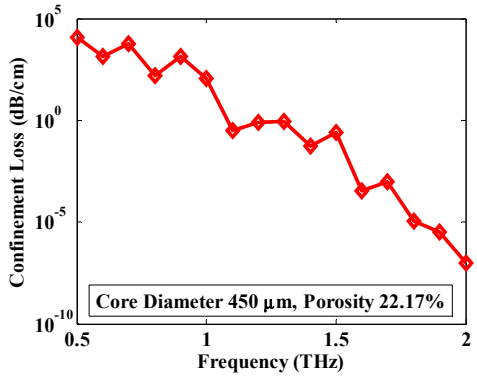


Fig. 8 Confinement loss as a function of frequency of the proposed design.

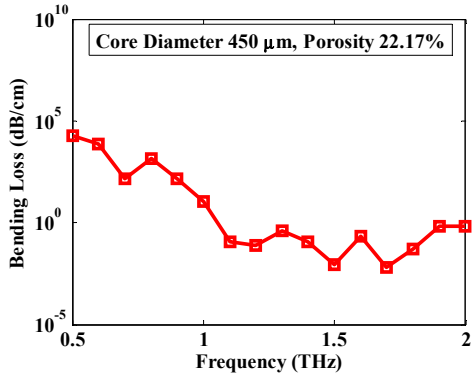


Fig. 9 Bending loss as a function of frequency of the proposed design.

Fig. 8 and 9 shows the confinement loss and bending loss of the reported fiber for the frequency range of 1.0-2.0 THz respectively. It is seen that the reported waveguide exhibits a negligible confinement and bending loss after 1 THz operating frequency. Fig. 10 shows the fraction of power through core air holes, cladding air holes and material region of the reported design. The highest power flow through the core region air holes (close to 60%) is at 0.9 THz operating frequency where the lowest EML response is calculated. From 1.0-2.0 THz frequencies, the core air holes show flatness in the power fraction which is responsible for the flat EML response. The lower power fraction in the cladding air holes resulted in a low confinement loss of the reported porous fiber.

The material dispersion is negligible for TOPAS material for 0.1-2.0 THz frequency bands [16, 17]. So we focused more on the physical structure of the waveguide design. Thus only the waveguide dispersion has been analysed for the proposed fiber. In this case, the dispersion of group velocity with respect to the frequency can be analysed by using the coefficient β_2 (ps/THz/cm). This coefficient corresponds to the second order term in the Taylor expansion of the modal propagation constant β. Here, β_2 can be expressed by [17, 18]:

$$\beta_2 = \frac{2}{c}\frac{dn_{eff}}{d\omega} + \frac{\omega}{c}\frac{d^2 n_{eff}}{d\omega^2} \qquad (4)$$

Where, $\omega=2f\pi$, and $n_{\text{eff}}=\text{Re}(\beta)c/\omega$, c is the velocity of light. Fig. 11 shows the dispersion properties of the reported hybrid core porous fiber. The fiber shows a near zero flat dispersion of 0.7±0.3 ps/THz/cm in the frequency range from 1.2-2.0 THz i.e., 0.8 THz flat band. This frequency band exhibits better dispersion characteristics with low losses for THz frequencies which is very beneficial to communication applications.

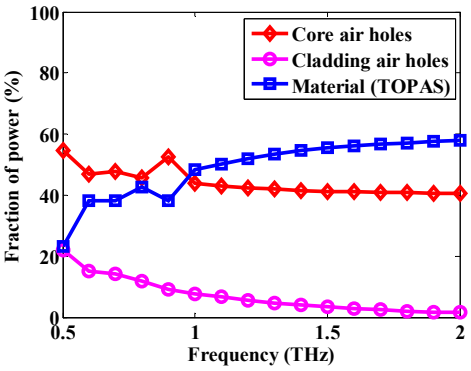


Fig. 10. Fraction of power of the reported hybrid core porous design.

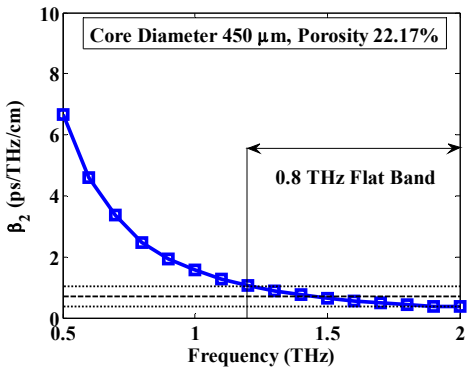


Fig. 11 Dispersion characteristics of the reported porous waveguide.

The practical fabrication process of porous core waveguides is very similar to that of PCFs in a manner though fabrication of fibers with high porosity ratio in core is a challenge. The preform fabrication can be done in several ways like drilling the hole pattern, staking of capillary tubes, casting into a mould, extrusion technique etc. [17]. The proposed porous fiber is consisting only circular shaped air hole which is a large advantage for practical fabrication. Considering the air hole shape, size and distribution in the reported waveguide, a capillary stacking method or also sol-gel method of fabrication will be proper for fabricating this hybrid core porous fiber [17].

IV. Conclusions

A novel hybrid core porous core fiber with ultra-low loss properties is proposed in this paper. The reported waveguide shows a low and flat EML of 0.57±0.0225 dB/cm from 1.0-2.0 THz band. The proposed waveguide exhibits an EML of 0.562 dB/cm at 1.0 THz operating frequency. At 0.9 THz frequency, the fiber core air holes transmit almost 60% of the total power flow and show a very low EML of 0.473 dB/cm with negligible confinement and bending losses. Numerical analysis shows a near zero flat dispersion of 0.7±0.3 ps/THz/cm in the frequency range from 1.2-2.0 THz with low losses. After observing the exclusive guiding properties and novel design methodology, it can be concluded that the proposed porous fiber is a noteworthy waveguide for THz wave propagation.

References

[1] J. Liang, Liyong Ren, Nana Chen, Changhe Zhou, "Broadband, low-loss, dispersion flattened porous core photonic band gap fiber for terahertz (THz) wave propagation," *Optics Communications,* vol. 295, pp. 257-261, 2013.

[2] B. Ung, Anna Mazhorova, Alexandre Dupuis, Mathieu Rozé, and Maksim Skorobogatiy, "Polymer microstructured optical fibers for terahertz wave guiding," *Optic Express,* vol. 19, p. B848, 2011.

[3] S. Atakaramians, Shahraam Afshar, Tanya M Monro, Derek Abbott, "Terahertz Dielectric Waveguides," *Advances in Optics and Photonics* vol. 5, pp. 169-215, 2013.

[4] S. Kaijage, Zhengbiao Ouyang, Xin Jin, "Porous-Core Photonic Crystal Fiber for Low Loss Terahertz Wave Guiding," *IEEE Photonic Technology Letters,* vol. 25, pp. 1454-1457, 2013.

[5] M. I. Hasan, S. M. Abdur Razzak, G. K. M. Hasanuzzaman and Md. Samiul Habib, "Ultra-Low Material Loss and Dispersion Flattened Fiber for THz Transmission," *IEEE Photonic Technology Letters,* vol. 26, pp. 2372-2375, 2014.

[6] R. Islam, S. Habib, GKM. Hasanuzzaman, R. Ahmad, S. Rana and S.F. Kaijage, "Extremely High-Birefringent Asymmetric Slotted-Core Photonic Crystal Fiber in THz Regime," *IEEE Photonic Technology Letters,* vol. 27, pp. 2222-2225, 2015.

[7] A. Hassani, A. Dupuis, and M. Skorobogatiy, "Low loss porous terahertz fibers containing multiple subwavelength holes," *Applied Physics Letters,* vol. 92, pp. 071101-071103, 2008.

[8] S. Atakaramians, S Afshar V, B M Fischer, D Abbott and T M Monro, "Low loss, low dispersion and highly birefringent terahertz porous fibers," *Optic Communication,* vol. 282, pp. 36-38, 2009.

[9] H. B. Chen, D. R. Chen, and Z. Hong, "Squeezed lattice elliptical-hole terahertz fiber with high birefringence," *Applied Optics,* vol. 48, pp. 3943-3947, 2009.

[10] K. Nielsen, H. K. Rasmussen, A. J. L. Adam, P. C. M. Planken, O. Bang and P. U. Jepsen, "Bendable, low-loss Topas fibers for the terahertz frequency range," *Optic Express,* vol. 17, pp. 8592-8601, 2009.

[11] N.-n. Chen, Jian Liang, and Li-yong Ren, "High-birefringence, low-loss porous fiber for single mode terahertz wave guidance," *Applied Optics,* vol. 52, pp. 5297-5302, 2013.

[12] S. Rana, GKM Hasanuzzaman, Samiul Habib, Shubi F. Kaijage, Raonaqul Islam, "Proposal for a low loss porous core octagonal photonic crystal fiber for T-ray wave guiding," *Optical Engineering,* vol. 53, pp. 1151071-1151074, 2014.

[13] L. Shaopeng, Hongjun Liu, Nan Huang and Qibing Sun, "Broadband high birefringence and low dispersion terahertz photonic crystal fiber," *Journal of Optics,* vol. 16, p. 105102, 2014.

[14] R. Islam, Sohel Rana, Redwan Ahmad, and Shubi Felix Kaijage, "Bend-Insensitive and Low-Loss Porous Core Spiral Terahertz Fiber," *IEEE Photonic Technology Letters,* vol. 27, pp. 2242-2245, 2015.

[15] R. Islam, Hasanuzzaman GKM, Habib Selim, Rana Sohel, Khan MAG, "Low-loss rotated porous core hexagonal single-mode fiber in THz regime," *Optical Fiber Technology,* vol. 24, pp. 38-43, 2015.

[16] S. Rana, Sharafat Ali, Nasim Ahmed, Raonaqul Islam, Syed A. Aljunid, "Ultra-High Birefringent and Dispersion-Flattened Low Loss Single-Mode Terahertz Wave Guiding," *IET Communications,* DOI: 10.1049/iet-com.2015.0629, 2016.

[17] Sharafat Ali, Nasim Ahmed, Syed Aljunid, Badlishah Ahmad, "Ultra-flat low material loss porous core THz waveguide with near zero flat dispersion," *Electronics Letters,* vol. 52, pp. 863-865, 2016.

Bruise Detection in Apples Using Infrared Imaging

Sourav Dey Roy,[1] Dipak Hrishi Das,[2] Mrinal Kanti Bhowmik[1,*] and Anjan Kumar Ghosh[3]

[1]Department of Computer Science and Engineering, Tripura University (A Central University), Suryamaninagar-799022, Tripura (W), India

[2]Department of Information Technology, Bir Bikram Memorial College, Agartala-799004, Tripura(W), India

[3]Tripura University (A Central University), Suryamaninagar-799022, Tripura (W), India

souravdeyroy49@gmail.com, dhdcse@gmail.com, *mrinalkantibhowmik@tripurauniv.in and anjn@ieee.org

Abstract— **Defects in apples cause food safety concerns touching the general public and strongly affect the commodity market. Due to the increasing incidence, the detection of bruises is a challenge now a day's especially when the bruises are not visible externally. Infrared imaging provides an important window for detection of bruises that are not visible externally. The study has been investigated on the infrared images of both fresh and bruised apples. For our investigation, a new dataset has been designed by maintaining standard acquisition protocol to improve the potentiality and accuracy of the thermograms. The contribution of the investigation includes asymmetric analysis of the acquired thermograms using first and second order statistical analysis and fractal analysis of pre-processed thermograms. Automatic detection of the bruised region was done using anisotropic diffusion and K-means clustering for investigation of the spread of the bruised region.**

Index Terms—**Food quality; Apple; Bruise; Infrared Imaging; Statistical Analysis; Fractal Analysis; Segmentation**

I. Introduction

Information regarding the quality of the food is most important before consumption. Assessing the quality parameters is a big challenge since differentiating impurities or freshness of food is difficult. Apple is one of the largely cultivated tree fruits today and is recognized as an important economic fruit worldwide. The Food and Drug Administration (FDA) has approved apples which are resistant to bruises could have unknown risks to human health [1]. Bruising is defined as damage of fruit tissue as a result of external forces which cause physical changes of texture and/or chemical changes of colour, smell and taste. The factors depending on which mechanical damages of apples occurs are soil cultivation, nutrition and weather conditions in the field during fruit growth. Owing to the physical properties of apples, the bruises are not visible externally [2]. The parameters which cannot be well detected by visual inspection and physical separation because of affected by psychological factors such as fatigue or acquired habits, there is a high risk of human error in classification processes. This is one of the most important drawbacks that can be prevented by automated inspection systems based on computer vision. It is a challenging task to detect bruises on the fruit which usually take place under the fruit skin. Infrared Thermography (IT) offers a potential non-invasive, non-contact and radiation free imaging modality for assessment of abnormal infrared radiation from objects. Up to 100% of apple bruises were detected using thermal imaging by discriminating surface temperature between bruised and non-bruised tissues [3]. There are generally two approaches for non-destructive bruise detection using infrared imaging: passive and active [4]. Passive thermography consists of measurement of infrared radiation from the objects under natural conditions between studied objects and surrounding air. On the other hand, active thermography consist of measurement of the infrared radiation after the object under study is heated by means of external source.

The paper mainly explores a non-invasive infrared imaging based automatic bruise detection using passive thermography. In our study, we have created our own dataset of bruise and fresh apples by following the standard thermogram capturing procedure. The acquisition setup is established at the Biomedical Infrared Image Processing Laboratory of Computer Science and Engineering Department of Tripura University (A Central University). The paper emphasizes on developing a statistical feature based analysis that can signify the presence of thermal asymmetry in the fresh and bruised apples. The developed method involves the extraction of eleven statistical features including both the first order and second order features. The structural asymmetry between both the types of apples (i.e. fresh and bruised apples) are analyzed using box counting method of fractal dimension. Automatic segmentation of the bruised region is carried out using anisotropic diffusion followed by k-means clustering and the spreadness of this area is measured in terms of area and fractal dimension.

The whole paper is organized as; Section II describes the literature survey on the computer aided techniques for detection of bruises in fruits. Section III elaborately describes the capturing procedure of the dataset used in the analysis. In Section IV methodology for asymmetric analysis and automatic bruised detection in the thermograms of apples has been illustrated. Section V reports the experimental results with discussion of the methodology. And finally, section VI concludes the paper.

II. Related Work

Detection of mechanical defects in apples plays a crucial role for quality inspection systems. In the recent years, the research communities have sparked off thunder for developing computer aided techniques to advance this area in terms of quality inspection and safety detection, and sorting of food. Some of these existing works are shown in TABLE I. Undergoing a number of research works, it has been found that detection of bruises is critical both to researchers and industry personnel in their work for development of the procedures to reduce this serious cause of loss. From the survey work it has been noticed that most of the researchers have concentrated on using hyperspectral imaging and very few work has been investigated on the infrared imaging based detection of bruises on apples. So in our work, we have created infrared imaging dataset of both bruised and non-bruised apples and performed statistical analysis for detecting the presence of asymmetry between both the types of apples.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

TABLE I
SURVEY TABLE OF DIFFERENT COMPUTER AIDED TECHNIQUES FOR QUALITY INSPECTION OF APPLES

Author/ Year	Method Used	Type of Image	Observation	Accuracy/ Result
J. Xing et.al./ 2005 [5]	Principal Components Analysis (PCA), Classification algorithm based on moments thresholding	Multi-spectral Imaging System	Found that among selected four bands of principal components (PCs), second and third PCs are suitable for detecting the presence of bruises in apples	86% accuracy
B.L. Upchurch et.al./ 1994 [6]	Feature extraction based on contrast of the images	Near-Infrared (NIR) Imaging	By comparing the contrast between two concentric circular profiles of bruised and unbruised areas, the changes in the NM reflectance were characterized.	Not Provided
J. Varith et.al. / 2013 [3]	Thermal Diffusivity	Thermal Imaging	The distinct temperature differences between bruised and sound tissues was due to the changes in thermal diffusivity.	Not Provided
B.S. Bennedsen et.al./ 2005 [7]	Thresholding Based Segmentation, Artificial Neural Network (ANN), Principal Components (PCs)	Near-Infrared (NIR) Imaging	The ability of the methods to find individual defects and measure the area ranged from 77 to 91% for the number of defects detected, and from 78 to 92.7% of the total defective area.	Bruise Detection Rate of 77 and 91.6% , with 1.1–3.6% false positives.
J. Xing et.al./ 2007 [8]	PCA, Contour Fill Algorithm, Morphological Operation	Hyperspectral Imaging	The stem-end regions were identified from the cheek surfaces from the contour features of the first PC score images and none of the sound tissue was misclassified	98.33% classification accuracy
G. ElMasry et.al./ 2008 [9]	Partial Least Squares method , Discrimination analysis	Hyperspectral Imaging	Based on the spectral data analysis, three wavelengths (750, 820, and 960nm)were choosen and observed that the selected group of wavelengths has a great discrimination power for bruise detection in different background colors.	93.95% of the variance between normal and bruised spectral data.
P. Baranowski et.al./ 2009 [4]	Pulsed-Phase Thermography (PPT), Thresholding based Segmentation	Thermal Imaging	The investigations have shown that active thermography is useful for detection of early bruises in apples.	Not Provided
P. Baranowski et.al./ 2012 [10]	PCA, Minimum Noise Fraction (MNF), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), Soft Independent Modelling of Class Analogy (SIMCA)	Hyperspectral Imaging and Thermal Imaging	They observed that Thermal MWIR imaging indicate the best prediction efficiency for bruise recognition.	LDA: 95% accuracy SVM: 92% accuracy SIMCA: 67% acccuracy

Also, structural analysis of the bruised apples are carried out by segmenting the bruised areas and measuring the spreadness in terms of area and dimension.

III. DATA COLLECTION

A. Environmental Condition and Camera Setup for Thermogram Acquisition

There are several factors need to be considered while the acquisition of thermograms that can influence the analysis and evaluation of infrared images. Practically it is difficult to control all the factors are very much difficult, but being aware of those factors is essential in many context. The factors that affects the acquisition of the thermograms are:

- Specification of Infrared Camera
- Environmental Condition
- Position of Infrared Imaging System and Objects

In our study, the protocol which has been followed to create a controlled environment in order to decrease the negative influence of the above mentioned three factors is shown in Fig.1. The camera has been mountain on a tripod stand with an alignment of 90° in between the camera and the area of imaging as shown in Fig. 2. To eliminate external infrared energy emitted from sources such as electric wires, pipes, outlets, etc., we made a background by black cloth which has been positioned behind the object. The size of the background has been fixed based on the Field of View (FOV) of the infrared camera and distance between the object and the camera.

B. Thermogram Dataset Description

Maintaining the above mentioned standard procedure, we have acquired the thermograms of apples in two different parts. In the first part, 50 fresh apples are bought from the market and thermograms of those apples are captured by stabilizing them at room temperature. After this, the apples are

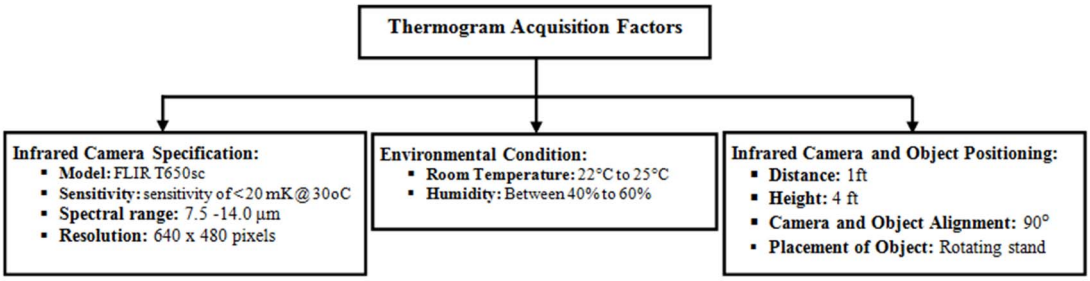


Fig. 1 Factors Affecting Thermogram Acquisition Standard

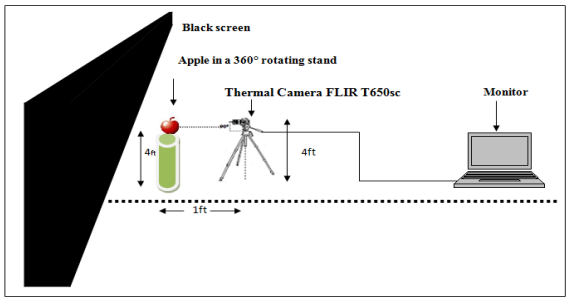


Fig. 2 Camera Setup and Positioning of the Object during Thermogram Acquisition

allowed to bruised by keeping them at normal temperature for seven days and again the thermograms of those apples are captured by maintaining the same procedure the thermograms of the similar apples are captured on the twenty first day. In the second part, 20 bruised apples that are almost rotten inside but are not visible externally are bought from the market as per instruction of the vendors and thermograms of those apples are captured by placing them in the rotating stand so that their bruised areas are detected. Some of the captured sample images are shown in Fig. 3. Here the first part of the database was used for asymmetric analysis of the thermograms and the second part of the database was used for automatic segmentation of the bruised area and to measure the spreadness of this area.

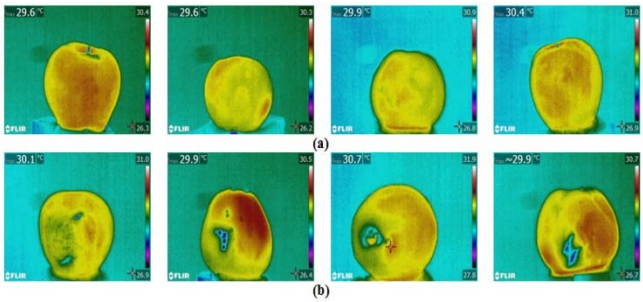


Fig. 3 Sample Images of Captured Thermograms (a) Fresh Apples; (b) Bruised Apples

IV. Methodology

A. Pre-processing of the thermograms

In the pre-processing stage, the thermograms of apples are cropped manually to a standard resolution of 200×200 pixels so as to remove all the unnecessary parts except the object of interest. Then these cropped images are converted into gray scale image for further analysis so as to extract statistical and structural features and to segment the bruised areas from the thermograms as described in section B and C.

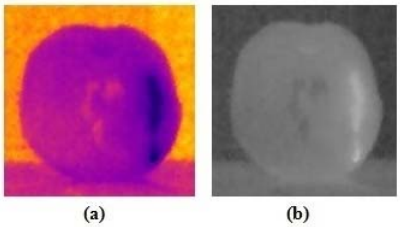


Fig. 4 Pre-Processing of the Thermograms (a) Original Image; (b) Pre-Processed Image

B. Asymmetric Analysis of the Thermograms

In this paper, the extraction of statistical features using first and second order statistical analysis and fractal dimension for asymmetry analysis between bruised and fresh apples is conducted based on the evaluation of intensity distribution.

1) First Order Statistical Feature Extraction: The first order statistical features are also known as the Gray Level Histogram Based Statistical Features as they are directly computed from probability of occurrence of the gray levels in the images [12]. For first order statistical analysis we have calculated mean, variance, standard deviation, skewness, kurtosis and entropy of the thermograms.

2) Second Order Statistical Feature Extraction: The second order statistical features are extracted from the Gray Level Co-occurrence Matrix (GLCM) that deals with the relative positions of the gray levels within an image [11]. On the other hand inter-pixel distance, orientation and the spatial resolution of an image also plays a crucial role in the computation of co-occurrence matrices. For second order statistical analysis we have calculated contrast, homogeneity, energy, entropy and correlation of the thermograms.

3) Fractal Feature Extraction using Box Counting method: Fractal features can be effective in analyzing the difference in textures. For asymmetric analysis between bruised and fresh apples, fractal dimension is calculated using the most widely used box-counting technique that quantifies the space filling capacity of a fractal pattern. In box counting method, fractal dimension is the slope of the line when we plot the value of log(N) on the Y-axis against the value of log(r) on the X-axis. Here N is the number of boxes that cover the pattern, and r is the magnification or the inverse of the box size. The equation to denote the fractal dimension is given by [12]:

$$D=\frac{\log(N)}{\log(r)} \tag{1}$$

C. Automatic Segmentation of the Bruised Region

Segmentation of the bruised region plays a significant role in the analysis of the infrared images. However, automated detection of bruised region from images is a very challenging task due to the lack of clear limits in these images. From this viewpoint, we have taken two ideas, i.e., anisotropic diffusion and K-means clustering for extraction of the bruised region from the whole thermograms. The methodology for automatic segmentation of the bruised region of the acquired thermograms of apples has been illustrated in Fig. 5. The brief details of all the steps are described below.

1) Smoothing using Anisotropic Diffusion: For accurate extraction of the Region of Interest (ROI), the pre processed images are smoothened using anisotropic diffusion filtering to avoid the blurring and localization problem in the acquired thermograms. The discretized equation of Perona and Mallik anisotropic diffusion is as follows [13]

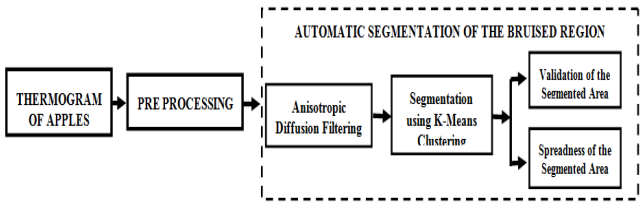


Fig. 5 Block Diagram of the Proposed Methodology for Automatic Segmentation of the Bruised Region

TABLE II
ASYMMETRIC ANALYSIS OF CAPTURED THERMOGRAMS IN FRESH CONDITION AND THERE AFTER BRUISING THEM AFTER 7 AND 21 DAYS

Sample No	First Order Statistical Features						Second Order Statistical Features				
	Mean	Variance	Standard Deviation	Skewness	Kurtosis	Entropy	Contrast	Correlation	Energy	Homogeneity	Entropy
Statistical Feature Analysis of Thermograms Captured on the First Day											
1	94.80	4407.64	66.39	0.13	1.40	0.0027	0.1197	0.1667	0.9305	0.9012	0.0027
2	92.73	4235.91	65.08	0.05	1.35	0.0019	0.1224	0.1641	0.9332	0.8971	0.9937
3	85.65	4525.2	67.27	0.41	1.40	0.0044	0.1029	0.2204	0.9360	0.9077	0.7856
4	81.15	4246.1	65.16	0.44	1.39	0.0039	0.1304	0.2312	0.9359	0.9050	0.6253
5	82.94	4094.0	63.98	0.20	1.41	0.0035	0.1034	0.2337	0.9452	0.9315	0.8960
Statistical Feature Analysis of the Similar Thermograms after Bruising them in Normal Temperature after 7 Days											
1	106.12	4861.1	69.72	0.14	1.49	0.0036	0.1710	0.3408	0.9435	0.9094	0.8113
2	114.18	6342.31	79.64	0.34	1.40	0.0077	0.1302	0.2499	0.9381	0.9373	0.0077
3	117.72	6525.73	80.78	0.43	1.42	0.0047	0.1306	0.3009	0.9409	0.9284	0.0047
4	112.85	6231.39	78.94	0.31	1.43	0.0042	0.1321	0.2788	0.9384	0.9274	0.0042
5	111.42	5915.47	76.91	0.35	1.44	0.0052	0.1130	0.2585	0.9471	0.9319	0.0070
Statistical Feature Analysis of the Similar Thermograms after Bruising them in Normal Temperature after 21 Days											
1	123.44	7127.06	84.42	0.49	1.58	0.0044	0.1886	0.3409	0.9439	0.9553	0.9284
2	121.34	6866.87	82.87	0.48	1.51	0.0099	0.1489	0.3240	0.9433	0.9530	0.9130
3	120.93	6918.83	83.18	0.51	1.57	0.0047	0.1548	0.3259	0.9518	0.9437	0.9652
4	124.39	7112.50	84.34	0.48	1.57	0.0056	0.1685	0.3055	0.9394	0.9389	0.9422
5	122.09	6936.61	83.29	0.41	1.53	0.0070	0.1219	0.3156	0.9501	0.9452	0.9284

$$I_{(t+1)}(s) = I_t(s) + \frac{\lambda}{|n_s|}\sum_{p\in N_s} g_k\left(e^{-\left(\frac{|\nabla I_{s,p}|}{k}\right)^2}\right)(I_t(P) - I_t(S)) \quad (2)$$

Here, I is the discretely sampled image, s denotes the pixel position in the discrete 2D grid, t denotes the iteration step, g is the conduction function, k is the gradient threshold parameter, λ is constant and $\lambda \in [0,1]$, n_s denotes the no of spatial neighbours considered for the pixel s.

2) *Segmentation of the bruised region:* To calculate the spreadness of the bruised region, segmentation of that region from the whole thermograms is carried out using k-means clustering [14]. On the basis of analysis of the pre-processed thermograms it has been found that there are ten number of clusters depending on the color palette division of temperature distribution and maximum intensity is present in the bruised area. So the main aim of this algorithm is to minimize the objective function given by:

$$J = \sum_{j=1}^{k}\sum_{i=1}^{n}\left\|x_i^{(j)} - c_j\right\|^2 \quad (3)$$

Here $\left\|x_i^j - c_j\right\|^2$ is a distance measure between a data point $x_i^{(j)}$ and the cluster centre c_j.

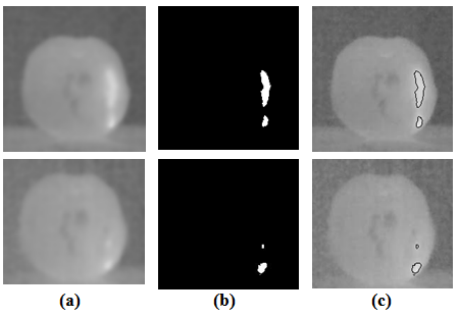


Fig. 6 Smoothening and Segmentation of the Bruised Area (a) Smoothened Image; (b) Segmented ROI; (c) Mapping of the Segmented ROI in the Original Image

V. EXPERIMENTAL RESULTS AND DISCUSSIONS

A. Statistical Feature analysis of thermograms

The statistical feature analysis was carried out on the first part of the database i.e. on 50 samples. After extracting these 11 feature values from both fresh (i.e. thermograms captured on the first day) and bruised (i.e. thermograms captures after 7 and 21 days) apples, an analysis of the feature values are done to find out any asymmetric characteristics. The first order statistical feature values obtained from 5 randomly selected thermograms of both the category of apples are shown in TABLE II. From this analysis it has been found that in most of the cases the statistical feature values of the fresh apples is less than that of the apples bruised externally after 7 days and less than that of 21 days. Among all first order features mean, standard deviation and variance shows a better discrimination power but the significant difference of skewness, kurtosis and entropy is less. It has been found that the mean temperature value is less than 100 in case of all the fresh apples whereas the value is more than 100 in case of the bruised apples. After analyzing the first order feature values, the second-order features are also extracted and analyzed in both the category of thermograms as shown in TABLE II. Although, there is a asymmetric difference between the fresh and bruised thermograms, but these differences are not decisive. However among all these second order features it has been found the value of homogeneity and correlation shows better discrimination ability between the fresh and bruised thermograms of apples but there is no such asymmetric difference in the values of contrast. Thus, from the statistical feature based asymmetry analysis of the thermograms, it has been observed that intensity based histogram plays a crucial role in asymmetric analysis of the thermograms.

B. Fractal features analysis of the thermograms

The fractal feature analysis of the thermograms has been done for asymmetric analysis between fresh and bruised apples. Because of the irregularity and heterogeneity present in the bruised apples they show higher values of fractal dimension. The analysis has been carried out on 10 fresh apples and 10 bruised apples and it has been observed that fractal feature values are higher in case of bruised apples than fresh apples as shown in TABLE III. In the given data it can be analyzed that the fractal dimension is higher, for example, less than 1.25 in case of fresh apples whereas more than 1.25 in case of bruised apples.

C. Objective Validation of the Extracted Bruised Region:

After extraction of the bruised region from the whole thermogram, the objective evaluation of the segmented bruised regions is carried out based on the four metrics i.e.

TABLE III
FRACTAL ANALYSIS AND OBJECTIVE VALIDATION AND SPREADNESS MEASUREMENT OF EXTRACTED BRUISED REGION

Sample No	Fractal Analysis of the Thermograms		Objective Validation of the Extracted Bruised Region				Spreadness Measurement	
	Fresh Apple	Bruised Apple	Probability Rand Index (PRI)	Global Consistency Error (GCE)	Variation of Information (VOI)	Boundary Displacement Error (BDE)	Area	Fractal Dimension
1	1.2252	1.2760	0.9665	0.0020	0.1200	0.0130	558	0.8293
2	1.2076	1.2679	0.9751	0	0.1529	0.0530	155	0.7386
3	1.2253	1.2566	0.9776	0.0056	0.1779	0.0177	2045	0.9423
4	1.2045	1.2715	0.9793	0.0053	0.1113	0.0790	1798	0.9154
5	1.1947	1.2778	0.9879	0.0067	0.1454	0.0167	2223	0.9629
6	1.1917	1.2941	0.9841	0.0045	0.1404	0.0444	5178	1.0507
7	1.2432	1.2758	0.9786	0.0017	0.1489	0.0119	1713	0.9072
8	1.2474	1.2935	0.9915	0.0019	0.0705	0.0782	512	0.8107
9	1.2367	1.2796	0.9863	0.0034	0.1227	0.0835	1043	0.8653
10	1.1230	1.2989	0.9553	0.0037	0.1353	0.0567	1763	0.9083

Probability Rand Index (PRI), Variation of Information (VoI), Global Consistency Error (GCE) and Boundary Displacement Error (BDE) [15]. Among this four performance metric, in case of PRI higher probability is better whereas in case of the remaining three metric the lower distance is better.

Each of this performance metrics are measured by comparing the automated detected bruised region with the manually segmented bruised region i.e. the ground truth images that are created by the human observer. Digitized manually drawn segmented region is obtained using the software GIMP (GNU Image Manipulation Program) [16]. The proposed results (i.e. automated extracted region) are compared with these reference ground truth and this comparison is given in TABLE III. From the objective evaluation it is clear that PRI is about 0.8827 on average which is almost close to 1 and other three metrics i.e. VoI, GCE and BDE on average are 0.0003, 0.1192 and 0.0397 respectively which seems to be more highly or nearly correlated with human hand segmentations i.e. the ground truth images.

D. Spreadness Measurement of the Extracted Bruised Region:

After validation of the detected ROI, the spreadness measurement of the extracted bruised region is carried out in terms of area and fractal dimension. Areas are calculated as the total number of pixels in the extracted region and fractal dimensions are calculated using the box counting method. The spreadness measurement of the extracted bruised region is shown in TABLE III. Here spreadness measurement of image no 1 and 2 are the corresponding outputs of the extracted bruised region shown in Fig. 5. So we can see that since area of the bruised region in image 1 is more as compared to image 2 so its corresponding fractal dimension is also more and with the increase of the bruised region its corresponding fractal dimension also increases as shown in TABLE III.

VI. CONCLUSION

Here we have presented a methodology for asymmetric analysis of the thermograms (i.e. bruised and fresh apples) and automatic segmentation of the bruised region and measure its spreadness. The temperature based statistical and fractal analysis shows that the feature values of the fresh apples is less than that of the bruised apples bruised. Also after automatic segmentation of the bruised region, validation of the extracted region is measured using four metric and spreadness of the extracted bruised region is measured in terms of fractal dimension.

ACKNOWLEDGMENT

The work presented here is being conducted in the Bio-Medical Infrared Image Processing Laboratory (B-MIRD), of Computer Science and Engineering Department, Tripura University (A Central University), Suryamaninagar-799022, Tripura (W). The research work is supported by the Grant No. BT/533/NE/TBP/2013, Dated 03/03/2014 from the Department of Biotechnology (DBT), Government of India.

REFERENCES

[1] FDA approves GMO apples, potatoes that don't bruise or brown as being safe to eat. [Online]. Available: https://www.rt.com/usa/243065-fda-approves-gmo-apples/.

[2] L. Qiang, and T. Mingjie, "Detection of hidden bruise on kiwi fruit using hyperspectral imaging and parallelepiped classification," in Procedia Environmental Sciences, 2012, Vol. 12, pp. 1172-1179.

[3] J. Varith, G.M. Hyde, A.L. Baritelle, J.K. Fellman and T. Sattabongkot, "Non-contact bruise detection in apples by thermal imaging," Innovative Food Science & Emerging Technologies, Vol. 4, No . 2, pp. 211-218, June 2013.

[4] P. Baranowski, W. Mazurek, B.W. Walczak and C. Slawinski, "Detection of early apple bruises using pulsed-phase thermography," Postharvest biology and technology, Vol. 53, No. 3, pp. 91-100, Sept. 2009.

[5] J. Xing, C. Bravo, P.T. Jancsok, H. Ramon, B. J.D. Baerdemaeker, "Detecting Bruises on 'Golden Delicious' Apples using Hyperspectral Imaging with Multiple Wavebands," Biosystems Engineering, Vol. 90, No. 1, pp. 27-36, Jan. 2005.

[6] B.L. Upchurch, J. A. Throop, and D. J. Aneshansley, "Influence of time, bruise-type, and severity on near-infrared reflectance from apple surfaces for automatic bruise detection," Transactions of the ASAE, Vol. 37, No. 5, pp. 1571-1575, 1994.

[7] B.S. Bennedsen and D. L. Peterson, "Performance of a system for apple surface defect identification in near-infrared image," Biosystems Engineering, Vol. 90, No. 4, pp. 419-431, April 2005.

[8] J. Xing, P. Jancsók, and Josse De Baerdemaeker, "Stem-end/calyx identification on apples using contour analysis in multispectral images," Biosystems Engineering, Vol. 96, No. 2, pp. 231-237, Feb. 2007.

[9] G. ElMasry, N. Wang, C. Vigneault, J. Qiao and A. ElSayed, "Early detection of apple bruises on different background colors using hyperspectral imaging," LWT-Food Science and Technology, Vol. 41, No. 2, pp. 337-345, March 2008.

[10] P. Baranowski, W. Mazurek, J. Wozniak and U. Majewska, "Detection of early bruises in apples using hyperspectral data and thermal imaging," Journal of Food Engineering, Vol. 110, No. 3, pp. 345-355, June 2012.

[11] N. Aggarwal and R. K. Agrawal, "First and second order statistics features for classification of magnetic resonance brain images," Journal of Signal and Information Processing, Vol. 3, No. 2, pp. 146-153, 2012.

[12] F. Kenneth, "Fractal geometry: mathematical foundations and applications," John Wiley & Sons, 2004.

[13] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion, " IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, No. 7, pp. 629-639, July 1990.

[14] N. Shi, L. Xumin, and G. Yong, "Research on k-means clustering algorithm: An improved k-means clustering algorithm," 2010 Third International Symposium on Intelligent Information Technology and Security Informatics (IITSI), IEEE, pp. 63-67, 2010.

[15] M. Mignotte, Max, "Segmentation by fusion of histogram-based-means clusters in different color spaces," IEEE Transactions on Image Processing, Vol. 17, No. 5, pp. 780-787, May 2008.

R.W. Solomon, "Free and open source software for the manipulation of digital images," American Journal of Roentgenology, Vol. 192, No. 6, pp. W330-W334, June 2009.

Impact of Scaling Channel Length on the Performances of Nanoscale FETs

Jibesh K. Saha,[1*] N. Chakma,[2] and M. Hasan[3]
[1,2] Electrical and Electronic Engineering (EEE), Shahjalal University of Science and Technology
Kumargaon, Sylhet-3114, Bangladesh.
[3] Department of Materials Science Engineering and Commercialization, Texas State University
San Marcos, TX 78666, United States.
*jksaha14@gmail.com

***Abstract*—The investigation of short-channel effects (SCE) due to channel length reduction for four different types of n-channel FETs: Bulk MOSFET, SOI MOSFET, DG MOSFET and CNTFET are carried out in this work. Simulators are used to investigate SCEs like threshold voltage (V_{th}) roll-off, subthreshold Swing (SS) and I_{on}/I_{off} ratio. Our study shows that DG MOSFET, SOI MOSFET and Bulk MOSFET reach their scalable limit at 30 nm, 50 nm and 100 nm channel length respectively due to elevation of leakage power consumption as they exhibit rapid degradation of V_{th}, SS beyond 100mV/decade and less I_{on}/I_{off} ratio. On the contrary, CNTFET can be scaled down below 10 nm as it shows negligible SCEs with stable V_{th}, ideal SS (60mV/decade) and high I_{on}/I_{off} ratio as channel length decreases. Our numerical analysis shows CNTFET creates only 1.11% V_{th} variation whereas DG MOSFET creates 39.33% V_{th} variation. CNTFETs advantages over MOSFET make it viable for faster and enhanced applications in nanoelectronics.**

***Index Terms*—Short-channel Effects, Bulk MOSFET, SOI MOSFET, DG MOSFET, CNTFET, Threshold Voltage, Sub-threshold Swing, I_{on}/I_{off} ratio.**

I. Introduction

Scaling down the size of transistor and increasing the integration of transistor in a single chip are the themes of semiconductor technology. So the field effect transistor (FET) devices with small size, low leakage current, near ideal sub-threshold swing, higher drive current and high performance are highly demanded. The advanced semiconductor devices have been scaled down to nanometer level and the device size is further shrinking as predicted by Moores law [1]. During last four decades semiconductor device technology has been changing with a stunning speed [2]. As channel length has gone down to lower nanometer level, conventional Bulk MOSFET suffers severe performance degradation due to short-channel effects. The short-channel effects limit the scaling capability of FET. After the scaling of Bulk MOSFET has reached its limit, many new devices such as Silicon on Insulator (SOI) MOSFET, DG MOSFET etc. have been proposed and constructed to overcome device scaling difficulty. SOI MOSFETs have many numerous advantages over Bulk MOSFET such as reduced parasitic capacitance and less short-channel effects. The MOSFET device scaling might not be extended to 10nm because silicon based technology will soon reach its limit when the channel length goes down to 10 nm [3]. Therefore, search for alternatives to MOSFETs have started through research. Now the semiconductor industry is looking for different materials and devices to integrate with the present silicon based technology. Carbon nanotube based FET has become necessary due to its unique electrical properties in this region. The CNTFET is widely expected to overcome all the limitations of silicon FETs in many ways [4].

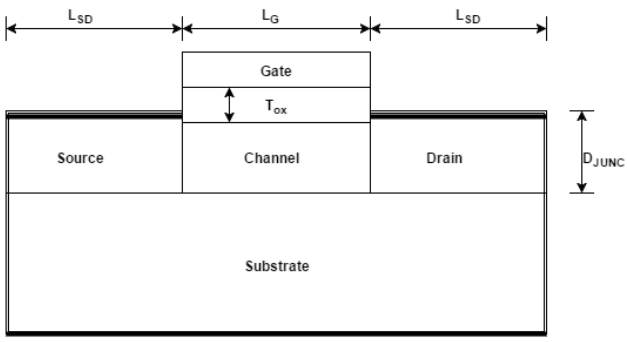


Fig. 1. Structure of a conventional n-type Bulk MOSFET

II. Device structure

A. Conventional Bulk MOSFET

Fig. 1 shows a conventional n-type Bulk MOSFET. It's three terminals are: Source, through which the carriers enter the channel. Drain, through which the carriers leave the channel and Gate, the terminal that modulates the channel conductivity. Source current, Drain current and drain to source voltage are designated by I_s, I_d and V_{DS} respectively. By applying voltage at the gate terminal we can control I_d.

B. SOI MOSFET

In SOI technology, a thin layer of active silicon is placed on top of a thick layer of insulator such as silicon dioxide

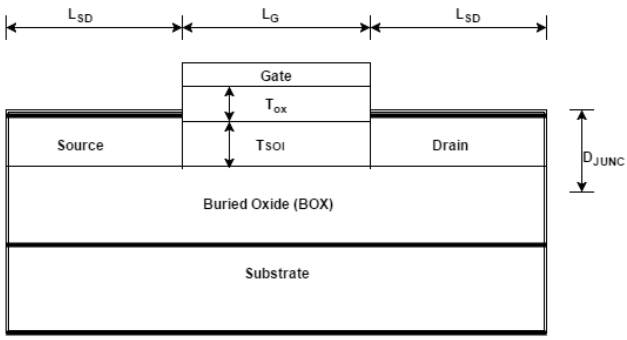


Fig. 2. Structure of a n-type SOI MOSFET

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

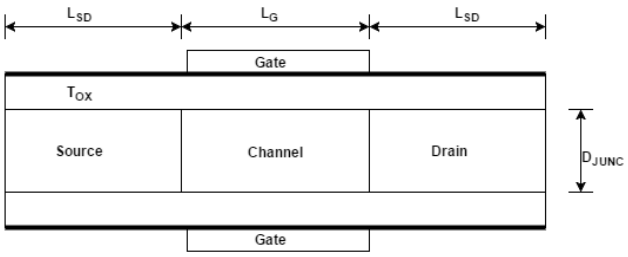


Fig. 3. Structure of a conventional n-type DG MOSFET

(SiO_2) or sometimes referred as Buried Oxide (BOX) [5] and this dielectric isolation of silicon reduces the parasitic or internal junction capacitance on components [6]. Fig. 2 shows an FDSOI MOSFET which has identical terminals as any conventional MOSFET. The buried oxide region and the junction depth is designated by (BOX) and T_{SOI}.

C. DG MOSFET

Fig. 3 shows a general structure of a DG MOSFET. The gate on the upper side is called the top gate and another one in the lower side is bottom gate. Due to these two gates the gate to channel coupling gets doubled and hence the short-channel effects drastically reduce.

D. CNTFET

CNTFET is a device consisting of a semiconducting nanotube or an array of nanotubes bridging source and drain contacts and acting as a carrier channel which is turned on or off via gate contact. The switching operation of CNTFET is controlled by the modulation of Schottky barrier formed at the nanotube and metal (drain/source) contacts [7]. The CNTFET shown in Fig. 4 is a single-walled CNTFET with intrinsic channel region and having chirality of (11,0).

III. Short-channel Effects

As channel length scaling of FET devices continues, many limitations are imposed upon the performance of the devices. These performance limiting factors are known as short-channel effects. The short-channel effects includes threshold voltage (V_{th}) roll-off , drain-induced barrier lowering (DIBL) i.e. the reduction of V_{th} induced by drain voltage and increase of subthreshold leakage current.

A. Threshold Voltage (V_{th})

$\mathbf{V}_{th}$ depends on material properties and device geometry. $\mathbf{V}_{th}$ variation is undesirable because circuit designers would like $\mathbf{V}_{th}$ to be invariant with transistor dimensions and biasing conditions. So to minimize leakage power, a stable $\mathbf{V}_{th}$ is necessary with scaled geometries. The $\mathbf{V}_{th}$ of DG MOSFET and CNTFET can be expressed by (1) and (2) respectively [8] [9].

$$V_{T_0} = \Phi_{GC} - 2 \cdot \varphi_F - \frac{Q_{B0}}{C_{ox}} - \frac{Q_{ox}}{C_{ox}} \tag{1}$$

Here, Φ_{GC} is the work function difference between the gate and the channel, $2\phi_f$ is the gate voltage necessary to change the surface potential, $\frac{Q_{B0}}{C_{ox}}$ is the gate voltage component to offset the depletion region charge and $\frac{Q_{ox}}{C_{ox}}$ is the gate voltage necessary to offset the fixed charges in the gate-oxide and Si-oxide junction.

$$V_{th} = \frac{aV_{\Pi}}{\sqrt{3}qD_{CNT}} \tag{2}$$

Here, a = 2.49 is the lattice constant, q = charge of an electron and V_{Π} is the carbon Π to Π bond energy, 3.033 eV.

B. Drain Induced Barrier Lowering (DIBL)

In long-channel devices, the potential barrier is flat along most of channel. The drain depletion region acts to reduce this barrier. In short-channel devices, boundary at the drain side of the channel becomes influential over a significant portion of the channel, in comparison with the gate contact. Thus lowering the potential barrier makes it harder for devices to fully turn off and can result in significant leakage currents in their off state. The DIBL of a FET can be expressed as (3) [10].

$$DIBL = \frac{\Delta V_{th}}{\Delta V_{ds}} \tag{3}$$

C. Subthreshold Swing (SS)

SS indicates how effectively the flow of drain current can be stopped when V_{GS} is less than $\mathbf{V}_{th}$ Small SS in devices leads to lower leakage current consumption. So small SS is required to provide a faster transition between off current (I_{off}) and on current (I_{on}). The SS for a FET can be expressed as (4) [11].

$$SS = \frac{dV_{GS}}{d(logI_{DS})} \tag{4}$$

Where V_{GS} is the gate voltage and I_{ds} is the drain current.

D. I_{on}/I_{off} Ratio

A low leakage power consumption (less I_{off}) is required to decrease static power dissipation. High on current reduces the gate switching delay. Thus a high on-off current ratio indicates improved FETs performance. The I_{on}/I_{off} ratio is expressed by (5) [12].

$$I_{on}/I_{off} ratio = \log_{10} I_{on}/I_{off} \tag{5}$$

IV. Simulation Analysis and Result

Numerous I_d-V_g curves and I_d-V_d curves of Bulk MOSFET, SOI MOSFET, DG MOSFET and CNTFET varying channel length are simulated using nanoHUB simulator tools: MOSfet, CNTFET Lab, FETToy etc. and influences of this parameter on $\mathbf{V}_{th}$, SS and I_{on}/I_{off} ratio are presented in this section. Moreover a detailed observation and reasons behind the effects of this parameter variations are also elucidated. The input parameters and different physical values of the materials of the devices are shown in TABLE I.

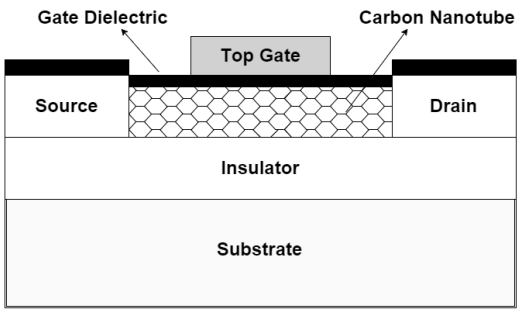


Fig. 4. Structure of a conventional n-type CNTFET

TABLE I
PARAMETERS USED IN SUMULATION

Parameters	**Bulk MOSFET, SOI MOSFET and DG MOSFET**	**CNTFET**
Channel length	Variable	Variable
Oxide thickness	2nm	2nm
Dielectric material	SiO_2	SiO_2
Dielectric constant	3.9	3.9
Source/Drain doping Concentration	$2e20/cm^3$	$2e20/cm^3$
Buried Oxide Thickness (For SOI only)	50nm	—
Chirality	—	(11,0)
Nanotube Diameter	—	0.8611nm
Temperature	300K	300K
Gate to source voltage	0.5-1 V	0.5-1 V

A. Bulk MOSFET, SOI MOSFET and DG MOSFET

The FET devices have become smaller to achieve higher packing density. The FET characteristics are changed with the variation of channel length. So it is expected that the channel length plays a key role in FET performance.

Fig. 5 shows drastic degradation of $\mathbf{V}_{th}$ in DG MOSFET below 30nm whereas SOI MOSFET and Bulk MOSFET show rapid degradation below 100nm and 200nm respectively. As the channel Length reduces, the ability of the gate electrode to solely control the charge distribution in the channel reduces and the field originating from the source/drain can influence the channel region. Thus $\mathbf{V}_{th}$ roll-off occurs.

In Fig. 6, the DG MOSFET shows SS above 100mV/decade below 40nm whereas Bulk MOSFET and SOI MOSFET exhibit comparatively high SS below 100nm and 50nm respectively. Because the sensitivity of the surface potential to gate voltage variation is evaluated by SS. The reduction of channel length produces SCE, which in turn reduces this sensitivity. Thus SS increases as channel length reduces.

In Fig. 7, The I_{on}/I_{off} ratio of DG MOSFET decreases below 70nm whereas the I_{on}/I_{off} ratio of Bulk MOSFET and SOI MOSFET decreases below 200nm and 150nm respectively. DG MOSFET shows rather satisfactory performance as channel length decreases by keeping the I_{on}/I_{off} ratio at 10^3 ranges due to its tolerence to SCE.

From these observations it is concluded that DG MOSFETs channel length can be scaled down more than SOI MOSFET and Bulk MOSFET respectively.

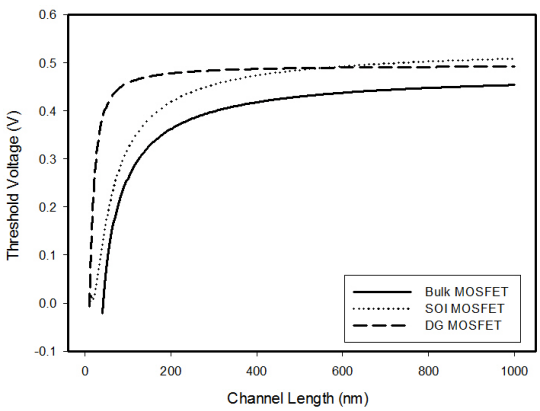


Fig. 5. V_{th} roll-off for Bulk MOSFET, SOI MOSFET and DG MOSFET

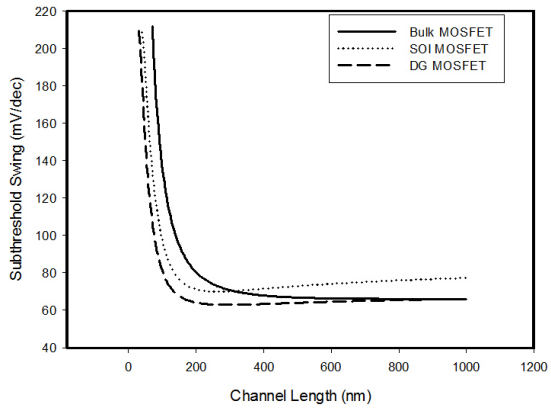


Fig. 6. SS vs channel length for Bulk MOSFET, SOI MOSFET and DG MOSFET

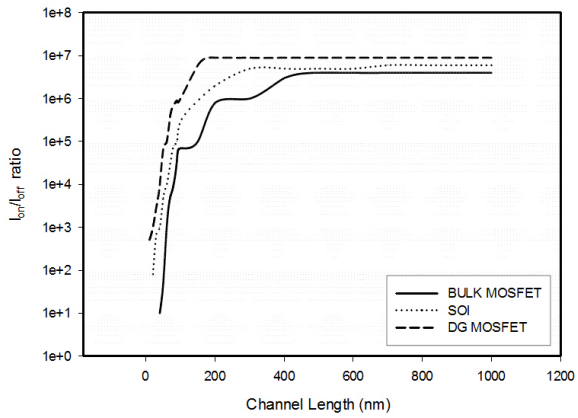


Fig. 7. I_{on}/I_{off} ratio vs channel length for Bulk MOSFET, SOI MOSFET and DG MOSFET

B. CNTFET vs DG MOSFET

Because of superior short channel performance of DG MOSFET compared to Bulk MOSFET and SOI MOSFET, we conducted our performance study to observe the SCEs only for DG MOSFET and CNTFET.

In Fig. 8, at below 10 nm, $\mathbf{V}_{th}$ is increased rapidly in CNTFET whereas in DG MOSFET $\mathbf{V}_{th}$ decreases drastically below 30 nm. For conventional MOSFETs in nanometer regime both quantum capacitance C_q and insulator capacitance C_i have

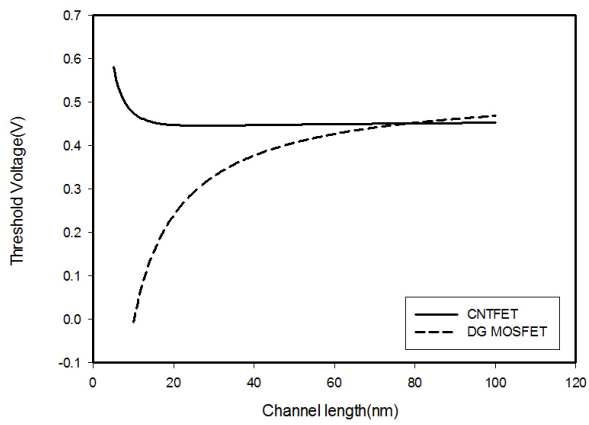


Fig. 8. V_{th} vs channel length for DG MOSFET and CNTFET

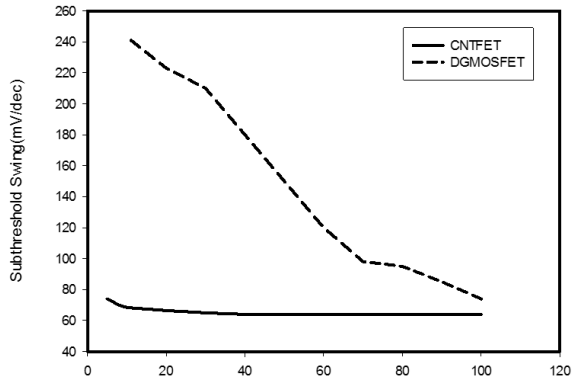


Fig. 9. SS vs channel length for DG MOSFET and CNTFET

significant impact on gate capacitance C_g and thus threshold voltage. So as C_q increases C_g increases and $\mathbf{V}_{th}$ decreases. But CNTFET has the unique property of quantum confinement at deep nanometer channel lengths. CNTFETs C_g drastically decreases with the reduction of the diameter of CNT. This in turn drastically increases the $\mathbf{V}_{th}$ with the reduction of channel length below 10nm.

Fig. 9 shows that DG MOSFET swings exponentially after 50-40nm whereas the SS of the CNTFET stays somewhat stable even after the channel length is decreased to 10nm. As there is no scattering in the channel formed using CNT and due to its excellent electrostatics and gate control, scaling of CNTFET offers much stable SS.

In Fig. 10, CNTFET shows better I_{on}/I_{off} ratio than DG MOSFET due to its non resistive ballistic conduction property of electrons at room temperature. As the scattering mechanism is totally ignored, the current density of CNTFET is much higher than DG MOSFET. The on-off ratio of CNTFET could be tuned between 10^6 and 10^9 order of magnitude, which is much higher than DG MOSFET.

C. Performance Evaluation

Summary of our findings is shown in TABLE II. Finally a CNTFET and a DG MOSFET are designed with 30nm channel length to evaluate the performance of the devices. Our numerical analysis using simulation results shows that when channel length is reduced a CNTFET creates only 1.11% V_{th} variation whereas DG MOSFET due to short-channel effe creates 39.33% V_{th} variations. Thus CNTFETs have $\mathbf{V}_{th}$ variations, higher drive current, better SS, lower leakage current and thus better scaling capability compared to DG MOSFET.

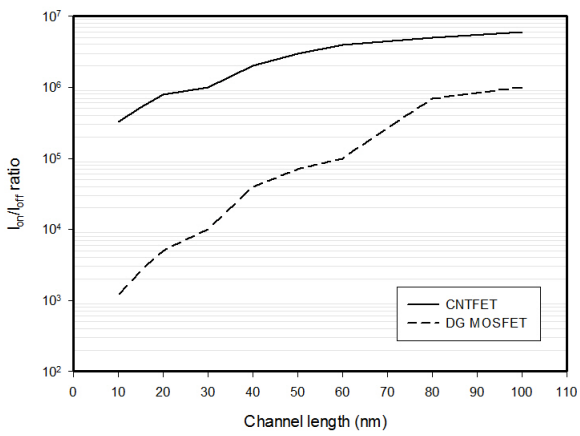


Fig. 10. I_{on}/I_{off} vs channel length for DG MOSFET and CNTFET

TABLE II
IMPACT OF SCALING CHANNEL LENGTH

FET Type	**Exhibits V_{th} roll-off below**	**SS increases above 100mV/dec. below**	**I_{on}/I_{off} ratio decreases below**
Bulk MOSFET	Lc=200nm	Lc=100nm	Lc=200nm
SOI MOSFET	Lc=100nm	Lc=50nm	Lc=150nm
DG MOSFET	Lc=30nm	Lc=40nm	Lc=70nm
CNTFET	Never (Shows rapid gradation below 10nm)	Increases but does not exceed 100mV/dec.	Lc=10nm

V. CONCLUSION

The SCEs like $\mathbf{V}_{th}$ roll-off, SS and I_{on}/I_{off} ratio which are very important issues related to performance analysis of short-channel MOSFETs were discussed in this research. Our analysis reveals that CNTFETs channel length is more scalable compared to DG MOSFET, SOI MOSFET and Bulk MOSFET respectively. Moreover the CNTFET provides zero V_{th} variation, a much less SS than DG MOSFET and thus consumes very low leakage power as channel length decreases. Thus CNTFET technology is found to offer devices with much better scalability and enhanced performance in terms of $\mathbf{V}_{th}$ variation, SS and I_{on}/I_{off} ratio compared to conventional Bulk MOSFET, SOI MOSFET and DG MOSFET thereby providing scope for further miniaturization of devices.

REFERENCES

[1] D. C. Brock and G. E. Moore, *Understanding Moore's Law*, Philadelphia, Pa.: Chemical Heritage Foundation, 2006.
[2] M. S. Chopade and M. S. Mane, "Design of DG-CNFET for reduction of short channel effect over DG MOSFET at 20nm", in *TENCON IEEE Region 10 Conference*, pp. 1−5, 2013.
[3] H. P. Wong, D. J. Frank and P. M. Solomon, C. H. J. Wann and J. Wenser, "Nanoscale CMOS," in *Proceedings of the IEEE,* 1999, vol. 87, no. 4. pp. 537−570.
[4] R. H. Baughman, A. A. Zakhidov, and W. A. de Heer, "Carbon nanotubes–the route toward applications," *Science vol. 297, no. 5582,* pp. 787−792, 2002.
[5] J.-W. Lee, H.-K. Kim, J.-H. Oh, J.-W. Yang, W.-C. Lee, J.-S. Kim, M.-R. Oh, and Y.-H. Koh, "A new SOI MOSFET for low power applications," in *Proceedings. IEEE International SOI Conference, IEEE,* 1998, pp. 65−66.
[6] S. K. Saha, "Device considerations for ultra-low power analog integrated circuits," in *4th International Conference on Computers and Devices for Communication, 2009 (CODEC 2009)*, 2009, pp. 1−6.
[7] Y-M. Lin, J. Appenzeller, J. Knoch, and P. Avouris, "High-Performance Carbon Nanotube Field-Effect Transistor with Tunable Polarities" in *IEEE Transactions on Nanotechnology,* vol. 4, no. 5, pp. 481−489, 2005.
[8] M. Zabeli, N. Caka, M. Limani and Q. Kabashi, "The threshold voltage of MOSFET and its influence on digital circuits," *Recent Advances In Systems, Communications & Computers*, pp. 6−8, 2008.
[9] Y. B. Kim, Y.-B. Kim, and F. Lombardi, "A novel design methodology to optimize the speed and power of the CNTFET circuits," in *52nd IEEE International Midwest Symposium on Circuits and Systems, MWSCAS'09*, 2009, pp. 1130−1133.
[10] H.-S. Huang, S.-Y. Chen, Y.-H. Chang, H.-C. Line, and W.-Y. Lin, "TCAD Simulation of Using Pocket Implant In 50nm N-Mosfets," *Institute of Mechatronic Engineering, National Taipei University of Technology, Taipei, TAIWAN*, 2004.
[11] S. M. Sze, and K. K. Ng, *Physics of semiconductor devices*, John Wiley & Sons, 2006.
[12] M. Stockinger, *Optimization of ultra-low-power CMOS transistors*, :Wien TU, 2000.

Optimization and Performance Analysis of PCBM Acceptor-Based Bulk Heterojunction Organic Solar Cells Using Different Donor Materials

Faysal Hakim and Md. Kawsar Alam
Department of Electrical & Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka-1205, Bangladesh
hakimfaysal@gmail.com; kawsaralam@eee.buet.ac.bd

***Abstract*— We present a comparative study on the bulk heterojunction (BHJ) organic solar cells using four different active layers: MDMO-PPV:PCBM, P3HT:PCBM, PSBTBT:PCBM and PBDTTT-C:PCBM. In this regard, firstly, we focus on optimizing the active layer thickness for different donor materials and assessing their open circuit voltages, short circuit current densities, fill factors and efficiencies. Then, we compare the photovoltaic parameters at optimized thickness for different donor materials. The study is further extended by investigating the reasons of enhanced efficiency at optimum thicknesses. It is found that PBDTTT-C:PCBM is the most efficient material among the group. The maximum efficiency for different active materials have been found at relatively lower blend thicknesses.**

***Keywords*— PCBM cell, MDMO-PPV, P3HT, PSBTBT, PBDTTT-C, BHJ.**

I. Introduction

Solar cell technology has emerged as an attractive source of renewable energy with increasing demand and scarcity of available resources. Latest researches have shown great potential for organic solar cells compared to the semiconductor based solar cells. Their flexibility, thin film architecture, properties and production methods promise a bright future for solar energy [1]. In organic photovoltaic (OPV) technology, polymer solar cells (PSC) have attained attraction in recent years. PSC devices are usually fabricated over a glass substrate (for protection) with a transparent electrode (anode). The electrode is then covered by a hole transport layer (HTL) over which the active material is deposited. Following the active layer, another electrode (with lower workfunciton than anode) is placed [1]. The active material is usually a blend composed of a conjugated polymer (donor) and small molecule acceptor. Bulk heterojunction structure, invented by Yu *et al.* [2], is one of the most promising combination of donor and acceptor materials where the blend is formed with bicontinual phase separation [2].

Fullerene and its derivatives are widely used as electron acceptors in BHJ organic solar cell. PCBM ([6,6]-phenyl-C61-butyric acid methyl ester) is a common and more practical choice as an electron acceptor than other fullerene derivatives due to its solubility in common solvent (chlorobenzene, toluene etc.) and strong absorption spectra [3]. Between two derivatives [60]PCBM and [70]PCBM, [70]PCBM usually yields higher efficiency although it has higher purification cost. The selection of donor material is, however, continuously changing with time. A wide range of donor materials, including polymers and small molecular compounds have been used over the past decades. At the very beginning of OPV research, PPV (p-phenylene vinylene) based polymers were dominant as electron donor with efficiency range of 2.5-3% [4]. Polythiophene (PT) and its derivatives were and are still very popular homoconjugated electron donors in BHJ architecture which yield 3.5-4.5% efficiency [5-6]. For achieving higher efficiencies of 5-7% [7-8], polymers with 2,1,3-benzothiadiazole (BT based polymer), benzo[1,2-b;4,5-b']-dithiophene (BDT based polymer) have been developed in recent past. There are also a number of different polymer groups (DPP etc.) which have been tried in D/A structures. Understanding the trend of change of materials and their utility in different aspects requires proper performance comparison. Most of the previous comparative studies include comparison among different inorganic materials, organic acceptor materials and supporting layers (e.g HTL) used in organic solar cells [9-11]. It is equally important to analyze the performance of different donor materials and their relative advantages and disadvantages.

In this paper, we have selected four donor materials from four different groups (PPV, PT, BT and BDT) while keeping the acceptor material fixed ([70] PCBM) for a comparative study and evaluating their performance. The donor materials are MDMO-PPV (Poly[2-methoxy-5-(3′,7′-dimethyloctyloxy)-1,4-phenylenevinylene]), P3HT (Poly(3-hexylthiophene-2,5-diyl)), PSBTBT(Poly[(4,4′-bis(2-ethylhexyl)dithieno[3,2-b:2′,3′-d]silole)-2,6-diyl-alt-(2,1,3-benzothiadiazole)-4,7-diyl]) and PBDTTT-C(Poly[(4,8-bis-(2-ethylhexyloxy)-benzo(1,2-b:4,5-b′)dithiophene)-2,6-diyl-alt-(4-(2-ethylhexanoyl)-, thieno [3,4-b]thiophene-)-2-6-diyl)]). We have simulated a conventional BHJ OSC structure with active materials mentioned above and varied their thicknesses. Once the optimized thicknesses of these materials are found we calculate device performance parameters such as short circuit current density (J_{sc}), open circuit voltage (V_{oc}), efficiency (η) etc. at those thicknesses and compare them.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

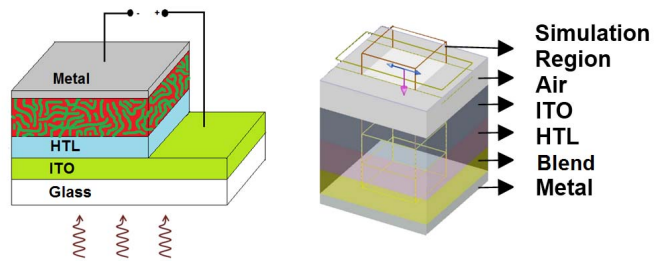
II. Modeling and Methodology

The structure we have considered is shown in Figure 1(a) (not in scale). Initially, the sun light is incident on the glass and then transmited through 110 nm thick layer of transparent conductive electrode, ITO (Indium Titanium Oxide) and a 100 nm thick hole transport layer PEDOT:PSS (poly (3,4-ethylenedioxythiophene): poly (styrenesulfonate)) to the active material. The optical power absorbed in this layer can be calculated from the divergence of pointing vector and expressed in terms of index and electric fields by Eq. (1)

$$P_{abs}(x,\lambda) = -\omega/2 \times |E(x,\lambda)^2| \times Im(\varepsilon(x,\lambda)) \quad \text{.....} (1)$$

where, Pabs, E and ε are position and wavelength dependent absorbed power, electric field and material permittivity, respectively. The position dependent generation rate profile for a particular wavelength λ, G(x, λ) can then easily be calculated from dividing the absorbed power by photon energy for each λ.

$$G(x,\lambda) = -\pi/h \times |E(x,\lambda)^2| \times Im(\varepsilon(x,\lambda)) \quad \text{.....} (2)$$



(a) (b)

Figure 1: (a) Conventional BHJ OSC structure (b) 3D setup for simulation

Finally, for broad spectrum sunlight (AM 1.5G), the position dependent generation rate can be calculated by integrating G(x, λ) over the spectrum.

$$G(x) = \int G(x,\lambda) d\lambda \quad \text{.....} (3)$$

Generated excitons are separated if they experience an internal electric field within their diffusion range. The motion of free charges (electron and hole) is described according to the continuity Eq. (4)

$$dc/dt = \pm 1/q \times \nabla J - R + G \quad \text{.....} (4)$$

where c is carrier/charge density, J is current density and R and G are recombination and generation terms, respectively. The distribution of current density needs to be calculated for deriving J-V characteristics.

In order to investigate the relationship between device performance and active layer thickness, a 2D device modelling has been done using Lumerical Inc. software in FDTD and DEVICE using "Solar Generation" and "Charge Transport Solver" module to calculate generation rate from absorption spectra and charge distribution from generation rate, respectively. We have used periodic boundary condition in XY plane and PML in the light propagating direction (Z axis). "Electrical contact" boundary condition has been employed in *J-V* analysis. The refractive index profiles (n, k), for different active materials (at optimum weight ratio) used in optical simulation, have been taken from experimental analysis [12-15]. The LUMO/HOMO levels of PCBM have been taken as -4.27 eV and -6.0 eV, respectively [5]. The electrical parameters and suitable metal electrodes used for different donor materials in the electrical simulation have been listed in Table 1.

Table 1: Electrical parameters used in simulation study

Donor Material	LUMO /HOMO (eV/eV)	Band Gap, (Donor: PCBM), E_g (eV)	Metal Electrode (Work function in eV)
MDMO-PPV	-3.2/-5.4	1.1	LiF/Al (2.6)
P3HT	-3.3/-5.3	1.0	Al (4.28)
PSBTBT	-3.27/-5.05	~0.8	Ca (2.82)
PBDTTT-C	-3.35/-5.16	~0.9	Ca/Al

III. Results and Discussions

The structures have been simulated using AM 1.5G spectrum in 300 nm to 900 nm wavelength range. We have not incorporated any recombination in device modelling since the electron to hole mobility ratio (μ_n/μ_p) has been kept under 50 throughout the simulation. The validity behind this assumption has been explicitly explained in other research works [5-6]. We are interested in relatively thin films since carrier recombination in thicker BHJ is not negligible and so thickness of active materials have been varied from 60 nm to 140 nm.

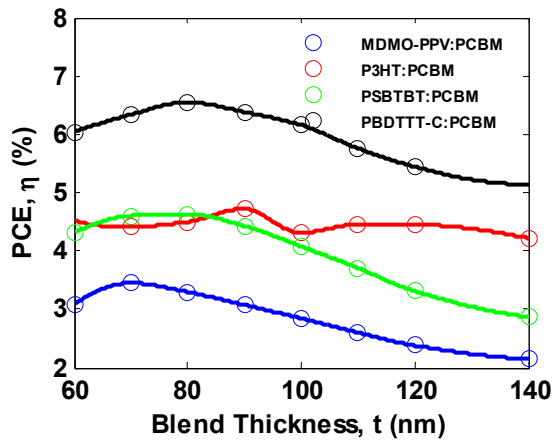


Figure 2: Power conversion efficiency versus the blend thickness for four active materials. The different colored circles represent the simulation results. The solid lines do not represent simulation results rather act as a guide to eye. They have been generated using suitable curve fitting method.

Figure 2 shows the variation of efficiency with active layer thickness for different active materials. We have not introduced any resistance in contact regions. Hence, values of photovoltaic parameters are slightly different than experimental values. No complete experimental dataset is available for the range of thickness considered here. However, we match with experimental values where data are available to benchmark our simulation environment.

A. MDMO-PPV:PCBM

As seen from Fig. 2, the optimum blend thickness for MDMO-PPV:PCBM has been found to be 70 nm where the efficiency is enhanced up to 3.45%. In our simulation environment, the PCE for 100 nm thickness has been found to be 2.89% with V_{oc} = 0.82 V, J_{sc} = 5.83 mA/cm^2 and FF=59.2%. Experimentally, at this thickness, the recorded efficiency is 2.5% where V_{oc} = 0.82 V, J_{sc} = 5.25 mA/cm^2 and FF=61% [4].

B. P3HT:PCBM

For P3HT:PCBM solar cell, the maximum efficiency has been noted as 4.72% at thickness of 90 nm. For benchmarking, we have compared our results with the analytical model proposed by Chowdhury *et al.* [5]. Results from their model suggest that the efficiency of a 100 nm thick cell is 4.31% whereas our results show 4.39% with V_{oc} = 0.66 V, J_{sc} = 10.02 mA/cm2 and FF=65.8%.

C. PSBTBT:PCBM

The maximum efficiency for PSBTBT:PCBM cell too, has been found at lower thickness. In this case, the optimum efficiency and thickness values are 4.64% and 80 nm respectively. A temperature dependent simulation study performed by Yang *et al.* [16] estimates a reduced efficiency of 4.24% at 80 nm thickness and 30°C temperature.

D. PBDTTT-C:PCBM

For PBDTTT-C:PCBM based solar cell with standard architecture, efficiency has been found to be in higher range than other active materials studied in this work. The optimum blend thickness, here too, has been noted as 80 nm where the efficiency reaches up to 6.55% with V_{oc} = 0.72 V, J_{sc} = 13.48 mA/cm2 and FF=67.2%. Experimental analysis at this thickness gives a reduced efficiency of 6.3% with V_{oc} = 0.711 V, J_{sc} = 14.1 mA/cm^2 and FF=63% [8].

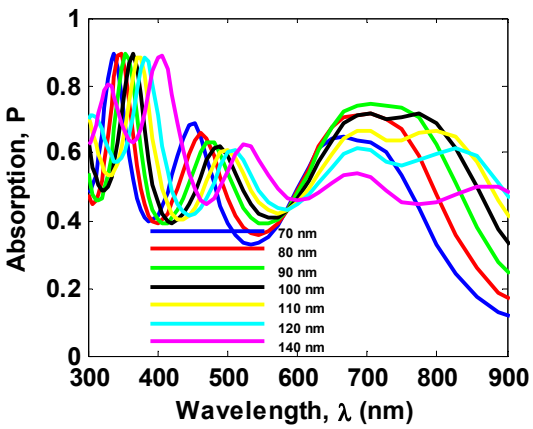


Figure 3: Absorption spectra of PBDTTT-C:PCBM solar cell with various blend thicknesses

To explain the reason of this oscillating behavior in the efficiency Vs thickness, we show the power absorption spectra for PBDTTT-C:PCBM solar cell at different blend thicknesses as an example (other cells show similar characteristics). As thickness increases, more power is absorbed from longer wavelengths as evident from Fig. 3. In particular, this increases the integrated absorption over the spectrum. Although integrated absorption increases continuously with thickness, the current density do not. Rather, there lies a thickness where the combined effect of charge generation, transportation and collection reaches an optimum state and only then efficiency is maximized. The above stated facts are illustrated in Fig. 4 where integrated absorption increases with blend thickness but the short circuit current densities decrease after 80 nm which is in our case actually the optimum thickness for PBDTTT-C:PCBM.

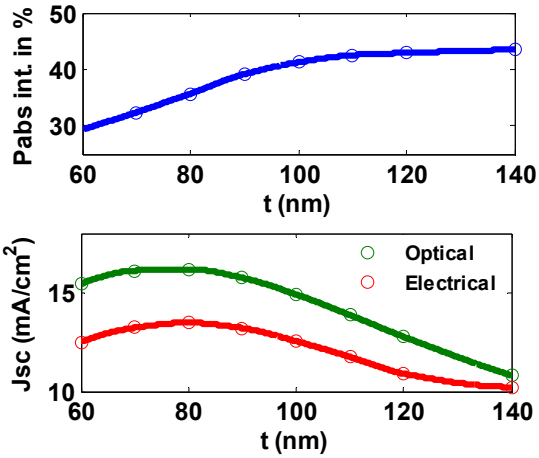


Figure 4: Variation of integrated absorption and short circuit current density with different PBDTTT-C:PCBM thickness. The blue circles in the inset represent calculated values. The green and red circles represent the optical and electrical simulation results respectively. The solid lines are provided as a guide to eye. The labels Pabs int., Jsc and t denote Integrated absorption, short circuit current density and blend thickness respectively.

We are also interested in investigating the dependence of short circuit current density (J_{sc}) and open circuit voltage (V_{oc}) of other active materials with layer thickness. Basically, V_{oc} is a strong function of built-in potential (V_{bi}) in active layer [17]. Since the V_{bi} remains same for all thicknesses, V_{oc} remains fairly constant. Variation of J_{sc} with thickness shows similar trend with that of efficiency due to proportional relationship. These facts can be manifested from figure 5 as well.

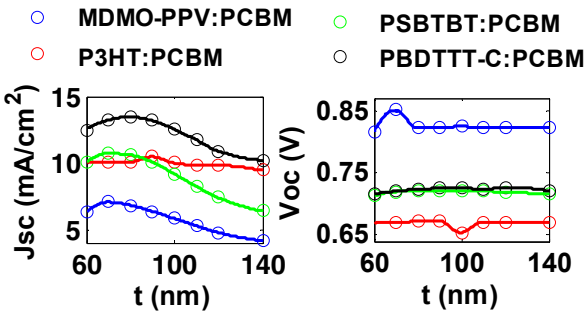


Figure 5: Variation of open circuit voltage and short circuit current density with blend thickness for different active materials. The circles represent simulation results. The solid lines are given as a guide to eye. The labels t, Jsc and Voc denote blend thickness, short circuit current density and open circuit voltage respectively.

The J-V characteristics at optimized thicknesses for four materials have been depicted in Fig. 6. The corresponding PCE (η), open circuit voltage (V_{oc}), short circuit current density (J_{sc}) & fill factor (FF) are listed in Table 2. There is a

gradual increase in maximum efficiency and J_{sc} from MDMO-PPV:PCBM to PBDTTT-C:PCBM. MDMO-PPV:PCBM also has lowest fill factor compared to others. In fact, the fill factor of the optimized cell has been found to be below 60 %. In spite of low efficiency, fill factor and J_{sc}, MDMO-PPV:PCBM based cell has maximum V_{oc} (0.851 V) among the group of materials studied. The lowest V_{oc} has been found in P3HT:PCBM cell with 0.673 V. The other two materials, PSBTBT:PCBM & PBDTTT-C:PCBM give V_{oc} of 0.721 V and 0.724 V respectively.

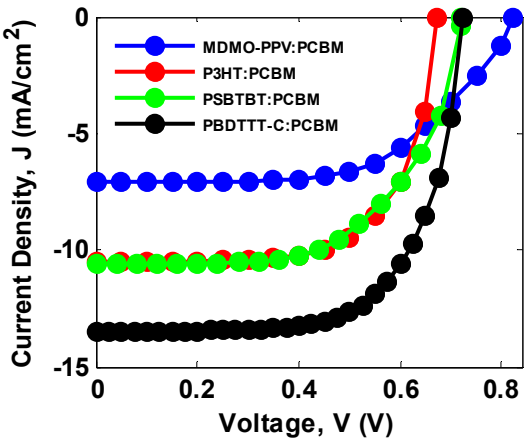


Figure 6: J-V characteristics for different active materials at optimum active layer thickness

Table 2: Photovoltaic properties of active materials studied at optimum thickness

Active Material	PCE (%)	V_{oc} (V)	J_{sc} (mA/cm^2)	FF (%)
MDMO-PPV:PCBM	3.45	0.851	7.071	59.5
P3HT:PCBM	4.71	0.673	10.079	66.3
PSBTBT:PCBM	4.64	0.721	10.613	60.5
PBDTTT-C:PCBM	6.55	0.724	13.478	67.2

Use of PBDTTT-C:PCBM as an active material is recommended for most of the systems where efficiency is a major criterion. PSBTBT:PCBM can be a good alternative if material cost is an issue. The efficiency of PSBTBT:PCBM can be increased close to 6 % using pre and post processing treatment [7]. If an application requires the availability of larger voltages from a cell at zero current then choice of MDMO-PPV:PCBM is obvious although it has a much lower efficiency. The open circuit voltage of highly efficient materials (e.g. PBDTTT-C:PCBM) can be increased by increasing the band gap at the cost of reduced short circuit current density.

IV. Conclusion

We have compared PCBM based BHJ OSC with four different donor materials by varying their thicknesses. The optimum thickness has been found to be 70nm, 90 nm, 80 nm and 80 nm for MDMO-PPV:PCBM, P3HT:PCBM, PSBTBT:PCBM and PBDTTT-C:PCBM, respectively. The highest efficiency, among the group of materials studied at optimum thickness, has been found for PBDTTT-C:PCBM. This recorded efficiency is 6.55%.

Acknowledgment

We would like to sincerely acknowledge the support from department of EEE, BUET.

References

[1] F. C. Krebs, T. Tromholt and M. Jorgensen, "Upscaling of polymer solar cell fabrication using full roll-to-roll processing," Nanoscale, vol. 2, no. 6, pp. 873-886, 2010.

[2] G.Yu,J.Gao, J.Hummelen, F.Wudl and A.Heeger,"Polymer Photovoltaic Cells: Enhanced Efficiencies via a Network of Internal Donor-Acceptor Heterojunctions," Science, vol. 270, pp. 1789-1791, 1995.

[3] J. Hummelen, B. Knight, F. Lepeq and F. Wudl, "Preparation and characterization of fulleroid and methanofullerene derivative," J. ORG CHEM, vol. 60, pp. 532-538, 1995.

[4] S. E. Shaheen, C. J. Brabec, S. Sariciftci, F. Padinger and T. Fromherz, "2.5% efficient organic plastic solar cells," APPL. PHYS. LETT. , vol. 78, no. 6, p. 841, 2001.

[5] M. Chowdhury and M. Alam, "An optoelectronic analytical model for bulk heterojunction organic solar cells incorporating position and wavelength dependent carrier generation," SOLMAT, vol. 132, pp. 107-117, 2015.

[6] S. Altazin, R. Clerc , R. Gwozieki, G. Pananakakis, G. Ghibaudo and C. Serbutoviez, "Analytical modeling of organic solar cells and photodiodes," APPL PHYS. LETT., vol. 99, p. 143301, 2011.

[7] R.C. Coffin, J.Peet, J.Rogers and G.C. Bazan, "Streamlined microwave-assisted preparation of narrow-bandgap conjugated polymers for high performance bulk heterojunction solar cells, " N. CHEM., vol 1, p. 657-661, 2009.

[8] P. Morvillo, R. Diana , R. Ricciardi, E. Bobeico and C. Minarini , "High-efficiency standard and inverted polymer solar cells based on PBDTTT-C:[70]PCBM blend," in IEEE, Naples, 2014.

[9] S. Ahmed, N.R. Mittal,A.B. Bhattacharya and M. Singh , "Simulation, output power optimization and comparative study of silicon and thin film solar cell modules," in IEEE, Taichung, 2010.

[10] X. Xi, W. Li, J. Wu, J. Ji, Z. Shi and G. Li , "A comparative study on the performances of small molecule organic solarcells based on CuPc/C60 and CuPc/C70, " SOLMAT, vol. 94, pp. 2435-2441, 2010.

[11] O. Abdulrazzaq, S. E. Bourdo, M. Woo and A.S. Biris, "Comparative Aging Study of Organic Solar Cells Utilizing Polyaniline and PEDOT:PSS as Hole Transport Layers", ACS APPL. MATER. INTER., vol. 7, no. 50, 2015.

[12] H. Hoppe, N. Sariciftci and D. Meissner , "Optical constants of conjugated polymer/fullerene based bulk-heterojunction organic solar cells," MOL. CRYST. LIQ. CRYST., vol. 385, no. 1, pp. 113-119, 2002.

[13] F. Monestier, J. J. Simon, P. Torchio, L. Escobus, F. Flory, S. Bailly, R. d. Bettignies, S. Guillerez and C. Defranoux, "Modeling the short-circuit current density of polymer solar cells based on P3HT:PCBM blend," SOLMAT, vol. 91, pp. 405-410, 2007.

[14] M. F. G. Klein, G. Q. Glasner de Medeiros , P. Kapetana, U. Lemmer and A. Colsmann, "Modeling approach to derive the anisotropic complex refractive index of polymer:fullerene blends for organic solar cells utilizing spectroscopic ellipsometry," J. PHOTON ENERGY, vol. 5, p. 057204, 2015.

[15] Y.Zhou, C.F.Hernandez, J.W.Shim, T.M.Khan and B. Kippelen, "High performance polymeric charge recombination layer for organic tandem solar cells," ENERG ENVIRON SCI., vol. 5, pp. 9827-9832, 2012.

[16] B. Yang , J. Cox, Y. Yuan, F. Guo and J. Huang, "Increased efficiency of low band gap polymer solar cells at elevated temperature and its origins," APPL. PHYS. LETT., vol. 99, p. 133302, 2011.

[17] P. Kumar, H. Kumar, S. C. Jain, P. Venkatesu, S. Chand and V. Kumar, "Effect of Active Layer Thickness on Open Circuit Voltage in Organic Photovoltaic Devices," JPN. J. APPL. PHYS., Vols. 48, no. 12R, 2009.

A Smart Metering System for Wireless Power Measurement with Mobile Application

Kazi Tajmul Islam, A. M. Jahirul Islam
Department of Electrical and Electronic Engineering
Chittagong University of Eng. & Tech., Bangladesh
Email:{tajmul.cuet, tawsif_eee}@yahoo.com
Email:{a.m.jahirulislam, pidim.srh}@gmail.com

Syed R. H. Pidim, Akramul Haque
Md. Tawsif H. Khan, Sayidul Morsalin
Chittagong University of Engineering & Technology
Email: {akramul073, sayidul.morsalin}@gmail.com

***Abstract*—At present, advanced meter enabled with wireless machine-to-machine (M2M) communication can play a pivotal role for energy measurement in smart grid. However, the traditional metering and billing system implemented in developing countries (e.g. Bangladesh) even some developed countries is too non-automated and complex that leads a wake-up call to establish a robust and smart metering system to measure energy for user convenience. In this paper, an automated smart metering system is presented to measure energy consumption wirelessly for residential/ industrial areas. The measurement system is an ARDUINO based master-slave wireless technology that incorporates with smart billing payment scheme via Android mobile application. This wireless technology requires a pair of RF transmitter and receiver for transmitting energy consumption data, and for subsequent billing and payment system. Here the authors also measure the performance of the proposed system analytically.**

***Index Terms*—AMR system, Android application, Master-slave communication, Smart metering**

I. Introduction

Due to the recent development of green energy technology and environmental friendliness issues, people are more likely to have a smart and user-friendly energy metering system. For metering, automated energy monitoring system is already evolved, in such Automatic Meter Reading System (AMR) is one of them . The *AMR* system is the remote collection of energy consumption data from consumers' utility e.g. electric meters [1]. This technology uses radio frequency, telephony, power-line or satellite communications technologies for data processing, afterward generates the bill [2]. AMR technologies include handheld, mobile and network technologies based on telephony platforms (wired and wireless), radio frequency (RF, or power line transmission) [3], [4]. In this research, the authors consider the AMR system because it has some advantages over traditional meter reading techniques such as Wi Fi AMR, Touch Technology AMR, Radio Frequency AMR, Power Line Communication AMR etc. [1] The shortcoming of touch-based AMR is that the meter reader carries a handheld computer or data collection device with a wand or probe [5]. In Powerline Carrier (PLC) communication, the AMR data is transmitted to the substation over power lines, then relayed to a central computer in the utility's main office [6]. However, line attenuation in PLC reduces the amplitude of the frequency of the line which can affect the carrier frequency. Then skin effect arises due to the weather issues especially heavy frost where frequency can change up to 5:1 depending on frequency size. In addition, modulator and demodulator circuits are needed in PLC for transmitting and receiving data which is very costly [7].

A lot of researches were conducted previously for energy metering. In [8], K. Ashna et. al develop a GSM based automatic energy meter reading system with instant billing system. In GSM based AMR system, the energy meter is continuously monitored and records are taken and transmit meter reading via SMS. However, the cost of GSM module is higher than RF module and recharging credit in the SIMcard is a manual process. Therefore, GSM system requires more operating cost than RF system. In [9], the authors also present wireless remote metering system with GSM/EDGE modem. However, both of these abovementioned works neglect the M2M operation with smart mobile application for billing and payment. In ZigBee-based system [10], energy consumption data is transmitted by a serial communication. Therefore, the interaction between server computer and microcontroller or Arduino is not possible. Moreover, power consumption and cost of ZigBee module is higher than RF module. In [11], Bluetooth-based energy meter has a lower data rate and lower power consumption. However, the coverage range of Bluetooth is less than a RF module [11]. Furthermore, the works neglect the AMR operation with smart mobile application for user convenience.

This proposed research overcomes the limitation of all aforementioned work. In this research, the authors present a wireless monitoring of energy meter and readings system with RF communication. The novelty of this research that the authors develop an Android mobile application for bill payment and explain the performance of the proposed system mathematically . The reasons for using RF is that the RF module can offer low-power operation, high security, robustness, and high scalability [12]. Here the authors use a master-slave communication technology to transmit energy consumption reading to a central unit where a server is located to make a handshake between the users and meter reading via mobile application. Therefore, due to the user-friendliness and automated operation, the proposed system can be a smart choice for residential/industrial areas.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

II. System Description

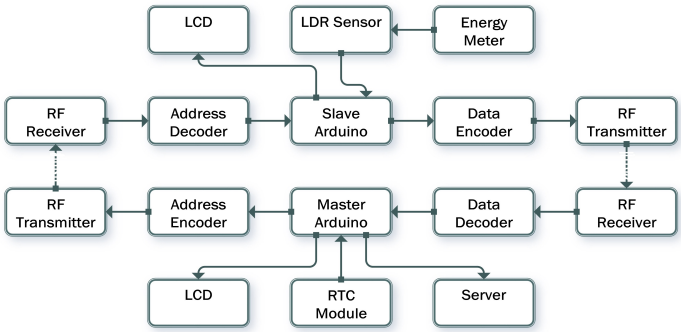


Figure 1: Block Diagram of the System

The block diagram of the overall system is illustrated by a block diagram as shown in Figure 1. The energy consumption is calculated by a LDR-photoresistor from the blinking of LED, used in the digital energy meter. The energy meter sends pulses to a ARDUINO when the LED blinks. A LCD display connected with the ARDUINO is used to display the energy consumption data and equivalent bills in the slave circuits. A pair of RF transmitter (JMRTX01) and receiver (RX806) are placed in master and each slave circuit for duplex Radio Frequency (RF) communication. The output data of ARDUINO are encoded to transmit when the address decoder receives address from the master circuit. The decoder of master circuit decodes the signal to make it readable for the server computer. A RTC (Real Time Clock) module is used in master circuit for time and date observation. Therefore, the user can see the bills and consumption unit, the paid bills log. After that, they can pay the bills using android application which is linked with the server through the IP4/IP6 internet network.

III. Hardware Implementation

A. Slave Unit

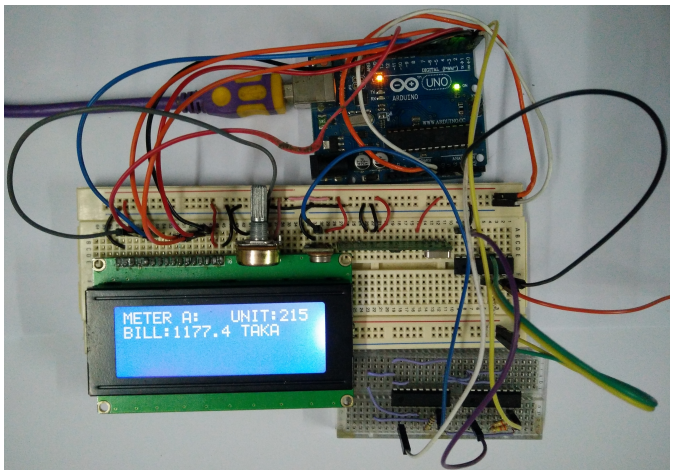


Figure 2: Hardware Implementation of the Slave Unit

This unit is connected with an energy meter. A LDR is coupled with the LED of the energy meter, in combine acts as an optocoupler and sends a signal when the LED blinks. Then the voltage signal is compared by a voltage comparator which sends a pulse to the ARDUINO board. The encoder receives a parallel data stream from the ARDUINO board and converts the data into serial data by ASK (Amplitude Shift Keying) modulation technique. A pair of RF transmitter and receiver is used for wireless communication. When receiver of the slave gets an address code, then it sends the data through the slave transmitter. The data are received by the RF transmitter module and transmitted through the air by using a quarter wavelength monopole antenna. The implemented slave unit for this research is shown in Figure 2.

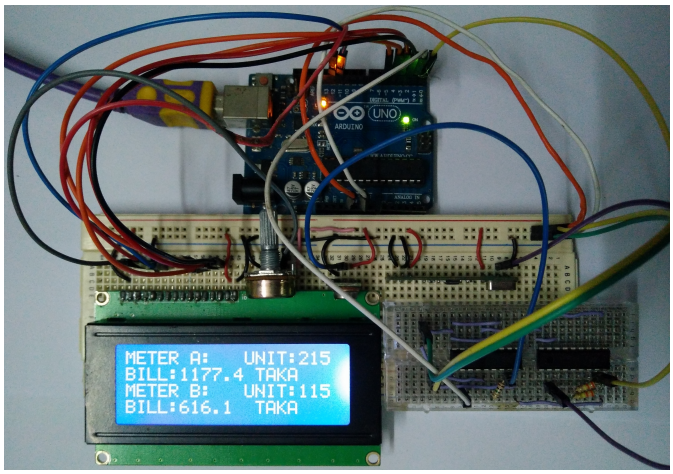


Figure 3: Hardware Implementation of Master Circuit

B. Master Unit

The master RF receiver module receives the serial data by an antenna. The decoder decodes the serial data to the parallel data and sends it to the ARDUINO which shows the data in an LCD display. The ARDUINO is connected with the server computer by USB cable for transferring data [13]. With the help of RTC module date and time of the received data are added to the server. The implemented master unit for this research is shown in Figure 3.

IV. Software Simulation

A complete circuit diagram to demonstrate the performance 5 has been simulated in the Proteus simulating software as shown in Figure 4 and 5. Proteus Software is generally used for microprocessor simulation, schematic capture and printed circuit board (PCB) design. According to the system algorithm, the code is written in ARDUINO which is INO file. This INO file is converted into HEX file with the help of compiler which is loaded into the ARDUINO to generate the appropriate output.

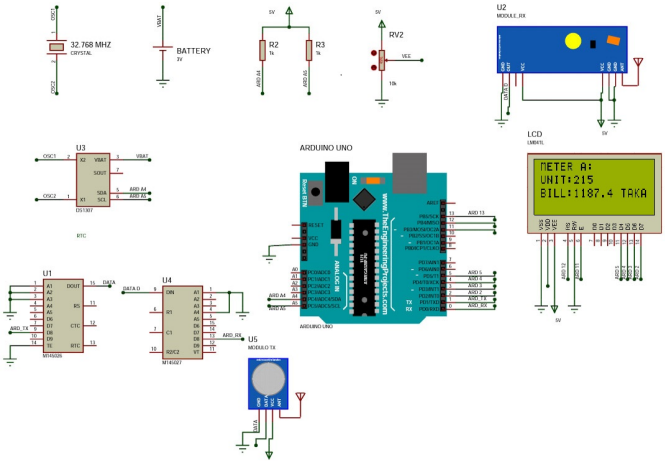


Figure 4: Proteus Simulation of Master Circuit

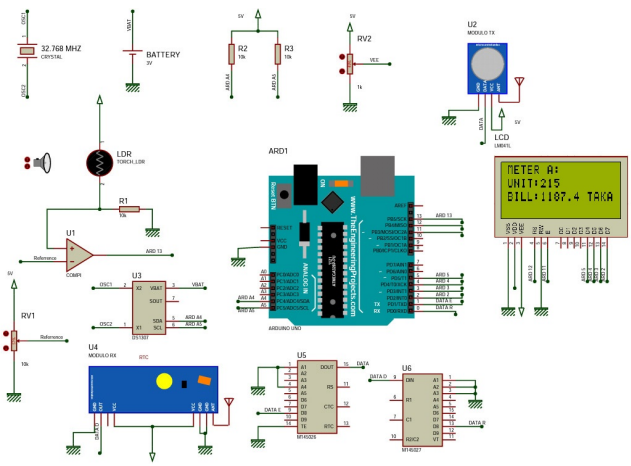


Figure 5: Proteus Simulation of Slave Circuit

V. Analytical Result

Let us consider, a cell of radius 7 kilometers. A BTS (Base Transceiver Station) can receive signals from all the transmitters placed at consumers' place. The BTS can also receive signal from all directions within the cellular area. Therefore, an Omnidirectional Antenna is suitable for the research purpose. The basic parameters which should be considered for designing a RF network are transmitting frequency (which the authors selected 2.4 GHz), Path Loss, Antenna Gain, Receiver Sensitivity; Effective Isotropic radiated Power (EIRP), Effective Received RF Signal Power, and System Operating Margin (SOM). The path loss for distance 7 Km is determined by,

$$\begin{aligned} P_l &= 100 + 20\log(R_{km}) \\ &= 100 + 20\log(7) \\ &= 100.85\ dB \end{aligned}$$

The EIRP (Effective Isotropic Radiated Power) is the antenna transmitted power. Therefore,

$$EIRP = P_{out} - C_t + G_t$$

Where,
P_{out}= transmitted output RF power to antenna in dB_m
C_t= transmitter cable attenuation in dB
G_t= transmitting antenna gain in dB_i
For this design, P_{out} = 17 dB_m (50 mW)

The cable length to connect T_x antenna to the transmitter module is 25 meters. So, T_x cable loss for 25 meter cable (0.6 dB/meter) are C_t = 15 dB and T_x antenna gain G_t = 18 dB_i. Therefore, Effective Isotropic Radiated power,

$$\begin{aligned} EIRP &= P_{out} - C_t + G_t \\ &= 20\ dB_m \end{aligned}$$

The Effective Received Signal Power can be calculated using the following equation:

$$S_i = EIRP - P_l + G_r - C_r$$

Where, P_l = Path loss in dB;
G_r = receiving antenna gain in dB_i;
C_r = receiving cable attenuation in dB;
The cable length to connect R_x antenna to the receiver module is 200 meters. So, T_x cable loss for 200 meter cable (0.6 dB/meter) are C_r = 120 dB and R_x antenna gain G_r = 18 dB_i. EIRP =20 dB_m and P_l = 100.85 dB. Therefore, the Effective Received Signal Power,

$$S_i = -182.85\ dB_m$$

The received signal power (S_i) should be at least 10 dB higher than receiver sensitivity (P_s). So, the receiver sensitivity should be at least equal to -172.85 dB_m. SOM (System Operating Margin), also known as fade margin, is the difference of the receiver signal level in dB_m minus the receiver sensitivity in dB_m. It is a measure of the safety margin in a radio link. A higher SOM means a more reliable over the air connection. The suitable value is minimum 10 dB, but 20 dB or more is the best. SOM is the difference between the signals received from radio and what it needs for good data recovery (i.e. receiver sensitivity). By using the transmitted and received RF signal power, the cable losses, the antenna gains, and the path losses as considered, SOM can be calculated . Thus the method can be used for designing and analyzing RF signal links used in wireless networking.

$$\begin{aligned} R_xSignal &= T_xPower - T_xCableLoss + T_xAntennaGain - \\ &\quad PathLoss + R_xAntennaGain - R_xCableLoss \\ SOM &= R_xSignalLevel - R_xSensitivity \\ &= 10\ dB_m \end{aligned}$$

The SOM expression can be modified to consider attenuation losses due to transmission through walls, etc., in an actual building wherein a home or office network would be installed. It is simply adding more loss terms to the SOM equation. But

first, the level of losses is considered through various materials. The signal attenuation loss for 2.4 GHz transmission through the following structures can be included in the R_x Signal Level equation for each pass-through in the straight line signal path (line of sight). The dB loss values will be subtracted from the transmitted signal power to reflect the loss of passing through the material structures.

VI. ANDROID APPLICATION

An android application (apps) is developed for observing bills log and paying the bills. This apps is developed in android studio software by JAVA coding [13], [14]. User have to log in with meter number and password. There are two window in this apps. In first user can see the previous bills and meter readings, bill status of previous months. By using second window user can pay the bills with debit card or credit card.

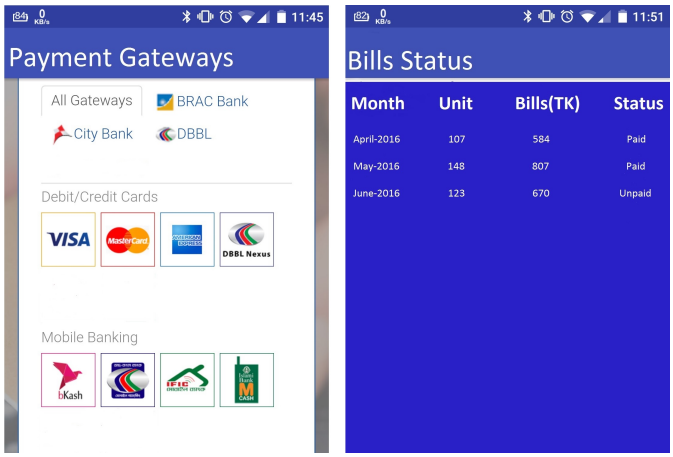


Figure 6: Screenshots of Android Application

VII. CONCLUSION

In this research, the developed system works very well without any basic error until it is physically damaged as well as it is very smart, reliable and robust. So the system can be implemented for measuring, billing and bills payment as it is very user-friendly. Moreover, it is very economic than any existing system. In spite of some limitations of the research it offers more flexibilities and facilities as the metering system has been turned into automatic and more secured for todays digitized world.

REFERENCES

[1] V. V. Das, "Wireless communication system for energy meter reading," in *Advances in Recent Technologies in Communication and Computing, 2009. ARTCom'09. International Conference on*. IEEE, 2009, pp. 896–898.

[2] M. Moghavvemi, S. Tan, and S. Wong, "A reliable and economically feasible automatic meter reading system using power line distribution network," *International Journal Of Engineering-Materials And Energy Research Center*, vol. 18, no. 3, pp. 301–318, 2005.

[3] A. Zaballos, A. Vallejo, M. Majoral, and J. M. Selga, "Survey and performance comparison of amr over plc standards," *IEEE transactions on power delivery*, vol. 24, no. 2, pp. 604–613, 2009.

[4] J. Newbury and W. Miller, "Multiprotocol routing for automatic remote meter reading using power line carrier systems," *IEEE transactions on power delivery*, vol. 16, no. 1, pp. 1–5, 2001.

[5] V. Vinu, "Wireless communication system for energy meter reading," in *International Conference on Advances in Recent Technologies in Communication and Computing Voltage Apparatus. Beijing*, vol. 9, 2009, pp. 1–9.

[6] T. Ahmed, M. Miah, M. Islam, M. Uddin *et al.*, "Automatic electric meter reading system: A cost-feasible alternative approach in meter reading for bangladesh perspective using low-cost digital wattmeter and wimax technology," *arXiv preprint arXiv:1209.5431*, 2012.

[7] S. Naiman, M. Kissaka, and N. Mvungi, "Energy meter reading and tampering protection through powerline communication channel," in *AFRICON, 2004. 7th AFRICON Conference in Africa*, vol. 2. IEEE, 2004, pp. 821–826.

[8] K. Ashna and S. N. George, "Gsm based automatic energy meter reading system with instant billing," in *Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), 2013 International Multi-Conference on*. IEEE, 2013, pp. 65–72.

[9] M. S. A. Rupok, M. Ahmed, and A. M. R. U. Haque, "Design and implementation of a novel remote metering system using usb gprs/edge modem," in *Electronics Computer Technology (ICECT), 2011 3rd International Conference on*, vol. 4. IEEE, 2011, pp. 237–240.

[10] P. G. Kate and J. R. Rana, "Zigbee based monitoring theft detection and automatic electricity meter reading?" in *2015 International Conference on Energy Systems and Applications*. IEEE, 2015, pp. 258–262.

[11] B. Koay, S. Cheah, Y. Sng, P. Chong, P. Shum, Y. Tong, X. Wang, Y. Zuo, and H. Kuek, "Design and implementation of bluetooth energy meter," in *Information, Communications and Signal Processing, 2003 and Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint Conference of the Fourth International Conference on*, vol. 3. IEEE, 2003, pp. 1474–1477.

[12] K. Lim, S. Pinel, M. Davis, A. Sutono, C.-H. Lee, D. Heo, A. Obatoynbo, J. Laskar, E. M. Tantzeris, and R. Tummala, "Rf-system-on-package (sop) for wireless communications," *IEEE Microwave Magazine*, vol. 3, no. 1, pp. 88–99, 2002.

[13] Y. Wang, R.-h. Chi, and M.-h. Liu, "Design of remote medical monitoring system based on android mobile platform and bluetooth hdp technology," in *Software Intelligence Technologies and Applications & International Conference on Frontiers of Internet of Things 2014, International Conference on*. IET, 2014, pp. 144–149.

[14] A. H. Sabani and R. Jailani, "Android based control and monitoring system for leg orthosis," in *Signal Processing & Its Applications (CSPA), 2015 IEEE 11th International Colloquium on*. IEEE, 2015, pp. 40–45.

Self-Consistent Determination of Quantum-Mechanical Threshold Voltage of Gate-All-Around Junction Less Nanowire Transistor

Nujhat Tasneem*, Md. Mohsinur Rahman Adnan, Md. Samzid Bin Hafiz, Quazi D.M. Khosru
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka-1000, Bangladesh
*e-mail: nujhat285@gmail.com

***Abstract* — Quantum definition based threshold voltage calculation of Gate-All-Around Junction Less Nanowire Transistor has been presented in this work. Employment of this approach has not yet been achieved on GAA-JLNT even though similar determination was previously established for TG FinFETs and GAA-MOSFETs in recent literature. The self-consistent solver, which takes wave function penetration and other quantum mechanical (QM) effects into account, has been utilized here to establish the Capacitance-Voltage characteristics, using which the threshold voltage has been calculated. Effect of silicon channel thickness on threshold has been studied and a modification of classical analytical formula is proposed based on a fitting parameter, using the extracted threshold voltages.**

***Index Terms* — Gate-All-Around Junction Less Transistor, Self- Consistent Schrödinger-Poisson Solver, Threshold Voltage, Fitting Parameter**

I. Introduction

The Junction Less Nanowire Field Effect Transistor (JLNWFET), or Junction Less Nanowire Transistor (JLNT) in short, is a novel multiple gate device whose operating principle is similar to a traditional FET device even though there is no distinct pn junction in its structure. The device mainly operates in Partial Depletion Mode. It contains a doped Silicon channel region enclosed by poly-gate structures, separated by an oxide layer on either side. Figure 1 illustrates the 2D cross section of Gate-All-Around JLNT (GAA-JLNT).

As the International Technology Roadmap for Semiconductors has foreseen [1], Silicon-On-Insulator (SOI) and multiple gate devices, such as double gate, tri-gate, pi-gate, Omega-gate and gate-all-around (GAA) JLNTs offer good electrostatic control necessary for gate lengths around 25 nm. GAA is the most favorable one among them, offering smallest Short Channel Effect (SCE) with device dimension reduction. The channel height and width must be kept lower than the effective gate length in order to maintain low SCE. GAA-JLNT also exhibits widening of first Eigen energy from conduction band edge with the decrease in channel thickness like other multi-gate structures. This phenomenon appreciably dominates the threshold voltage which cannot be held accounted in classical approach. For this reason, the incorporation of quantum effects is essential for accurate calculation of threshold voltage of such devices.

Threshold voltage determination is an important step in acquiring an in-depth understanding of electrostatics and dynamic operation of novel devices [2]-[4]. Several analytical models of threshold voltage calculation of JLNT have been reported in previous works [5]-[7]. However, the quantum mechanical effect was not taken into consideration in those works. Although a compact analytical model for determining

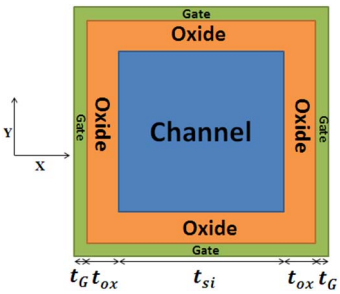


Fig. 1: 2D cross sectional view of GAA-JLNT perpendicular to the transport direction. The oxide (SiO_2 has been used in this work) thickness is 2 nm.

threshold voltage comprising the quantum mechanical effect was proposed [8], it was done only for the Double Gate JLNT. Yun *et al.* propose a quantum definition of threshold voltage that takes into account the quantum effects [9], however, detailed examination of how this approach of computing threshold voltage can be applied for a wide range of device families is not presented. A physically based analytical model for the threshold voltage in JLNT has been proposed in [6]. There are possibilities of modification in this model if the quantum mechanical definition of threshold voltage, given in [9], is taken into consideration.

In this paper, we have adopted the approach used in [3-4] where determination of the threshold voltages for TG-FinFET and GAA-MOSFET were discussed, whereas we have used the similar concept for GAA-JLNT. Using the similar method, the threshold voltage equation proposed in [6] has been modified by incorporating the quantum mechanical effects. Self-Consistent simulation using coupled Schrödinger-Poisson solver has been carried out taking wave function penetration and other quantum mechanical effects into account. The self-consistently calculated gate capacitance value has been utilized in the calculation of threshold voltage of GAA-JLNT.

II. Simulator Model

The evaluation of threshold voltage is performed by calculating the 2D charge profile at a cross section perpendicular to the channel as proposed in [3-4]. The ratio of peak value of the electron concentration profile (n_{peak}) and average electron concentration ($n_{average}$) is plotted against gate voltage (V_{gate}). The $n_{average}$ is defined as [10]:

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

$$n_{average} = \frac{1}{A_{si}} \iint_{A_{si}} n(x,y)dxdy \ \ (m^{-3}) \quad (1)$$

Threshold in JLNT can be considered as the gate voltage at which bulk conduction starts to occur. This condition is met when there is a point at the center of the channel that is not depleted and the device enters the partial depletion mode. The $n_{peak}/n_{average}$ vs. V_{gate} curve shows a distinct transition point at the threshold voltage which suggests the start of this non-depleted region buildup. The main idea of this approach is to find out the voltage at which the electron concentration exceeds the space charge concentration just enough to create a non-depleted point at the center of the channel.

The charge concentration is estimated by self- consistent Schrödinger-Poisson solver which assumes wave function penetration effects and other QM effects into account. After convergence, electron concentration per unit width (N) is obtained using:

$$N = \iint_{A_{si}} n(x,y)dxdy \ \ (m^{-1}) \quad (2)$$

Multiplication of N with electron charge (-1.6×10^{-19}) yields charge concentration (Q) and the Capacitance-Voltage characteristics is obtained as:

$$C(V_{gate}) = \frac{dQ}{dV_{gate}} \ \ (Fm^{-1}) \quad (3)$$

The analytical model for determining threshold voltage in [6] was mainly proposed for a Tri Gate JLNT, but can also be used for a GAA-JLNT having a rectangular gate structure. For GAA-JLNT, the expression of threshold voltage is as follows:

$$V_{TH} = V_{FB} - qN_D\left[\frac{A}{C_{ox}} + \frac{1}{\varepsilon_{si}}\left(\frac{A}{P}\right)^2\right] + \frac{\Delta E_0}{q} \quad (4)$$

Where, A and P are the device cross-sectional area and gate perimeter respectively and $\Delta E_0$is the variation of the minimum energy level in the conduction band due to quantum confinement.

Now, $\Delta E_0 = \frac{\pi^2 h^2}{m_e^* t_{si}^2}$ where m_e^* is the electron effective mass.

In case of GAA-JLNT, $A = t_{si}^2$ and $P = 4t_{si}$ where $t_{si}$is the silicon channel thickness. So, equation 4 becomes as follows:

$$V_{TH} = V_{FB} - qN_D[\frac{t_{si}^2}{C_{ox}} + \frac{1}{\varepsilon_{si}}(\frac{t_{si}}{4})^2] + \frac{\pi^2 h^2}{qm_e^* t_{si}^2} \quad (5)$$

The above expression can be simplified to,

$$V_{TH} = V_{FB} - n * [\varphi_{surface} + V_{ox}] + E_{G1} \quad (6)$$

Here,

V_{FB}= Flat Band Voltage (0.526 eV).

$\varphi_{surface}$= the surface potential at which the partial depletion regime starts to build up.

V_{ox}= $\frac{Q}{C_{ox}}$, the voltage drop in the MOS capacitor.

E_{G1}= Lowest subband energy above conduction band minima i.e. first Eigen Energy, a fully quantum mechanical term.

$[\varphi_{surface} + V_{ox}]$ is an entirely classical term. To integrate the quantum effect, the fitting parameter (n), which varies with device geometry, is introduced. Another adjustment is made by making C_{ox} equal to C_{final} [3-4], which is the saturation gate capacitance obtained from the self-consistent C-V. In this paper, we calculated the threshold voltage of GAA-JLNTs having various silicon channel thickness (t_{si}) and then determined the values of the fitting parameter (n) for all thicknesses. The GAA-JLNT device used in simulation has concentration (N_D) of $10^{19}cm^{-3}$. The work function of the gate metal used is 5.27 eV.

III. Results And Discussions

Silicon channel thickness is one of the most important process parameters which has a prodigious impact on the threshold voltage of MuGFETs [9]. In our simulations, t_{si} in the range of $2nm$ to $12nm$ has been considered. At first, JLNT is in Full Depletion mode, with no mobile carrier in the channel region. The induced depletion region reduces as gate voltage is increased, and when voltage exceeds threshold, the channel conductance initiates. The device operation switches from Depletion to Accumulation mode for voltages beyond flat-band. As depicted in the Figure 2, the charge concentration along the semiconductor region is nearly zero when the device is fully depleted. A channel is established where electron concentration rises from zero when the gate voltage is raised and reaches the state where a non-depleted region is present. This is the partial depletion region. In the accumulation region, the electron concentration in semiconductor region near the oxide is greater than that at the centre which is due to the accumulated negative charge near

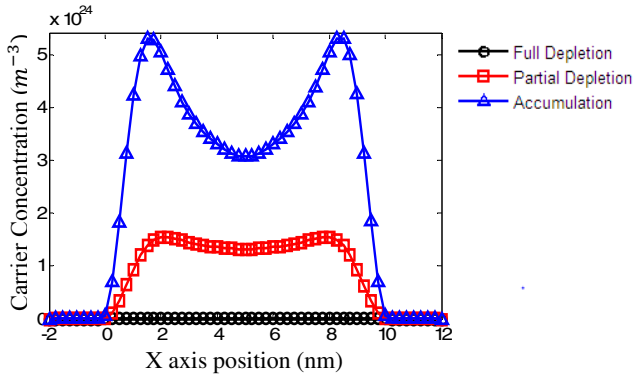


Fig. 2: Carrier Concentration along the X axis (channel cross section) for three different modes of operation of GAA-JLNT with Silicon channel thickness of 1nm.

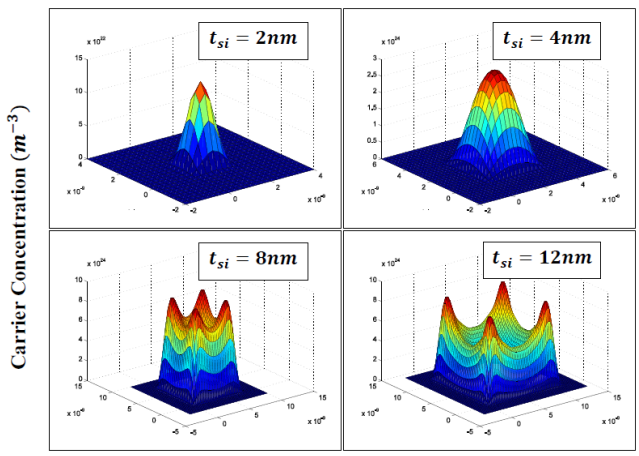


Fig. 3: Carrier concentration profile at $V_{gate} = 1V$ for four different geometries of GAA-JLNT.

substrate-oxide interface. This implies that there will be volume inversion in full and partial depletion mode of the device [12], whatever the silicon channel thickness (t_{si}), whereas in the accumulation mode there will be increased charge confinement at the corners and sidewalls of the channel. In Figure 3, the charge profiles of the device with different t_{si} are presented at $V_{gate} = 1\text{V}$. GAA-JLNT with lower t_{si} has volume inversion at 1V, which means it is still in the depletion region, and higher t_{si} gives rise to charge confinement at surface at the same voltage; meaning the device has entered the accumulation region. This implies that the threshold voltage of GAA-JLNT decreases with t_{si} [12].

It is evident from the band diagrams shown in Figure 4 that potential wells narrow down with reduced device dimensions. This, in effect, results in the increased quantum confinement of electrons and also causes the increase in splitting of energy levels into subbands. Quantum mechanical effects resulting from this increased quantum confinement has a marked effect on V_{TH}. From Figure 4 it can be seen that, as Eg1 increases, the threshold voltage also increases with the decrease of silicon channel thickness which is clear from equation (6). Based on the quantum definition of V_{TH} [9], $\frac{n_{peak}}{n_{average}}$ ratio vs. gate voltage (V_{gate}) variation curves are used to extract these threshold voltages (Figure 5). The threshold voltages extracted from Figure 5 have been used to demonstrate the dependence of V_{TH} on t_{si}(Figure 6). Our result conforms to the curve in [9] shown in Figure 7 in this paper. As anticipated, the threshold voltage increases considerably as the silicon film thickness is reduced, indicating the requirement of extra amount of charge for the buildup of non-depleted region in low-dimension devices. From Figure 8, it can be seen that the threshold voltages obtained in this work match closely with the results obtained in another work [11]. The small discrepancy is due to the fact that the quantum effect arising from small t_{si} has been ignored in [11] whereas it has been considered in this work. It should also be noted that the threshold voltage of GAA-JLNT with t_{si} below 6nm, when the quantum effect becomes more profound, was not determined in [11], but it has been considered in this paper.

The C-V characteristics, shown in Figure 9, of the simulated devices suggest that the rate of charge buildup is improved by the increase of channel thickness. The value of the saturation gate capacitances (C_{final}) can be extracted from this figure for various gate voltages. The use of these self-consistently obtained capacitive values surely gives a better approximation than the use of fixed oxide capacitance in equation (4-5). Figure 10 shows the variations of first Eigen energy and $-[\varphi_{surface} + V_{ox}]$ vs. channel. Using these data, the fitting parameter n is determined from equation (6), as described in previous section. The result is illustrated in Figure 11.

An effect of scaling in silicon film thickness on GAA-JLNT can be perceived from Figure 11. As the devices are scaled down from 12nm to 6nm, the value of n increases consequently. Furthermore, as scaling continues from 6nm to 2nm, the value of n drops sharply indicating the prominent effect of quantum mechanical term E_{G1} on V_{TH}. Threshold voltage in this level of scaling shows strong dependence on quantum mechanical effects. This huge impact of quantum confinement in low dimensional devices has been illustrated earlier in this paper. Thus, calculation of threshold voltage of GAA-JLNTs using the novel approach in literature [9] demonstrates the quantum mechanical effects methodically.

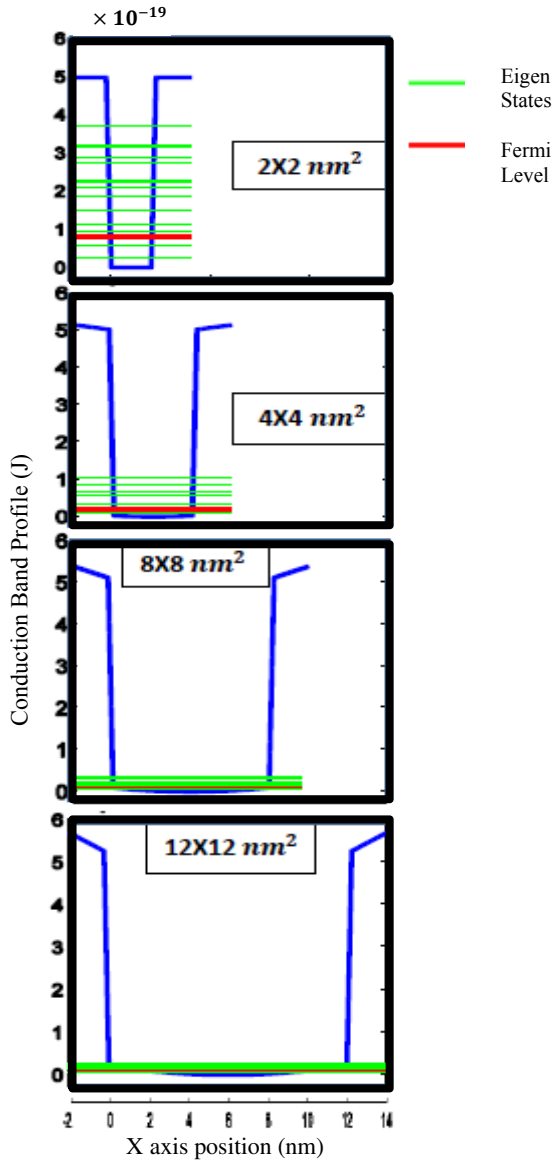


Fig. 4: Conduction band profile along the X axis (channel cross section), at $V_{gate} = V_{TH}$.

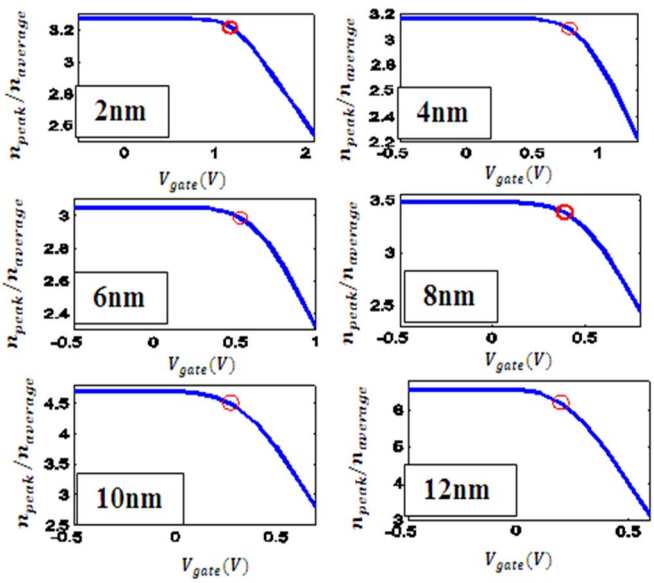


Fig. 5: $n_{peak}/n_{average}$ vs. V_{gate} for different silicon channel thickness. The point of transition (marked by red circle) in the curves indicate the threshold voltage.

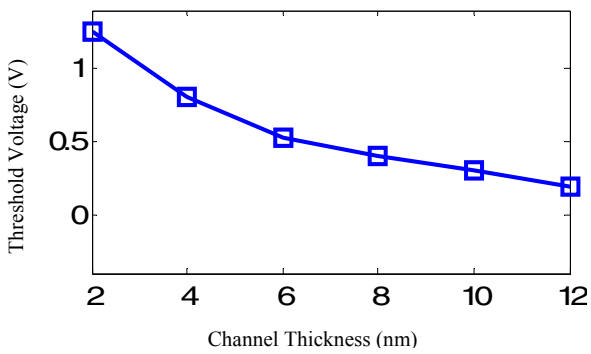


Fig. 6: Threshold voltage vs. channel thickness, as extracted from Figure 5.

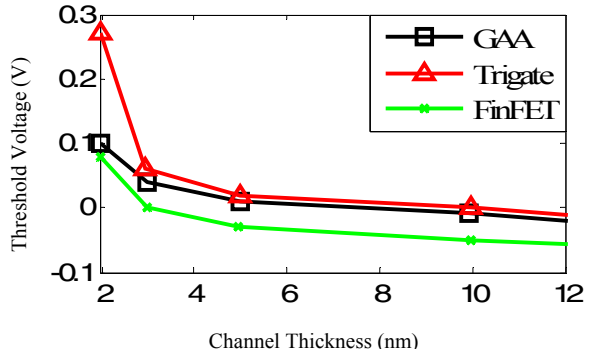


Fig. 7: Threshold voltages vs. channel thickness for different device geometries as shown in [9]. The threshold voltage obtained in this work follows the trend shown here.

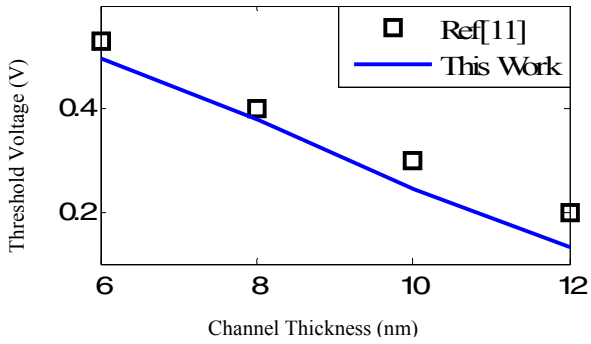


Fig. 8: Comparison between the threshold voltages obtained in this work and in [11] for silicon channel thickness varying from 6nm to 12nm.

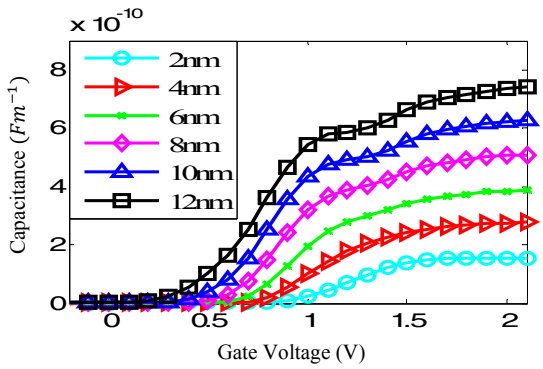


Fig. 9: Capacitance-Voltage characteristic varying silicon thickness

IV. Conclusion

Threshold voltage calculation of GAA-JLNT based on a quantum definition is performed using self-consistent Poisson and Schrödinger equation coupling. The result of threshold voltage characteristic with varying channel thickness matches satisfactorily with that predicted in previous literature. As

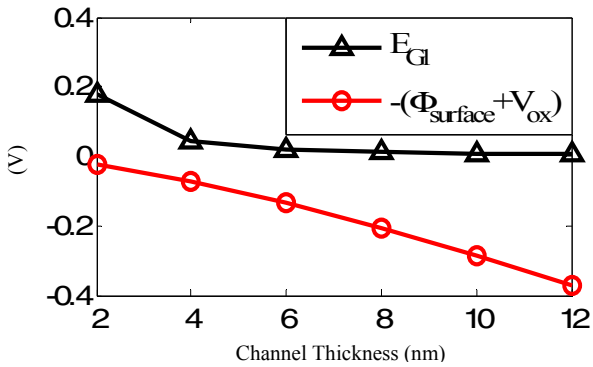


Fig. 10: Variation of First Eigen Energy and the purely classical term of equation (6) with silicon channel thickness.

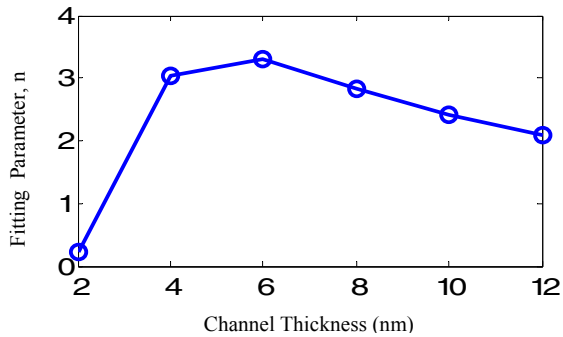


Fig. 11: Fitting Parameter vs. silicon channel thickness.

expected, the threshold voltage increases for downscaling silicon channel thickness, with a sharp transition below 6nm. Accurate estimation of the threshold voltage demands comprehensive incorporation of QM effects, which has been the focus of our study in this paper. Finally, fitting parameter n is introduced which stands for the quantum mechanical correction needed in the purely classical term ($\varphi_{surface} + V_{ox}$).

References

[1] "The international technology roadmap for semicoductors," 2008. [On-line]. Available: http://www.itrs.net/Links/2008ITRS/Home2008.htm

[2] A. Ortiz-Conde, F. J. Garc´ıa S´anchez, J. J. Liou, A. Cerdeira, M. Estrada,Y. Yue, "A review of recent MOSFET threshold voltage extraction methods," Microelectronics Reliability , vol. 42, no. 4-5, pp. 583-596,Apr.-May 2002.

[3] Emran Md. Amin , Md. Zunaid Baten, Raisul Islam and Quazi D. M.Khosru, "Quantum Mechanical Effect on Determining Threshold Voltage of Trigate FinFET Using Self-Consistent Analysis," TENCON 2009 -2009 IEEE Region 10 Conference, pp. 1-5.

[4] Rifat Zaman, Saeed Uz Zaman Khan, Md. Shafayat Hossain, Fahim Ur Rahman, Md. Obaidul Hossen, Quazi D. M. Khosru, ICECE, pp. 678-681, Dec 2012.

[5] R. D. Trevisoli, R. T. Doria, M. de Souza, J. P. Colinge and M. A. Pavanello, EUROSOI, 2012.

[6] R.D. Trevisoli *et al.*, Semi. Sci. Tech, Vol. 26, No. 10, pp. 105009(1-8), 2011.

[7] H.S. Wong *et al.*, Solid State Elec.,30, p. 953, 1987.

[8] S. Gupta., B. Ghosh, and S. Balmukund Rahi, Journal of Semiconductors, Vol. 36, No. 2, 2015.

[9] Se Re Na Yun, Chong Gun Yu, Jong Tae Park, Chi-Woo Lee, D.Lederer, A. Afzalian, Ran Yan, J.P. Colinge, "A quantum definition of threshold voltage in MuGFETs," in IEEE International SOI ConferenceProceedings, 2007.

[10] J. P. Colinge, J. C. Alderman, Weize Xiong and C. R. Cleavelin,"Quantum mechanical effects in trigate SOI MOSFETs," IEEE Trans.Electron Devices, vol. 195, no. 5, pp. 1131-1136, May 2006.

[11] Sung-Jin Choi, Dong-Il Moon, Sungho Kim, Juan P. Duarte, and Yang-Kyu Choi, IEEE Electron Device Letters, Vol. 32, No. 2, Feb 2011.

[12] J. P. Colinge, C. W. Lee, N. Dehdashti Akhavan, R. Yan, I. Ferain, P. Razavi, A. Kranti and R. Yu, "Junctionless Transistors: Physics and Properties", Semiconductor-On-Insulator Materials for Nanoelectronics Applications, pp. 187-200, Dec 2010.

Microcontroller Based Earthquake Detection System for Spontaneous Cut-off of Domestic Utility Lines for Safety Measures

Swapnil Sayan Saha[1], Shekh Md. Mahmudul Islam[2], Anindita Mashsharat[3]
Department of Electrical & Electronic Engineering, University of Dhaka
Dhaka, Bangladesh
swapnilsayansaha@ieee.org[1], mahmud@du.ac.bd[2], aninditamashsharat@gmail.com[3]

***Abstract*— Bangladesh is a country with a high frequency of earthquakes. Since the country lies at the junction of three tectonic plates, the intensity of earthquakes felt in this region is moderate. Surprisingly, the number of deaths and financial loss in this region by earthquakes is not due to building crashes or being crushed under homes. Rather, major reasons of losses are due to indirect phenomena such as induction of fear, as well as fire induced from a cracked gas line or faulty electrical transmission line damaged by earthquakes. As a result, a low cost automated microcontroller based system has been designed and implemented using inexpensive locally sourced electronic components, which senses earthquakes and gas leaks via force sensitive resistors and gas sensors respectively. The microcontroller operates a relay and a motor that cuts off electricity and gas supplies respectively during the event of an earthquake, helping to prevent associated potential disasters.**

Keywords- Microcontroller, Force Sensitive Resistor, Gas Sensor, Relay, Servo Motors

I. INTRODUCTION

Intermittent and frequent earthquakes are a common cause of terror and concern amongst the citizens of Bangladesh. As a country situated at the intersection of the Indian Plate, the Eurasian Plate and the Burma Plate, Bangladesh has been affected by some large earthquakes; some of them had magnitudes greater than 7.0 on the Richter scale and had epicenters inside the country [1]. Several major historical earthquakes that have affected Bangladesh are shown in Table I.

TABLE I. LIST OF MAJOR HISTORICAL EARTHQUAKES AFFECTING BANGLADESH (>7.0 MAGNITUDE) [1][2][3]

Date	Name	Magnitude (Richter)	Epicentral distance from Dhaka (in km)
8th Jul 1918	Srimangal Earthquake	7.6	150
2nd Jul 1930	Dhubri Earthquake	7.1	250
23rd Oct 1943	Numaligarh Earthquake	7.2	438
12th Sep 1946	Mawlaik Earthquake	8.0	407
15th Aug 1950	Assam Earthquake	8.5	-

Earthquakes over the last few years that have affected this region have stayed within the 4-7 magnitude range, with the 2015 Nepal Earthquakes as exceptions. Figure 1 shows the magnitudes of earthquakes near Dhaka, Bangladesh that have affected the fast emerging megacity in the last ten years (2006-2016).

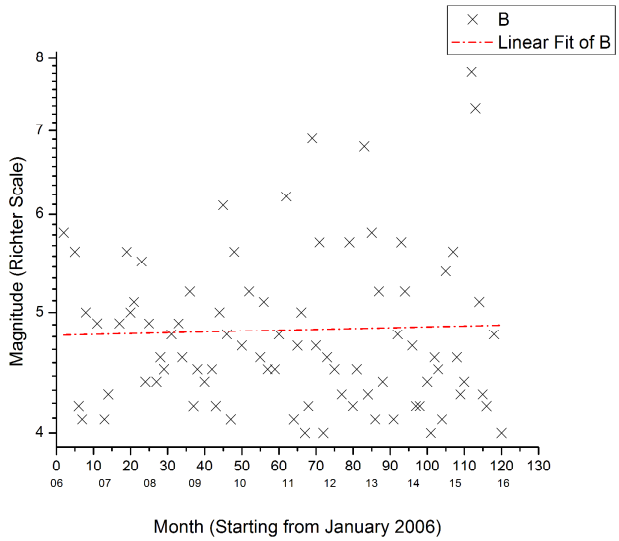


Figure 1. Scatter plot of Earthquakes affecting Bangladesh in the last ten years (Data taken from earthquaketrack.com [3]).

The red line in Figure 1 shows the mean earthquake magnitude to be 4.8 with a slight postive correlation.

News of deaths due to earthquakes caused by being crushed under homes is quite rare in Bangladesh, thanks to the epicentrial distances of extremely large earthquakes being 400-800 km away from the capital city. However, other than injuries and deaths caused by induction of panic, another subsidiary cause of fatalities might be fire caused by a fissured gas supply line or short circuits from faulty electricity transmission lines. Examples of severe destruction of lives and properties caused by fires from gas lines include fires caused by the San Franscisco (1906), the Tokyo (1923), The Loma

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

Prieta (1989), Northridge (1994) and Kobe (1995) earthquakes [4]. Since gas and electrical utility lines of Bangladesh especially are prone to damage due to earthquakes, it is advisable that these utlity lines should be cut-off during the interval during which the tremors are felt.

Unfortunately, the natural instinct of humans would be to take cover or run for their lives rather than doing something intricate like switching off circuit breakers or turning off gas valves during emergencies. Hence, this paper presents the design and implementation of an automated earthquake detection system that automatically cuts of the discussed two utlity lines using the AVR family of microcontrollers.

II. Literature Review

Several US patents exist which claim to shut off utility lines during earthquakes [5] [6]. This system, compared to the existing systems, uses a microcontroller as the central hub. Force sensitive resistors tuned to resonate at seismic wave frequencies have been utilized for vibration detection. The system also encompasses several modern and convenient features such as integration with an android application, turning on emergency lights to guide the user to an emergency evacuation route and using an attention grabbing alarm.

Compared to existing systems using mechanical means or relays, this system uses a single high current relay and a servo motor to turn off the entire domestic gas and electricity supply.

III. Materials and Methods

A. *Operational Principle*

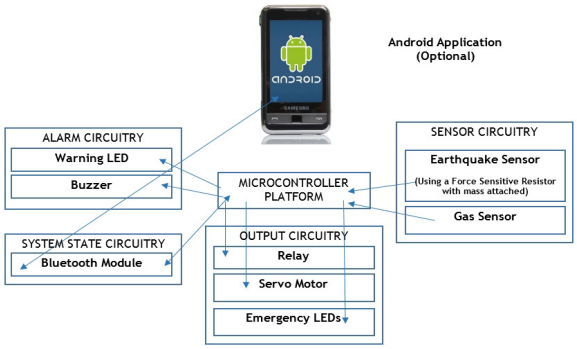


Figure 2. Block diagram of the designed automated earthquake detection system. The relay and the servo motor are connected to the electrcity and gas supplies respectively.

As shown in Figure 2, the designed system consists of primarily five parts. The microcontroller platform controls the operation of the entire system. The sensor circuitry detects tremors caused by earthquakes as well as leaks in the gas line. Once an earthquake is detected, the microcontroller platform shuts down the electricity supply through a relay and causes the servo motor attached to the gas line valve to rotate. It also turns on the emergency lights which illuminates the evacuation path for the user. The alarm circuitry is also activated simultaneously.

If no earthquakes are detected but a gas leak is detected by the gas sensor, then the system shuts down the gas valve only.

The entire operation can be monitored via an android application, which facilitates remote operation and control of the system.

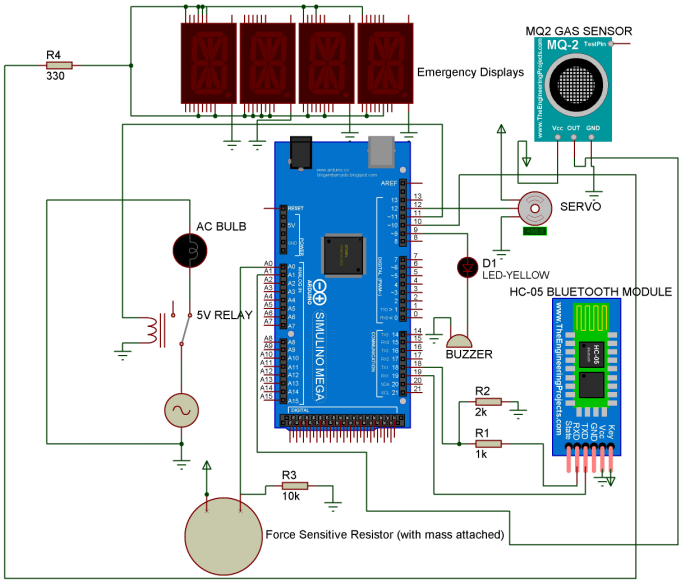


Figure 3. Schematic Diagram of the Entire System.

B. *The Microcontroller Platform*

This system has been designed using the Arduino Mega microcontroller development platform, which features the Atmel Atmega 2560 microcontroller. The commercial prototype should not be built using the Arduino if one wants to lessen production costs and space [7].

C. *The Sensor Circuitry*

Two sensors have been utilized in this designed system for two different purposes.

In this designed system, the MQ-2 gas sensor module has been utilized. This particular sensor has been chosen due to its non-exorbitant cost, low power requirements, wide detecting scope, fast response, high sensitivity and a long life [8], all of which promotes its use for engineering applications in Bangladesh. Most vitally, the sensor is well known for its use in embedded systems in detecting combustible gases [8].

Figure 4 shows the sensitivity curve of this gas sensor [9].

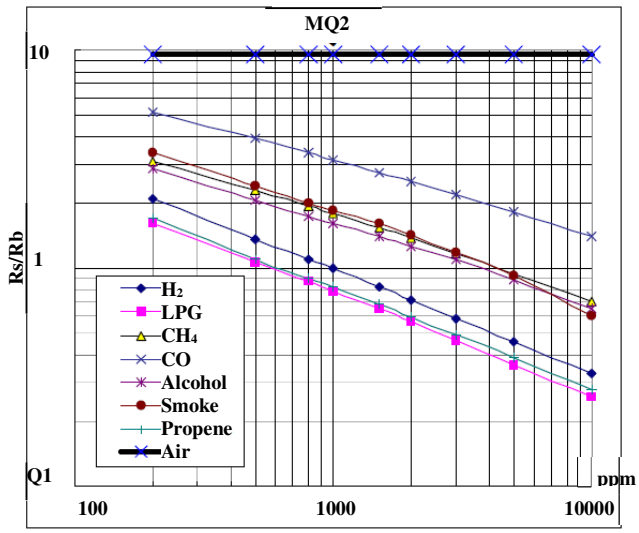


Figure 4. Typical Sensitivity Characteristics of the MQ-2 Gas Sensor [9]. We are interested in the sensitivity curves of methane, alcohol, carbon monoxide and smoke in this designed system.

Concentrations of atmospheric methane is 0.722 ppm [10] and carbon monoxide is 0.2 ppm, which are way beneath the minimum sensitivity of the sensor as show in Figure 6. Hence, the sensor would not produce false values due to atmospheric concentration of these gases.

A force sensitive resistor is a variable resistor whose resistance decreases with an increase in force applied on the resistor [7].

In order for the force sensitive resistor to oscillate at surface wave frequencies of earthquakes, a mass should be attached to the force sensitive resistor such that the resonant frequency of oscillation of the resistor matches with the resonant frequency of the surface waves.

Several factors should be considered in choosing the resonant frequency of oscillation for the sensor:

- *Resonant Frequency of the buildings:* As a rule of thumb, the resonant frequency of civil structures is found using the following empirical relationship [11]:

  $$1/\phi_B = 0.1N \qquad (1)$$

  where,
  ϕ_B = Resonant frequency of the building
  N = Number of stories of the building (1, 2...)
- *Seismic Microzonation map of the area:* Not all structures within a given area may have the same resonant frequency. Thus, seismic microzonation needs to be applied in order to determine which structures are at the maximum risk. For this particular project, the city of Dhaka was chosen as the intended area. Figure 5 shows the seismic microzonation map of Dhaka, divided into 4 bands [12].

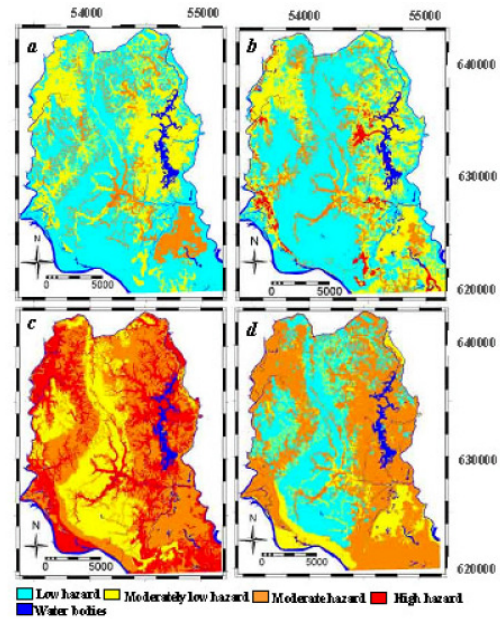


Figure 5. Seismic Microzonation Map of Dhaka City. Band a is for structures 2-5 storeys high, Band b is for structures of 6-8 stories high, Band c is for structures 9-14 stories high, Band d is for structures 15-20 stories high [12].

It is evident from Figure 5 that buildings 9-20 stories high are at the maximum risk of being affected by earthquakes, whose resonant frequencies lie in the range of 0.5-1 Hz using eqn. 1.

As a result, the magnitude of the mass that needs to be attached to the system is given by:

$$m = k/(2\pi\phi_B)^2 \qquad (2)$$

where,
ϕ_B = Resonant frequency of the building in Hertz
k = Spring constant of the oscillating system
m = mass of the oscillating system

The spring constant of an end-loaded cantilever is given by:

$$k = Et^3w/4l^3 \qquad (3)$$

where,
E = Young's Modulus, which can be found from Figure 6
t = thickness, w = width, l = length of the system [13].

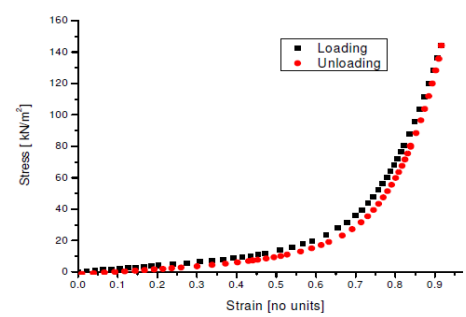


Figure 6. Stress-strain characteristics across two faces of a carbon filled PDMS sponge, the material typically used for crafting force sensitive resistors [14].

The high amplitude oscillations of the mass-sensor system causes a force to act on the sensor, which alters its resistance and thus, voltage output from the sensor increases.

D. The Output Circuitry:

As soon as an earthquake is detected, the microcontroller activates a relay that cuts off electricity supply. It also sends pulses to a servo motor attached to the gas valve which rotates in order to turn a gas valve off, cutting off gas supply. If there is no earthquake event but a gas leak is detected, then the microcontroller platform simply turns off the gas supply. The system also activates a set of emergency alphanumeric displays that read "E X I T" once an earthquake is triggered to mark the evacuation route for the user.

E. The Alarm Circuitry:

The alarm circuitry consists of a DC buzzer and an LED to warn the user of a possible earthquake event through the use of light and sound.

F. The System State Circuitry:

The system state circuitry consists of the Bluetooth module and the android application for serial communication. Once an earthquake is detected, a warning message is displayed on the android application. However, the application's role is prominent when a gas leak is detected and the system turns off gas supply completely. It displays a message "Gas leak detected! Please consult a qualified repair personnel" Once the leak has been fixed by the technician, the user can reactivate the gas supply by typing in "A" from the application.

IV. Implementation and Cost Analysis

TABLE II. Bill of Materials for Prototyping the System

Feature	Quantity	Unit Price (USD)	Cost (USD)
Arduino Mega 2560	1	14.74	14.74
MQ-2 Gas Sensor Module	1	12.10	12.10
406 Force Sensitive Resistor	1	11.70	11.70
HC-05 Bluetooth Module Breakout	1	7.13	7.13
4 Channel 5V Relay Module	1	3.59	3.59
Buzzer	1	0.26	0.26
Alphanumeric Displays	4	0.60	2.40
Servo Motor	1	3.90	3.90
Miscellaneous (LED, Light Bulb, Power Cord, Power supplies)	-	-	3.00
TOTAL			**58.82**

Table II shows that the entire system was built under $60 (all materials were bought from techshopbd.com). A commercial prototype of this system would cost around $100-120. Note that, a device that mimics the operation of the designed system is not available in this country. Seismic sensors are available on eBay, lacking capabilities to turn off utility lines and only supporting the detection of earthquakes.

V. Results, Conclusion & Future Work

The designed system was a just a proof of qualitative concept and was not tested on a large scale during actual earthquakes. The prototype did manage to detect simulated earthquake-like tremors with magnitudes of 4.0+.

Several improvements may be brought to this system:

1. A 3 axis accelerometer can replace the FSR for detection in all planes independent of frequency.
2. 2D Multi-gas mapping techniques with cascading module system over a wireless sensor network can be used to monitor the gas leaks at multiple points [15].
3. Pattern recognition algorithms based on positive decision logic may be integrated into this system [16].

References

[1] T. Al-Hussaini, I. Nawrin Chowdhury and M. N. A. Noman, "Seismic Hazard Assessment for Bangladesh - Old and New Perspectives", in First International Conference on Advances in Civil Infrastructure and Construction Materials, Military Institute of Science & Technology, Dhaka, 2015.

[2] Md Hossain Ali. "Earthquake Database and Seismic Zoning of Bangladesh." INCEDE Report 11, 1998: 59-73.

[3] "Today's Earthquakes in Bangladesh > 4.0M", Earthquaketrack.com, 2016. [Online]. Available: http://earthquaketrack.com/p/bangladesh/recent?mag_filter=4. [Accessed: 24- Jun- 2016].

[4] Mohammad Sayeed Hossain, "Estimate of Damage of Burried Gas Pipelines in Dhaka City Due to an Earthquake" MSc. Thesis, Dept. of Civil Eng., Bangladesh University of Engineering & Technology, Dhaka, 2013.

[5] R. A. Pelliccia, "Vibration sensor and electrical power shut off device", US4390922 A, 1983.

[6] A. Y. Flig, Paul Regan, "Earthquake utilities cut-off control system", US4841287 A, 1989.

[7] Swapnil Sayan Saha, Shekh Md. Mahmudul Islam, "Microcontroller Based Automated Domestic Security System for Bangladesh Perspective", in Student Conference on Informatics, Electronics & Vision, University of Dhaka, Dhaka, 2016.

[8] V. Ramya, and B. Palaniappan. "Embedded system for Hazardous Gas detection and Alerting." International Journal of Distributed and Parallel Systems (IJDPS) Vol 3 (2012): 287-300.

[9] "MQ-2 Gas Sensor Datasheet", DFRobot, 2016. [Online]. Available: http://www.dfrobot.com/image/data/SEN0127/MQ-2.pdf. [Accessed: 25- Jun- 2016].

[10] T. J. Blasing, "Recent greenhouse gas concentrations." (2009).

[11] Rakesh K. Goel, and Anil K. Chopra. "Period formulas for moment-resisting frame buildings." Journal of Structural Engineering 123.11 (1997): 1454-1461.

[12] A. S. M. M. Kamal, and S. Midorikawa. "Geomorphological approach for seismic microzoning within Dhaka city area, Bangladesh." International Association for Engineering Geology and the Environment 457 (2006): 1-2.

[13] J. P. Cleveland, et al. "A nondestructive method for determining the spring constant of cantilevers for scanning force microscopy." Review of Scientific Instruments 64.2 (1993): 403-405.

[14] M. G. King, et al. "Porous PDMS force sensitive resistors." Procedia Chemistry 1.1 (2009): 568-571.

[15] A. Al Arabi, B. Barman Shubha, A. Diana Corraya, "2D Multi-gas Mapper with Cascading Module System along with Wireless Sensor Network" in 5th International Conference on Informatics, Electronics and Vision, University of Dhaka, Dhaka, 2016

[16] Manfred Joswig, "Pattern recognition for earthquake detection." Bulletin of the Seismological Society of America 80.1 (1990): 170-186.

Reduction of Detection Volume in Total Internal Reflection Fluorescence Microscopy Using Graphene

Shiekh Zia Uddin and Muhammad Anisuzzaman Talukder*
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka 1205, Bangladesh
*anis@eee.buet.ac.bd

***Abstract*—We show that it is possible to decrease the thickness of the detection volume of total internal reflection fluorescence microscopy (TIRFM) by ∼35% using a graphene layer in the interface between glass and water layers of a typical TIRFM structure without sacrificing the fluorescence intensity. The highly mobile surface bound electrons of a graphene mono-layer quenches the fluorophores that are less than ∼40 nm away from it. The decreased detection volume using the proposed structure will increase the resolution of a typical TIRFM technique. We find that the results are qualitatively similar for different incidence angles and polarizations of the excitation field. So, the proposed structure will also find applications in variants of TIRFM techniques, e.g., where incidence angles and polarizations are varied.**

***Index Terms*—Total Internal Reflection Fluorescence Microscopy, Graphene, and Detection Volume.**

I. Introduction

When light propagating through a medium of high index of refraction (glass) encounters a planar interface with a medium of low index of refraction (water), it undergoes total internal reflection for incidence angles greater than the critical angle. Although, totally reflected, the incident beam creates an evanescent electromagnetic field that penetrates into the medium of low index of refraction and decays exponentially from the interface [1]. Total internal reflection fluorescence microscopy (TIRFM) employs this evanescent field to selectively excite fluorophores only in a very thin layer near the substrate, and thereby, achieving resolution much greater than that possible by the diffraction limited optics [2], [3]. TIRFM has been used for the last three decades in a wide range of applications, e.g., to image single molecules attached to planar surfaces and to study the position, orientation and dynamics of molecules, and organelles in living culture cells near the contact regions with the glass [4]–[6].

Graphene is a single atomic layer of carbon arranged in a hexagonal lattice [7], which has drawn intense interest after being discovered in 2004 [8], [9]. Graphene has excellent mechanical strength, chemical stability, electro-optical tunability as well as the highest mobility of carriers (both electrons and holes) due to the unique conical band structure. Although, initially, graphene was widely explored for applications in electronics [10], in recent years, it has also shown promises as a key element in photonics [11], optoelectronics [12], plasmonics [13], metamaterials [14], and biology [15].

In this work, we propose a graphene-based TIRFM technique that will decrease the detection volume significantly and thus, will increase the resolution of microscopy significantly. In particular, we propose to place a graphene mono-layer in the interface between glass and water layers of a typical TIRFM technique. The mobile electrons of the graphene quench the fluorescence of the molecules that are close to the graphene layer. This quenching effectively decreases the thickness of the detection volume by ∼35% of that of a typical TIRFM technique without decreasing fluorescence intensity. The decreased detection volume will increase the resolution of TIRFM, and thus, furthering the study of biology. We find that the detection volume decreases for all angles of incidence greater than the critical angle and all polarizations of the excitation field. So, the proposed structure can also be used to reduce detection volume in modified versions of TIRFM where incidence angle and polarization variations are used to acquire images, e.g., in polarized TIRFM.

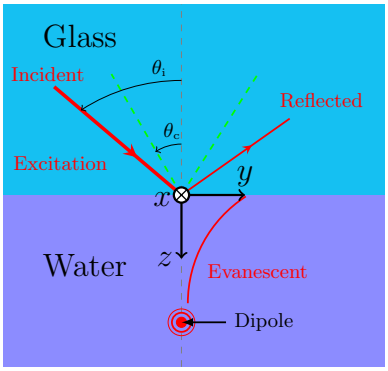


Fig. 1. Typical TIRFM structure where a glass-water interface is used to generate the evanescent excitation wave.

II. Proposed Structure

In Fig. 1, we show a schematic illustration of a typical structure that is used in TIRFM. The planar interface between a glass prism (dielectric constant $\epsilon_p = 2.304$) and water ($\epsilon_w = 1.769$) is used to generate the evanescent field. The sample labelled with fluorophores is placed in water near the interface. The excitation light incident from the prism side creates an exponentially decaying field profile in water from the interface. When fluorophores are excited by this evanescent light, they become fluorescent and emit light, which is collected in the prism side and converted to a microscopy image using appropriate optics [4]. In this work, we propose to place a graphene mono-layer in the glass-water interface

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

of a typical structure. The excitation and collection schemes used in a typical TIRFM technique can also be used in the proposed technique.

III. Theoretical Background

A. The Optical Properties of Graphene Sheets

The widely used model for optical response of graphene is a very thin, two-sided surface, which is characterized by a surface conductivity $\sigma(\omega, \mu_c, \Gamma, T)$, where ω is the angular frequency, μ_c is the chemical potential, Γ is the phenomenological scattering rate, and T is the temperature. Graphene's complex conductivity can be determined from the Kubo formula [16]–[18]

$$\sigma(\omega, \mu_c, \Gamma, T) = \sigma_{\text{intra}}(\omega, \mu_c, \Gamma, T) + \sigma_{\text{inter}}(\omega, \mu_c, \Gamma, T), \quad (1)$$

where σ_{intra} and σ_{inter} are intra-band and inter-band conductivity, respectively. The simplified intra-band contribution can be written as [18]

$$\sigma_{\text{intra}}(\omega, \mu_c, \Gamma, T) = i\frac{1}{\pi\hbar^2}\frac{e^2 k_B T}{(\omega + 2i\Gamma)} \left\{\frac{\mu_c}{k_B T} + 2\ln\left[\exp(-\frac{\mu_c}{k_B T}) + 1\right]\right\}, \quad (2)$$

and the inter-band contribution can be approximated as ($\mu_c >> k_B T$)

$$\sigma_{\text{inter}}(\omega, \mu_c, \Gamma, T) = i\frac{e^2}{4\pi\hbar}\ln\left[\frac{2|\mu_c| - \hbar(\omega + 2i\Gamma)}{2|\mu_c| - \hbar(\omega + 2i\Gamma)}\right], \quad (3)$$

where e is the electronic charge, $\hbar$ is the reduced Planck's constant, and k_B is the Boltzmann constant.

The chemical potential μ_c can be tuned by an application of transverse gate voltage, electric field, magnetic field, and chemical doping. Typically chemical potential can be varied in the range of 0 to 1 eV. The phenomenological relaxation rate Γ can be approximated to be 0.1 meV, as reported in Ref. [19]. Because of the 2D nature of the graphene sheet, it behaves as a uni-axially anisotropic material, whose dielectric tensor can be written as

$$\overleftrightarrow{\epsilon}_{\text{gr}} = \begin{bmatrix} \epsilon_{\text{g}\|} & 0 & 0 \\ 0 & \epsilon_{\text{g}\|} & 0 \\ 0 & 0 & \epsilon_{\text{g}\perp} \end{bmatrix}, \quad (4)$$

where $\epsilon_{\text{g}\|}$ and $\epsilon_{\text{g}\perp}$ are the ordinary and extraordinary relative permittivities of graphene. Due to the two dimensional nature of graphene, the electric field that is polarized in the normal direction to the graphene layer cannot excite any current in it. So, the normal component of the permittivity or the extraordinary permittivity of graphene is given by $\epsilon_{\text{g}\perp} = 1$ [20]. The ordinary or the tangential permittivity can be defined by [17]

$$\epsilon_{\text{g}\|} = 1 + i\frac{\sigma(\omega, \mu_c, \Gamma, T)}{\omega\epsilon_0 t_{\text{gr}}}, \quad (5)$$

where $t_{\text{gr}} = 0.264$ nm is the thickness of a graphene monolayer. The real and imaginary parts of the graphene tangential permittivity $\epsilon_{\text{g}\|}$ are shown in Figs. 2(a) and 2(b) for varying free-space wavelength (λ) and chemical potential (μ_c). We note that the resonance wavelength of the graphene blue-shifts as the chemical potential increases. In the analysis presented in this paper, we choose $\mu_c = 0.9$ meV.

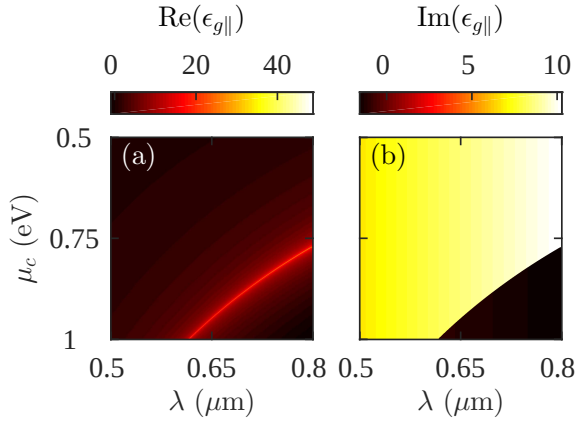


Fig. 2. (a) Real and (b) imaginary parts of the tangential permittivity of graphene ($\epsilon_{\text{g}\|}$) for varying wavelength (λ) and chemical potential (μ_c).

B. Total Internal Reflection

The critical angle for incident beam from glass is $\theta_c = \sin^{-1}\left(\sqrt{\epsilon_w/\epsilon_p}\right)$, where ϵ_w and ϵ_p are the dielectric constants of water and glass prism, respectively. For any polarization, if the incidence angle (θ_i) of an incident wave of wavelength λ is greater than the critical angle, then the transmitted wave is exponentially decaying with the distance from the interface. For both typical and proposed structures, the p-polarized excitation field can be written as [25]

$$\mathbf{E_{ex}} = E_0 t_p(\hat{\mathbf{x}}\sqrt{\epsilon_p \sin^2\theta_i - \epsilon_w} + j\hat{\mathbf{z}}\sqrt{\epsilon_p}\sin\theta_i)\exp(-z/2d), \quad (6)$$

and the s-polarized excitation field can be written as

$$\mathbf{E_{ex}} = E_0 t_s \hat{\mathbf{y}}\exp(-z/2d), \quad (7)$$

where $d = (\lambda/4\pi)\sqrt{\epsilon_p \sin^2(\theta_i) - \epsilon_w}$ is the penetration depth, E_0 is the amplitude of incident light, and t_p and t_s are Fresnel transmission coefficients for p- and s-polarized light, respectively. The Fresnel coefficients of anisotropic layered media are calculated using the 4×4 transfer matrix method [21]–[23].

C. Fluorescence Near Planar Stratified Media

Flurophores are molecules that absorb light at a specific frequency and emit at a different frequency. The emitted fluorescence is collected by appropriate optics. If the fluorophores are randomly oriented and the concentration $C(z)$ is only z dependent, then the collected fluorescence intensity from a pixel in a planar microscope can be written as [24], [25]

$$\begin{aligned} F &= k\int dz\ C(z)[w^{\perp}(z)Q^{\perp}(z) + w^{\|}(z)Q^{\|}(z)] \\ &= k\int dz\ C(z)g(z), \end{aligned} \quad (8)$$

where k is a proportionality constant consisting of conversion factors, normalization and arithmetic constants, $w^{\perp,\|}(z)$ are the weighting terms, and $Q^{\perp,\|}(z)$ are the collection efficiencies for vertical and horizontal dipoles. The weighing function $g(z) = [w^{\perp}(z)Q^{\perp}(z) + w^{\|}(z)Q^{\|}(z)]$ acts as a mask on the concentration profile and selects the region of sample volume whose concentration contributes to the collected fluorescence intensity. The shape of the weighing function $g(z)$ determines the thickness of detection volume. Calculation of $g(z)$ requires evaluation of collection efficiencies and weighing terms.

Collection efficiency is the ratio of fluorescent energy collected by the imaging system to the total emitted energy, and it depends on the position, orientation and environment of the fluorophore. The theory for calculating collection efficiencies for fluorophores in any stratified media is outlined in Refs. 24 and 25. We use this theory to calculate the collection efficiencies for horizontal and vertical dipoles in typical and proposed structures and show them in Fig. 3. In the proposed structure, the collection efficiencies decrease significantly if the fluorophores are near the interface.

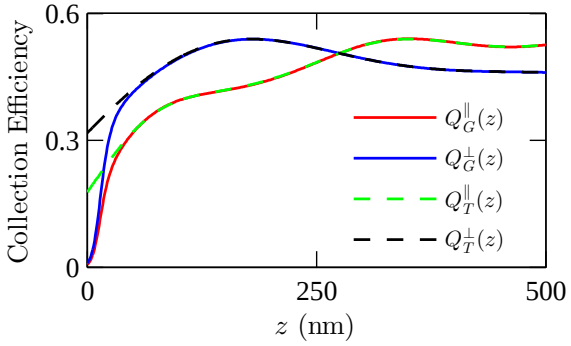


Fig. 3. Collection efficiency for horizontal and vertical dipoles in the proposed structure [$Q_G^{\parallel}(z)$ and $Q_G^{\perp}(z)$] and the typical structure [$Q_T^{\parallel}(z)$ and $Q_T^{\perp}(z)$], respectively.

If $\mathbf{E^{ex}} = \hat{\mathbf{x}}E_x^{ex} + \hat{\mathbf{y}}E_y^{ex} + \hat{\mathbf{z}}E_z^{ex}$ is the phasor representation of the excitation electric field, then the weighting terms can be written as

$$w^{\perp,\parallel}(z) = |E_x^{ex}|^2 w_x^{\perp,\parallel}(z) + |E_y^{ex}|^2 w_y^{\perp,\parallel}(z) + |E_z^{ex}|^2 w_z^{\perp,\parallel}(z). \tag{9}$$

The detail forms of the parameters $w_{x,y,z}^{\perp,\parallel}(z)$ are given in Ref. 24. The parameters $w_{x,y,z}^{\perp,\parallel}(z)$ are functions of lifetime ratio $\eta(z) = \tau^{\perp}(z)/\tau^{\parallel}(z)$, where $\tau^{\parallel}(z)$ and $\tau^{\perp}(z)$ are the lifetimes of horizontal and vertical dipoles situated at height z.

IV. Results

The lifetimes for horizontal and vertical dipoles in the typical and proposed structures are shown in Fig. 4(a). The lifetime ratio function $\eta(z)$ for the typical and proposed structures is shown in Fig. 4(b). We note a difference in $\eta(z)$ in a typical and the proposed structure when the fluorophores are near the interface. The weighing terms, which are functions of the lifetime ratio $\eta(z)$, are shown in Fig. 4(c).

Figure 5 shows the weighing function $g(z)$ that selects a region of the sample, where the fluorophores are excited to emit light and contribute to the collected fluorescence. Figure 5(a) shows the weighing function for different values of incidence angle of p-polarized excitation field for the typical TIRFM structure. We note that, for any incidence angle, the weighing function decreases with height (z) from the glass-water interface. The weighing function in the proposed structure is shown in Fig. 5(b) for p-polarized excitation. We note that, for a region close to the interface, the weighing function is zero due to quenching of fluorophores by the graphene layer. The quenching of fluorescence decreases the effective thickness of sample layer excited by the p-polarized field. We observe similar results when the excitation light is s-polarized, as shown in Figs. 5(c) and 5(d). The amplitude

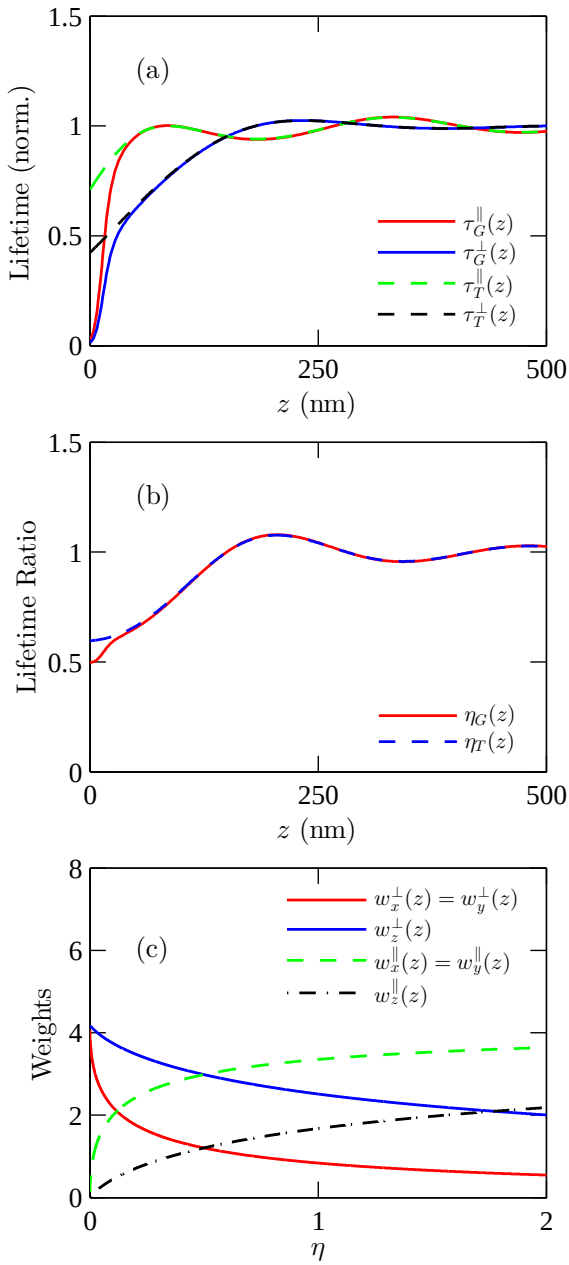


Fig. 4. (a) Lifetime of horizontal and vertical dipoles with position in typical and proposed structures, (b) Lifetime ratio with position in typical and proposed structures, and (c) weighing terms with lifetime ratio.

levels of the weighing functions in the typical and proposed structures are similar.

We calculate the full width at the half maximum (FWHM) of the weighing function $g(z)$ to determine the thickness of the detection volume in the z-direction. Figure 6(a) shows the FWHM thickness of the detection volume for p-polarized excitation in the typical (δ_T^p) and the proposed structures (δ_G^p). We find that the FWHM thickness of $g(z)$ of the proposed structure decreases by ∼80 nm and ∼40 nm for an angle of incidence of 65° and 85°, respectively. Similar quantitative results are found for s-polarized excitation.

V. Conclusion

We used mono-layer graphene in the glass-water interface to modify the behaviour of TIRFM and found decreased detection volume without decreasing fluorescence intensity. We found

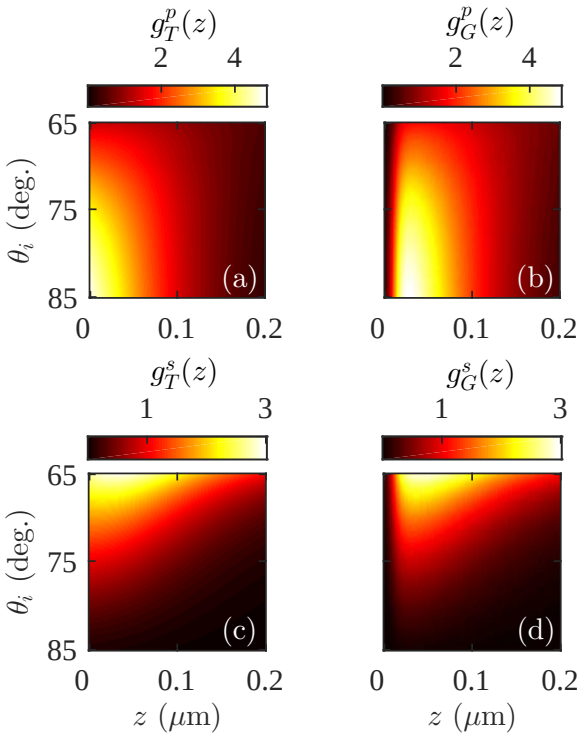


Fig. 5. Weighing function $g(z)$ for varying incidence angle and position in the (a) typical and (b) proposed structure for p-polarized light and (c) typical and (b) proposed structure for s-polarized light.

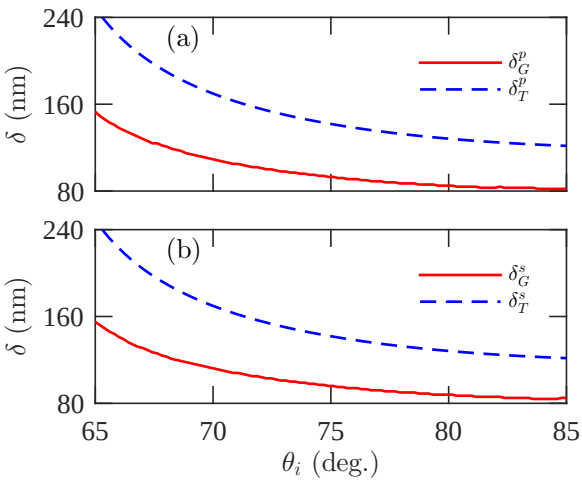


Fig. 6. Detection volume thickness for (a) p- and, (b) s-polarized light in the typical and proposed TIRFM structure.

that the detection volume in TIRFM can decrease as much as ∼35% with the use of a graphene mono-layer. The results remain qualitatively the same for both p- and s-polarized incident light.

VI. Acknowledgement

We gratefully acknowledge the support received from the Department of Electrical and Electronic Engineering of Bangladesh University of Engineering and Technology (BUET) in carrying out this work.

References

[1] S. J. Orfanidis, *Electromagnetic waves and antennas*, Rutgers University, New Brunswick, NJ, 2002.
[2] D. Axelrod, "Total internal reflection fluorescence microscopy," *Methods in Cell Biology*, vol. 30, pp. 245–270, 1989.
[3] D. Axelrod, E. H. Hellen, and R. M. Fulbright, *Topics in fluorescence spectroscopy*, Springer, 2002.
[4] A. L. Mattheyses, S. M. Simon, and J. Z. Rappoport, "Imaging with total internal reflection fluorescence microscopy for the cell biologist," *Journal of Cell Science*, vol. 123, pp. 3621–3628, 2010.
[5] D. S. Johnson, J. K. Jaiswal, and S. Simon, "Total internal reflection fluorescence (TIRF) microscopy illuminator for improved imaging of cell surface events," *Current Protocols in Cytometry*, vol. 61, pp. 12–29, 2012.
[6] D. Axelrod, "Total internal reflection fluorescence microscopy in cell biology," *Traffic*, vol. 2, pp. 764–774, 2001.
[7] A. C. Neto, F. Guinea, N. Peres, K. S. Novoselov, and A. K. Geim, "The electronic properties of graphene," *Reviews of Modern Physics*, vol. 81, pp. 109, 2009.
[8] K. S. Novoselov, A. K. Geim, S. Morozov, D. Jiang, Y. Zhang, S. Dubonos, , I. Grigorieva, and A. Firsov, "Electric field effect in atomically thin carbon films," *Science*, vol. 306, pp. 666–669, 2004.
[9] K. Novoselov, A. K. Geim, S. Morozov, D. Jiang, M. Katsnelson, I. Grigorieva, S. Dubonos, and A. Firsov, "Two-dimensional gas of massless dirac fermions in graphene," *Nature*, vol. 438, pp. 197–200, 2005.
[10] A. K. Geim, K. Novoselov, and S. Konstantin, "The rise of graphene," *Nature Materials*, vol. 6, pp. 183–191, 2007.
[11] F. Bonaccorso, Z. Sun, T. Hasan, and A. Ferrari, "Graphene photonics and optoelectronics," *Nature Photonics*, vol. 4, pp. 611–622, 2010.
[12] C.-H. Liu, Y.-C. Chang, T. B. Norris, and Z. Zhong, "Graphene photodetectors with ultra-broadband and high responsivity at room temperature," *Nature Nanotechnology*, vol. 9, pp. 273–278, 2014.
[13] A. Grigorenko, M. Polini, and K. Novoselov, "Graphene plasmonics," *Nature Photonics*, vol. 6, pp. 749–758, 2012.
[14] M. A. Othman, C. Guclu, and F. Capolino, "Graphene-based tunable hyperbolic metamaterials and enhanced near-field absorption," *Optics Express*, vol. 21, pp. 7614–7632, 2013.
[15] K. Kostarelos and K. S. Novoselov, "Exploring the interface of graphene and biology," *Science*, vol. 344, pp. 261–263, 2014.
[16] L. Falkovsky and S. Pershoguba, "Optical far-infrared properties of a graphene monolayer and multilayer," *Physical Review B*, vol. 76, pp. 153410, 2007.
[17] L. Falkovsky, "Optical properties of graphene," *Journal of Physics: Conference Series*, vol. 129, pp. 012004, 2008.
[18] G. W. Hanson, "Dyadic greens functions and guided surface waves for a surface conductivity model of graphene," *Journal of Applied Physics*, vol. 103, pp. 064302, 2008.
[19] A. Vakil and N. Engheta, "Transformation optics using graphene," *Science*, vol. 332, pp. 1291–1294, 2011.
[20] B. Zhu, G. Ren, S. Zheng, Z. Lin, and S. Jian, "Nanoscale dielectric-graphene-dielectric tunable infrared waveguide with ultrahigh refractive indices," *Optics Express*, vol. 21, pp. 17089–17096, 2013.
[21] D. W. Berreman, "Optics in stratified and anisotropic media: 4×4-matrix formulation," *Journal of the Optical Society of America*, vol. 62, pp. 502–510, 1972.
[22] P. Yeh,"Optics of anisotropic layered media: a new 4×4 matrix algebra," *Surface Science*, vol. 96, pp. 41–53, 1980.
[23] M. Schubert, "Polarization-dependent optical parameters of arbitrarily anisotropic homogeneous layered systems," *Physical Review B*, vol. 53, pp. 4265, 1996.
[24] E. H. Hellen and D. Axelrod, "Fluorescence emission at dielectric and metal-film interfaces," Journal of the Optical Society of America B," vol. 4, pp. 337–350, 1987.
[25] L. Novotny and B. Hecht, *Principles of nano-optics*, Cambridge university press, 2012.

3-D Vertical Solenoid Inductors by Electrodeposition Techniques

Fazle Rabbi*,[1], Jannatun Naher*, Zhijian Xie*, *Senior Member, IEEE*, Numan S. Dogan*, *Senior Member, IEEE*
*RF Lab, Electrical and Computer Engineering Department, North Carolina A&T State University, United States.
[1]email: frabbi@aggies.ncat.edu

***Abstract*—On-chip spiral inductors suffer from substrate loss and parasitics, also have large footprints. This paper presents three-dimensional (3-D) vertical solenoid inductors by electro deposition technique. Metal ground shield inserted in the oxide layer to terminate electric field to enter inside substrate. Compared to on-chip spiral inductors, capacitive energy loss can be reduced about 60% and power loss about 90% in our proposed method.**

***Index Terms*—vertical solenoid inductor, electro deposition technique, quality factor, ground shield.**

I. Introduction

High quality inductors are known as performance and cost limiting elements in radio frequency (RF) integrated circuits such as LNA, voltage-controlled oscillator and impedence matching networks [1], [2]. Unwanted energy loss in planar spiral inductor were reported in many literatures [3]. Moreover, planar spiral inductors use large footprint of the chip area. To minimize these limitations many techniques were implemented. Ashby suggested use of higher substrate resistance and lower series resistance [4]. Yue proposed patterned ground shield for Si-based RF IC's [5]; but patterned ground shield increase coupling capacitance due to reduced distance between the metal and ground.

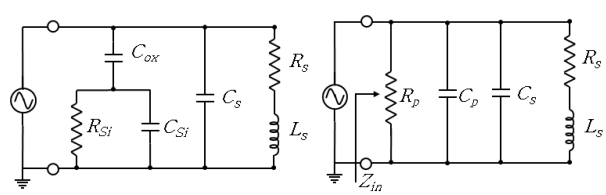
The improvement techniques of inductors performance can be categorize into two main branches; one is micromachining technology [6] and the other is post fabrication assembly process [7]. Micromachining technology can be implemented in two ways; one is bulk micromachining technology and another is surface micromachining technology [8]. In bulk micromachining technology substrate etching can be done completely which was reported in [9], [10]. In surface micromachining technology thick insulating layer or a sacrificial layer is deposited on substrate, has been applied in solenoid inductors [11], [12], spiral inductors [13], [14] and suspended spiral inductors [15], [16]. Although, good Q value is achievable by micromachining techniques at higher frequency, it raises concern in reliability and long term mechanical stability. In post fabrication assembly process inductor keep away from substrate by tilting [17], [18]. Completely different types of improvement techniques, neuoro-genetic algorithm may apply to improve inductor performance [19], [20].

In this paper, we have proposed three-dimensional (3-D) inductor fabricated by electro deposition technique. This technique will be compatible with existing IC foundry technology, because it will be an addition to foundry process and the inductor coil can be added on IC chips like a superstructure. In Section II we have discussed design considerations, which has included background and modeling of inductors, electrodeposition techniques, vertical solenoid inductor and effect of solid ground shield in vertical solenoid inductor. In section III results and comparison has presented. A brief conclusion has drawn in section III.

II. Design Considerations

A. Background and Modeling of Inductors

Spiral inductor model is shown in Fig. 1a [3], where series branch consists of R_s, C_s and L_s. R_s is the metal series resistance which signify skin effect, C_s is mainly overlap capacitance between spiral and underpass, L_s is the spiral inductance. The shunt branch parasitics can be modeled as C_{ox}, C_{Si} and R_{Si}. C_{ox} represents oxide capacitance between spiral and substrate, C_{Si} is silicon substrate capacitance and R_{Si} is substrate resistance. Fig. 1b is the equivalent representation of Fig. 1a, where C_{ox}, C_{Si} and R_{Si} is substituted by R_p and C_p (Eqn. 1 and 2), while R_s, C_s and L_s remain unchanged. Fig. 1b helps to understand parasitics impact easily, their values can be calculated from simulation and measurement data in a easy way.



(a) Lumped physical model (b) Equivalent circuit of Fig. 1a

Figure 1: Model of one terminal spiral inductor

$$R_p = \frac{1}{\omega^2 C_{ox}^2 R_{Si}} + \frac{R_{Si}\left(C_{ox}+C_{Si}\right)^2}{C_{ox}^2} \tag{1}$$

$$C_p = C_{ox}.\frac{1+\omega^2\left(C_{ox}+C_{Si}\right)C_{Si}R_{Si}^2}{1+\omega^2\left(C_{ox}+C_{Si}\right)^2 R_{Si}^2} \tag{2}$$

Although (1) and (2) shows R_p and C_p are frequency dependent, to avoid complexity in calculation we have considered these as fixed values. Parameters extraction of equivalent model (Fig. 1b) can be done in the following ways:

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

$$L_s = Im\left(Z_{in}\right)/\omega \tag{3}$$

$$R_s = Re\left(Z_{in}\right) \tag{4}$$

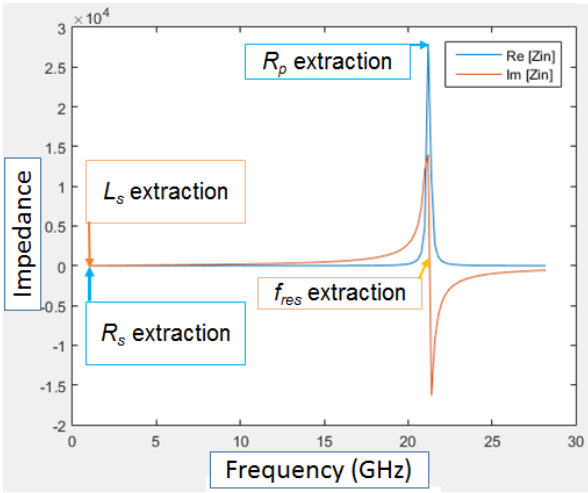


Figure 2: Inductor parameter extraction

$$R_p = Re\left(Z_{in}\right) \tag{5}$$

$$f_{res} = \frac{1}{2\pi\sqrt{L_s\left(C_s + C_p\right)}} \tag{6}$$

Eqn. (3) and (4) needs to apply much lower frequency than the resonance frequency (f_{res}) for L_s and R_s calculation (Fig. 2). The reason is, at lower frequency the impedence arise from L_s and R_s dominate over impedence arise from other parameters: C_p, C_s and R_p. At f_{res}, R_p dominates over R_s, so (5) can apply for R_p calculation at f_{res}. The simulation was performed for one terminal for parameter extraction, in that process C_p and C_s cannot extract separately. Electric and magnetic energy is equal at f_{res}, so (6) can apply for $(C_s + C_p)$ calculation.

The quality factor (Q) of an inductor can the defined as

$$Q = 2\pi.\frac{peak\ magnetic\ energy - peak\ electric\ energy}{energy\ loss\ in\ one\ oscillation\ cycle} \tag{7}$$

Energy equations can be derived in terms of circuit elements in Fig. 1b:

Peak magnetic energy,

$$E_{peak\ magnetic} = \frac{V_0^2 L_s}{2.\left[\left(\omega L_s\right)^2 + R_s^2\right]} \tag{8}$$

Peak electric energy,

$$E_{peak\ electric} = \frac{V_0^2\left(C_s + C_p\right)}{2} \tag{9}$$

Energy loss in one oscillation cycle,

$$E_{energy\ loss} = \frac{2\pi}{\omega}.\frac{V_0^2}{2}.\left[\frac{1}{R_p} + \frac{R_s}{\left(\omega R_s\right)^2 + R_s^2}\right] \tag{10}$$

V_0 is peak voltage across inductor terminal. Using (8)-(10) into (7), Q can be derived as follows:

$$Q = \frac{\omega L_s}{R_s}.\frac{R_p}{R_p + \left[\left(\omega L_s/R_s\right)^2 + 1\right]R_s}$$
$$.\left[1 - \frac{R_s^2\left(C_p + C_s\right)}{L_s} - \omega^2 L_s\left(C_s + C_p\right)\right]$$

$$= \frac{\omega L_s}{R_s}.substrate\ loss\ factor.\ self\ resonance\ factor \tag{11}$$

Three governing factors shows in (11) drives Q of inductors. $\omega L_s/R_s$ increases with the increase in frequency, but the other two factors, substrate loss factor (represent energy dissipation in silicon substrate) and self-resonance factor (represent peak electric energy) begins to dominate after certain frequency and thus Q degrades.

B. Meniscus-confined 3-D Electro-deposition Technique

Although 3-D electro-deposition technique was descrived in details in [21], a brief discussion needs to place here. Metalwire growing follows meniscus-confined 3D electrodeposition technique which works in an ambient air environment. Dispensing nozzle holding by micropipette is the main part of the electrodeposition technique (Fig. 3). Diameter of the nozzle can be used from 100nm to several micrometer depending on thickness of metal wire. Electrodeposition initiated by the appropriate potential difference between micropippete and substarte. Direct writing by Cu, Pt and even other metals can be electro-chemically deposited with the help of electrolyte solutions. Although, by electrodeposition technique it is possible to grow metal as minimum as 100nm diameter, 20um dia metal coil for inductor is optimum size which increases mechanical stability as well as reduce series resistance.

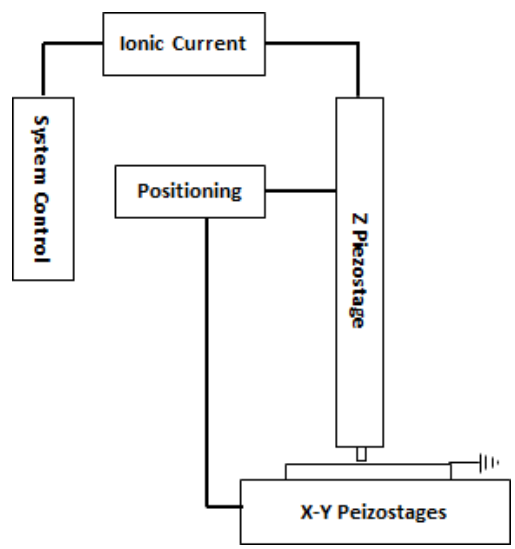


Figure 3: Schematic diagram of general setup for meniscus-confined 3D electrodeposition technique [21].

C. Vertical Solenoid Inductors

On-chip solenoid inductors has been reported in [6, 11, 12, 22]. In [6] copper windings are deposited around manually

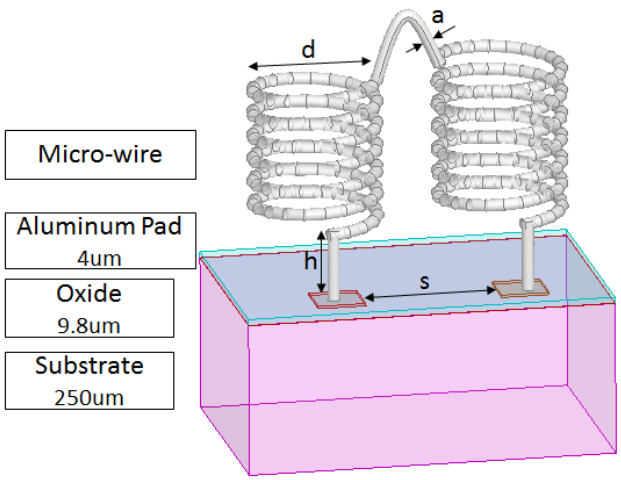


Figure 4: Schematic diagram of vertical solenoid inductor using Electro-deposition technique

placed alumina core using 3-D laser photography. In [11] and [12] 3-D solenoid inductors were fabricated by multiple layer of electroplated metals. Micromachining technique were applied to remove the substrate to form straight or curved solenoid configuration in [22]. On-chip vertical solenoid inductors utilizes all metal layers support by the technology. The number of turns in inductor is limited by the number of available metal layers.

In this paper, we proposed an alternative way to form vertical solenoid inductors with arbitrary number of turns and dimensions (Fig. 4). At first, we can make a structure supported by foundry technology consists of Si-substrate (thickness $250\mu m$), oxide layer (thickness $9.8\mu m$) and two metal pads (each dimension is $78\mu m$ x $78\mu m$ x $4\mu m$). Spacing between two pads depend on the diameter of the solenoid (d), in our case we keep this $220\mu m$. In the second stage, we have to apply 3-D electrodeposition techniques on that structute to fabricate vertical solenoid inductor. Copper can be used as electrolyte solution in micropipette (see Fig. 3) and the nozzle size was $20\mu m$ as we used conductor diameter (a) $20\mu m$. Micropipette can move horizontal plane and upward direction only, we have to grow microwire on each pad separately upto the top height level. Pad needs to keep in ground potential to make certain potential difference between nozzle and the pad to ensure ionic current flow through them. The solenoid coil starts growing about $100\mu m$ above the pad level (h) to maintain certain distance from substrate which minimizes substrate effects. The diameter of solenoid coil (d) is kept about $160\mu m$.

D. Impact of Solid Ground Shield

Fig. 5 shows substrate effect of vertical solenoid inductors, where magnetically induced eddy current and electrically induced conduction and displacement current is the main source of substrate loss. We can symbolize substrate loss as R_{Si}. From (1) it can be seen that R_p can be infinity if R_{Si} become either zero or infinity. A solid metal layer can be inserted inside oxide layer to provide a short to ground. Physically, the electric field will terminate before reaching the substrate. As small image current will be induced in the solid ground shield by the magnetic field of the vertical solenoid inductor, R_{Si} will be small value instead of zero, as well as R_p will be large instead of infinity according to (1).

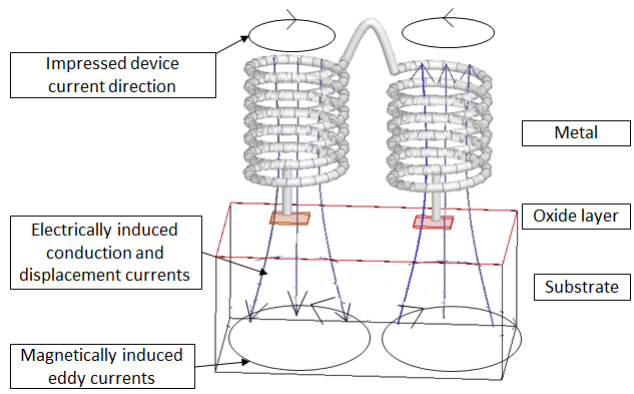
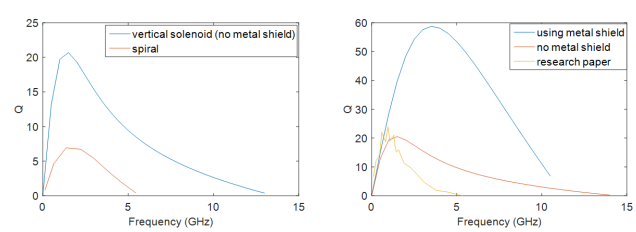


Figure 5: Substrate effect of vertical solenoid inductor

III. Results and Discussions

High Frequency Structure Simulator (HFSS) was used to simulate our proposed vertical solenoid inductor. Fig. 6a shows comparison of Q between vertical solenoid inductor (without metal shield) and spiral inductor (source IBM PDK). The extracted element set $\{L_s, R_s, C_{ps}, R_p, f_{res}, Q_{max}\}$ for our proposed model and spiral inductor are equal to $\{10nH, 2\Omega, 14fF, 3.5k, 13.5G, 20\}$ and $\{10nH, 4.88\Omega, 63fF, 2.4K, 6.3G, 6.5\}$ respectively. It shows C_{ps} is much less in our proposed inductor model, because the metal coil is far distant from substrate plane. In that way capacitive energy loss can be reduced by 60% compared to spiral inductor. By keeping metal diameter large ($20\mu m$) we are getting small serias resistance, R_s. Although metal coil of vertical solenoid inductor kept almost $100\mu m$ above the substrate, magnetic field still can penetrate inside the substrate.



(a) vertical solenoid and spiral

(b) vertical solenoid with metal shield, without metal shield and results from [18]

Figure 6: Comparison of Q of inductors

Fig. 6b shows comparison of Q among vertical solenoid inductor using metal shield, without using metal shield and 3D inductor by PDMA technique [18] of 10nH inductance value.

The extracted element set $\{L_s, R_s, C_{ps}, R_p, f_{res}, Q_{max}\}$ for vertical solenoid inductor with metal shield and without metal shield are $\{10nH, 2\Omega, 23.5fF, 31K, 11.3G, 58\}$ and $\{10nH, 2\Omega, 14fF, 3.5k, 13.5G, 27\}$ respectively. From the element set we see that R_p is much better by using metal shield. Increased R_p reduces substrate loss almost 90% compared to spiral inductor, which increases Q. From Fig. 6b it is also seen that frequency at Q_{max} is higher for vertical solenoid inductor in comparison with the other two types, which states it's performance at higher frequency.

IV. Conclusions

3D vertical solenoid inductor was presented. A nobel technique was proposed to reduce parasitic capacitance and substrate loss which is very important for THz operation. In our proposed method capacitive energy loss can be reduced almost 60% and substrate loss can be reduced by 90% in comparison to spiral inductors.

ACKNOWLEDGMENT

The authour would like to thank Army Research Office (ARO) to provide necessary financing to conduct this research.

References

[1] T. H. Lee, "The design of cmos radio-frequency integrated circuits," *Cambridge University Press*, 2002.

[2] F. Rabbi, Y. Arafat, and M. Z. R. Khan, "Analytical modelling of early voltage and current gain of sige heterojunction bipolar transistor," *Electronic Devices, Systems and Applications (ICEDSA)*, pp. 53–58, 2011.

[3] C. P. Yue, J. L. C. Ryu, T. H. Lee, and S. S. Wong, "A physical model for planar spiral inductors on silicon," *Int. Electron Devices Meet. Tech. Dig.*, pp. 155–158, December 1996.

[4] K. B. Ashby, I. A. Koullias, W. C. Finley, J. J. Bastek, and S. Moinian, "High q inductors for wireless applications in a complementary silicon bipolar process." *IEEE Journal of Solid-State Circuits*, pp. 4–9, 1996.

[5] C. P. Yue and S. S. Wong, "On-chip spiral inductors with patterned ground shields for si-based rf ic's," *IEEE Journal of Solid-State Circuits*, May 1998.

[6] D. J. Young, V. Malba, J. J. Ou, A. F. Bernhardt, and B. E. Boser, "Monolithic high-performance three-dimensional coil inductors for wireless communication applications," *IEEE Int. Electron Devices Meeting Tech. Dig.*, 1997.

[7] V. M. Lubecke, B. Barber, E. Chan, D. Lopez, M. E. Gross, and P. Gammel, "Seif-assembled mems variable and fixed rf inductors," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 11, pp. 2093–2098, November 2001.

[8] J. Y. Park and M. G. Allen, "Micromachined high q inductors for high frequency applications," *Proc. SPIE*, vol. 3514, pp. 218–228, September 1998.

[9] L. H. Lu, G. E. Ponchak, P. Bhattacharya, and L. P. B. Katehi, "High-q x-band and k-band micromachined spiral inductors for use in si-based integrated circuits," *IEEE Topical Silicon Monolithic Int. Cir. in RF Syst.*, 2000.

[10] H. Jiang, Y. Wang, J. L. A. Yeh, and N. C. Tien, "On-chip spiral inductors suspended over deep copper-lined cavities," *IEEE Trans. Microwave Theory Tech*, vol. 48, pp. 2415–2423, 2000.

[11] Y. J. Kim and M. G. Allen, "Surface micromachined solenoid inductors for high frequency applications," *IEEE Trans. Comp. Packag. Manufact. Technol.*, vol. 21, pp. 26–33, 1998.

[12] J. B. Yoon, B. K. Kim, C. H. Han, E. Yoon, and C. K. Kim, "Surface micromachined solenoid on-si and on-glass inductors for rf applications," *IEEE Electron Device Lett.*, vol. 20, pp. 487–489, 1999.

[13] B. K. Kim, B. K. Ko, and K. Lee, "Monolithic planar rf inductor and waveguide structures on silicon with performance comparable to those in gaas mmic," *IEEE Int. Electron Devices Meeting Tech. Dig.*, pp. 717–720, December 1995.

[14] J. W. M. Rogers, V. Levenets, C. A. Pawlowicz, N. G. Tarr, T. J. Smy, and C. Plett, "Post-processed cu inductors with application to a completely integrated 2-ghz vco," *IEEE Trans. Electron Devices*, vol. 48, pp. 1284–1287, June 2001.

[15] R. P. Ribas, J. Lescot, J. L. Leclercq, J. M. Karam, and F. Ndagijimana, "Micromachined microwave planar spiral inductors and transformers," *IEEE Trans. Microwave Theory Tech.*, vol. 48, pp. 1326–1335, August 2000.

[16] K. J. Chan, W. C. Hon, J. Zhang, and L. L. Leung, "Cmos-compatible micromachined edge-suspended spiral inductors with high q-factors and self-resonance frequencies," *IEEE Electron Device Lett*, vol. 25, no. 6, pp. 363–365, June 2004.

[17] K. Grenier, B. P. Barber, V. Lubecke, M. Zierdt, H. Safar, P. Pons, and P. L. Gammel, "Integrated rf mems for single chip radio," *Int. Solid-State Sensors and Actuators Conf.*, pp. 1528–1531, 2001.

[18] J. Zou, C. Liu, D. R. Trainor, J. Chen, J. E. Schutt-Ainé, and P. L. Chapman, "Development of three-dimensional inductors using plastic deformation magnetic assembly (pdma)," *IEEE Tran. on Microwave Theory and Tech.*, vol. 51, no. 4, pp. 1067–1075, April 2003.

[19] M. Aidoo, M. Harouna, A. Homaifar, N. D. Dogan, Z. Xie, H. Savci, and P. Roblin, "Multi-objective optimization of rotary travelling wave oscillator (rtwo) with neuro-genetic nondominated sorting algorithm," *Wireless Symp.*, pp. 1–4, 2013.

[20] M. Kirkman-Bey, R. Sebsebie, M. Aidoo, L. Wu, N. Dogan, and Z. Xie, "Modeling coupled lines in 65 nm cmos for millimeter-wave applications." *IEEE Southeastcon*, pp. 1–4, 2014.

[21] J. Hu and M.-F. Yu, "Meniscus-confined three-dimensional electrodeposition for direct writing of wire bonds," *American Assoc. for the Adv. of Science*, 2010.

[22] L. Gu and X. Li, "High-q solenoid inductors with a cmos-compatible concave-suspending mems process," *Journal of Microelectromechanical Sys.*, vol. 16, no. 5, pp. 1162–1172, October 2007.

A Single Phase Two Switch SEPIC Power Factor Corrected Rectifier

Kazi Khurshidi Haque Dia[1*], Mohammad Ali Choudhury[2]
[1]Assistant Professor, Dept. of EEE, Ahsanullah University of Science and Technology
141-142, Love road, Tejgaon, Industrial Area, Dhaka, Bangladesh
[2]Professor, Dept. of EEE, Bangladesh University of Engineering and Technology
Dhaka, Bangladesh
*kkh.dia@gmail.com

***Abstract*—A single phase two switch SEPIC output regulated rectifier is proposed in this paper. The proposed AC-DC converter has power factor correction and input current shaping facility. The rectifier produces buck-boost voltage/current gain. Taking advantage of this characteristic the design of the proposed two switch rectifier has been made separately to operate as buck converter and boost converter. A feedback power factor corrected controller is used for evaluating performance parameters of efficiency, input current total harmonic distortion and input power factor. The simulated results show, the use of separate design to operate the proposed converter in buck and boost region separately, give very high conversion efficiency, very high input power factor and very low input current total harmonic distortion in both buck and boost mode operation of the proposed single phase SEPIC power factor corrected rectifier.**

***Keywords*— Ĉuk; single-ended primary inductance converter (SEPIC); Total harmonic distortion (THD); Proportional Integral (PI) Controller; Power factor correction (PFC); Power factor (P.F).**

I. Introduction

Conventional controlled and uncontrolled AC-DC converters have non-sinusoidal input current. Improvement of their input current shapers and power factor at good power conversion efficiency is a desirable task. With the development of power electronics input power factor and total harmonic distortion of diode bridge rectifiers can be improved by passive and active methods. The common active method of power factor correction (PFC) and input current improvement of a single phase diode bridge rectifier is the boost PFC circuit. Other circuit topologies like buck, buck-boost, Ĉuk and SEPIC also have been used for the same purpose [1-8]. In an active PFC technique either the output DC or the input AC is switched at high frequency to make the input current switched sine wave in phase with the input sinusoidal voltage. Proper feedback control of PFC circuit with feedbacks from output and input makes the rectifier power factor corrected and output voltage regulated with small input current THD and maintaining good conversion efficiency [1, 2].

In this paper, a new two switch single phase SEPIC topology based output capacitor split type configuration is proposed in which input AC and output DC switches to provide path for the input current to flow and hence make it continuous. The rectifier connected to the topology through a combination of inductor and capacitor keeps the input current smooth and in-phase with the supply voltage. The SEPIC topology of switching has the advantages of input boost topology stage, boost-buck gain characteristics and single ended capacitively isolated positive output [1-3].

II. The Principle of Operation of the Proposed 1-Phase Two Switch Sepic Output Capacitor Split Rectifier

The proposed input-output switched (two switch) SEPIC AC-DC converter configuration is shown in Figure 1. It has four states as shown from Figure 2 to Figure 5 where, when switch S1 is ON, switch S2 is OFF and vice-versa, during both positive and negative cycle of the input. During positive half cycle, when S1 is ON and S2 is OFF, the input inductor L1 gets charged and the capacitor C discharges and energizes the inductor L2. When S1 is OFF and S2 ON, the inductor current i_{L1} and i_{L2} transfer energy to the load Ro. and Figures 2 and 3 represent the positive half cycle operation with switch ON and OFF positions, whereas, the operation under negative half cycle is shown in Figures 4 and 5 with switch ON and OFF positions respectively. The direction of current flowing through the load is same during each half cycle and thus the output voltage gets rectified by the bridgeless (capacitor-split) type configuration.

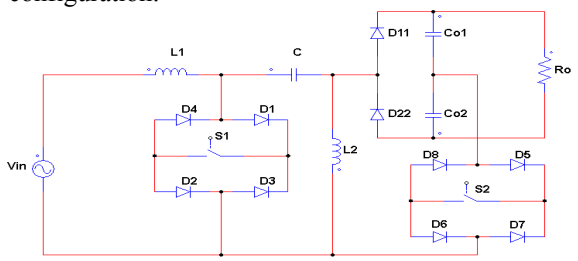


Figure 1: Input-Output Switched SEPIC Converter Bridge Configuration

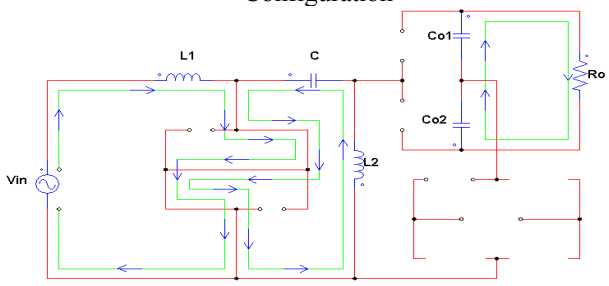


Figure 2: Equivalent circuit of the circuit of Figure 1 for positive cycle switch S1-ON, S2-OFF

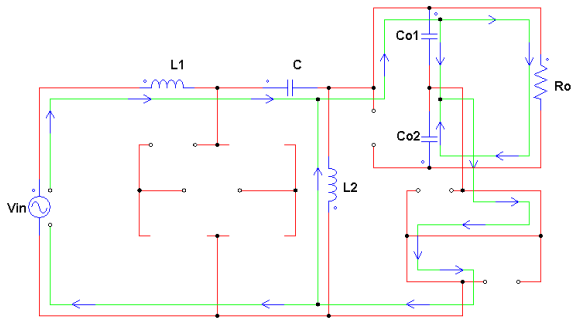


Figure 3: Equivalent circuit of the circuit of Figure 1 for positive cycle switch S1-OFF, S2-ON

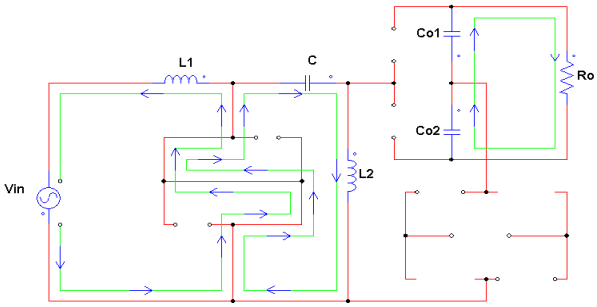


Figure 4: Equivalent circuit of the circuit of Figure 1 for negative cycle switch S1-ON, S2-OFF

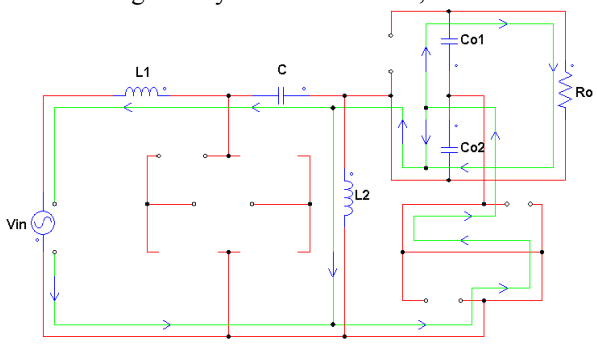


Figure 5: Equivalent circuit of the circuit of Figure 1 for negative cycle switch S1-OFF, S2-ON

III. The Conventional 1-Phase Output Swithed Sepic Rectifier

To have understanding of the performance of proposed new topology over the conventional output switched SEPIC rectifier, the circuit diagram of the conventional output switched SEPIC rectifier in shown in Figure 6, where input AC is converted to DC by bridge rectifier and then this DC voltage is fed to SEPIC DC-DC Converter. The single-ended primary inductance converter (SEPIC) can produce an output voltage that is either greater or less than the input but with no polarity reversal.

The relationship between input and output voltage of the DC-DC SEPIC converter is,

$$V_o = \frac{D}{1-D} V_{in} \quad (1)$$

$$I_o = \frac{1-D}{D} I_{in} \quad (2)$$

Where, Vo, I_o and V_{in}, I_{in} are output voltage, current and input voltage, current respectively and D is the duty cycle of the gate pulses of the switch.

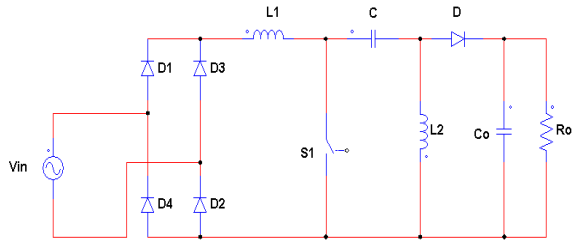


Figure 6: Conventional single phase SEPIC output switched AC-DC converter

III. PFC Feedback Controller

SEPIC topology based single phase AC-DC converters provide step-up/down voltage gain, output voltage control and regulation. In the open loop control the power factor and efficiency of the proposed circuits are reasonably good for a wide range of duty cycle variation. However, the input current total harmonic distortions are not in acceptable range for proposed and conventional circuit. It is therefore necessary that appropriate feedbacks are adopted in the circuit for performance parameters to be in the acceptable ranges. Feedback is also necessary for output voltage regulation around a constant value.

In controlling a PFC rectifier, the main objective is to draw a sinusoidal current, in-phase with the utility voltage. The reference inductor current must be of the full wave rectified form. The requirements on the form and the amplitude of the inductor current lead to two control loops: voltage control loop and inner current control loop [1].

The two control loops can be designed separately and can be realized by two simple operational amplifier working as a PI controller, each for two different loops. The inner current control loop must be faster than the outer voltage control loops.

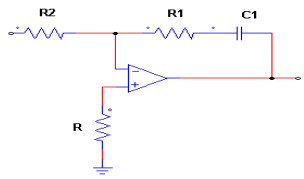


Figure 7: PI controller

The PI current controller parameters are:
$R_1 = 220k\Omega$ $R_2 = 100\Omega$ $R = 10k\Omega$ $C_1 = 0.022\mu F$
The PI voltage controller parameters are:
$R_1 = 10k\Omega$ $R_2 = 15k\Omega$ $R = 10k\Omega$ $C_1 = 0.1\mu F$

Switching frequency of the PWM generator is taken 10KHz. With properly designed PFC feedback controller, increase/decrease of the reference voltage (V_{REF}) of the feedback controller can increase/decrease the output voltage of the power stage respectively.

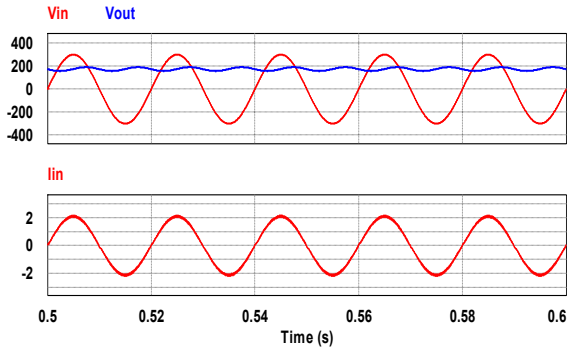
IV. SIMULATION AND RESULTS

High efficiency and power factor in AC-DC converter is not possible in both boost and buck operation modes, keeping the circuit parameters same. That is why the proposed topology is designed separately for buck and boost gain operations with different circuit parameters. The circuit parameters used to investigate and evaluate the performance of the proposed topology is given in Table 1. These parameters are also used in evaluating the performance of the conventional output switched SEPIC rectifier. All simulated results are obtained for proposed and conventional circuits using same PFC feedback controller.

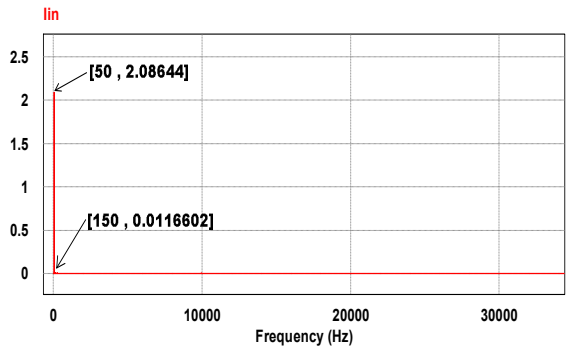
TABLE I: CIRCUIT PARAMETER SPECIFICATIONS OF PROPOSED TOPOLOGY

Parameter name	Parameter value	Used to operate as
L1	10 mH	Buck converter
L2	12 mH	
L1	6 mH	Boost converter
L2	5 mH	
Co1, Co2	330 µF	Both buck and boost converter
Ro	100 Ω	
C	1.5 µF	

Simulated results of the typical input-output voltages, input current and spectrum of input current waveform of proposed circuit in Figure 1operated in buck and boost regions are shown in Figures 8and 9 respectively.

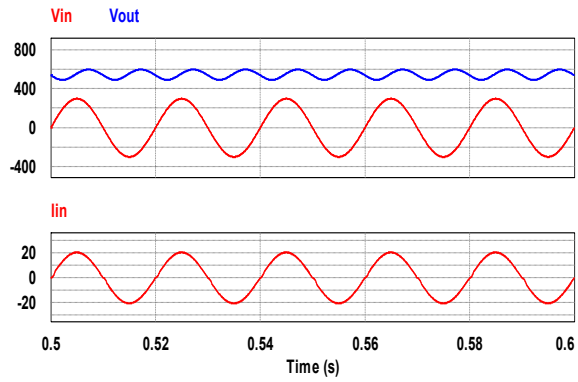


(a)

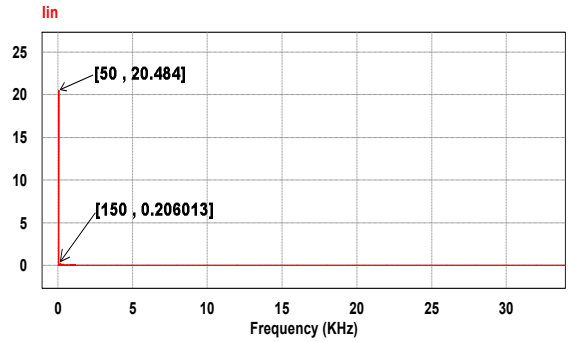


(b)

Figure 8: Input-output switched single phase SEPIC-Bridgeless configuration operating as buck converter (a) Input-output voltage (Upper picture) and input current (Lower picture), (b) Spectrum of input current in (a)



(a)



(b)

Figure 9: Input-output switched single phase SEPIC-Bridgeless configuration operating as boost converter (a) Input-output voltage (Upper picture) and input current (Lower picture), (b) Spectrum of input current in (a)

In the typical waveforms of Figure 8 and 9 of the proposed two switch single phase SEPIC rectifier, it is evident that the AC-DC converter give buck/boost voltage gain and switched input current follows the input voltage giving near unity power factor and the spectrum of the input currents indicate that the harmonic components of the fundamental frequency 50 Hz is very lower producing very low input current THD. The output voltages of Figure 7 and 8 are 180V DC and 550V DC respectively with an input voltage of 212.12 V R.M.S. value.

Dynamic response of the proposed circuit to regulate the output voltage and input current with load variation is shown in Figure 10 and 11 for buck and boost operation respectively. ± 20% variation of 100Ω load is considered for simulation that is from 100Ω to 120 Ω to 80 Ω to 100 Ω to 120 Ω to 80 Ω at .50, 1, 1.50, 2.0 and 2.50 seconds. The proposed circuit of Figure 1give well regulated output voltage of 180V DC (Figure 10-upper picture) and 550V DC (Figure 11-Upper picture).

Figures 12 to 14 depict the comparison of simulation results of the proposed circuit and conventional circuit in terms of efficiency, input current THD, input power factor with variation of reference voltage (V_{REF}) of PFC feedback controller [1] under buck operation when V_{REF} is from 0.15V (corresponds to 70V o/p DC) to 0.42V (Corresponds

to 193.2V o/p DC) and boost operation when V_{REF} is from 0.5V (corresponds to 245.3V o/p DC) to 1.5V(corresponds to 670V o/p DC).

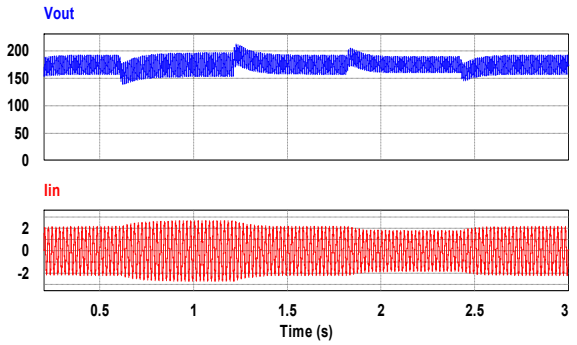


Figure 10: Dynamic response of output voltage (upper picture) and input current (lower picture) with load variation when the proposed circuit is under buck operation

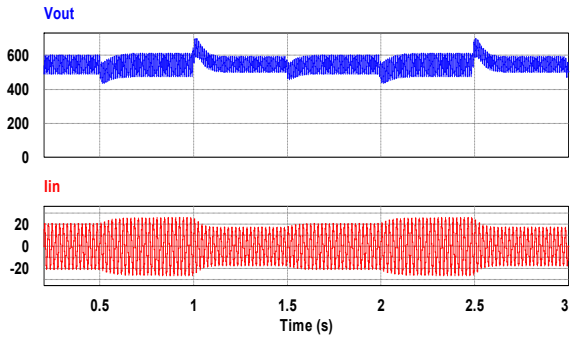


Figure 11: Dynamic response of output voltage (upper picture) and input current (lower picture) with load variation when the proposed circuit is under boost operation

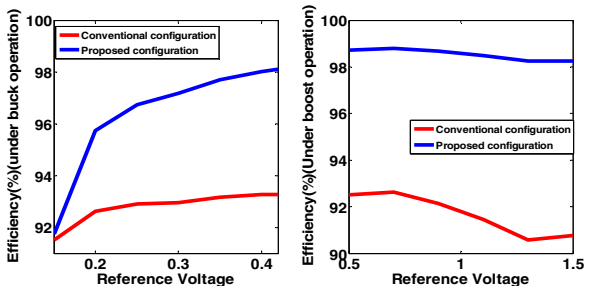


Figure 12: Comparison of efficiency between conventional and proposed circuits operating as buck converter (Left), boost converter (Right)

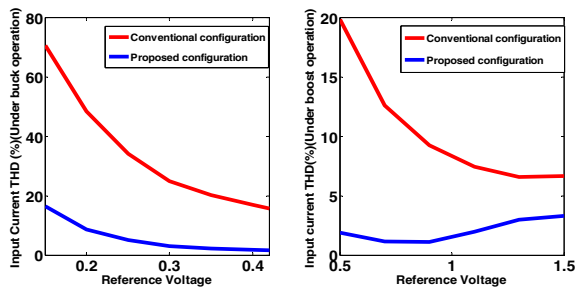


Figure 13: Comparison of input current THD between conventional and proposed circuit operating as buck converter (Left), boost converter (Right)

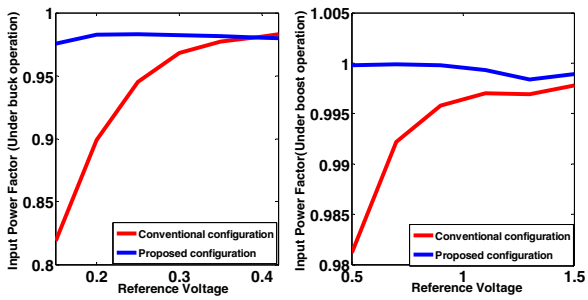


Figure 14: Comparison of input power factor between conventional and proposed circuit operating as buck converter (Left), boost converter (Right)

The efficiency (Figure 12) of the proposed circuit is above 90 percent under buck operation, above 98 percent under boost operation, which is greater efficiency than the conventional output switched SEPIC rectifier under both operation mode. The proposed circuit has very low input current THD (Figure13) as low as below 10% under buck operation and below 5% under boost operation whereas using the same feedback controller and same circuit parameter, the conventional circuit gives very high input current THD which is not in acceptable range. In terms of input power factor (Figure 14), the proposed circuit gives p.f. above 0.97 under buck operation and above 0.99 under boost operation which is better than the conventional circuit.

V. Conclusion

The proposed two switch (input-output switched) single phase SEPIC PFC rectifier gives buck-boost voltage gain capability, very high conversion efficiency above 90%, high power factor above 0.97 and very low input current THD below 10% under buck and boost region of operation than the conventional single phase output switched SEPIC rectifier.

References

[1] N. Mohan, Power Electronics. John Wiley and Sons, Inc, USA, 2012.

[2] D. W. Hart, Power Electronics. McGraw-Hill Companies, Inc, New York, 2011.

[3] O.Garcia, J. A.Cobos, "Single Phase Power Factor Correction: A Survey." IEEE Transactions on Power Electronics, VOL. 18, NO. 3, pp. 749-755, May 2003.

[4] A. De Bustiani Lange,T. B. Soeiro,M. Silveria Ortmann,M. Lobo Heldwein, " Three Level Single Phase Bridgeless PFC Rectifiers.", IEEE Transactions on Power Electronics, VOL. 30, issue-6, pp. 2935-2949, June 2015.

[5] L. Huber, Jang Yungtack, M. Jonanovic Milan, "Performance Evaluation of Bridgeless PFC Boost Rectifiers." IEEE Transactions on Power Electronics, VOL. 23, NO. 3, May 2008.

[6] I. Siyana,V. A. Manjusha, "Single Phase Single Stage High Power Factor Ĉuk Rectifier.", International Journal of Advanced Research in Electrical Electronics and Instrumentation Engineering, VOL. 3, issue-9, pp. 2320-3765, September, 2014.

[7] F. Musavi, "A High-Performance Single Phase Bridgeless Interleaved PFC Converter for Plug-in Hybrid Electric Vehicle Battery Chargers", IEEE Transactions on Industrial Applications, VOL. 47, NO. 4, pp. 1833-1843, July/August. 2011.

[8] E. H. Ismail, "Bridgeless SEPIC Rectifier With Unity Power Factor and Reduced Conduction Losses", IEEE Transactions on Industrial Electronics, VOL. 56, NO. 4, pp. 1147-1157, April. 2009.

Sensitivity of a nanodoped polymer based X-ray image detector

Naimul Hassan[1,*] and Shaikh Asif Mahmood[1]

[1]Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology
Dhaka-1000, Bangladesh
[*] naim074@gmail.com

Abstract—**The X-ray sensitivity of direct conversion X-ray image detectors based on poly(triarylamine) (PTAA) doped with bismuth oxide (Bi_2O_3) nanoparticle is determined for mammography. Using COMSOL Multiphysics 5.0, the continuity equations are solved considering drift of charge carriers in the presence of deep traps. The recombination between drifting charge carriers and nonuniform electric field across the photoconductor layer is also considered. Poole–Frenkel mechanism has been considered for electric field dependent mobility of charge carriers. A very promising X-ray sensitivity of 1.3295μC/cm^{2}R has been found at an applied electric field of 10V/μm. The calculated sensitivity has been compared with that of the commercially available amorphous selenium (a-Se) based direct conversion X-ray image detector.**

Index Terms—**X-ray, nanoparticle, recombination, sensitivity, PTAA.**

I. Introduction

The invention of X-ray in 1895 was a significant scientific advancement that benefitted various fields (e.g., crystallography, medical imaging). X-ray imaging has been regarded as a very important tool for patient diagnosis. In the early years, it was largely a film-based analog technology. In the late 20th century, flat panel X-ray image detectors have been introduced to replace the conventional analog technology with digital system [1]. In these digital detectors, image can be detected either by direct approach or by indirect approach. In the indirect approach, a phosphor screen converts X-ray photons into visible light which is then converted to electrical signal by photodiodes. In the direct detection approach, an active photoconductor layer converts X-ray photons into proportionate amount of charge carriers in a single step. Research has shown that the direct conversion flat panel X-ray image detectors provide excellent images. These detectors are based on inorganic materials such as amorphous selenium (a-Se), cadmium zinc telluride (CdZnTe), and mercuric iodide (HgI_2). They are commercially available and provide high energy resolution and high detection efficiency at room-temperature [2]. However, from the perspectives of large scale fabrication, cost efficiency, and device flexibility, these inorganic materials based detectors are not the best choice.

A very few research works have been done using organic polymer based direct conversion X-ray image detector. Intaniwet *et al* have reported on properties of poly(triarylamine) (PTAA) used in thick film diode structures which are shown to detect X-ray radiation [3]. To be used as an X-ray photoconductor, a material should have high chemical and electronic stability, high charge carrier mobility, low ionization energy to create electron hole pairs, and low dark current at operating electric field. The PTAA requires a comparatively low ionization energy (~6.606eV) and its relatively high hole mobility (10^{-3}-10^{-2}cm^2/Vs) and long charge carrier lifetime (1.7ms) are very promising comparing with the other polymers [3]-[6]. It has also been mentioned that PTAA based X-ray image detectors show a very good stability with no observable performance degradation [3]. However, PTAA has a very low X-ray attenuation coefficient and hence high thickness of the active layer is required to obtain an X-ray sensitivity comparable with the inorganic material based X-ray image detectors [7]. Recently Intaniwet *et al* have demonstrated a way to increase X-ray attenuation efficiency of PTAA by introducing high density bismuth oxide (Bi_2O_3) nanoparticles [8]. In this paper, the X-ray sensitivity of PTAA/Bi_2O_3 composite based direct conversion detector has been determined.

II. Theoretical Model

The direct conversion X-ray image detector structure is built by sandwiching an active layer of photoconductor material between two parallel plate electrodes. One of the electrodes is the charge collecting electrode and the other one is exposed to X-ray radiation. A very high voltage is applied between these electrodes. Both of the electrodes are ideally non-injecting to photoconductor material [5]. However, both of the metal-photoconductor junctions pass X-ray generated charge carriers from the photoconductor to the outside circuitry. Hence, current only flows due to X-ray generated charge carriers.

The interaction of the X-ray photons with the photoconductor creates an exponentially decaying charge carrier profile. Then these charge carriers start to move due to the very high applied electric field across the photoconductor layer and result induced current in the external circuit. This current is then integrated by the external integrator circuit to provide signal for the digital interfacing. In this paper, we consider only the drift of charge carriers and ignored the diffusion process due to high applied electric field.

The term X-ray sensitivity is defined as the collected charge per unit area of the charge-collecting surface per unit X-ray exposure [4]. In case of disordered semiconductor (e.g., PTAA) the recombination and trapping have adverse effect on the charge collection efficiency. Hence, these effects are considered here. We have considered a constant X-ray exposure for a duration greater than charge carrier transit time through the photoconductor layer. If the X-ray radiation receiving electrode is positively biased with respect to the charge collecting electrode, the X-ray generated holes will move away from the exposed electrode due to drift mechanism. Taking the radiation receiving electrode at $x = 0$, the continuity equation for holes is

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

$$\frac{\partial p}{\partial t} = -\mu_p \frac{\partial}{\partial x}(pF) + g_p - \frac{p}{\tau_p} - C_r pn - C_p pn_t \quad (1)$$

where μ_p is the hole mobility, τ_p is the hole lifetime, p is the free hole concentration, p_t is the trapped hole concentration, F is the electric field, C_r is the bimolecular recombination coefficient between free electrons and holes, C_p is the capture coefficient between free holes and trapped electrons, and t is the time. It has been stated that bimolecular recombination and recombination through traps in polymers can be explained by Langevin recombination mechanism [9], [10]. Therefore, $C_r = e(\mu_p+\mu_n)/\varepsilon$ and $C_p = e\mu_p/\varepsilon$, where μ_n is the electron mobility, e is the elementary charge and ε is the permittivity of the photoconductor. The hole generation rate is

$$g_p = \frac{\alpha E \phi_0(E)}{T_{ex} W} \exp(-\alpha x) \quad (2)$$

where T_{ex} is the X-ray exposure time, W is the ionization energy, and α is the X-ray attenuation coefficient at the X-ray photon energy, E [11]. The incident photon fluence in photons/cm^2 is

$$\phi_0(E) = \frac{5.45\times10^{13} X}{\left(\alpha_{air}(E)/\rho_{air}\right)E} \quad (3)$$

where X is the X-ray exposure in roentgen, $\alpha_{air}(E)$ is the energy dependent X-ray attenuation coefficient (in per unit cm) of air, ρ_{air} is the density of air in g/cm^3 and E is in eV [12].

Initially, the hole concentration is zero inside the photoconductor. The boundary condition for hole is $p(0, t) = 0$. The hole current density, $J_p(t)$ is calculated as

$$J_p(t) = \frac{e}{L}\int_0^L \mu_p F p dx \quad (4)$$

where L is the thickness of the photoconductor layer [11]. In case of positively biased X-ray detectors, electrons will travel from the charge collecting electrode to the X-ray receiving electrode. Hence, the continuity equation for electrons is

$$\frac{\partial n}{\partial t} = \mu_n \frac{\partial}{\partial x}(nF) + g_n - \frac{n}{\tau_n} - C_r pn - C_n p_t n \quad (5)$$

where C_n is the capture coefficient between free electrons and trapped holes, and $C_n = e\mu_n/\varepsilon$. All the terms used in equation (5) bear the same significance for electrons as described before for holes. Using similar approach used for holes, we determine electron current density.

The rate equations for deep trapping are [4],

$$\frac{\partial n_t}{\partial t} = -C_p n_t p + \frac{n}{\tau_n} \quad (6)$$

and

$$\frac{\partial p_t}{\partial t} = -C_n p_t n + \frac{p}{\tau_p} \quad (7)$$

The electric field inside the photoconductor deviates from initially applied electric field due to the trapped and free charge carriers. The Poisson equation is

$$-\frac{\partial^2 V}{\partial x^2} = \frac{e}{\varepsilon}(p + p_t - n - n_t) \quad (8)$$

where V is the voltage. The boundary condition is $V(0,t) = V_0$, $V(L,t) = 0$ where V_0 is the initially applied voltage between the two electrodes. Hence, the initially applied electric field is $F_0 = V_0/L$.

It has been reported that charge carrier conduction through PTAA based diode occurs via a Poole–Frenkel mechanism [3]. According to this mechanism mobility is

$$\mu(T,F) = \mu_0 \exp\left(\frac{(\beta F)^{1/2}}{k_B T}\right) \quad (9)$$

where μ_0 is the mobility at zero electric field, k_B is the boltzmann constant, T is the absolute temperature and Poole-Frenkel coefficient $\beta = e^3/\pi\varepsilon$ [13].

Neglecting the post X-ray exposure current, collected charge per unit area due to hole movement is determined by

$$q_p = \int_0^{T_{ex}} J_p(t) dt \quad (10)$$

Similarly, the collected charge per unit area due to electron movement, q_n can be determined. Then, the sensitivity, S of the X-ray detector is,

$$S = (q_p + q_n)/X \quad (11)$$

III. Results And Discussion

The X-ray photon energy used for mammography is around 20keV with a mean X-ray exposure of 12mR. At this energy, the mass attenuation coefficient of PTAA doped with 56 weight% of Bi_2O_3 has been found to be 45.21cm^2/gm [14]. The volume percentage of Bi_2O_3 in PTAA/Bi_2O_3 composite is 12.5 [8]. X-ray attenuation coefficient of this composite can be derived by multiplying its mass attenuation coefficient with its density. The equivalent density of the composite can be determined by dividing the mass to the volume of the composite.

The density of PTAA/Bi_2O_3 composite with 56 weight% of Bi_2O_3 has been calculated to be 1.99gm/cm^3. Therefore, the X-ray attenuation coefficient of the composite is α = 8996.8/m [14]. The electromechanical property of the doped polymer mostly originates from the organic PTAA, as inorganic Bi_2O_3 nanoparticles occupy a very small volume compared to the total volume of the composite [8]. However, in addition, the advantage of high attenuation coefficient is gained.

The continuity equations for holes and electrons are solved using COMSOL Multiphysics 5.0. The thickness of the photoconductor layer is taken to be L = 100μm. The charge carrier lifetime for both the hole and electron is taken to be 1.7ms [5]. The hole and electron mobilities at zero electric field are μ_p=0.01cm^2/Vs and μ_n=1.3×10^{-6}cm^2/Vs, respectively [3], [15]. Initially, the applied electric field is F_0= 10V/μm for which charge carrier schubweg (i.e. the length a charge carrier travels in duration of its lifetime) is greater than the photoconductor thickness. Hence, the initial hole and electron transit times are T_{pi}= 0.1416μs and T_{ni}= 1.1ms, respectively. The X-ray exposure time is taken to be T_{ex}= 10ms.We have considered constant temperature of operation at T=300K.

In Fig. 1, the hole concentration distribution along the photoconductor thickness is shown at different times up to the initial hole transit time. The peak hole concentration increases with time and moves towards the charge collecting electrode. When initial hole transit time is reached, the peak hole concentration occurs nearly at the charge collecting electrode. Then, the hole concentration along the photoconductor thickness remains nearly a steady state distribution till the X-ray exposure is on.

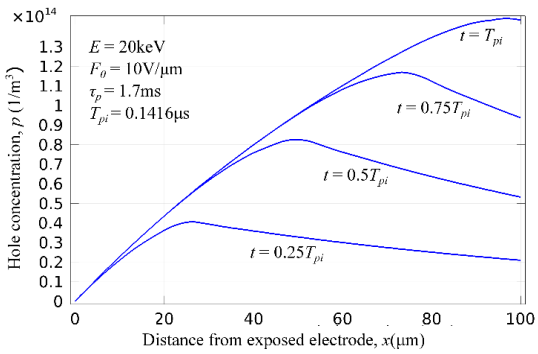


Fig. 1 Hole concentration vs distance from X-ray radiation receiving electrode at different times at 20keV X-ray photon energy for PTAA/Bi_2O_3 composite.

From the hole concentration distribution along the photoconductor layer thickness, mean hole current densities at different times are calculated. Fig. 2 shows the hole current density as a function of time. From this figure, it is evident that the hole current increases from zero to its peak value abruptly as hole transit time is very small. Then, it changes slowly during the rest of the exposure due to the variation of electric field distribution. After exposure time, this current rapidly goes to zero.

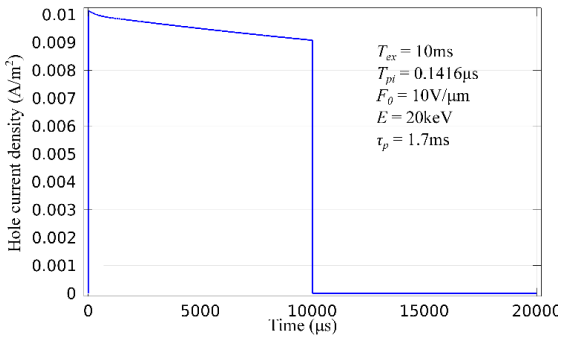


Fig. 2 Hole current density vs time plot for PTAA/Bi_2O_3 composite at 20KeV.

In Fig. 3, the electron concentration distribution along the photoconductor thickness is shown at different times up to the initial electron transit time. The peak electron concentration is always at the radiation receiving electrode as this electrode is positively biased. From Fig. 1 and 3, it is evident that the electron concentration is higher than the hole concentration due to higher electron transit time.

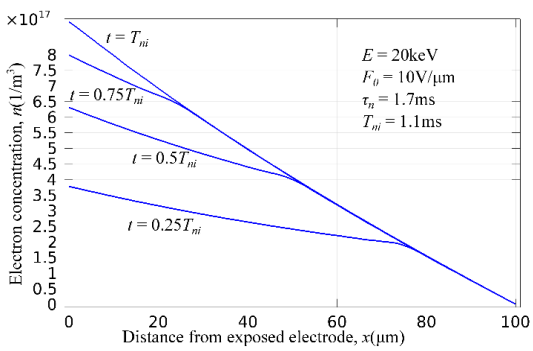


Fig. 3 Electron concentration vs distance from X-ray receiving electrode plot at different times at 20keV X-ray photon energy for PTAA/Bi_2O_3 composite.

Fig. 4 shows the electron current density as a function of time. Electron current increases very quickly up to electron transit time, then it reaches a nearly steady value. After the electron transit time, it changes due to the electric field variation. From Fig. 2 and 4, it is evident that electron current density is small comparing to hole current density. This is due to very low electron mobility of PTAA comparing to the hole mobility.

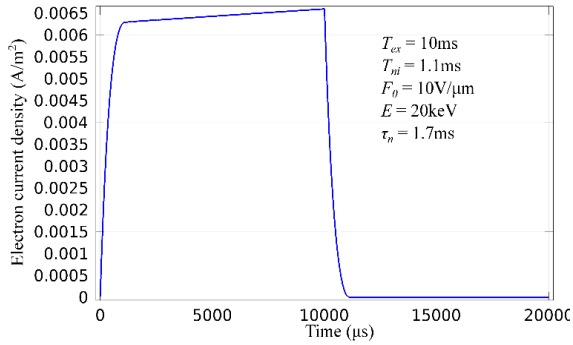


Fig. 4 Electron current density vs time plot for PTAA at 20keV for PTAA/Bi_2O_3 composite.

The electric field distribution across the photoconductor at different times up to initial electron transit time (T_{ni}) is shown in Fig. 5. The electric field across the photoconductor deviates from the initially applied electric field due to the trapping of charge carriers. The electron transit time is greater than the hole transit time. So, after electron transit time both free electron and hole distribution across the photoconductor changes very little. We observe that electric field increases 1.555% of initially applied electric field at the exposed electrode and decreases 0.77% of the initially applied electric field at the charge collecting electrode at the initial electron transit time.

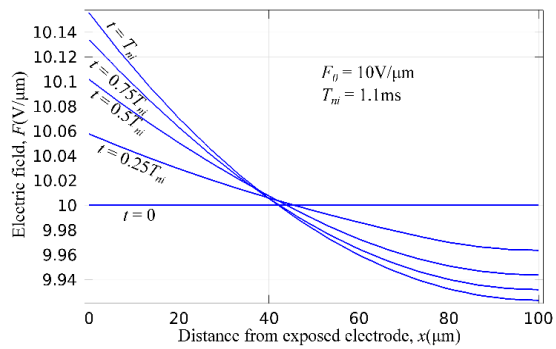


Fig. 5 Electric field vs distance from X-ray receiving electrode plot at different times at 20keV X-ray photon energy for PTAA/Bi_2O_3 composite.

According to the Poole–Frenkel mechanism charge carrier mobilities are dependent on electric field. The hole and electron mobility distributions across the photoconductor are shown in Fig. 6 and 7, respectively. Initially, hole and electron mobilities are 0.706cm^2/Vs and 9.18×10^{-5}cm^2/Vs, respectively. At the electron transit time, hole and electron mobilities increase to 0.73 cm^2/Vs and 9.49×10^{-5}cm^2/Vs, respectively at the exposed electrode. However, these mobilities decrease to 0.6945 cm^2/Vs and 9.029×10^{-5}cm^2/Vs, respectively at the charge collecting electrode.

Integrating both hole and electron current densities up to exposure time, collected charge per unit area is determined. Dividing the total collected charge per unit area by mean X-ray exposure, we get the X-ray sensitivity of PTAA/Bi_2O_3 composite based detectors to be 1.3123μC/cm^{2}R at 20keV. The sensitivities due to hole and electron current are 0.7928μC/cm^{2}R and 0.5195μC/cm^{2}R, respectively. The

contribution of electron current to total sensitivity is lower than that of hole current due to lower electron mobility. From Fig. 2 and 3, it is evident that both the hole and electron currents don't go to zero just at the end of exposure. Integrating the current after the exposure time (T_{ex}), sensitivity increases to 1.3295μC/cm^{2}R.

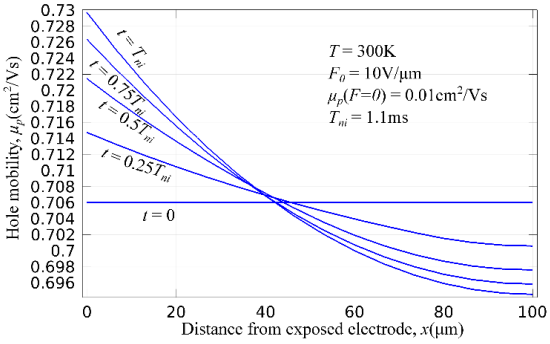


Fig. 6 Hole mobility vs distance from X-ray receiving electrode plot at different times at 20keV X-ray photon energy for PTAA/Bi_2O_3 composite.

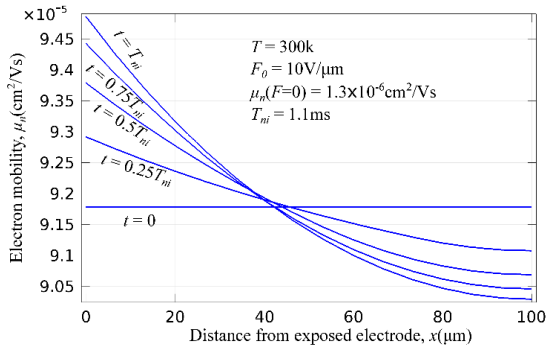


Fig. 7 Electron mobility vs distance from X-ray receiving electrode plot at different times at 20keV X-ray photon energy for PTAA/Bi_2O_3 composite.

We have calculated the sensitivity of the commercially available a-Se based X-ray image detectors for mammography considering only drift of charge carriers in the presence of deep traps [7]. Taking μ_p= 0.12cm^2/Vs, μ_n= 0.003cm^2/Vs, F = 10V/μm, τ_p = 50μs, τ_n= 500μs, α = 21151/m, L = 200μm, and W = 60eV, we get the sensitivity of 0.2643μC/cm^{2}R for mammography. Therefore, for mammography, X-ray image detectors based on PTAA/Bi_2O_3 composite provides more sensitivity than the a-Se based detectors while the required thickness of the active layer is half of the later material. It has also been found that PTAA/Bi_2O_3 composite based diode shows a very low leakage current density (0.25–0.9 nAcm^{-2}) at 5V/μm applied electric field [8]. Hence, dark current density at operating electric field can be assumed to be negligible compared with the photocurrent density. Therefore, it can be a better alternative for a-Se as a photoconductor in direct conversion X-ray image detector at low range X-ray application.

IV. Conclusions

The sensitivity of direct conversion X-ray image detector based on PTAA doped with 56 weight% of Bi_2O_3 nanoparticle has been determined to be 1.3295μC/cm^{2}R for mammography at an applied electric field of 10V/μm. Due to the presence of trapped carriers, electric field increases at the exposed electrode and decreases at the charge collecting electrode. The carrier mobilities also increase at the exposed electrode and decrease at the opposite electrode arising from the Poole-Frenkel effect. Peak electron concentration has been found to be greater than the peak hole concentration. However, hole current density is greater than electron current density due to higher hole mobility. For commercially available a-Se based detector, we found the sensitivity to be 0.2643μC/cm^{2}R at an applied electric field of 10V/μm for mammography. Therefore, from sensitivity perspective, nanodoped polymer composite PTAA/Bi_2O_3 is a very good alternative of a-Se to be used as X-ray photoconductor material for mammography.

References

[1] J. A. Rowlands and J. Yorkston, *Handbook of Medical Imaging*, J. Beutel, H. L. Kundel, and R. L. V. Metter, Ed. Washington: SPIE Press, 2000, vol. 1, chap. 4.

[2] H. Han, S. Lee, J. Seo, C. Mahata, S. H. Cho, A. Han, K. Hong, J. Park, M. Soh, C. Park, and T. Lee, "Enhancement of X-ray detection by single-walled carbon nanotube enriched flexible polymer composite," *Nanoscale Research Letters*, 9, 610, 2014.

[3] A. Intaniwet, C. A. Mills, M. Shkunov, H. Thiem, J. L. Keddie and P. J. Sellin, "Characterization of thick film poly(triarylamine) semiconductor diodes for direct x-ray detection," *J. Appl. Phys.*, vol. 106, no. 064513, 2009.

[4] M. Z. Kabir, "Modeling of x-ray photoconductors for x-ray image detectors", 2005, University of Saskatchewan.

[5] A. Intaniwet, J. L. Keddie, M. Shkunov, and P. J. Sellin, "High charge-carrier mobilities in blends of poly(triarylamine) and TIPS-pentacene leading to better performing X-ray sensors," *Organic Electronics* 12, 1903-1908, 2011.

[6] C. R. Newman, H. Sirringhaus, J. C. Blakesley, and R. Speller, *Appl. Phys. Lett.*91, 142105, 2007.

[7] N. Hassan and S. A. Mahmood, "X-ray sensitivity of poly(triarylamine) (PTAA)," in *ICTP*'15, 2015.

[8] A. Intaniwet, C. A. Mills, M. Shkunov, P. J. Sellin, and J. L. Keddie, "Heavy metallic oxide nanoparticles for enhanced sensitivity in semiconducting polymer x-ray detectors," *Nanotechnology* 23, 235502, 2012.

[9] P. W. M. Blom and M. J. M. de Jong, "Electrical characterization of polymer light emitting diodes," *IEEE J. Select. Topics Quantum Electron.*, vol. 4, pp. 105-112, 1998.

[10] Kuik, M., L. J. A. Koster, G. A. H. Wetzelaer, and P. W. M. Blom, "Trap-assisted recombination in disordered organic semiconductors." *Physical review letters* 107, no. 25 (2011): 256805.

[11] S. Siddiquee and M. Z. Kabir, "Modeling of photocurrent and lag signals in amorphous selenium x-ray detectors," *J. Vac. Sci. Tech. A*, 33, 041514, 2015.

[12] J. M. Boone, *Handbook of Medical Imaging*, J. Beutel, H. L. Kundel, and R. L. V. Metter, Ed. Washington: SPIE Press, 2000, vol. 1, chap. 1.

[13] Moliton, André, *Optoelectronics of molecules and polymers*. Vol. 104. Springer, 2010.

[14] NIST XCOM: Element/Compound/Mixture - Physics Laboratory. [Online]. Available: http://physics.nist.gov/PhysRefData/Xcom/html/xcom1.html

[15] S. Schols, S. Verlaak, C. Rolin, D. Cheyns, J. Genoe, and P. Heremans, "An Organic Light - Emitting Diode with Field - Effect Electron Transport," *Advanced functional materials* 18, no. 1: 136-144,2008.

Analysis of a Silica Cladded Silicon Nanowire Having a Deep Sub-Wavelength Central Air Hole

Nazifa Rumman[1], Maisha Mesbah[1], Tasmina Mahmud[1], and M. Shah Alam[2]

[1]Dept. of Electrical Electronic and Communication Engineering
Military Institute of Science and Technology, Mirpur Cantonment, Dhaka-1216

[2]Dept. of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology (BUET)
Dhaka 1000, Bangladesh
[2]Email: shalam@eee.buet.ac.bd

***Abstract*—This paper presents different properties of silicon nanowire when a hole of sub-wavelength diameter is introduced in the central region. The analysis has been carried out using edge element based vector finite element method (FEM). It has been found that the light enhancement and the average mode intensity are higher within smaller holes. The nonlinear coefficient and the group velocity dispersion are also calculated which appear to be very promising for varieties of applications.**

Keywords—Silicon nanowire; sub-wavelength air hole; light enhancement; nonlinearity; zero dispersion.

I. Introduction

In recent years an explosion of interest has been witnessed in research of optical devices with sub-wavelength dimension because of their unique material and structurally dependent properties. A great deal of research has focused on silicon nanowire waveguide considering facts like its low-cost and mature fabrication technology and also for its low optical losses at the optical telecommunication wavelengths. Also silicon wires with sub-micrometer diameter or nanometer diameter have been successfully fabricated and its electrical and optical properties are well studied [1]-[5]. Besides, a tight confinement allowing efficient scaling and close packing of photonic components is possible owing to its high index contrast [6]. In waveguide applications, the fabrication of quantum wire and the initiation of number of studies in the silicon photonics field has been allowed with the small size and high index contrast of silicon core with silica cladding [7]-[9].

With the introduction of an air hole with sub-wavelength diameter within the core region, a remarkable impact on the characteristics of the nanowire can be obtained [10]. Concentration of optical energy within a sub-wavelength air hole running along the length of an optical fiber has been demonstrated in [8]. Enhancement of light within nanoscale air holes in the core region of F2 lead silicate glass and As_2S_3 nanowires has also been presented in [10]. High index silicon-on-insulator (SOI) waveguide structures [11] were proposed based on the principle that light intensity on the low index side of the interface of the air hole scales with the square of refractive index contrast at the interface [12]. A numerical study of the impact on fiber group velocity dispersion owing to tactful arrangement of air holes has been studied in [13]. More recently, the effects of incorporating nanoholes inside the core of an As_2Se_3 on its mechanical and optical properties have been investigated in [14]. However, silicon nanowire with a deep sub-wavelength air hole has not been characterized in details.

In this work, a cylindrical nanowire composed of a silicon core and a silica cladding having an air hole in the core region has been numerically characterized by using the FEM. To explore the impact of the small air hole on silicon nanowire, the confinement factor, nonlinear coefficient of the structure and the dispersion features have been analyzed. The dependence of light enhancement and average mode intensity onhole diameter has also been studied.

II. Results and Analysis

The nanowire structure consists of silicon ($n_{silicon}$=3.45) core and silica (n_{silica}=1.44) cladding and a nanoscale air hole in the center of the nanowire. A vector FEM is applied to obtain mode field distributions over the cross sections of the nanowires using the COMSOL Multiphysics. The light wavelength is fixed to be 1.55 μm unless stated otherwise.The outer radius of the nanowire is *R*, and the nanoscale air hole is of radius *r*. During modal analysis, a cylindrical PML has been set as the boundary of the silica cladding layer.

Before analyzing the nanowire with the central air hole we provide the power distribution, nonlinearity, and dispersion of the silicon nanowire without air hole in Figs. 1, 2, 3 and 4, respectively. It can be seen in Fig. 1 that the power is confined at the center of the nanowire. In this case the nanowire diameter is very small (200 nm) and the power confinement area looks elliptical though it is circular at higher values of the diameter. The nonlinearity versus core radius for fundamental mode and higher order modes are shown in Fig. 2.

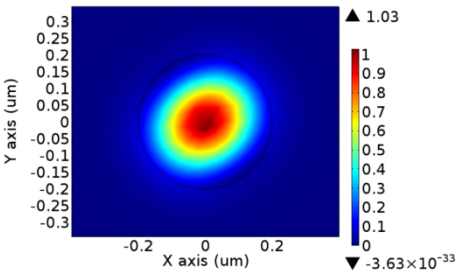


Fig.1. Power distribution of H^x_{01} mode over the cross section.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

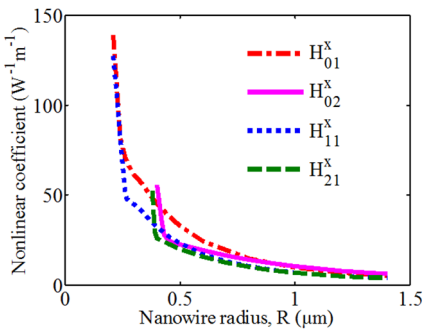


Fig. 2. The nonlinear coefficient versus the nanowire radius.

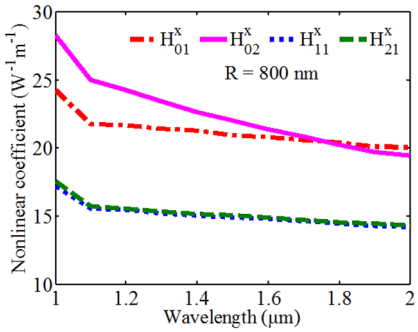


Fig. 3. Variation of the nonlinearity with the wavelength.

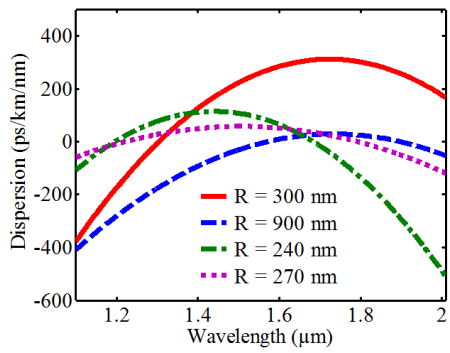


Fig. 4. The dispersion versus the wavelength for different core radius.

The nonlinearity coefficient can be defined as

$$\gamma = 2\pi n_2 / \lambda A_{eff} \quad , \qquad (1)$$

where $n_2 = 4\times10^{-18}$ m^2/W is the Kerr nonlinear coefficient of the material and A is the cross-sectional area of the nanowire. The structure shows higher nonlinearity when effective area becomes smaller. It is observed from Fig. 3 that the nonlinearity decreases as the wavelength increases because of the reduction of light confinement in the core. The lower the value of light confinement in the core region the larger will be the value of the effective area and this results in decrease of nonlinearity.

Fig. 4 shows the group velocity dispersion (GVD) properties of the nanowire. The GVD values are achieved from the spectral dependence of the effective indices (n_{eff}) as defined by

$$D = -\frac{\lambda}{c}\frac{d^2 n_{eff}}{d\lambda^2} \quad , \qquad (2)$$

where c is the velocity of light and λ is the operating wavelength. For smaller values of wavelength field remains confined in the core, but with larger values of wavelength, the modal field spreads into the silica cladding significantly and thus the effective index reduces at a faster ratefor a nanowire with smaller diameter [6]. Consequently, a zero dispersion can be seen around λ=1.43 μm for R=300 nm. The accretion of power in cladding is likely to be large when core radius is smaller, in order that, maximum value of dispersion is achieved at a particular higher wavelength. Next, we consider the nanowire with the hole at the central region and examine various properties from our numerical results.

A. Effective index and the mode field distribution

The cross section of the structure and the variation of index over the radial axis of the cross section are shown in Fig. 5. The core effective index (β/k_0) will be reduced because of the presence of the air hole at the center and this is shown in Fig. 6, which shows the effective index versus the nanowire radius, R.

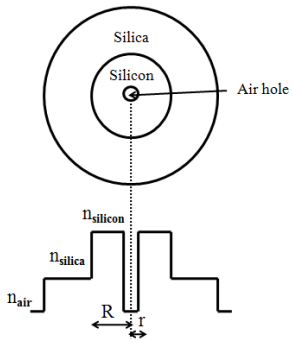


Fig. 5. Cross section of nanowire and the variation of refractive index.

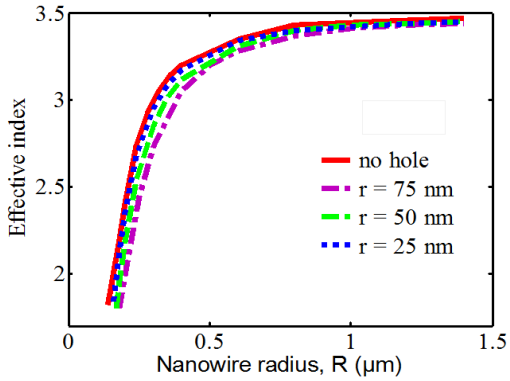


Fig. 6. The effective index of silicon nanowire with subwavelength hole.

Fig. 7 shows the electric field and power distribution of the fundamental mode over the cross section of the structure. It can be seen that the electric field in the hole region (E_h) is higher than that in the silicon (E_s) on the low index side of the interface. Because of the presence of air hole, E_h slightly decays within the hole but maintains a higher intensity in the hole region. It has been found in Fig. 7(a) that the intensity inside the hole is much stronger than that of the silicon core due to high index at the boundary. The transverse electric field intensity $|E_t|^2$ and the total electric field intensity$|E_{xyz}|^2$ are

also shown in Figs. 7(b) and (c),respectively, where the transverseelectric field intensity is circularly symmetric but thetotal electric field is not because $|E_z|$ is not negligible when the nanowire size is undersized in scale. The z-component of the time-averaged power flow is shown in Fig. 7(d). The similar type of phenomenon is also discussed in [10].

B. The mode effective area and the confinement factor

The structure allows the strongest confinement of light with the highest intensity in the center of core. Here the modeeffective area (A_{eff}) and the normalized effective area (A_{norm}) have been calculated as

$$A_{eff} = \frac{\left(\iint \left|\sqrt{E_x + E_y}\right|^2 dxdy\right)^2}{\iint \left|\sqrt{E_x + E_y}\right|^4 dxdy} \qquad (3)$$

and

$$A_{norm} = A_{eff} \big/ R^2 . \qquad (4)$$

As shown in Fig. 8, the normalized A_{eff} increases as the core radius, R decreases and the values reach maximum when R arrives at its cutoff value. But most importantly, it can be seen that the value of A_{norm} increases with the introduction of the hole and it increases further with the increases in hole radius. Though the confinement factor slightly reduces with the introduction of a hole just slightly above the cutoff R_C, but for larger values of R the factor remains unchanged.

C. Light enhancement and average mode intensity

The light enhancement, P is defined as the ratio of the power confined in the hole of the nanowire to that confined in the same area of the nanowire without the hole and can be expressed as [10]

$$P = \int_A S_{z1} dA \Big/ \int_A S_{z2} dA , \qquad (5)$$

where S_{z1} is the power flow in the nanowire with the hole and $S_{z2}$is that in the nanowire without hole. Fig. 9 shows the light enhancement with respect to hole diameter. It is observed that enhancement of light increases as hole diameter decreases and is equal to 1 when the hole diameter is about 15nm-20 nm. The power confinement is higher when the hole diameter is smaller than the critical value.

Fig. 9 also shows the average mode intensity (I_{hole}) which is calculated by integrating the normalized power as [10]

$$I_{hole} = \int_H S_z dA / \pi r^2 , \qquad (6)$$

where S_z is normalized power flow over the hole and r is the radius of the hole. The average mode intensity also increases when the hole size is very small. The reason behind higher light enhancement and higher average mode intensity in smaller holes can be explained by field decay coefficient inside the hole as is explained in [10].

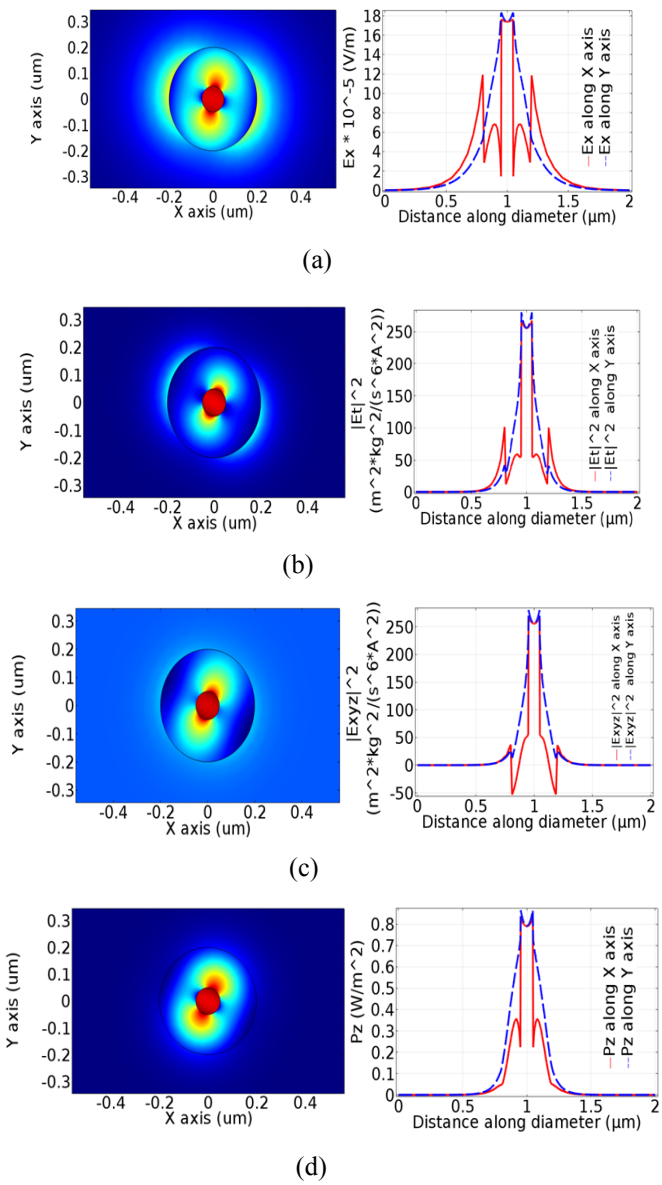


Fig. 7. Fundamental mode field distribution, (a) Electric field E_x component. (b)Transverse electric field intensity$|E_t|^2$. (c) Total electric field $|E_{xyz}|^2$ and (d) z--component of the Poynting vector (power flow) along X and Y axis.

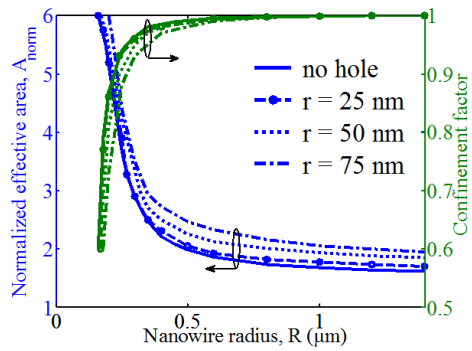


Fig. 8. The normalized effective area and the confinement factor.

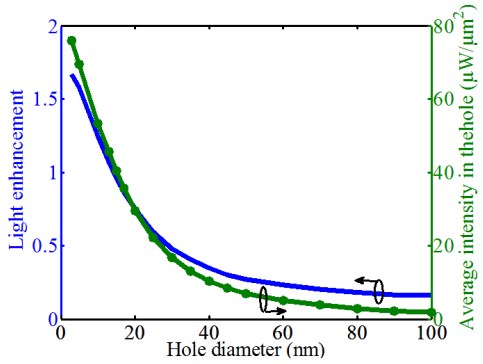


Fig. 9. The Light enhancement and the average intensity versus hole diameter.

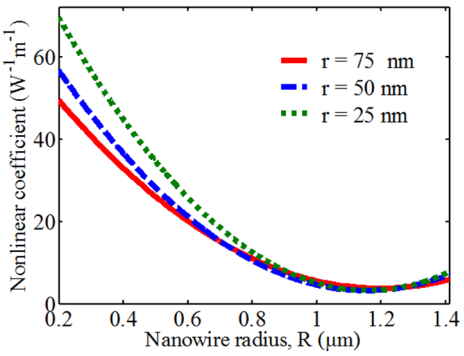


Fig. 10. The nonlinear coefficient of the silicon nanowire with hole.

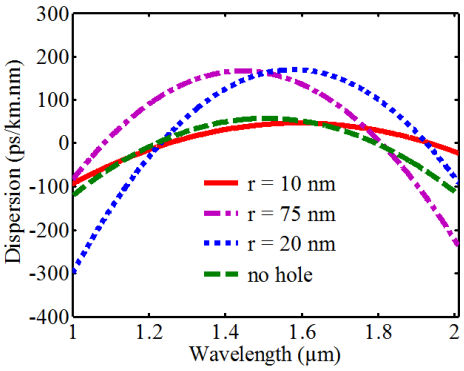


Fig. 11. The dispersion of silicon nanowire for different hole size.

D. Effect of hole on nonlinearity and dispersion

The nonlinear coefficient γ is calculated using (1). Fig. 10 shows the variation of nonlinear coefficient with the nanowire radius. We have considered three hole diameters, e.g., 50 nm, 100 nm, and 150 nm. Because of the presence of hole in the core region, the mode area of the fundamental mode increases which has been observed in Fig. 8 and thus the nonlinearity decreases with the increase in effective area. It can be seen from Fig. 10 that nonlinearity decreases as hole diameter increases.Dispersion profiles of silicon nanowire in the presence of hole are shown in Fig. 11. It can be observed that dispersion in silicon nanowire is greatly influenced by the presence and the size of air holes. The zero dispersion wavelengths are also shifted. The results mentioned above may present opportunities for engineering nonlinear applications like supercontinuum generation.

III. Conclusion

The silicon nanowire with a hole of sub-wavelength diameter located at the center is analyzed by using a vector FEM. The normalized effective area, confinement factor, nonlinear coefficient, and the light enhancement have been calculated. It has also been presented that the light enhancement and the average mode intensity are higher within smaller holes. From the GVD properties, it has been found that the zero dispersion wavelength shifts because of the presence of air hole. Thus, the results may provide ample opportunities in various applications like sensing, supercontinuum generation, dispersion compensation, etc.

References

[1] A. M. Morales and C. M. Lieber, "A laser ablation method for the synthesis of crystalline semiconductor nanowires," *Science,* vol. 279, pp. 208–211, Jan. 1998.

[2] J. Appenzeller, J. Knoch, M. T. Bjork, H. Riel, H. Schmid, and W. Riess, "Toward Nanowire Electronics,"*IEEE Trans. Electron Devices*, vol. 55, no. 11, pp. 2827-2845, Nov. 2011.

[3] L. Tong, J. Lou, and E. Mazur, "Single-mode guiding properties of subwavelength-diameter silica and silicon wire waveguides," O*ptics Express*, vol. 12, no. 6, pp. 1025-1035, Feb. 2004.

[4] J. B. Driscoll, X. Liu, S. Yasseri, I. Hsieh, J.I. Dadap, and R. M.Osgood, Jr., "Large longitudinal electric fields (E_z) in silico nanowire waveguides," *Optics Express*, vol. 17, no. 4, pp. 2797-2803, Feb. 2009.

[5] A. L. Henneghien, B. Gayral, Y. Desieres, and J. M. Gerard, "Simulation of waveguiding and emitting properties of semiconductor nanowires with hexagonal or circular sections," *J. Opt. Soc. Am*., vol. 26, no. 12, pp. 2396-2402, Dec. 2009.

[6] N. Kejalakshmy, A. Agrawal, Y. Aden, D. M. H. Leung, B.M.A. Rahman, K.T.V. Grattan, "Characterization of silicon nanowire by use of full-vectorial finite element method," *Applied Optics*, vol. 49, no. 16, pp. 3173-3181, May 2010.

[7] M. Lipson, "Guiding Modulating and Emitting Light on Silicon Challenges and Opportunities," *J. Lightwave Technol*., vol. 23, no. 12, pp. 4222-4238, Dec. 2005.

[8] G. S. Wiederhecker, C. M. B. Cordeiro, F. Couny, F. Benabid, S. A. Maier, J. C. Knight, C. H. B. Cruz, and H. L. Fragnito, "Sub-Wavelength Intensity Profiles and Field Enhancement within an Optical Fiber," *in Proceedings of conference on Lasers and Electro-Optics/Quantum Electronics and Laser Science Conference and Photonic Applications Systems Technologies*, 2007.

[9] Y. K. Lize, E. C. Magi, V. G. Taeed, J. A. Bolger, P. Steinvurzel, and B. J. Eggleton, "Microstructured optical fiber photonics wires with subwavelength core diameter," *Optics Express*, vol. 12, no. 14, pp. 3209-3217, July 2004

[10] Y. Ruan, S. Afshar, and T. M. Monro, "Light Enhancement Within Nanoholes in High Index Contrast Nanowires," *IEEE Phot. Soc., v*ol. 3, no. 1, pp. 130-139, Feb. 2011.

[11] V. R. Almeida, Q. Xu, C. A. Barrios, and M. Lipson, "Guiding and confining light in void nanostructure," *Opt. Lett*., vol. 29, no. 11, pp. 1209–1211, June 2004.

[12] C. R. Pollock, Fundamentals of Optoelectronics, Marietta, OH: Ceramic Book & Literature Service, 2003.

[13] A. Agrawal, N. Kejalakshmy, B. M. A. Rahman, and K. T. V. Grattan, "Soft Glass Equiangular Spiral Photonic Crystal Fiber for Supercontinuum Generation," *Photon. Technol. Lett.*, vol. 21, no. 22, pp. 1722 -1724, Nov. 2009.

[14] D. Hasan and M. S. Alam, "Ultra-Broadband Confinment in Deep Sub-Wavelength Air Hole of a Suspended Core Fiber," *J. Ligthwave Technol.*, vol. 32, no. 8, pp. 1434-1441, Apr. 2014.

On the Staircase like Structures in Gate C-V Characteristics of InSb/$In_{0.8}Al_{0.2}Sb$ MOS-HEMT Devices

Nugaira Gahan Mim* and Anisul Haque
Department of Electrical and Electronic Engineering, East West University, Dhaka 1212, Bangladesh.
*Jahanewu@gmail.com

***Abstract*— Gate C-V characteristics of InSb/$In_{0.8}Al_{0.2}Sb$ MOS-HEMT devices show staircase like behavior for both positive and negative gate voltages. We investigate the origin of this staircase nature. It is found that at low magnitudes of the gate voltages, the carriers are confined within the channel and the top InAlSb barrier behaves as a dielectric layer. Thus the gate oxide capacitance is determined by the series combination of the capacitance of the gate dielectric layer and the InAlSb barrier layer, as expected. However, when the gate voltage magnitude increases above a critical value, carriers are not confined within the channel region only. The InAlSb barrier region is also populated with carriers. As a result, InAlSb layer does not contribute to gate capacitance. The gate capacitance increases sharply and exhibits the staircase like structure.**

Keywords- Gate capacitance,* InSb, *MOS-HEMT, Staircase structure, Carrier concentration.

I. Introduction

As Si CMOS scaling has approached the scaling limits [1], III-V semiconductor is emerging as a potential alternative [2]. MOS-HEMT device is an attractive choice because of its superior electron mobility and flexible heterostructure engineering [3]. MOS-HEMT is one kind of buried-channel MOSFET which has a quantum well epi-structure with a wide band-gap barrier layer between gate insulator and channel to reduce the effects of interface trapstates at the dielectric semiconductor interface, as well as to improve carrier mobility. In recent years, antimony based electronic devices have become attractive for field effect transistors because of low effective mass and high mobility of antimonide semiconductors [4]. InSb channel MOS-HEMT devices have been investigated widely [5]. The gate capacitance of InSb MOS-HEMT exhibits a staircase like behavior for both positive and negative gate voltages [6]. In this study we investigate the origin of such staircase structure of the gate capacitance. Simulations are performed using the Silvaco-ATLAS TCAD tool [7].

II. Device Structure and Simulation Model

The structure of III-V MOS-HEMT consists of a buried channel known as quantum well in between two wide band-gap III-V materials. A semi-insulating substrate supports an undoped wide band-gap buffer layer, then an intrinsic channel layer and finally another barrier layer. Fig 1 shows a layer diagram of a MOS-HEMT. Our considered structure, InSb/$In_{0.8}Al_{0.2}Sb$ MOS-HEMT, consists of a InSb quantum well channel where undoped $In_{0.8}Al_{0.2}Sb$ is used as buffer and n+ doped barrier. Platinum (Pt) is used as gate contact with Hafnium di-oxide (HfO_2) as gate oxide.

The energy band diagram InSb MOS-HEMT with channel thickness T_{ch} = 4 nm, top barrier thickness T_b = 8 nm, conduction band offset ΔE_c = 0.84 eV and valance band offset ΔE_v = 0.406 eV, between channel and barrier is shown in Fig.2 under flatband condition.

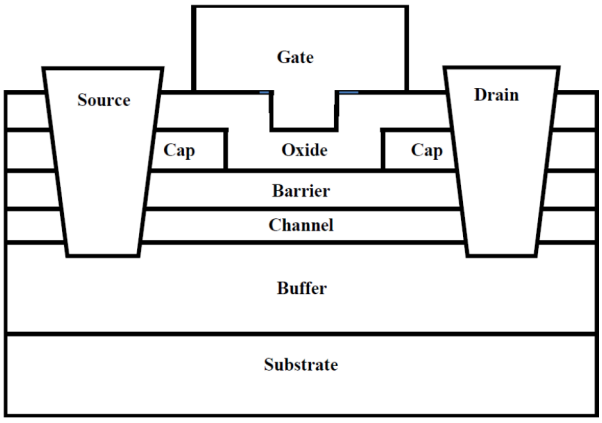


Fig. 1 Schematic diagram of a III-V MOS-HEMT.

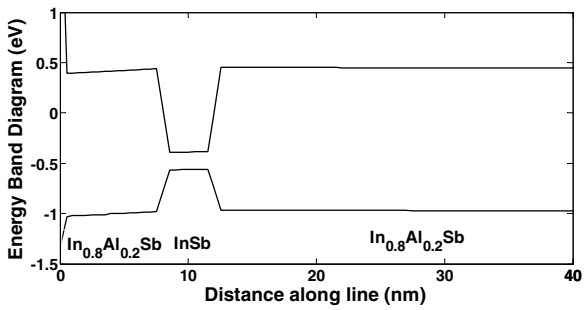


Fig. 2 Energy diagram of a $In_{0.8}Al_{0.2}Sb$/InSb MOS-HEMT under flatband condition.

We have used the Silvaco-ATLAS TCAD tool to simulate the device. To get more accuracy in our simulation, quantization effect is accounted for by using 1D self-consistent Schrödinger-Poisson equations solver for both electrons and holes that self-consistently solves Poisson's equation for potential and Schrödinger equation for bound state energies and carrier wavefunctions. Quasi-static gate capacitance is calculated from the converged potential and charge distributions considering Fermi-Dirac statistics. Solution is obtained for one Γ-valley with isotropic electron effective mass, heavy hole and light hole bands with isotropic

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

hole effective mass also. Important parameters of materials specification are shown in TABLE I .

TABLE II
PROPERTIES OF MATERIALS USED IN SIMULATION AT 300K

Parameters	Material		
	HfO_2	InAlSb	InSb
Relative permittivity ϵ_r	22	13.2	17.7
Bandgap E_g (eV)		1.42	0.174
Electron affinity χ (V)		3.75	4.59
N_c (cm^3)		2.57×10^{18}	3.94×10^{16}
N_v (cm^3)		2.02×10^{19}	7.12×10^{18}
n_i (cm^3)		8.7×10^{6}	1.84×10^{16}
m_c/m_0		0.0388	0.0135
m_{hh}/m_0		0.532	0.430
m_{lh}/m_0		0.04	0.015

III. RESULTS AND DISCUSSIONS

We have previously reported the quasi-static gate C-V characteristics of InGaAs/InAlAs and InSb/InAlSb MOS-HEMT devices [6]. The C-V curves of InSb devices are shown in Fig. 3 which exhibit staircase like behavior in both inversion and accumulation regions. The gate voltages at which the stairs appear depend on the thickness of the top barrier layer.

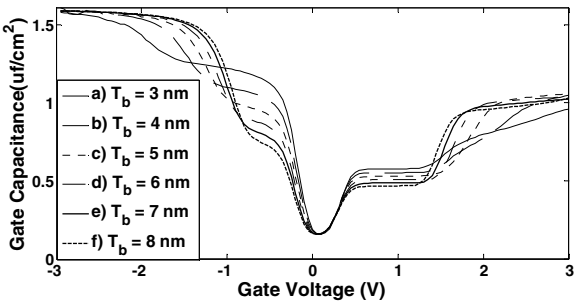


Fig. 3 Gate C-V characteristics of a InSb MOS-HEMT with EOT = 1.773 nm and T_{ch} = 4 nm for different barrier thicknesses T_b = 3-8 nm at 300 K.

To investigate the origin of this staircase like behavior, we calculate the carrier density as a function of distance from semiconductor-dielectric interface for two different values of T_b under inversion and accumulation bias voltages. Fig. 4 and Fig. 5 show results for T_b = 8 nm and Fig. 6 and Fig. 7 show results for T_b = 5 nm, respectively. Fig. 4 is for inversion bias. We have observed the electron density at two different gate voltages, first one is at a gate voltage 1.2 V on the first plateau before start of the second stair and the second one is at a gate voltage 2.4 V at the second plateau. At lower positive gate voltage of 1.2 V, curve (a) shows small carrier density in the top barrier region demonstrating that this layer serves as another dielectric layer in series with the gate dielectric. The gate capacitance is thus determined by the series combination of the two dielectrics, in addition to the semiconductor capacitance. At higher positive gate voltage, curve (b) shows high electron concentration within the top barrier layer whose density is comparable to that within the channel. Consequently the device is behaving like a surface channel transistor and the top barrier does not contribute to the gate capacitance in this bias condition. The gate capacitance exhibits a sharp increase when the top barrier begins to get populated by carriers.

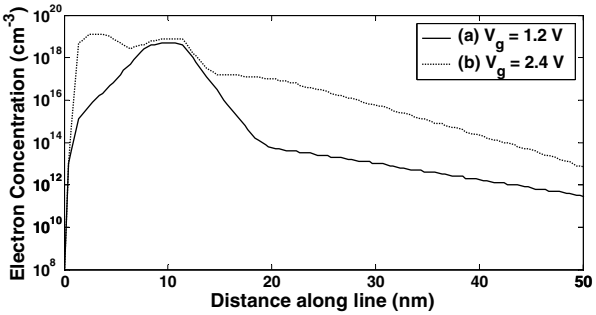


Fig. 4 Electron density as a function of distance from semiconductor-dielectric interface. (a) gate voltage 1.2 V and (b) gate voltage 2.4 V for T_b = 8 nm.

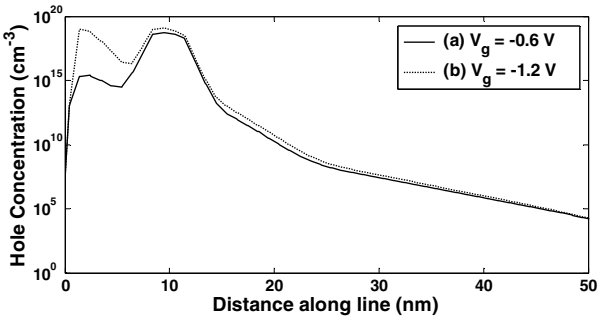


Fig. 5 Hole density as a function of distance from semiconductor-dielectric interface. (a) gate voltage -0.6 V and (b) gate voltage -1.2 V for T_b = 8 nm.

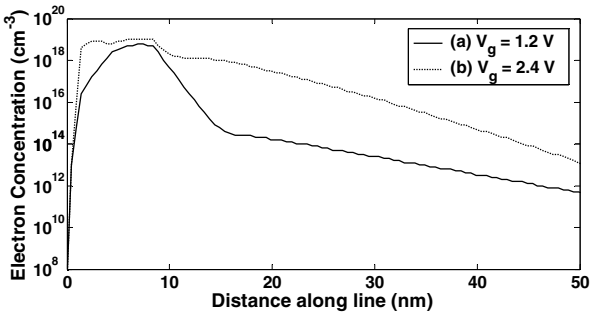


Fig. 6 Electron density as a function of distance from semiconductor-dielectric interface. (a) gate voltage 1.2 V and (b) gate voltage 2.4 V for T_b = 5 nm.

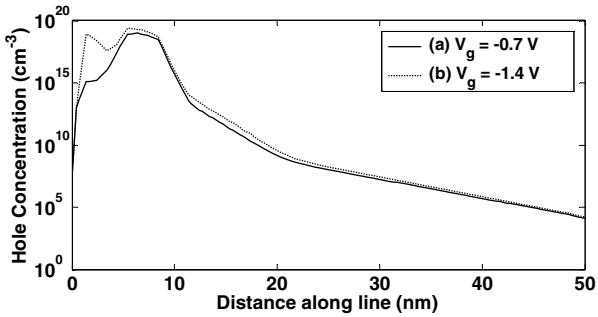


Fig. 7 Hole density as a function of distance from semiconductor-dielectric interface. (a) gate voltage -0.7 V and (b) gate voltage -1.4 V for T_b = 5 nm.

Similarly, we have also calculated the hole density at two different gate voltages -0.6V and -1.2V, respectively for the same device under the accumulation bias as shown in Fig.5. For lower magnitude of the gate voltage, carriers are again confined within the channel region. The top barrier serves as a dielectric layer that contributes in gate capacitance On the other hand, for higher magnitude of the gate voltage, the top barrier is populated by carriers and this layer does not contribute to the gate voltage causing a sharp rise of gate capacitance. Fig. 6 and Fig. 7 show results similar to those presented in Fig. 4 and Fig. 5, respectively for a device with T_b = 5 nm. Again the top barrier does not get populated with carriers at lower magnitudes of gate voltage but above a certain value of the gate voltage, the top barrier is not able to keep the carriers confined to the channel region. As a result the gate capacitance shows sharp rise or staircase like behavior.

IV. Conclusions

The C-V characteristics have been widely used as a diagnostic tool for device characterization that contains a wealth of information about the semiconductor and the device properties. In this work, the origin of the staircase like behavior in the gate capacitance of InSb MOS-HEMT has been investigated through simulation using Silvaco-ATLAS TCAD tool. The sharp rise in the gate capacitance is attributed to the inability of the top barrier to keep the carrier concentration away from the semiconductor surface and within the channel region. At the second plateau, the devices behave more like surface channel transistors.

References

[1]. K.C. Saraswat, C.O. Chui, D. Kim, T. Krishnamohan, A. Pethe, "High Mobility Materials and Novel Device Structures for High Performance Nanoscale MOSFETs," in *Electron Devices Meeting, 2006.IEDM '06.International*, pp.1-4, 11-13 Dec. 2006.

[2]. R Chau, S. Datta, and A. Majumdar, “Opportunities and challenges of III-V nanoelectronics for future high-speed, low-power logic applications,” in *Proc. IEEE CSIC Symp.,* 2005*,* pp. 17–20.

[3]. S. Datta et al., “85nm gate length enhancement and depletion mode InSb quantum well transistors for ultra high speed and very low power digital logic applications,” *in Proc. IEEE IEDM*, 2005, pp. 763-766.

[4]. B. Bennett, R. Magno, J. Boos, W. Kruppa and M. Ancona, "Antimonide-based compound semiconductors for electronic devices: A review", *Solid-State Electronics*, vol. 49, no. 12, pp.1875-1895, 2005.

[5]. T. Ashley, A.R. Barnes, L. Buckle, S. Datta, A.B. Dean, M.T. Emeny, M. Fearn, D.G. Hayes, K.P. Hilton, R. Jefferies, T. Martin, K.J. Nash, T.J. Phillips, W-H.A. Tang, P.J. Wilding and R. Chau, “Novel InSb-based Quantum Well Transistors for Ultra-High Speed, Low Power Logic Applications”, *Proc. ICSICT 3*, pp. 2253-2256 , 2004.

[6]. N. Mim, S. Kutub and A. Haque, "Characterization of top barrier thickness from gate capacitance of high mobility III-V semiconductor MOS-HEMT devices", Results in Physics, vol. 6, pp. 233-234, 2016.

[7]. "Atlas - Device Simulation Framework", *Silvaco.com*, 2016.[Online]. Available:http://www.silvaco.com/products/tcad/device_simulation/atlas/atlas.html.[Accessed: 21- Jul- 2016].

A Coordinated Approach for Frequency Control of Zero Emission Based Smart PV-Wind-Battery Power System

Saeed Sepasi, Abdul Motin Howlader*, Ehsan Reihani and Leon R. Roose

Hawaii Natural Energy Institute, University of Hawaii, Manoa
1680 East-West Road, Honolulu, Hawaii, USA
*motin@ieee.org

Abstract—**Power demand of isolated communities or islands primarily relies on heavy oil based diesel generator systems where these fossil fuel based power generators are not only costly but also create adverse effect on the environment. Photovoltaic generator system (PV system), wind energy conversion system (WECS) and battery energy storage system (BESS) are clean sources of alternative energy. These energy sources can be the alternative of diesel generator based power system for an isolated communities such as islands. However, both PV systems and WECSs are uncertain and fluctuating sources of energy causing power fluctuations and frequency deviations. This paper proposes a coordinated approach of the frequency control for a zero emission based small power system with renewable energy penetration. PV system, WECS and BESS have been utilized to control frequency of the proposed zero emission based power system. Efficacy of the proposed frequency control system is shown with simulations results.**

Index Terms— **Photovoltaic system, wind energy conversion system, battery energy storage system, frequency control.**

I. Introduction

Electric power demand of isolated communities and islands largely depends on the fossil fuel based diesel generator systems. Burning of fossil fuel emits greenhouse gases which cause global warming and have adverse effects on the environment [1, 2]. In order to reduce these emissions, world is now searching for emission free clean sources of energy. The green renewable energy has been getting attention from the past few decades as a clean source of energy [3, 4]. Integration of wind turbines and solar energy sources into the power grid is increasing exponentially. Though wind and solar energy considered as the most promising clean source of energy but these renewable sources have some shortcomings. Power output of wind turbine generator (WTG) depends on wind velocity, and photovoltaic (PV) generator system power output relies on the solar insulation. The weather conditions effect wind velocity and solar insulation and cause fluctuations on their output power [5, 6]. On the other hand, load of a power system is not constant. All these fluctuations could lead to oscillation in frequency of power systems. When renewable power generators are integrated within the power system, frequency has to be controlled within the tolerable range.

This paper proposes a frequency control approach for the zero emission based small power system. Wind energy conversion system (WECS), PV system and battery energy storage system (BESS) are considered as the power generator sources of the electrical system.

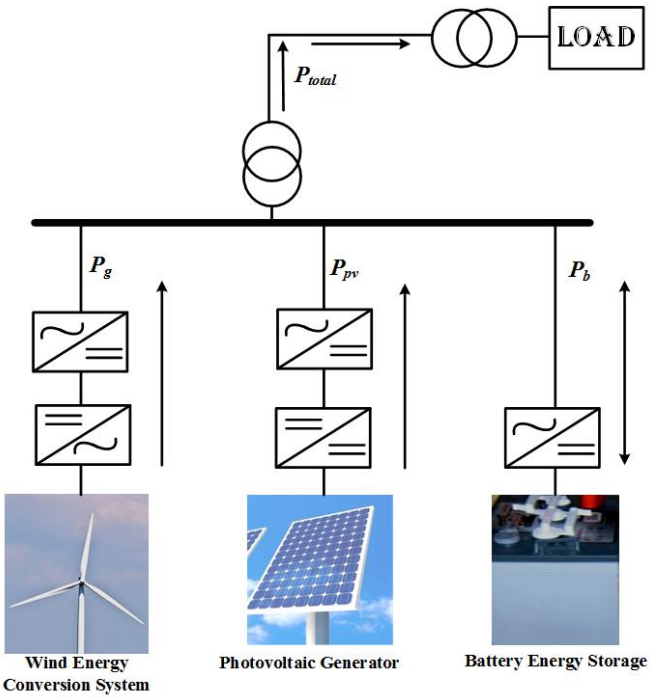


Fig. 1 Block diagram of proposed power system

A coordinated approach is applied to control the power system frequency. Low frequency components of the frequency deviation are controlled by pitch control of the WECS. Mid and high frequency components of the frequency deviation are reduced by PV system and BESS, respectively. Output of the proposed control system has been compared with a power system that doesn't have frequency control strategy. Simulations have been executed in the MATLAB/Simulink environment.

II. Description of Proposed Power System

Fig. 1 shows a schematic view of the proposed power system. WECS, PV system and BESS are parts of this small power system. The rated load of the grid is 687.5 kW. Besides, rated powers of the WECS, PV system and BESS are considered as 275 kW, 450 kW and BESS 400 kW, respectively. The mathematical model of the PV system is given in Ref. [5] and WECS's mathematical model is demonstrated in Ref. [7]. The frequency control system for the proposed power system model has been shown in Fig. 2. The power command of the BESS, P^*_b has been determined from

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

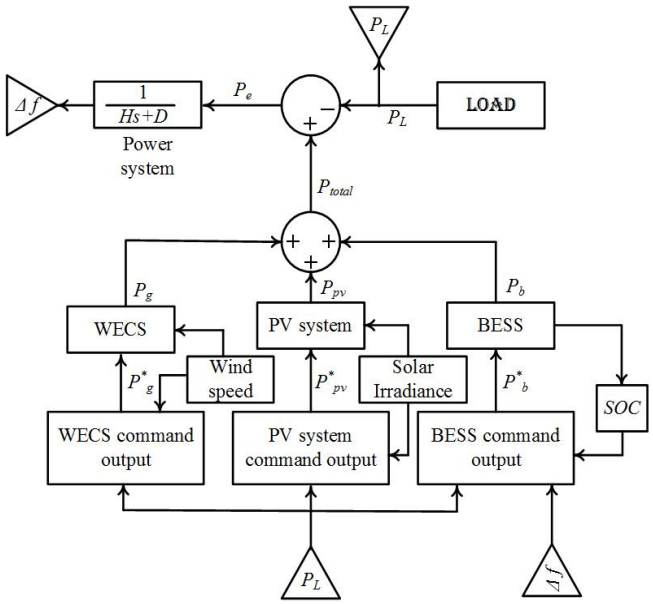


Fig. 2 Frequency control system

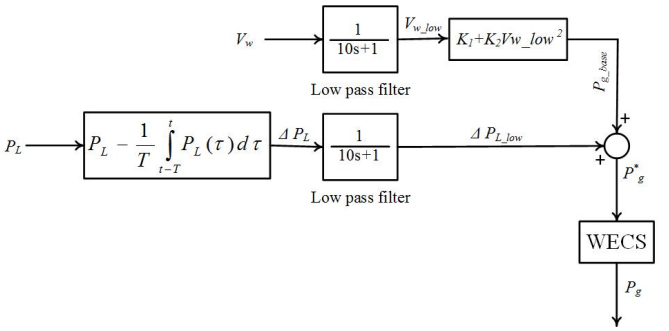
the frequency deviation, load fluctuation index and the state of charge (SOC) of the battery. WECS's output power command, P^*_g, has been estimated from load fluctuation index and wind velocity. PV system's output power command, P^*_{pv}, has been calculated from solar irradiance and load fluctuation index. Total injection power of the WECS, PV system and BESS is subtracted from the load demand and determined the frequency deviation, Δf, using the rotating mass of the power system. Symbols H and D, are inertia and damping constants of the power system, respectively.

Control strategies for the WECS, PV system and BESS are shown in Fig. 3. Output power command, P^*_g, for the WECS is illustrated in Fig. 3(a). The output power command of the WECS, is determined from the load fluctuation index, ΔP_L, and WECS base power P_{g_base}. The load fluctuation index, ΔP_L is calculated from the following equation [8]:

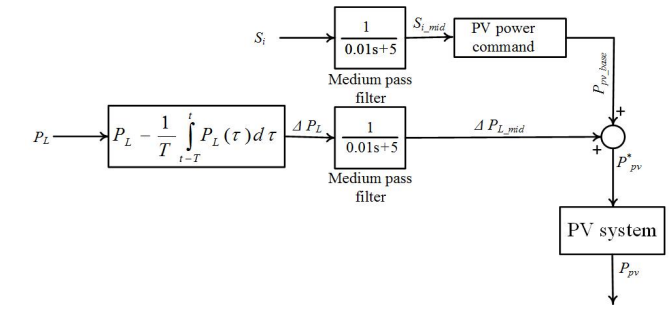
$$\Delta P_L = P_L - \frac{1}{T}\int_{t-T}^{t} P_L(\tau)d\tau \qquad (1)$$

where t is the current simulation time, T=50s is an integral time interval and τ is the sampling period.

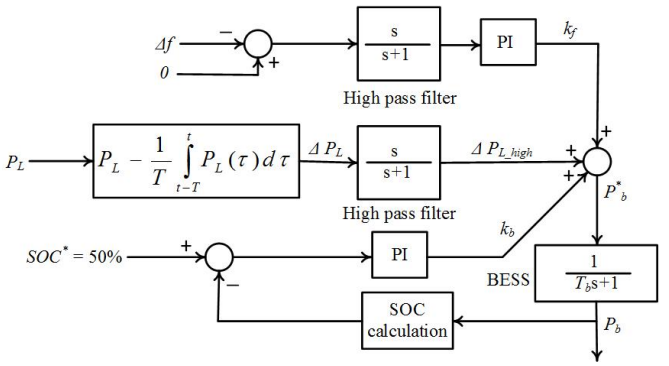
Both load fluctuation index, ΔP_L and wind velocity V_w are passed through a low pass filter to determine the output power command of WECS. Output power command of the PV system is shown in Fig. 3(b), medium level frequency deviation is reduced by the PV system. Load fluctuation index, ΔP_L and solar insulation S_i are passed through a medium pass filter which delivered the power command of the PV system. Output power control command for the BESS is illustrated in Fig. 3(c). The BESS command output is determined from the load fluctuation index, frequency deviation and SOC of the BESS. Load and frequency deviation are passed through a high pass filter. This filter determines command factors of k_f and ΔP_{L_high} using a PI controller. SOC of the BESS is controlled at 50% level. Finally, all three factors, k_f, ΔP_{L_high}, and k_b deliver the output power command of BESS. Parameters of the WECS and PV system are given in Refs. [5, 8].



(a) WECS control system

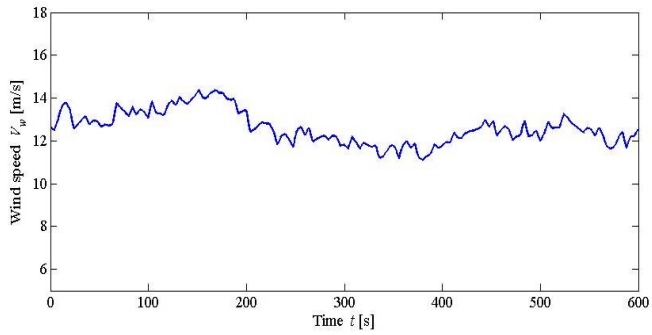


(b) PV system control system

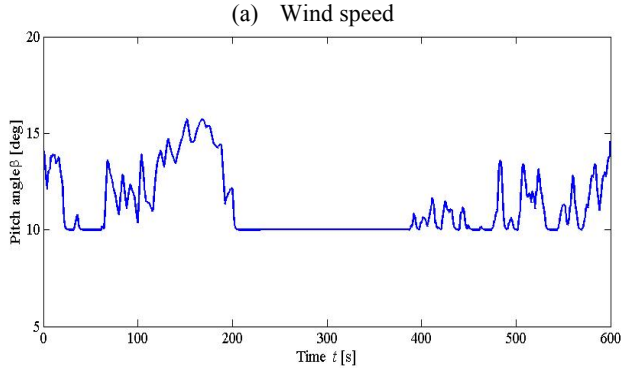


(c) BESS control system

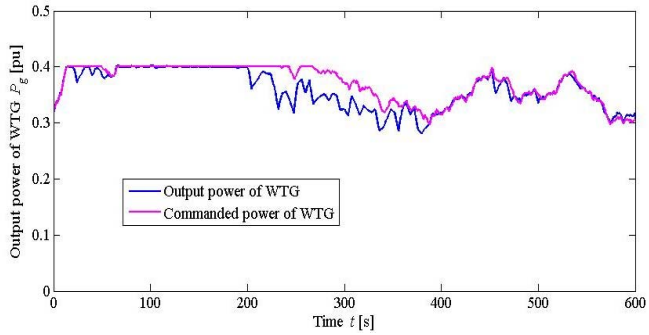
Fig. 3 Control strategies



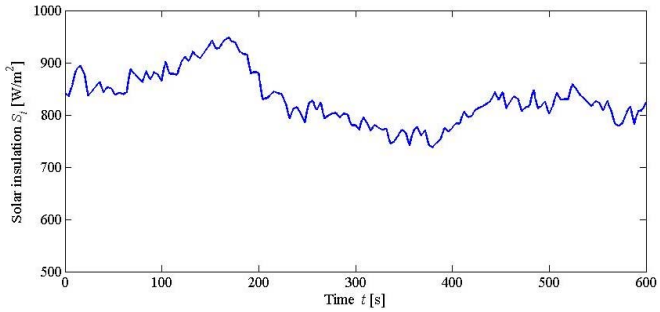
(a) Wind speed



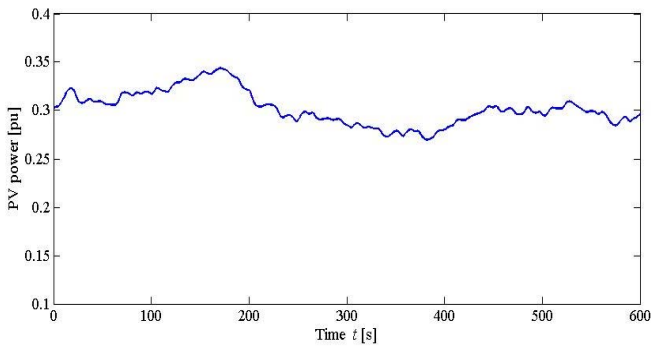
(b) Pitch angle



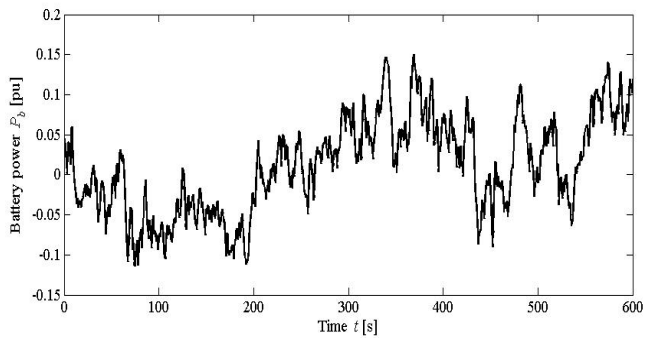
(c) WTG Power output



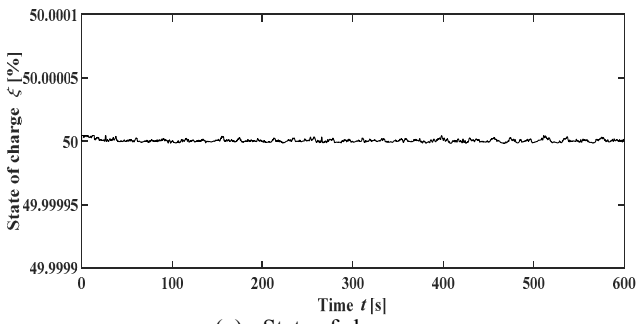
(d) Solar insulation



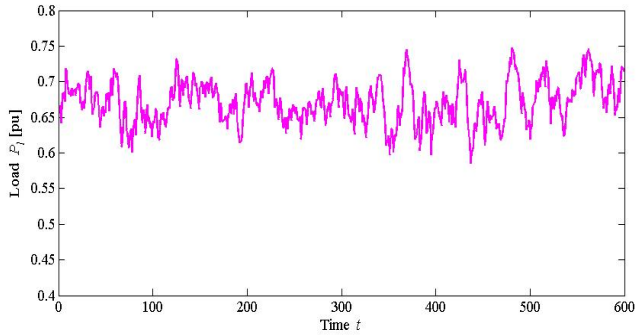
(e) PV power output



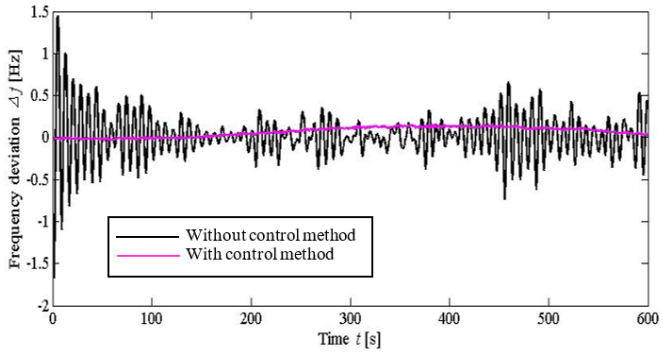
(f) BESS output



(g) State of charge



(h) Power system load



(i) Frequency deviation

Fig. 4 Simulation Results

III. RESULTS AND DISCUSSIONS

Simulation results of the proposed PV-Wind-BESS based power system have been shown in Fig. 4. Results compare outputs of a system which has the proposed control method with a system which doesn't have this control system. Fig. 4(a) illustrates wind velocity for the WECS. The rated wind speed is considered as 12.5m/s for the WECS. In Fig. 4(a), wind velocity is considered as above and below the rated speed. Outcome of the pitch action is shown in Fig. 4(b). When wind velocity is above the rated point (i.e. 12 m/s), the pitch angle is triggered to control the rated power of the WECS, moreover, below the rated wind velocity, the pitch angle is constant 10° for extracting the maximum power from wind velocity. Output power of the WECS is reflected in Fig. 4(c). From this figure, when wind velocity is above rated point, WECS maintains the rated power (i.e. 0.4 pu) using pitch angle controller. Solar insulation for the PV system is shown in Fig. 4(d). As it can be seen, fluctuations of solar insulation is very similar to actual insulations. Output power of the PV system is illustrated in Fig. 4(e) where power output of the PV system follows solar insulation. Power output and SOC of the BESS are illustrated in Figs. 4(f) and (g), respectively. The power system load is shown in Fig. 4(h). From this figure, load is being fluctuated as an actual load. Frequency deviation of the proposed power system is shown in Fig. 4(i). As it can be seen in Fig 4(i), the proposed frequency control method reduces the frequency deviation extensively.

IV. CONCLUSION

A coordinated method for the frequency control for an isolated renewable energy based small power system is shown in this paper. WECS, PV system and BESS are applied for the coordinated approach. Low, medium and high frequency components of the frequency deviation are controlled by pitch angle of WECS, PV system and BESS, respectively. These control strategies have been described in this paper. A

comparison of simulation results is done with the non-controller strategies. These results ensured that the proposed method can reduce the frequency deviation extensively leading to a stable power system.

REFERENCES

[1] A. M. Howlader, N. Urasaki, A. Yona, T. Senjyu, and A. Y. Saber, "A review of output power smoothing methods for wind energy conversion systems," *Renewable and Sustainable Energy Reviews*, vol. 21, pp. 135146, Oct. 2013.

[2] A. M. Howlader, N. Urasaki, and A. Y. Saber, "Control strategies for wind-farm-based smart grid system," *IEEE Transactions on Industry Application*, vol. 50, no. 5, pp. 3591 - 3601, Feb. 2014.

[3] S.E. Hosseini, M.A. Wahid, "Hydrogen production from renewable and sustainable energy resources: Promising green energy carrier for clean development," *Renewable and Sustainable Energy Reviews*, vol. 57, pp. 850-866, May. 2016.

[4] S. Sepasi, R. Ghorbani and B.Y. Liaw "A novel on-board state-of-charge estimation method for aged Li-ion batteries based on model adaptive extended Kalman filter," Journal of Power Sources, 245, pp. 337-344.

[5] M. Datta, T. Senjyu, A. Yona, H. Sekine, T. Funabashi, "Smoothing Output Power Variations of Isolated Utility Connected Multiple PV Systems by Coordinated Control," *Journal of power electronics*, vol. 9, no. 2, pp. 320-333, Mar. 2009.

[6] A. M. Howlader, T. Senjyu, and A. Y. Saber, "An Integrated Power Smoothing Control for a Grid-Interactive Wind Farm Considering Wake Effects," *IEEE System Journal*, vol. 9, no. 3, pp. 954 - 965, Dec. 2014

[7] T. Senjyu, R. Sakamoto, N. Urasaki, T. Funabashi, H. Fujita, and H. Sekine, "Output Power Leveling of Wind Turbine Generator for All Operating Regions by Pitch Angle Control", *IEEE. Trans. on Energy Conversion*, vol. 21, no. 2, pp. 467-475, 2006.

[8] T. Kakeko, A. Uehara, T. Senjyu, A. Yona, and N. Urasaki, "An integrated control method for a wind farm to reduce frequency deviations in a small power system", *Applied Energy*, vol. 88, no. 4, pp. 1049-1058, Apr. 2011.

Impact of Interface Charges on the Performance of Dual Material Double Gate Tunnel FET

Samantha Lubaba Noor,[1,*] Samia Safa,[2] and Md. Ziaur Rahman Khan[3]
[1−3]Department of Electrical and Electronic Engineering,
Bangladesh University of Engineering and Technology
Dhaka-1205, Bangladesh
*lubaba.noor@gmail.com

***Abstract*—In this work, effect of interface charge on the performance of a Dual Material Double Gate (DMDG) Tunnel Field Effect Transistor (TFET) device has been studied. Due to the presence of high electric field in the device, interface charges are generated at oxide-semiconductor interface by hot carrier effect. A 2-D model of the device with interface charge has been developed in TCAD simulator and impact of both positive and negative interface charges has been observed in potential distribution, band profile, carrier concentration, tunneling current and threshold voltage. Since high electric field mainly exists in the source-body junction of the device, interface charges are mainly localized at that region. Effect of localized charge spreading over varying length has also been analyzed on surface potential and threshold voltage of the device. The study will be useful in fabrication and application of DMDG TFET devices.**

***Index Terms*—DMDG Tunnel FET, interface charge, threshold voltage, TCAD model**

I. Introduction

The physical limit of 60mV/dec Subthreshold Slope (SS) of MOSFET poses a challenge for further scaling of its power supply voltage. The power crisis has enforced researchers to search for steeply switching devices for MOSFET replacement. Tunnel FET devices that rely on band-to-band tunneling mechanism have recently shown great promise in overcoming the SS limit [1], [2]. In Tunnel FET devices tunneling occurs in the region of high electric field at source-body interface where local band bending reduces the width of energy barrier. Band bending is achieved by applying reverse bias at gate which pushes the energy band downwards (n type TFET) and turn on the device. The source-body high field is often in the order of MV/cm or higher. Such electric fields are above the threshold value (which is in the order of KV/cm) that causes carrier velocity saturation and hot electron generation [3]. Experimental studies of TFET based flash memories have shown the presence of hot carrier effects in TFET [4]. Some of the hot electrons may gain sufficient energy to overcome the semiconductor-oxide energy barrier. This will not only produce gate leakage current but also may permanently damage the oxide creating oxide-semiconductor interface charges like MOSFETs [5].

Dual Material Double Gate structure is a recent addition to TFET device which shows promising features like improved on current, better subthreshold slope [6]. Although a number of research have been carried out on this device [7]–[10], no work has yet been reported to include the effect of interface charges. With technologies going to nanoscale domain, it is increasingly important to study overall performance of novel devices with the presence of the hot-carrier degradation effects.

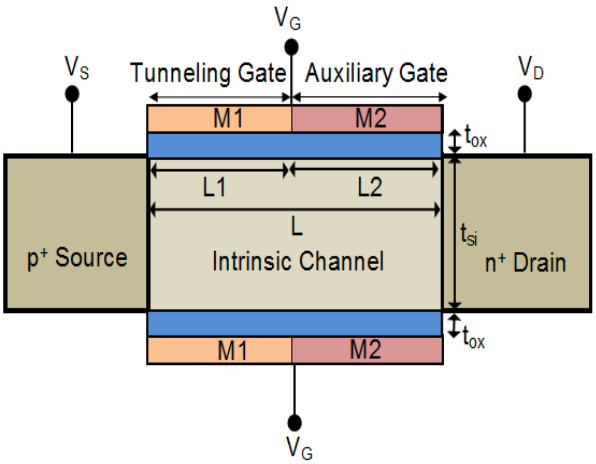


Fig. 1. Schematic representation of n type DMDG TFET

In this work, a detail numerical study of the performance of a n type DMDG TFET device has been performed with the presence of both positive and negative interface charges.

II. Simulated Device Structure

A schematic structure of n type DMDG TFET is shown in Fig. 1. Like other TFET structure, this device is a gated P-I-N diode. Its main feature is that the gate is made of two metals of different work-functions. The metal gate close to the source side is called tunnel gate, and it controls the tunneling between source-body junction. The metal gate near the drain side is auxiliary gate and its work-function variation mainly controls the OFF state current. Researchers have found that when tunneling gate has lower work-function than the auxiliary gate, it results in a higher I_{ON}, lower I_{OFF} and better SS [6].

The values of metal work-functions used in this work are: $\Phi_{tunnel} = 4eV$ and $\Phi_{auxiliary} = 4.4eV$. The source and drain are made of highly doped p type and n type dopant with doping levels of $N_{source} = 10^{20}cm^{-3}$ and $N_{drain} = 10^{19}cm^{-3}$ respectively. The body region is lightly n type doped with doping concentration of $10^{17}cm^{-3}$. Other device parameters are: channel length, $L = 50nm$ where $L_1 = L_2 = 25nm$, silicon film thickness, $t_{si} = 10nm$ and oxide thickness, $t_{ox} = 2nm$. Electron affinity, $\chi_{si} = 4.17eV$ and silicon bandgap, $E_g = 1.1eV$ are used. To study the effect of interface charge, positive and negative charges with concentration (N_{inf}) varying from $10^{12}cm^{-3}$ to $10^{13}cm^{-3}$ has been considered during simulation. Later, effect of localized charge

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

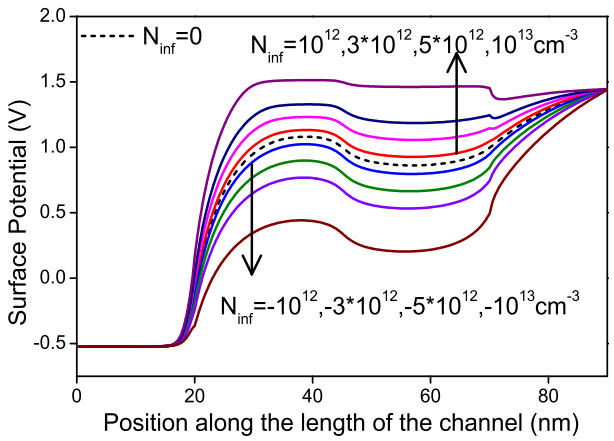


Fig. 2. Effect of interface charges on surface potential profile of the device. The dotted curve represents the potential profile with no interface charge. $V_{GS} = 0.5V$, $V_{DS} = 1V$

has also been studied.

III. Numerical Simulation

The DMDG n-type TFET investigated here has been simulated using Silvaco Atlas [11]. In all simulations, junctions were quasi-perfectly abrupt with junction width of $0.5nm$. To minimize the tunneling barrier, the high source doping needs to fall off to intrinsic level of the body in as short a width as possible. This requires a change in the doping concentration of $4 - 5$ orders of magnitude within a distance of only few nanometers [12], [13]. That is why abrupt doping profile has been considered here. Kane's band to band tunneling model has been included during simulation. Besides, to introduce non-linear effects, other models are included. They are: concentration dependent mobility, electric field dependent mobility, SRH recombination, concentration dependent lifetime SRH, auger recombination and bandgap narrowing model.

IV. Results and Discussion

Surface potential profile along the channel length direction has been studied in Fig. 2 for device with interface charge and fresh device. Effect of both positive and negative interface charges has been studied. It can be observed that as the negative charge concentration increases, surface potential becomes lower. Increment of positive charge has reverse effect. Besides, as positive interface charge is increased, potential difference between region under the two metal gate decreases. For example, at $N_{inf} = 10^{13}cm^{-3}$, the change of potential at the metal junction point almost diminishes.

The band profile for affected device has been compared with fresh device in Fig. 3 for $V_{GS} = 0.5V$ and $V_{DS} = 1V$. Positive interface charge pushes the energy band downward. For example, Fig. 3(a) shows that interface charge density increases by 10 order in magnitude, energy band is pushed downwards by $0.4eV$. This in turn decreases the shortest tunneling length in the source side. Contrary to the positive charges, negative interface charge pushes the overall band profile upward. It can be seen from Fig. 3(b) that as negative interface charge density has been increased from $10^{12}cm^{-3}$ to $5 \times 10^{12}cm^{-3}$, height of energy band overlap changes

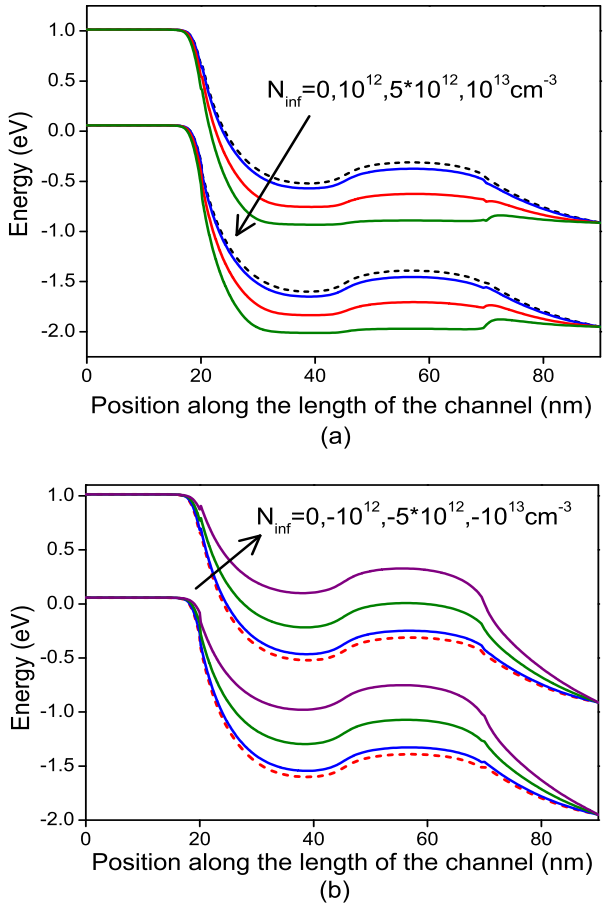


Fig. 3. Change of band profile with the presence of (a) positive interface charge and (b) negative interface charge. $V_{GS} = 0.5V$, $V_{DS} = 1V$

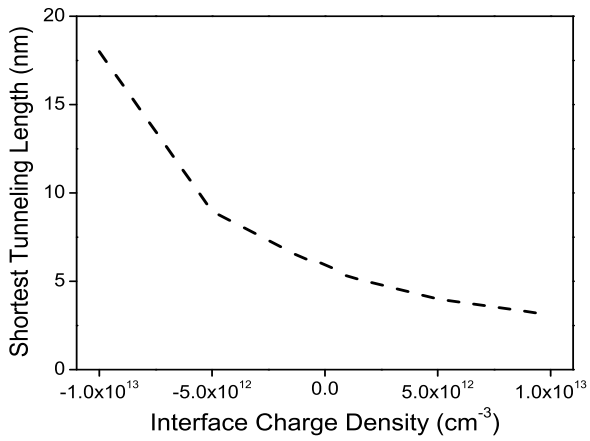


Fig. 4. Variation of shortest tunneling length for varying interface charge density. $V_{GS} = 0.5V$, $V_{DS} = 1V$

from $0.5eV$ to $0.25eV$. This results in an increase in shortest tunneling length.

Drain current in a TFET device is a function of shortest tunneling length (L_T) near the Fermi levels [14]. That is why, variation of L_T has been observed in Fig. 4. It can be seen that L_T decreases by $3nm$ for increasing the positive interface charge density by 10 order. For the same increment of negative charge density, L_T increases by $12nm$.

Electron concentration in Si body at $1nm$ depth from top interface has been observed in Fig. 5. For a fixed bias, inversion

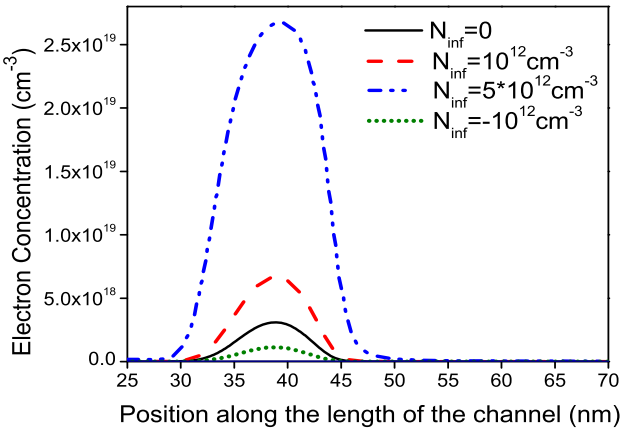


Fig. 5. Variation of surface electron concentration for varying interface charge density. $V_{GS} = 0.5V$, $V_{DS} = 1V$

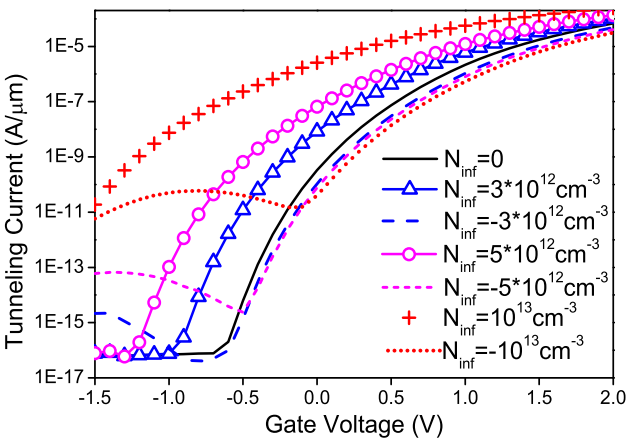


Fig. 6. Transfer characteristics for varying interface charge density. $V_{DS} = 1V$

charge is higher for higher positive N_{inf} (13.5 times higher for a 10 order change in N_{inf}). The scenario is opposite is case of negative charge. For instance, for $N_{inf} = 5 \times 10^{12} cm^{-3}$, maximum inversion charge is $2.67 \times 10^{19} cm^{-3}$ in case of positive charge, while the channel is depleted in case of negative charge. So it can be inferred that inversion occurs at lower voltage in the presence of positive fixed charge.

Transfer characteristics curve in Fig. 6 reflects the observations of band profile and tunneling length. For a fixed bias, positive interface charge results in higher tunneling current due to reduced tunneling length. As N_{inf} is increased, the device requires more and more negative gate bias to turn off. At $V_{GS} = -0.6V$, current, $I_D = 2.2 \times 10^{-15} A/\mu m$ when $N_{inf} = 10^{12} cm^{-3}$ which means that the device is already off. But in the same biasing condition $I_D = 1.8 \times 10^{-10} A/\mu m$ and $1.3 \times 10^{-7} A/\mu m$ for $N_{inf} = 5 \times 10^{12} cm^{-3}$ and $10^{13} cm^{-3}$ respectively. The negative interface charge has little effect for positive gate bias while it has noticeable effect in negative gate bias. Increasing negative charge density results in increasing off current and also enhanced ambipolarity of the device. This happens due to increased band overlap in the drain side at negative V_{GS}. Fig. 7 shows the change in threshold voltage (ΔV_T) with varying positive and negative interface charge density. Here, threshold voltage has been calculated as the

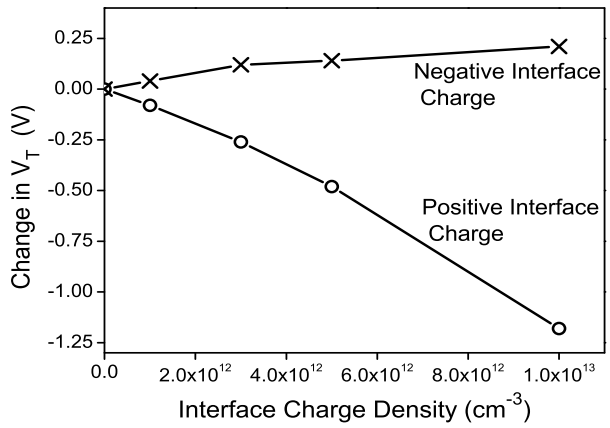


Fig. 7. Threshold voltage change for varying interface charge density. $V_{GS} = 0.5V$, $V_{DS} = 1V$

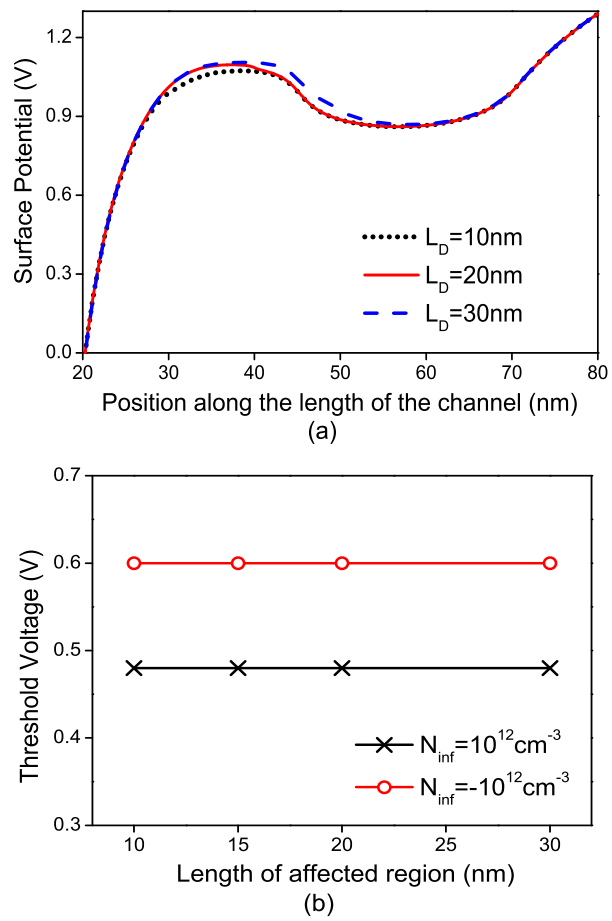


Fig. 8. (a) Potential profile for varying length of localized charge. Only the channel region has been shown. $N_{inf} = 10^{12} cm^{-3}$ (b) Threshold voltage for varying length of localized charge. $V_{DS} = 1V$

gate bias for which I_D reaches $10^{-7} A/\mu m$, which has been used in earlier works of TFET [15], [16]. ΔV_T represents the difference between V_T of fresh device and the damaged device. It can be seen that $|\Delta V_T|$ increases with increasing charge density. This happens because increasing N_{inf} changes the flat band voltage (V_{FB}). This in turn changes the effective gate voltage ($V_{GS} - V_{FB}$) resulting in a change in threshold voltage from the fresh device. Fig. 7 also reveals that this change is more prominent in case of positive interface charge.

Since in the DMDG TFET device, electric field is high mainly in source-body region; interface charges are more likely to be localized at that region. In MOSFET, length of the localized charges affect the overall characteristics of the device [17], [18]. To study the effect of localized charge, length of the damaged region (L_D) has been changed from $5nm$ to $30nm$ taking source edge as reference. $N_{inf} = 10^{12}cm^{-3}$ has been considered. Fig. 8 shows that although surface potential changes with changing L_D, V_T remains almost constant. This can be easily understood by closely examining the surface potential profile of Fig. 8(a). It reveals that changing L_D does not have any effect on the surface potential near source region. Increasing L_D only pulls up the potential at channel region. Since current in the DMDG TFET is dominated by carrier tunneling from source, not by the carrier transport in channel; interface charges away from the source end has negligible effect on tunneling current or threshold voltage of the device.

V. Conclusion

In this work, DMDG TFET with interface charges induced by hot carrier effect has been analyzed. It has been found that the presence of positive or negative interface charge modify the surface potential profile resulting in a change in the shortest tunneling length. This subsequently changes the tunneling current and also the threshold voltage of the device. The interface charges are mainly localized at the source-body junction and tunneling current or threshold voltage is not at all affected by the presence of interface charge anywhere except at the source end. The analysis of this work can be helpful in designing DMDG TFET based charge-trapped memory devices.

Acknowledgment

The authors would like to thank the Department of EEE, East West University, Bangladesh for the support with Atlas, Silvaco software.

References

[1] W. Y. Choi, B. -G. Park, J. D. Lee, and T. -J. K. Liu, "Tunneling field effect transistors (TFETs) with subthreshold swing (SS) less than 60 mV/dec," *IEEE Electron Device Lett.*, vol. 28, pp. 743-745, Aug. 2007.

[2] A. Mallik, and A. Chattopadhyay, "Tunnel Field-Effect Transistors for Analog/Mixed-Signal System-on-Chip Applications," *IEEE Trans. Electron Devices*, vol. 59, no. 4, pp. 888-894, Jan. 2012.

[3] M. S. Lundstrom, and D. A. Antoniadis, "Compact models and the physics of nanoscale FETs," *IEEE Trans. on Electron Devices*, vol. 61, no. 2, pp. 225-233, 2014.

[4] H. Wu, S. Qin, Y. Cai, Q. Huang, and R. Huang, "High-gate-injection tunneling field effect transistor for flash memory applications," in *Proc. ICSICT*, 2012, pp. 1-3.

[5] E. G. Ioannidis, A. Tsormpatzoglou, D. H. Tassis, C. A. Dimitriadis, G. Ghibaudo, and J. Jomaah, "Effect of localized interface charge on the threshold voltage of short-channel undoped symmetrical double-gate MOSFETs," *IEEE Trans. on Electron Devices*, vol. 58, no. 2, pp. 433-440, 2011.

[6] S. Saurabh, and M. J. Kumar, "Novel attributes of a dual material gate nanoscale tunnel field-effect transistor," *IEEE Trans. Electron Devices*, vol. 58, no. 2, pp. 404-410, 2011.

[7] R. Vishnoi, and M. J. Kumar, "Compact analytical model of dual material gate tunneling field effect transistor using interband tunneling and channel transport," *IEEE Trans. Electron Devices*, vol. 61, no. 6, pp. 1936-1942, 2014.

[8] P. Pandey, R. Vishnoi, and M. J. Kumar, "A full-range dual material gate tunnel field effect transistor drain current model considering both source and drain depletion region band-to-band tunneling," *J. Comput. Electron.*, vol. 14, no. 1, pp. 280-287, 2015.

[9] S. L. Noor, S. Safa, and M. Z. R. Khan, "Dual-material double-gate tunnel FET: gate threshold voltage modeling and extraction," *J. Comput. Electron*, 2016.

[10] S. L. Noor, S. Safa, and M. Z. R. Khan, "Two-dimensional model of Dual Metal double gate tunnel FETs," in *Proc. WIECON-ECE*, 2015, pp. 70-73

[11] ATLAS Device Simulation Software, Silvaco Int., Santa Clara, CA, USA, 2010.

[12] A. M. Ionescu, K. Boucart, K. E. Moselund and V. Pott, *Small Swing Switches*, Cambridge Univ. Press, 2016.

[13] A. M. Ionescu, and H. Riel, "Tunnel field-effect transistors as energy-efficient electronic switches," *Nature*, vol. 479, no. 7373, pp. 329-337, 2011.

[14] L. De Michielis, L. Lattanzio, K. E. Moselund, H. Riel, and A. M. Ionescu, "Tunneling and occupancy probabilities: How do they affect tunnel-FET behavior," *IEEE Electron Device Lett.*, vol. 34, no. 6, pp.726-728, 2013.

[15] K. Boucart, and A. M. Ionescu, "Double-gate tunnel FET with high-k gate dielectric," *IEEE Trans. Electron Devices*, vol. 54, no. 7, pp. 1725-1733, 2007.

[16] S. Saurabh, and M. J. Kumar, "Impact of Strain on Drain Current and Threshold Voltage of Nanoscale Double Gate Tunnel Field Effect Transistor: Theoretical Investigation and Analysis," *Jpn. J. Appl. Phys.*, vol. 48, no. 6, p. 064503, 2009.

[17] Y. S. Jean, and C. Y. Wu, "The threshold-voltage model of MOSFET devices with localized interface charge," *IEEE Trans. Electron Devices*, vol. 44, no. 3, pp. 441-447, 1997.

[18] T. K. Chiang, "A compact model for threshold voltage of surrounding-gate MOSFETs with localized interface trapped charges," *IEEE Trans. Electron Devices*, vol. 58, no. 2, pp. 567-57, 2011

Dy doped $BiFeO_3$: A Multiferroic with Bulk Structural and Ferroelectric Properties Comparable with Nano Counterparts

Sayeed Shafayet Chowdhury,[1*] Abu Hena Mostafa Kamal,[2] Bashir Ahmmad[3], and M. A. Basith[2]
[1]Department of Electrical and Electronic Engineering,
Bangladesh University of Engineering and Technology, Dhaka-1000, Bangladesh.
[2]Department of Physics, Bangladesh University of Engineering and Technology, Dhaka-1000, Bangladesh.
[3]Graduate School of Science and Engineering, Yamagata University, 4-3-16 Jonan, Yonezawa 992-8510, Japan.
*turjo11@gmail.com

***Abstract*—The structural and ferroelectric properties of Dy doped $BiFeO_3$ multiferroic materials are reported. The bulk polycrystalline $Bi_{0.9}Dy_{0.1}FeO_3$ samples were synthesized using solid state reaction technique and their nano counterparts were prepared using ultrasonic probe sonication. XRD patterns demonstrated significant improvement of phase purity due to Dy substitution both in bulk sample and corresponding nanoparticles. Again, FESEM imaging was carried out to analyze the surface morphology of bulk sample as well as to determine the size and distribution of the synthesized nanoparticles. Furthermore, leakage current density measurements were performed where, interestingly, we found comparable performance for $Bi_{0.9}Dy_{0.1}FeO_3$ bulk and its corresponding nanoparticles. Moreover, the polarization versus electric field measurements showed higher remanent polarization and coercive field for nanoparticles compared to their bulk samples which might be associated with leaky behavior generated from oxygen related defects of the nanoparticles. However, saturation polarization within the applied field range was observed for $Bi_{0.9}Dy_{0.1}FeO_3$ bulk only indicating inherent ferroelectric behavior. Overall, this investigation provides evidence of $Bi_{0.9}Dy_{0.1}FeO_3$ multiferroic having enhanced structural purity and ferroelectric properties over undoped $BiFeO_3$. In particular, it opens up the pathway of a material with equivalent bulk and nano leakage current performance which may be useful in many applications.**

***Index Terms*—Multiferroics, XRD, FESEM imaging, Leakage Current Density, P–E loop.**

I. Introduction

Multiferroics, which possess concurrent ferromagnetic, ferroelectric, and/or ferroelastic orderings, have been at the core of advanced functional and nanostructured materials research in recent times [1]–[3]. Among these, $BiFeO_3$ (BFO) has been an epitome due to its technological promise in the novel multifunctional devices [4]. Notably, the suitability of undoped BFO in many cases becomes limited because of its spiral modulated spin structure (SMSS) [5]. Moreover, the formation of different impurity phases [6] further deteriorates the properties of undoped BFO. Thus, to overcome these hindrances and to achieve superior multiferroic properties, several attempts have been made by substituting Bi and Fe in BFO with rare-earth and transition metal elements, respectively. The enhancements in magnetism and ferroelectricity of the parent material have been achieved due to substitution of Bi by ions such as Gd, La, Nd etc. or simultaneous substitution of Bi and Fe in BFO by ions such as La and Mn, La and Ti, Nd and Sc, Gd and Ti etc. , respectively, only a few to mention from numerous examples [7]–[10]. Although, these investigations reported enhanced ferroelectric/ferromagnetic properties, rigorous synthesis procedure had to be adopted to attain nanoparticles with very low leakage current densities. Comparable ferroelectric performance for bulk and their nano counterparts was yet to be obtained.

In the present investigation, we report the synthesis as well as structural and ferroelectric characterization of Dy doped BFO, $Bi_{0.9}Dy_{0.1}FeO_3$. The objective of performing partial substitution of Bi^{3+} ions by Dy^{3+} ions is twofold: (i) to suppress the impurity phases of BFO and thus, (ii) to reduce leakage current density to obtain alike ferroelectric properties of bulk sample and nano counterparts. For comparison, we also prepared undoped BFO samples. The structural properties were investigated using X-ray diffraction (XRD) patterns. Following this, the morphological study was carried out using field emission scanning electron microscope (FESEM) imaging to determine surface morphology and particle size. Then, we performed polarization versus electric field hysteresis loops as well as leakage current density measurements with a view to obtaining electrical properties of the prepared samples. Our results indicate formation of highly phase pure $Bi_{0.9}Dy_{0.1}FeO_3$ materials and also formation of nanoparticles was confirmed from the FESEM images. However, the most interesting properties were observed from the leakage current density experimentation which demonstrated that bulk Dy doped BFO and their corresponding nanoparticles provided almost equivalent leaky behavior. In addition, the saturation polarization was observed only for Dy doped bulk BFO sample.

II. EXPERIMENTAL DETAILS

The polycrystalline samples with compositions of undoped $BiFeO_3$ (referred to as BFO) and Dy doped $Bi_{0.9}Dy_{0.1}FeO_3$ (referred to as BDFO) were synthesized using standard solid state reaction technique [10]. Then the nanoparticles were fabricated directly from their bulk powder materials using the ultrasonic probe dispersion technique (direct sonication) with varying amount of sonication time by the application of ultrasonic energy generated from a probe [11]. For the synthesis of nanoparticles, a portion of the micro-meter sized powder was subsequently mixed with isopropanol. We then put the mixtures of isopropanol and powder with a mass per-

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

centage of ∼0.5% [11] into an ultrasonic probe and sonicated for 30, 60 and 90 minutes respectively, for three different samples. X-ray diffraction (XRD) analysis was performed using a diffractometer (Rigaku Smart Lab) with CuK_α (λ = 1.5418 $\AA$) radiation to study the crystalline structures of the as synthesized materials. To measure electrical properties, pellets were prepared by pressing the powders using a hydraulic press and annealing at 750°C with high heating rate (20°C/min). A ferroelectric loop tracer in conjunction with an external amplifier (10 kV) was used to trace the leakage current density and ferroelectric polarization of the pellet shaped samples.

III. RESULTS AND DISCUSSION

A. Structural Characterization

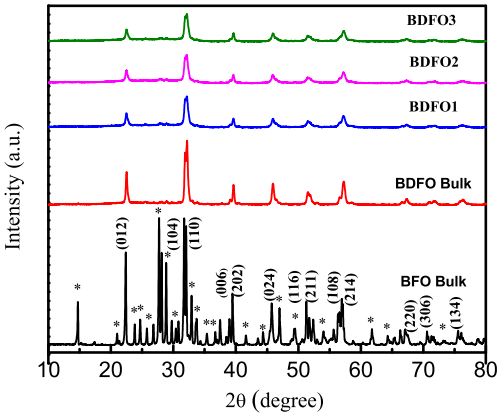


Fig. 1. XRD patterns of BFO bulk, BDFO bulk and its corresponding nanoparticles sonicated for 30, 60 and 90 minutes respectively. The asterisk marks in BFO bulk patterns represent impurity phases.

Figure 1 shows the XRD patterns of the BFO bulk, BDFO bulk and its nano counterparts respectively. We denote the BDFO nanoparticles sonicated for 30, 60 and 90 minutes by BDFO1, BDFO2 and BDFO3 respectively. As we can clearly see from Fig. 1, the bulk BFO contains significant impurity phases which will affect its multiferroicity. However, the substitution of Bi with Dy enabled the removal of most of these unwarranted phases in BDFO. The BDFO bulk comprises mostly of the characteristic phases with negligible impurities present. Again, when these BDFO bulk samples were sonicated for various amount of time, we did not observe any further significant suppression of the impurity phases as evidenced by Fig. 1. This clearly provides an indication of identical phase purity of the bulk samples and their nanoparticles. However, in case of our direct sonication probe dispersion technique [12], we observed a tendency of conversion from crystalline to amorphous phase indicated by the gradual decrease of the intensity of (104) and (110) peaks with increasing sonication time as shown in Fig. 2. Interestingly, for ultrasonication bath technique this type of problem was not found [11]. In bath method, only slight amorphization is observed mainly due to indirect interaction of the waves with the particles. The dead layer formation due to grinding for long periods might also be attributed to such phase conversion of the ultrasonically prepared nanoparticles using probe technique. Moreover, the diffraction (104) and (110) peaks of BDFO bulk as well as nanoparticles shifted to

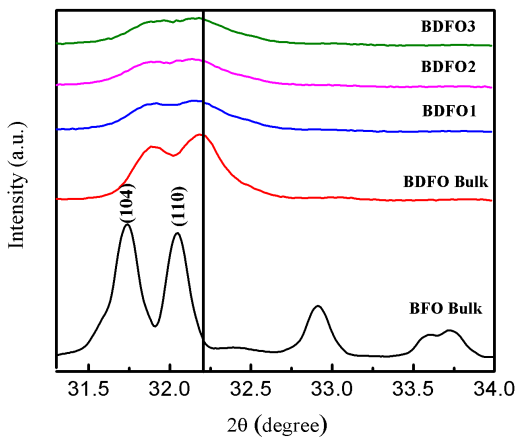


Fig. 2. Enlarged view of the XRD patterns of BFO bulk, BDFO bulk and its corresponding nanoparticles sonicated for 30, 60 and 90 minutes respectively around the peaks at (104) and (110). The vertical line is added to highlight the peak shifting from BFO to BDFO samples.

TABLE I
CALCULATED STRUCTURAL PARAMETERS FROM XRD OF BDFO BULK AND NANOPARTICLES

Sample specification	FWHM (radians)	Particle size (nm)	Microstrain
Bulk	0.630		0.0226
BDFO1	0.704	35	0.0247
BDFO2	0.694	28	0.0363
BDFO3	0.694	24	0.0380

higher angle as compared with that of BFO bulk indicating the substitution of Bi^{3+} ions by Dy^{3+} ions which has affected the original structure of BFO. Another indication of amorphization is peak broadening. To investigate this effect, the full width at half maxima (FWHM) values of the peaks at (104) and (110) are presented in Table I. While performing this calculation, instrumental peak broadening factor was taken into consideration. For this purpose, the XRD peaks of amorphous Al_2O_3 were analyzed. We calculated the corresponding FWHM of the peaks of Al_2O_3 and thus adjusted the FWHM of the peaks of our samples accordingly. The results demonstrate the broadening of the peaks due to Dy substitution. The broad and diffuse peaks once again indicate crystalline to amorphous like phase transition. From here, we may conclude that crystallinity of the samples gets affected while formation of nanoparticles. However, the improvement in terms of phase purity from BDFO bulk to nano is not remarkable. Thus, the bulk samples provide satisfactory impurity phase reduction while preserving the crystalline structural attributes.

To further explore the structural properties of the samples, we calculated the particle size using modified Scherrer equation [13] from the XRD peaks as given in Table I. As expected, the size decreases with sonication time. Again, we calculated the microstrain of the synthesized samples. From the values of Table I, we infer that with increasing sonication period, the strain gradually increases. This once again reinforces the distortion in crystallinity induced by energy impact from ultrasonication for a longer duration.

B. Morphological Study

To investigate the microstructure and the surface morphology of the BDFO bulk pellets and nanoparticles, FESEM

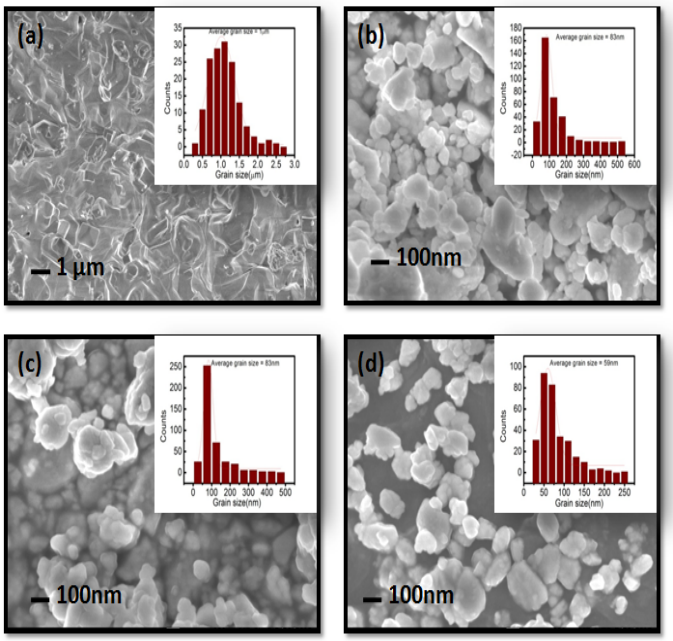


Fig. 3. FESEM images and their corresponding histograms of (a)BDFO, (b) BDFO1, (c) BDFO2 and (d) BDFO3 samples.

imaging was carried out. In Fig. 3, FESEM images and their corresponding histograms of the BDFO, BDFO1, BDFO2 and BDFO3, respectively are shown. The average particle sizes calculated from the histograms for BDFO, BDFO1, BDFO2 and BDFO3 are 1000, 83, 78 and 59 nm, respectively. It is observed that the average particle size decreases significantly with increase of sonication time. Previous investigation showed the particle size of the BFO to be 5μm$\sim$15μm [10]. However, from the FESEM images, average grain size of BDFO bulk pellet was 1000 nm. The reduction of the particle size may be attributed to the addition of Dy ions with smaller ionic radius (0.912 Å) than Bi (1.03 Å) into BFO [14]. The ultrasonic process reduced the particle size of synthesized nanoparticles further. This is due to the fact that ultrasound irradiation generates many localized hot spots in the solution with the particles [15] and during the process, the implosive collapse of the bubble causes an inward rush of liquid known as 'microstreaming' in which high velocity is produced. Due to the high velocity, the bubbles bombard the agglomerations with high momentum producing splitting and causing rapid dispersion.

In our previous investigation [16], high quality monodisperse Gd and Ti co-doped $BiFeO_3$ nanoparticles were produced using ultrasonic energy through bath (indirect process) directly from their bulk powder materials. Noticeably, in our current investigation for ultrasonic probe dispersion (direct process), 30 minutes are enough to produce nanoparticles within the range of 80 nm to 90 nm, whereas, within the same timeframe, bath sonication/indirect process is able to produce 150 to 250 nm grains only. Although the doping elements in this investigation and those of Ref. [16] are different, we get an idea of the scale of particle size reduction for the two ultrasonication procedures from these two cases. Since for the direct process, lesser time is required to produce the same amount of decline in particle size, we may infer that for ultrasonic probe dispersion (direct process), higher mechanical energy is bombarded at the particles from the ultrasonic wave in direct process compared to bath technique.

It has been suggested that the grain growth depends upon the concentration of oxygen vacancies and diffusion rate of the ions [17] . Due to highly volatile nature of Bi, its evaporation generates large number of oxygen vacancies in pure BFO. This makes it easier for the ions to diffuse, resulting in a very large grain size as compared to the Dy doped BFO samples. But, when the Dy ion is substituted within BFO at Bi site, we anticipate reduction in oxygen vacancies since some amount of volatile Bi has been substituted. If indeed oxygen vacancies are reduced, it will result in reduction in the leakage current density. So, in our next section, we perform leakage current density as well as polarization measurements.

C. Electrical Measurements

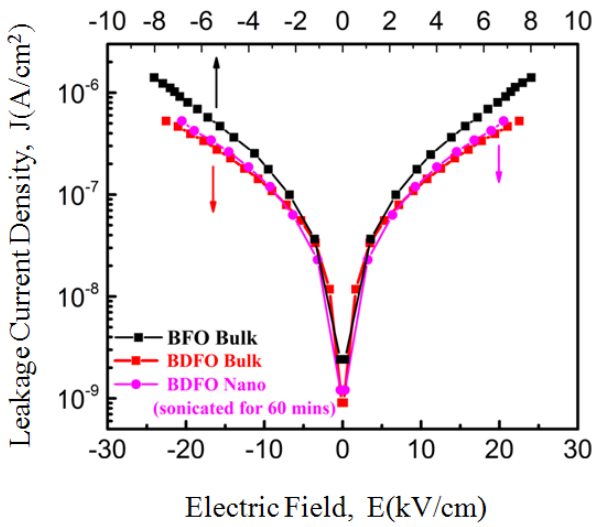


Fig. 4. Leakage current density of BFO, BDFO bulk materials and their nanoparticles sonicated for 60 minutes as a function of applied electric field.

To compare the electrical properties of BFO and BDFO bulk along with corresponding nanoparticles, leakage current density, J versus electric field, E measurements were performed for an applied field of up to $\pm$20 kV/cm as shown in Fig. 4. It is clear that leakage current density reduced from BFO to BDFO bulk. However, we observed no significant difference between leakage current densities of BDFO and BDFO2. This phenomenon may be explicated from the fact that very few impurity phases were present in BDFO samples as shown in the XRD patterns in Fig. 1. Thus, further sonication to form nanoparticles did not enhance the leakage current property a great deal. This presents a very notable result as bulk and nano samples provide comparable performance in this case. This could be very promising for applications with low leakage current requirements where we no longer need to go for nanoparticles. Rather the bulk samples would provide us equally acceptable solution. Although the leakage current densities of BDFO bulk and their nanoparticles are identical, still they are much lower than the parent BFO ceramics. As a result, the P-E loops of BDFO bulk materials and their corresponding nanoparticles are expected to exhibit enhanced ferroelectric properties. To confirm this, next we investigate the polarization versus electric field (P–E) loop measurements.

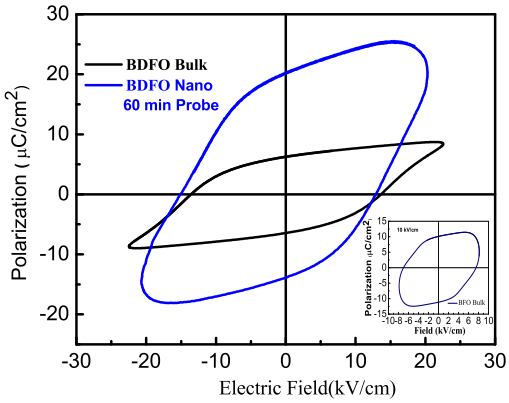


Fig. 5. Room temperature polarization vs. electric field loops of BDFO bulk materials and their nanoparticles prepared by ultrasonication sonicated for 60 minutes. The inset shows the P–E loop of BFO bulk for comparison.

Figure 5 shows the P–E loops for BDFO and BDFO2 carried out at room temperature with a frequency of 50 Hz. The P–E loop of bulk BFO is also given in the inset for comparison. We found that the remanent polarization (P_r) and the coercive field (E_c) of BDFO2 are higher compared to that of bulk BDFO. But, the electrical resistivity (ρ) of the nanoparticles reduced on the contrary. P_r for BDFO was 6.24 μC/cm^2 while for nano it increased to 16.98μC/cm^2. The larger polarization in BDFO nanoparticles may be associated with higher leakage current, reduced electrical resistivity and hybridization of the 6s^2 lone pair electrons of Bi^{3+} with the 2s/2p orbitals of O_2 [14]. In bulk counterparts, P_r is lower, however, the loop is nearly saturated. The P-E hysteresis loops for BFO multiferroics are often dominated by leakage current affected by mixed valence of Fe^{2+} and Fe^{3+} ions, or by oxygen vacancy related defects, or by both [18]. The interpretation of ferroelectric behavior of materials with such P-E loops can often be misleading which may lead to erroneous conclusions. It has been asserted in Ref. [19] that true ferroelectrics exhibit saturation in polarization and have concave region in P-E loop. We find that BDFO bulk provides saturated polarization while BFO bulk and BDFO2 do not, indicating more reliable ferroelectric characteristics for BDFO bulk. The oxygen vacancies due to volatility of Bi and hopping of itinerant electrons between Fe^{2+} and Fe^{3+} ions result in high conductivity and leakage current in undoped BFO. Hence replacement of highly volatile Bi with Dy reduces the concentration of oxygen vacancies giving reduced leakage current and improved P-E loop.

IV. Conclusions

The structural and ferroelectric characterizations of BDFO bulk and corresponding nanoparticles were performed in this investigation. XRD results indicated that BDFO bulk and their nano counterparts possess similar phase purity. However, the nano samples suffer from amorphization. Thus, a bulk material like BDFO bulk would be more suitable in many applications requiring negligible impurity phases and better crystallinity. Moreover, the experimental difficulties associated with synthesis of nanoparticles as well as degradation of surface morphology are evaded too in case of suitability of BDFO bulk samples. In addition, FESEM images were used to ensure formation of desired bulk and nano states and to determine their sizes. Leakage current density measurements showed that BDFO bulk samples had even lower leakage current compared to their corresponding nanoparticles. This further asserts our idea of BDFO being a multiferroic with comparable or even superior bulk properties in comparison to nano. In our future investigations, we intend to explore the magnetic properties of these samples where we anticipate identical improved properties of BDFO bulk samples compared to their nano counterparts.

Acknowledgment

This work was financially supported by Infrastructure Development Company Limited (IDCOL), Dhaka, Bangladesh.

References

[1] N. A. Spaldin, and M. Fiebig, "The renaissance of magnetoelectric multiferroics," *Science*, vol. 309, pp. 391–392, 2005.
[2] W. Eerenstein, N. D. Mathur, and J. F. Scott, "Multiferroic and magnetoelectric materials," *Nature*, vol. 442, pp. 759–765, 2006.
[3] R. Ramesh, "Materials science: Emerging routes to multiferroics," *Nature*, vol. 461, pp. 1218–1219, 2009.
[4] G. S. Lotey, N. K. Verma, "Structural, magnetic, and electrical properties of Gd-doped $BiFeO_3$ nanoparticles with reduced particle size," *J. Nanopart Res.*, vol. 14, pp. 742, 2012.
[5] P. Fischer, M. Polomska, I. Sosnowska, and M. Szymanski, "Temperature dependence of the crystal and magnetic structures of $BiFeO_3$," *J. Phys. C: Solid State Phys.*, vol. 13, pp. 1931, 1980.
[6] M. M. Kumar, V. R. Palkar, K. Srinivas, S. V. Suyanarayana, "Ferroelectricity in a pure $BiFeO_3$ ceramic," *Appl. Phys. Lett.*, vol. 76, pp. 2764–2766, 2000.
[7] G. Anjum, R. Kumar, S. Mollah, D. K. Shukla, S. Kumar, and C. G. Lee, "Structural, dielectric, and magnetic properties of $La_{0.8}Bi_{0.2}Fe_{1-x}Mn_xO_3$ (0.0≤x≤0.4) multiferroics," *J. App. Phys.*, vol. 107, pp. 103916, 2010.
[8] R. A. Agarwal, S. S. Ashima, and N. Ahlawat, "Structural transformation and improved dielectric and magnetic properties in Ti-substituted $Bi_{0.8}La_{0.2}FeO_3$ multiferroics," *J. Phys. D: Appl. Phys.* , vol. 45, pp. 165001, 2012.
[9] T. D. Rao, and S. Asthana, "Evidence of improved ferroelectric phase stabilization in Nd and Sc co-substituted $BiFeO_3$," *J. Appl. Phys.*, vol. 116, pp. 164102, 2014.
[10] M. A. Basith, O. Kurni, M. S. Alam, B. L. Sinha, and B. Ahmmad, "Room temperature dielectric and magnetic properties of Gd and Ti co-doped $BiFeO_3$ ceramics," *J. Appl. Phys.*, vol. 115, pp. 024102, 2014.
[11] M. A. Basith, D.-T. Ngo, A. Quader, M. A. Rahman, B. L. Sinha, B. Ahmmad, F. Hirose and K. Molhave, "Simple top-down preparation of magnetic $Bi_{0.9}Gd_{0.1}Fe_{1-x}Ti_xO_3$ nanoparticles by ultrasonication of multiferroic bulk material," *Nanoscale*, vol. 6, pp. 14336, 2014.
[12] H. Zhang, L. Wang, Y. Chen, D. Wang, Y. Yingbang, "Observation of room temperature saturated ferroelectric polarization in Dy substituted $BiFeO_3$ ceramics," *J. Appl. Phys.*, vol. 111, pp. 074105, 2012.
[13] A. Monshi, M. R. Foroughi, and M. R. Monshi, "Modified Scherrer Equation to Estimate More Accurately Nano-Crystallite Size Using XRD," *World Journal of Nano Science and Engineering*, vol. 2, pp. 154–160, 2012.
[14] M. Muneeswaran, and N. V. Giridharan, "Effect of Dy-substitution on the structural, vibrational, and multiferroic properties of $BiFeO_3$ nanoparticles," *J. Appl. Phys.*, vol. 115, pp. 214109, 2014.
[15] F. Liang, L. Jian, J. Sheng, Z. Fengang, D. Wen, and S. Mingrong, "Experimental and theoretical evidence of enhanced ferromagnetism in sonochemical synthesized $BiFeO_3$ nanoparticles," *Applied Physics Letters*, vol. 97, pp. 242501, 2010.
[16] M. A. Basith, F. A. Khan, B. Ahmmad, S. Kubota, F. Hirose, D.-T. Ngo, Q. H. Than, and K. Molhave, "Tunable exchange bias effect in magnetic $Bi_{0.9}Gd_{0.1}Fe_{0.9}Ti_{0.1}O_3$ nanoparticles at temperatures up to 250 K," *J. Appl. Phys.*, vol. 118, pp. 023901, 2015.
[17] K. Brinkman, T. Iijima, K. Nishida, T. Katoda, and H. Funakubo, "The influence of acceptor doping on the structure and electrical properties of solgel derived $BiFeO_3$ thin films," *Ferroelectrics*, vol. 357, pp. 35–40, 2007.
[18] S. S. Ali, W. J. Li, K. Javed, D. W. Shi, S. Riaz, Y. Liu, Y. G. Zhao, G. J. Zhai, and X. F. Han, "Utilizing the anti-ferromagnetic functionality of multiferroic shell to study exchange bias in hybrid core-shell nanostructures," *Nanoscale*, vol. 7(32), pp. 13398–13403, 2015.
[19] J. F. Scott, "Ferroelectrics go bananas," *Journal of Physics: Condensed Matter*, vol. 20(2), pp. 021001, 2007.

Incorporating Advanced Power Electronic Concepts with Hands-on Emphasis in a Low-cost Power Engineering Curriculum

M.M. A. Rahman

Electrical Engineering
Grand Valley State University
Grand Rapids, Michigan, USA 49504

rahmana@gvsu.edu

Abstract—**The contents of a advanced power electronic course have been developed for a power engineering curriculum that is evolving despite of having limited resources. These course contents can be segmented and inserted into junior, senior, and graduate level courses to benefit students. Some of the techniques involve hands-on teaching using project-based approach. Some of the project outputs will be presented in this paper. Even though the currently available resources do not allow administering new courses this approach will help students learn advanced topics practiced in power industries and taught in resource-rich academic institutes.**

***Index Terms*- Power electronics, education, hands-on.**

I. Introduction

The concept of power electronics technology has evolved from simple and small consumer electronics such as battery chargers to complex and integrated systems such as grid connected renewable energy sources. The newest generation of power electronics based systems brings with it a new generation of design challenges as well. They must be smaller in size, lighter, more efficient (i.e., take in less but give out more power compared to current systems), cooler without a cooling fan (i.e., produce minimal or no heat), minimal or no interference (EMI) to communication systems, etc. The semiconductor industry has retorted by inventing new power devices, microcontrollers, and accessories that are smaller in size and have higher capability to meet these design challenges. As shown in Figure 1, advanced devices (e.g. COOLMOS, wide-bandgap MOSFET) increase efficiency by reducing losses (low conduction loss due to low on-resistance, $R_{ds(on)}$ in milli-Ohm range) when compared to standard power devices (e.g. MOSFET). Now, our resources (faculty expertise and curriculum) should reciprocate by ensuring that students can model, design, and implement systems utilizing this contribution from industry.

In order to meet the above mentioned design challenges and keep up with modern semiconductor technology, it is desirable to reduce the size of the power electronic systems by reducing the size of other components, especially reactive components, such as inductor, capacitor, and transformer. As a result, high frequency power conversion that increases power density (Watts per meter^3) has attracted much interest [1-5] in recent years. However, with the increase in switching frequency the switching power losses increase due to the cross-over of voltage across switch, v_T and current through the switch, i_T during transitions as depicted in Figure 2. This is known as hard-switching which causes significant EMI issues in addition to switching power loss. Hence, an efficient and high-density power electronic system needs to minimize switching power losses by soft-switching [1-5] technologies (zero-voltage switching ZVS, zero-current switching ZCS, zero-voltage transition ZVT, etc.). In these techniques cross-over of voltage across switch, v_T with the current through it, i_T during transition is eliminated. This reduces switching loss contributing to higher efficiency. Another predicament of high frequency switching is utility pollutions (generation of unwanted harmonics, electro-magnetic interference, etc.) when power is supplied from commonly available AC source such as wall outlet. This must be minimized by using active power factor correction (PFC) technology [6] for AC-to-DC converters. A recent trend to increase efficiency further is to replace diodes (having a higher on-state conduction power loss, P_{cond}) in power converters with advanced MOSFETs (having lower conduction power loss) called synchronous rectifiers. This topology is mainly available only for DC-to-DC converters [9], not for AC-to-DC power converters (especially for the single-stage AC-to-DC converter topology which is more efficient but requires more advanced and complex switching scheme).

Properly overcoming the above mentioned shortcomings by implementing appropriate technology is a design challenge especially under today's strict performance constraints. The majority of the research work done in this area deals with some of the shortcomings separately, not all of them in a single topology. Most of the synchronous rectification work is done on DC-to-DC converters, not on AC-to-DC converters. Therefore, this curriculum development work will take on all of these challenges together leading to power engineering education on the design and development of DC-toDC and AC-to-DC converter topologies that incorporates several advanced concepts such as: (i) active power factor correction, (ii) AC-to-DC power conversion, (iii) application of advanced devices, (iv) soft-switching techniques, and (v) synchronous rectification.

This paper is presented as outlined as follows: course objectives and contents are described in Section II. Section III presents enrolled history in power engineering courses and means to optimize delivery of course materials for better student learning. Some examples of project-based delivery

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

are given in Section IV. Section V presents some student feedbacks and Section VI concludes this paper.

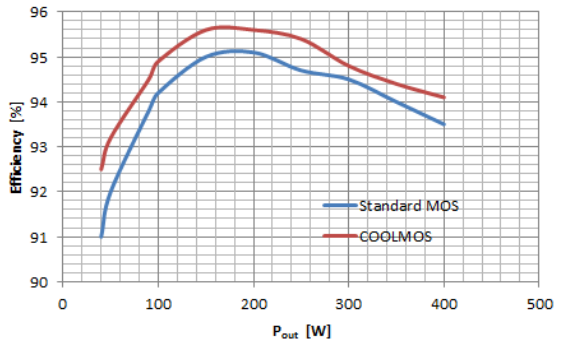


Fig. 1 Efficiency of recent vs standard devices in hard-switching [7].

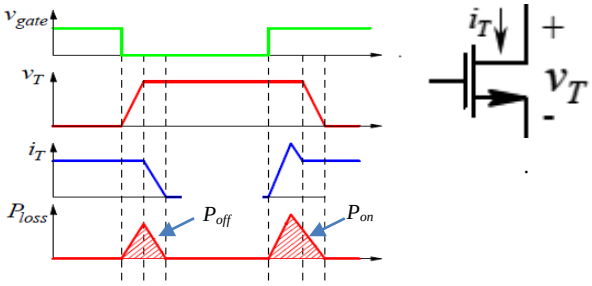


Fig. 2 Switching power loss, P_{loss} in hard-switching technology

II. Design of the Course

This course is designed to serve students following the Electrical Energy content area in the Masters of Science in engineering program (Electrical and Computer Engineering emphasis). This course is also designed to serve junior and senior Electrical Engineering students who wish to take this course as an elective outside of their content area. It emphasizes student learning through tradition lecture and discussions, hands-on laboratory activities, and sponsored design projects resolving real life engineering problems. In completing this course, students will be able to: (i) model recent power converter topologies, (ii) analyze and design power converters, (ii) model power systems using renewable energy sources, (iv) design dc transmission and distribution systems, and (v) describe the challenges and remedies associated with utility pollution and stability. These objectives also satisfies some of the criteria set by ABET [8]. To fulfill these objectives following topics are included in this course: Active power factor correction, interleaving techniques, resonant converters and applications, soft-switching techniques and applications, bi-directional converters, application of power electronics in renewable energy systems, high voltage dc transmission system, dc distribution system, dc-to-ac inverters and applications, static VAR compensators and grid stability.

III. Course Enrolment

Power Engineering is a relatively new curriculum in our relatively new but evolving engineering program. The very first course in our power engineering curriculum was added in Summer 2005 starting with EGR 330 Power Systems Analysis being the first course. Based on this course several other power engineering courses were gradually developed such as: EGR 430/530 Electro-mechanics, EGR 655 Power Electronics, EGR 656 Electrical Drive Systems, and EGR 665 Advanced Power Electronics. In line with the tradition in academia, courses with higher course number contain more advanced level of contents. Figure 3 shows the enrolment patterns in these courses over the last 11 years starting since 2005. Even though EGR 665 was developed and approved by university graduate council few years back, there was not enough enrolment to formally offer it in classroom setting. However, the materials of this course are very relevant to current industry practices and taught in most research-I universities. Hence efforts were made to teach as much materials as possible in lower level courses using hands-on approach. Table I shows how some of the materials can be distributed and taught in other courses. It should be noted here that this not the perfect solution and must be co-ordinated with laboratory activities to accommodate in the semester schedule in an appropriate and timely manner. Active power factor correction can be taught right after passive power factor correction traditionally taught in first power system analysis course (EGR 330). However, students need to have the fundamental understanding of boost regulators to conceive this topic. High voltage (HV) transmission and distribution systems can be taught when students learned about transformer applications in EGR 430. Resonant converters and soft-switching technologies can be taught in the first power electronics course in design project

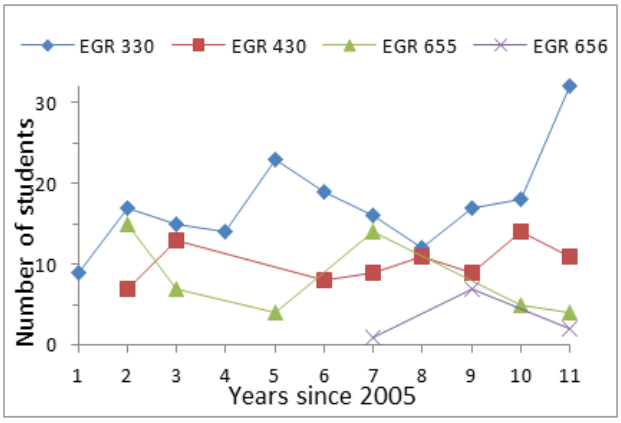


Fig. 3 Enrolment in Power Engineering courses during last 11 years

TABLE I
Course Content Distribution

Materials	Taught in courses			
	EGR 330	EGR 430	EGR 655	EGR 656
Active power factor correction	X			
HV dc transmission and distribution		X		
Resonant and soft-switching converters			X	
Interleaving techniques			X	X
Renewable energy applications				X

component of the course. Interleaving technique is an effective method to reduce harmonic contents and can be taught in power electronics and drive systems courses. Application of power electronics is an integral part of renewable energy systems and can be taught along with slip-energy recovery systems in electrical drive systems course.

IV. Project Based Delivery

Both solid scientific knowledge and rich hands-on experience are essential to teaching effectiveness in power engineering courses since students learn by passively hearing and actively relating to their earlier experiences [11-20]. This can be accomplished in traditional class lectures and

interactive laboratory activities. Project based teaching in laboratory setting is a dynamic approach and has become an effective method of engineering education since last decades. This approach of knowledge delivery has been adopted to teach advanced materials when the cost of offering new course is prohibitive due to low enrolment and our limited resources. This section describes a list of projects can be administered and some of projects already implemented in our program.

A. Course Projects

Course project is an essential component of course delivery in our engineering program. We administer project based teaching at all levels of our program as shown in Table II with varying levels of rigor and emphasis. Many of these projects are sponsored by heavily surrounded local industries in this area. Students who are unable to attract industry sponsors are provided with a list of tentative projects as shown in Table III. Some of the projects (Figure 4) are also sponsored by our faculty teaching and learning center.

TABLE IIII
TENTATIVE PROJECTS LIST

	Project titles
1	Solar charger for different types of smart phones
2	Multi-port battery charging station
3	Dimmable power LED string with Bluetooth control
4	Electric motor fault detection and prognostics
5	AC-to-AC converter for aerospace applications
6	Inductor coupled Synchronous Rectifier Based Wireless battery charger

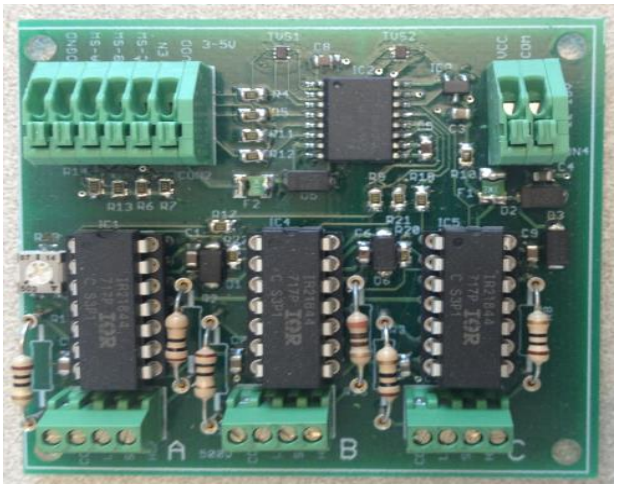


Fig. 4 A 3-phase MOSFET driver implemented in PCB.

B. Active Power Factor Correction using DCM

In this project a boost-integrated, single-phase, pulse width modulated (PWM), fixed frequency ac-to-dc boost converter with active power factor correction was implemented. Here the boost inductor was operating in the discontinuous current mode (DCM) to ensure natural power factor correction. PWM control is used to control the output voltage at different load conditions and it is implemented using ATMEGA8 micro-controller (µC) from Atmel. A 110 V, 60 Hz to 150 V DC, 450 W, 100 kHz ac-to-dc converter was designed and implemented in laboratory. A laboratory prototype has been developed and tested. Some prototype experimental results acquired to verify the DCM operation and high power factor is shown in Figure 5.

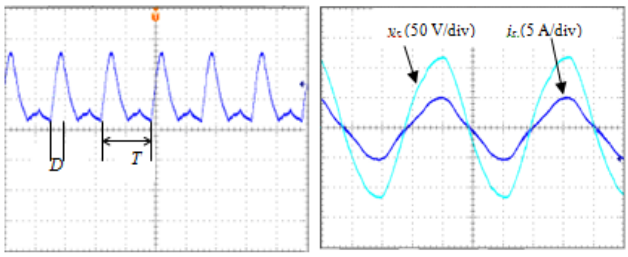


Fig. 5 Boost inductor current (left) and source voltage and current (right).

C. The Hertz Factor

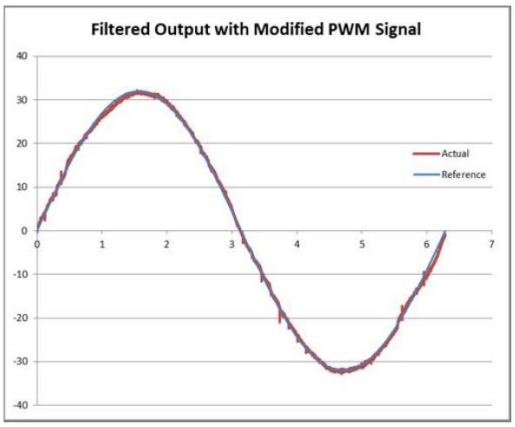
AC power used on aircraft is typically at a frequency of 400 Hz. However, available wall power in the United States is at 60 Hz. This presents a lot of difficulties in testing aircraft equipment in a lab. This project provides a conversion from 60 Hz wall power to 400 Hz aircraft power for use in lab testing. An H-Bridge implementation was selected, using a microcontroller to provide 2 Pulse Width Modulated (PWM) signals. The hardware design can be separated into 4 areas: 170 VDC regulation, 12 VDC regulation, PWM generation, and H-Bridge output. To obtain 170 VDC, a full-bridge rectifier was used with a capacitor to filter the rectified power. To obtain 12 VDC, a step-down transformer was used, followed by a full-bridge rectifier and filter capacitor. To ensure a steady 12 V output, a linear regulator and filter capacitor were also used. The PWM signals were generated using the Arduino Nano microcontroller board. Two digital output pins were used, one for each side of the H-Bridge. The H-Bridge output consisted of 4 MOSFETs, 2 on each side of the load. A MOSFET driver (Fig. 4) was used for each side of the H-Bridge, using a single digital signal to drive the high and low MOSFETs on each side. An LC low-pass filter was used on the output to turn the PWM signal into a sine wave.

D. VSI Using Car Battery

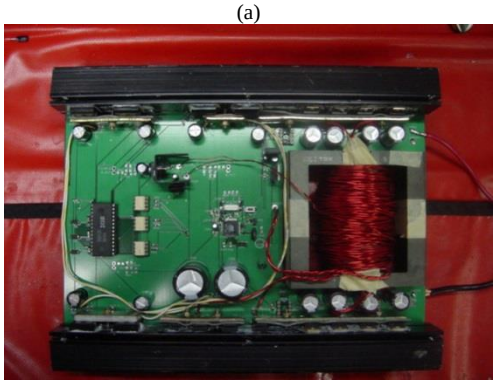
In this power converter project a three-phase, ac-to-dc voltage source inverter (VSI) running on a car battery for uninterruptible power supply applications was developed. Some of the features of this VSI includes: high frequency transformer isolation, micro-controller based control circuitry, and application of recent switching devices. The developed converter produces a 3-phase variable frequency sinusoidal ac output voltage (converted from a 12 V dc voltage source) while output voltage magnitude and frequency can both be controlled digitally using a micro-controller. A 3-leg full bridge converter topology with sinusoidal-triangle pulse width modulation (SPWM) was used. This converter used a switching frequency of 50-200 kHz and produced 3-phase 200 V at 60 Hz for testing purpose. It was also tested to run a variable frequency drive application. A output voltage waveform and picture of the finished product is given in Figure 6.

TABLE II
COMPARISON OF PROJECT ATTRIBUTES [10]

	Freshman Year	Sophomore Year	Junior Year	Senior Year
Duration	Multiple, small projects	1 half-semester project	1 whole-semester project	One two-semester project
Mentorship	Faculty	Faculty	Faculty, some industry	Industry, some faculty
Design Challenge	Low	Medium	High	Very High
Diversity of Projects	All projects the same	All projects the same	All projects different	All projects different
Proj. Failure impact	Low	Medium	High	No graduation



(a)



(b)

Fig. 6 VSI (a) output voltage and (b) developed prototype

V. Student Feedback And Lessons Learned

The administered project based delivery scheme specifically developed the students' intuition for modern power electronics. Students were able to intelligently integrate their intuition and skills to develop the project prototype. Semester-end student feedbacks on this delivery scheme were generally positive with "not enough time" and "not enough power courses" being the most frequent complaints. The positive comments are very encouraging as they testify that the goals of this non-traditional course design are met. Some of the comments are: (i) "This project provided an excellent platform for future power-centered projects, as it allowed hands on experience in many intricate circuit design areas." (ii) "Learning about the complexities of switching high current at high frequencies, filtering signals, and loading effects was very enjoyable and had a gaining effect on the possessed power systems knowledge of our group." (iii) "I now have a greater respect for those who design power system related products including inverters, amplifiers, and the multitude of other convenience enhancing products that are on the market to date." (iv) "This project provided a great learning experience and a window into the power engineering world." (v) "The course project clarified many things that were learned in the lecture sessions which were very helpful." (vi) "Needs more electives in Power Engineering, because this is one of the big industries and is applicable to many industries."

VI. Conclusions

This paper presented how advanced power electronic concepts can be taught in a low-cost power engineering curriculum. Contents of an advanced course were taught in multiple courses using hands-on project based teaching method. This non-traditional pedagogical scheme satisfies the students, helps them understand advanced power electronics concepts effectively, and promotes a wide variety of competences. We believe that with a better time management and control of student workloads, this scheme could fill an important gap modernize our curriculum and help better educational development of our students.

References

[1] N. Altintas, A.F. Bakan, and I. Aksoy, "A Novel ZVT-ZCT PWM Boost Converter," *IEEE Trans. on Power Electroics*, vol. 29, no. 1, pp. 256–265, Jan. 2014.

[2] J. M. Burkhart, R. Korsunsky, and D. J. Perreault, "Design methodology for a very high frequency resonant boost converter," *IEEE Trans. on Power Electroics*, vol. 28, no. 4, pp. 1929–1937, Apr. 2013.

[3] M. Pahlevaninezhad, P. Das, J. Drobnik, P. K. Jain, and A. Bakhshai, "A ZVS interleaved boost AC/DC converter used in plug-in electric vehicles," *IEEE Trans. Power Electron*., vol. 27, no. 8, pp. 3513–3529, Aug. 2012.

[4] F. Nejabatkhah, S. Danyali, S. H. Hosseini, M. Sabahi, and S. M. Niapour, "Modeling and control of a new three-input dc-dc boost converter for hybrid PV/FC/battery power system," *IEEE Trans. Power Electron.*, vol. 27, no. 5, pp. 2309–2324, May. 2012.

[5] T. F.Wu, Y. D. Chang, C. H. Chang, and J. G Yang, "Soft-switching Boost Converter with a Flyback Snubber for High Power Applications," *IEEE Trans. Power Electron*., vol. 27, no. 3, pp. 1108–1119, Mar. 2012.

[6] P. F. Melo, R. Gules, E. F. R. Romaneli, and R. C. Annunziato, "A Modified SEPIC Converter for High-power-factor Rectifier and Universal Input Voltage Applications," *IEEE Trans. Power Electron*., vol. 25, no. 2, pp. 310–321, Feb. 2010.

[7] Super-junction MOSFET datasheet from www.infineon.com

[8] Accreditation board for engineering and technology www.abet.org

[9] E. Adib and H. Faranehfard, "Zero-Voltage-Transition PWM Converter with Synchronous Rectifier", *IEEE Trans. Power Electronics*, vol. 25, no. 2, pp. 105-110, 2010.

[10] M.M.A Rahman and *et.al.*; "A Project-Based Approach to Teach Introductory Circuit Analysis," *ASEE/IEEE Frontier in Education Conference* , NY, 2008.

[11] M. Jalali, J. Marti, A. Kirchhoff, F. Lawrenz, and S. Campbell, "A Low-cost Hands-on Laboratory to Introduce Lithography Concepts," *IEEE Transaction on Education*, vol. 55, no. 4, November 2012, pp. 517-524.

[12] Z. Zhang, C. T. Hansen, and A.E. Andersen, "Teaching Power Electronics With a Design-Oriented, Project-Based Learning Method at the Technical University of Denmark," *IEEE Transaction on Education*, vol. 59, no. 1, February 2016, pp. 32-38.

[13] N. Hoic-Bozic, V. Mornar, and V. Boticki, "A Blended Learning Approach to Course Design and Implementation," *IEEE Transaction on Education*, vol. 52, no. 1, February 2009, pp. 19-30.

[14] F. Martinez, L.C. Herrero, and S. de Pablo, "Project-based Learning and Rubrics in the Teaching of Power Supplies and Photovoltaic Electricity," *IEEE Transaction on Education*, vol. 54, no. 1, February 2011, pp. 87-96.

[15] D. G. Lamar *et al*., "Experiences in the Application of Project-based Learning in Switching-mode Power Supplies Course," *IEEE Transaction on Education*, vol. 55, no. 1, 2012, pp. 69-77.

[16] K. M. Y. Law and K. B. Chuah, "What motivates engineering students: A study in Taiwan," *Int. Journal of Engineering Education*, vol. 25, no. 5, 2009, pp. 1068-1074.

[17] R. H. Chu, D. D. Lu, S. Sathiakumar, "Project-Based Lab Teaching for Power Electronics and Drives," *IEEE Transsaction on Education*, vol. 51, no. 1, February 2008, pp. 19-30.

[18] M. A. Selver, "Design and Configuration of a Medical Imaging Systems Computer Laboratory Syllabus," *IEEE Transaction on Education*, vol. 59, no. 2, May 2016, pp. 129-136.

[19] F. Blaabjerg, "A Power Electronics and Drives Curriculum with Project-oriented and Problem-based Learning: A Dynamic Teaching Approach for the Future," J. Power Electronics, vol. 2, no. 4, October 2002.

[20] N. Hosseinzadeh and M. R. Hesamzadeh, "Application of Project-based learning (PBL) to the teaching of Electrical Power Systems Engineering," *IEEE Transaction on Education*, vol. 55, no. 4, November 2012, pp. 495-501.

Chebyshev Type-I Low Pass Filter Using Annular Ring Resonator: A Comparative Performance Analysis For Different substrates

Shakhawat Parvez, * Nazmus Sakib,* Md. Nurunnabi Mollah
Department of EEE, Khulna University of Engineering & Technology
Khulna-9203, Bangladesh
*shakhawat_parvez@yahoo.com,*sakib483@yahoo.com

***Abstract*— In this letter, compact microstrip Chebyshev function low pass filters (LPFs) with sharp roll-off based on annular ring resonators (ARRs) on various substrates and it's design theory are demonstrated. To fabricate ARRs, Chebyshev type – I function with fixed side lobe level (SLL) and method of moments (MOM) for numerical analysis are applied. The ARRs inscribed on the ground plane, behave as shunt connected capacitances with low radiation, nevertheless the 50 Ω transmission line on the upper plane provide the total transmission to realize the Chebyshev function response. Consequently, RT/Duroid 5880 prevails with low insertion loss (<0.6 dB), wide stopband, very sharp roll-off (>90 dB/GHz) and isolation up to 70 dB in the stopband. The frequency responses of the device from the two different software's are in good agreement with the known theoretical data.**

***Keywords*— Microstrip; Chebyshev function; Annular ring resonators (ARRs); Side lobe level.**

I. Introduction

Recently, Electromagnetic Band Gap structures (EBGs) which is a subset of metamaterials have many different patterns, extraordinary and unique characteristics. Furthermore, in both microwave and millimeter-wave engineering electromagnetic bandgap (EBG) and defected ground structure (DGS) have been frequently implemented with various structures [1]-[6]. Not only EBGs display sharper bandpass properties but also wider stopband properties. In various communication systems, microwave EBG assisted filters are highly demandable because of their formulation technique [7] which led to small size and desired resonant frequencies. Nevertheless, the harmonics in the transmission line has an adverse effect in ultra-wide band RF system. Different approaches employing slow wave properties [8], [9] and linear phase resonators were presented to suppress these spurious transmission so as to enhance the band rejection properties. To fabricate microstrip filter on a substrate, some representative parameters need to be secured effectively. Namely, spacing between DGS structure, center frequency, thickness, width, height of the substrate, filling factor (FF) of the structure and design pattern of the elements which have huge impact on filter sizes, bandpass and band rejection properties. However, some approach applying circular [10] and square [11] patterns in various substrates were reported so as to determine the high performance filter with appropriate substrate.

In fact, many difficulty arises in microwave filter due to the radiation problem as a result of improper selection of substrate.

In this exposition, conventional annular ring patterned DGS etched on the ground plane of a 50-Ω transmission line applying polynomial distribution is proposed. However, polynomial distribution namely, Chebyshev distribution has some great features like sharper cut-off, equal ripple and wider stopband characteristics. Likewise, the side lobe level (SLL) which is a unique measure to the Chebyshev co-efficient is exploited to mitigate the insertion loss in the passband and improve the return loss in the stopband. Besides, if such filters were designed with improper formulation technique then, it would create an adverse effect on roll-off [12] and insertion loss. Still, not only substrates height and width but also effective permittivity has proportional relationship with the course of the guided wavelength.

In this paper, applying polynomial distribution in Microstrip low pass filter not only has specific ground of originality but also some great features, which comprises the following:

- The low-pass filter architecture is simple.
- For all substrates, Chebyshev type-1 approach (where radius of the pattern will vary according to the amplitude co-efficient) is used.
- Traditional annular ring resonator inscribed on the ground plane, taken as the design and simulated for all the substrates.
- Wider rejection bandwidth found at 20-dB magnitude.
- Not only lower insertion loss but also sharper cut-off found in the passband.
- Size of the low-pass filter could be varied according to the effective dielectric constant (ϵ_{eff}).

Finally, this design of the microstrip low-pass filter are simulated in various substrates by planar MOM based EM software *Zeland IE3D* and *ANSYS v16.0*.

II. Design theory of structures

Basically, the side lobe level (SLL) which is an exclusive characteristics of Chebyshev array. Moreover, it also signifies the distinction between uniform and non-uniform arrays. The recursion formula [12] for Chebyshev polynomial is:

$$T_m(z)=2zT_{m-1}(z)-T_{m-2}(z) \quad (1)$$

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

Each polynomial can be evaluated employing the formulas:

$$T_m(z)=\cos[m\cos^{-1}(z)];\ -1\leq z\leq +1 \quad (2)$$

$$T_m(z)=\cosh[m\cosh^{-1}(z)];\ z\langle -1,\ z\rangle +1 \quad (3)$$

For this particular structure, magnitude of side lobe level fixed at 26-dB. Besides, the radiation intensity will vary according to the 26-dB SLL. Consequently, the amplitude co-efficient for the 26-dB SLL with Chebyshev array for 10 elements will be:

$$0.36\ 0.49\ 0.71\ 0.89\ 1\ \ 1\ 0.89\ 0.71\ 0.49\ 0.36$$

As illustrated in Fig. 1, the size of the ring resonator will depend on the amplitude co-efficient of 26-dB SLL. Likewise, the center frequency of the microstrip structure can be evaluated by:

$$\beta a=\pi \quad (4)$$

where, "*a*" is the period of ring pattern and "*β*" is the wave number of dielectric slab.

However, to attain the rejection band characteristics, the spacing between the each annular ring resonator need to be restrained sensitively. Nevertheless, the filling factor (FF) of the 50Ω transmission line architecture are designed from this equation:

$$FF=\frac{r_1}{a} \quad (5)$$

where, outer circle radius (r_1), filling factor (*FF*), spacing between two resonators (*a*).

Furthermore, aspect ratio which plays a vital role for resolving the inner circle radius as well as hollow between annular ring.

$$AR=\frac{r_2}{r_1} \quad (6)$$

where, aspect ratio (*AR*), inner circle radius (r_2).

This microstrip structure has been designed on different substrates while preserving the proportional relationship between the effective dielectric constant of the substrate which has an effect on guided wave length (λ_g) and side lobe level which influences the insertion loss (S_{21}) in the passband and rejection band.

III. Filter Design and Results

Considering the design of an order-10 Chebyshev LPF with passband ripple 0.1 dB, a center frequency of (f_0) 10 GHz. The amplitude co-efficient corresponding to this LPF are given by the Chebyshev array of 10 elements. Maintaining the CPW technology the LPF has been realized through ten magnetically coupled shunt connected annular ring resonators (ARRs) etched on the ground plane with 50Ω transmission line

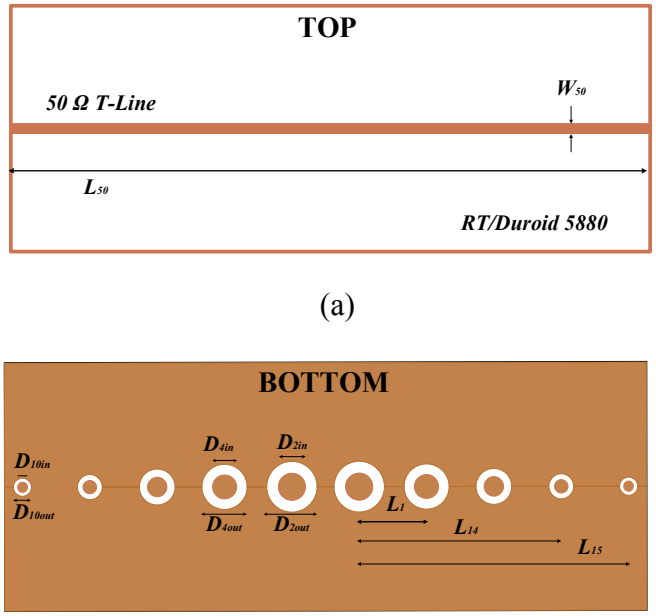


Fig. 1: Detailed layout (a) top plane and (b) bottom plane; of the ten stage Chebyshev low pass filter. The dimensions of the ten stage filter are L_{50} = 98 mm, W_{50} = 1.57 mm, L_1 = 10.908 mm, L_{14} = 32.724 mm, L_{15} = 43.632 mm, D_{2in} = 4.378 mm, D_{2out} = 8.182 mm, D_{4in} = 3.896, D_{4out} = 7.282 mm, D_{10in} = 1.576 mm, D_{10out} = 2.946 mm.

sections inscribed on upper substrate side. These ARRs and transmission line are magnetically coupled with each other.

By realizing the S-parameters of the 50Ω transmission line with ARRs below it (From EM simulations), the required value of inductance and capacitance can be calculated for designing the low pass filter (LPF).

The dimension of the ARRs have been resolved by implementing the amplitude co-efficient of SLL with aspect ratio. The inter-ring distance can be found by forcing the amplitude co-efficient to the required values of Chebyshev function model. Thus the equations given in [12] that relate the geometry of ARRs and amplitude co-efficient are only used as reference.

The layout of the fabricated Chebyshev type-1 LPF are illustrated in Fig. 1. The filter has been fabricated on different substrates which are (a) RT/Duroid 5880, ε_r = 2.20, *h* = 20mil = 0.508 mm; (b) RT/Duroid 5870, ε_r = 2.33, *h* = 20mil = 0.508 mm; (c) FR4, ε_r = 4.40, *h* = 20mil = 0.508 mm; (d) RT/Duroid 6010 LM, ε_r = 10.2, *h* = 25mil = 0.635 mm; (e) RT/Duroid 6010 LM, ε_r = 10.2, *h* = 50mil = 1.27 mm and results has been demonstrated on Table I. Among the substrates, the filter exhibits superior performance on RT/Duroid 5880, with the dielectric constant, (ε_r) = 2.20 and thickness, (*h*) = 0.508 mm or 20 mil. The stop band response of the filter has broaden by etching ARRs on the ground plane and the two transmission zeros can be clearly identified. Moreover, these magnetically

TABLE I : COMPARATIVE ANALYSIS AMONG SUBSTRATES WITH CHEBYSHEV TYPE-I FUNCTION.

εr, h(mm)	*Selectivity [dB/GHz]*	*S_{21max} [dB]*	*20-dB IL [GHz]*	*10-dB RL [GHz]*	*Fc [GHz]*
2.20;0.508	94.44	69.12	5.51	7.09	6.75
2.33;0.508	39.54	71.51	5.18	7.92	6.68
4.40;0.508	85.0	63.0	5.30	7.80	6.70
10.2;0.635	44.74	35.87	3.42	8.42	6.20
10.2;1.270	N/A	18.78	N/A	8.71	6.23

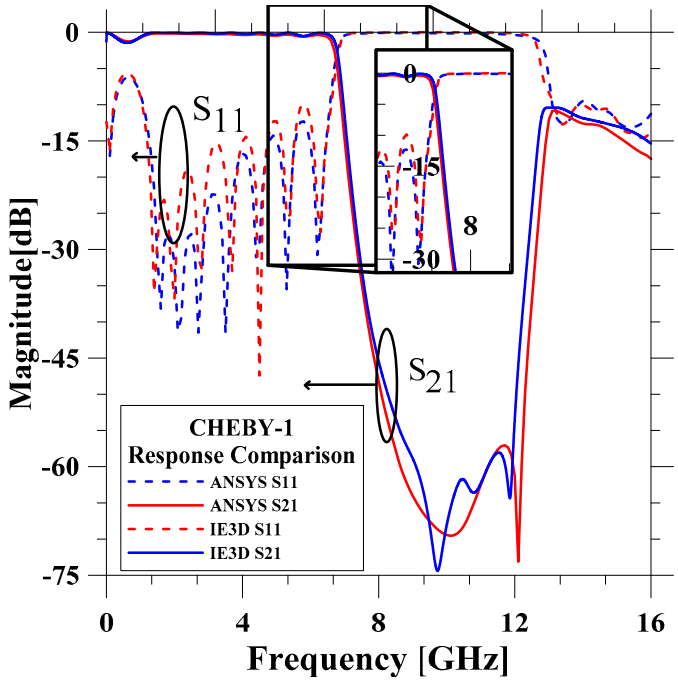


Fig. 2: Simulated transmission responses of Chebyshev type-I ten stage low pass filter.

coupled ARRs also provide the shunt capacitance for the filter. However, the filter dimensions is shown in Fig. 3(a) and 3(b).

The simulated (inferred from *ANSYS EM v16.0* and *Zeland IE3D v14*) frequency responses of the device are illustrated in Fig. 2. The effect of the substrates in the performance of the device with above indicated values are also shown in Table I, Fig. 3(a) and Fig. 3(b). The filter cut-off frequency is at 6.75 GHz. As shown in the Fig. 2, simulated insertion losses (IL) in the pass band less than 0.6 dB at 5.9 GHz and return losses (RL) are better than 12 dB from 0.8 to 6.3 GHz, over 81% of its passband. Besides, the maximum stopband rejection up to 70 dB which in agreement with both of the simulations. Moreover, the two transmission zeros located at upper and lower stopband over 6.75 to 12.7 GHz.

Based on Fig. 3(a), 20-IL bandwidth is proportional to maximum isolation over the five substrates. However, RT/Duroid 5880 presents more depth of rejection than any other substrates, which indicates a lot more efficient substrate against

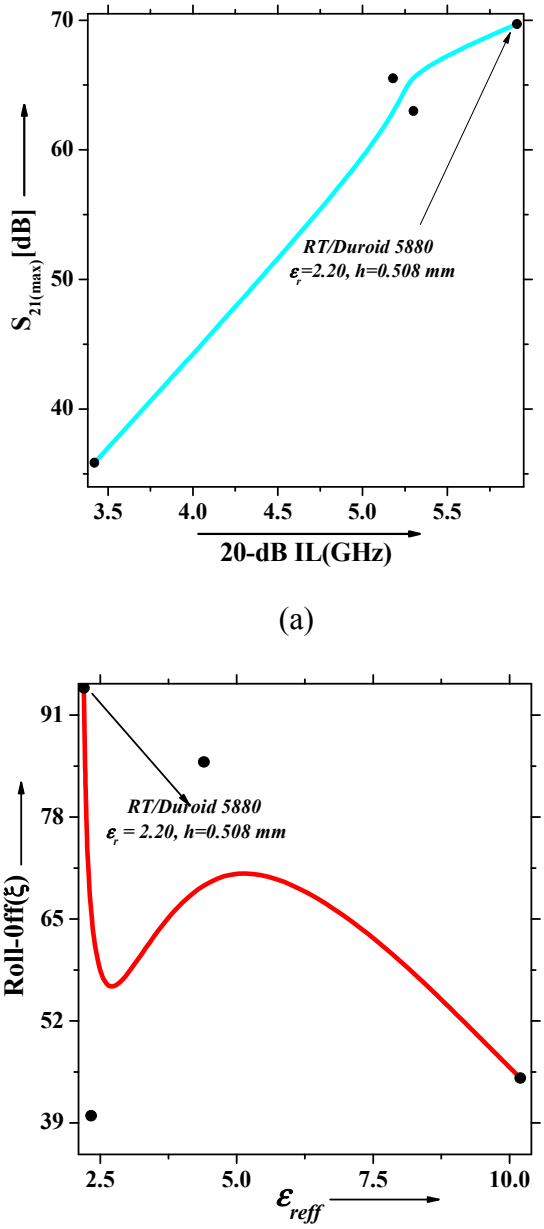


Fig. 3: Fluctuation of (a) maximum isolation vs. 20-dB IL bandwidth and (b) roll-off vs. effective dielectric constant (ε_{reff}) of Chebyshev function based LPF over five distinct substrates.

dielectric loss. Though the ARRs has small electrical length and sinusoidal decaying coupling over this length, the device exhibits an asymmetrical decreasing of roll-off over the effective dielectric constant as shown in Fig. 3(b). As illustrated in Fig. 4, the simulated group delay for the device, which is 1.18 nS at 6.75 GHz with a variations of 0.45 nS from 6.1 to 12.6 GHz, over 96% of its 3-dB pass bandwidth. For the device as shown in Fig. 1, the device exhibits excellent skirt performance up to 95 dB/GHz from 3 dB to 20 dB range. Likewise, the 20-dB stopband rejection level better than 55 dB from 1.1 f_c (7.4 GHz) to 1.9 f_c (12.8 GHz) which is roughly 5.50 GHz and also

TABLE II: PERFORMANCE COMPARISON

REF.	Roll-off [dB/GHz]	S_{21max} [dB]	IL/PB [dB]	F_C [GHz]
[7]	55	<25	<0.8	8.70
[10]	50	<62	<0.7	6.74
[11]	40	<65	<0.7	6.38
This work	94.45	<70	<0.6	6.75

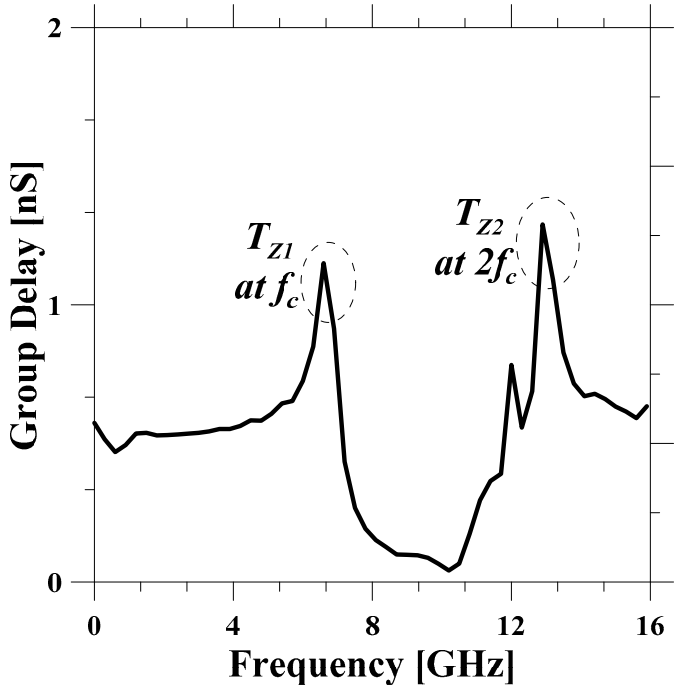


Fig. 4: Simulated frequency responses of group delay for the Chebyshev function based ten stage LPF with ARRs on RT/Duroid 5880 substrate

the 10-dB RL bandwidth is up to 7 GHz. The dimension of the ARRs in the proposed device, starting from $0.13\lambda_g$ to $0.05\lambda_g$, which responsible for sharp roll-off and compact size of the device. Likewise, the height of the device is measured to be $0.19\lambda_g$ (6 mm) [where, λ_g is the guided wavelength at the cut-off frequency i.e. 6.75 GHz].

The performance and dimensions of such filters and this filter are compared in Table II. As illustrated in Table II, this device presents sharper roll-off, good rejection and low insertion loss over the published designs.

IV. CONCLUSIONS

Chebyshev type-I distributed low pass filter of order 10 based on annular ring resonators (ARRs) with 50 Ω transmission line have been presented in this paper. To sum up, RT/Duroid 5880 presents superior performance than other substrates, however the basic filter model remains the same. This device exhibits not only skirt roll-off attributes but also wide stopband up to 2.1 f_C and low insertion loss in the passband. Besides, this magnetically coupled ARRs have very small electrical length ranging which provide the compactness of the filter. The resultant design provides two transmission zeros at the upper stopband. Meanwhile, this transmission zeros diminishes at the further stopband as a result of weak capacitive coupling between ARRs, which can be improved by adding stub in the design. Moreover, the group delay of the filter is really impressive (<1.5 nS). The presented LPF model is validated by two software simulation performances.

REFERENCES

[1] Fei-Ran Yang, Kuang-Ping Ma, Yongxi Qian and T. Itoh, "A uniplanar compact photonic-bandgap (UC-PBG) structure and its applications for microwave circuit," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 47, no. 8, pp. 1509-1514, Aug 1999. doi: 10.1109/22.780402

[2] S. Parvez, N. Sakib and N. Mollah, "Dumbbell annular ring with circular patterned microstrip low-pass filter with higher selectivity, wider stopband and lower insertion loss," *Electrical Information and Communication Technology (EICT), 2015 2nd International Conference on*, Khulna, 2015, pp. 373-377. doi: 10.1109/EICT.2015.7391979

[3] D. Kaddour, E. Pistono, J. M. Duchamp, L. Duvillaret, A. Jrad and P. Ferrari, "Compact and selective lowpass filter with spurious suppression," in *Electronics Letters*, vol. 40, no. 21, pp. 1344-1345, 14 Oct. 2004. doi: 10.1049/el:20045969

[4] S. Parvez, N. Sakib and N. Mollah, "The 1x1 annular ring: Bandwidth enhancement of microstrip asymmetric couple-line bandpass filter with superior matching and lower insertion loss," *Electrical Information and Communication Technology (EICT), 2015 2nd International Conference on*, Khulna, 2015, pp. 368-372. doi: 10.1109/EICT.2015.7391978

[5] D. Ahn, J. S. Park, C. S. Kim, J. Kim, Y. Qian and T. Itoh, "A design of the low-pass filter using the novel microstrip defected ground structure," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 1, pp. 86-93, Jan 2001. doi: 10.1109/22.899965

[6] F. C. Chen, H. T. Hu, J. M. Qiu and Q. X. Chu, "High-Selectivity Low-Pass Filters With Ultrawide Stopband Based on Defected Ground Structures," in *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 5, no. 9, pp. 1313-1319, Sept. 2015. doi: 10.1109/TCPMT.2015.2461663

[7] N. C. Karmakar and M. N. Mollah, "Investigations into nonuniform photonic-bandgap microstripline low-pass filters," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 51, no. 2, pp. 564-572, Feb. 2003. doi: 10.1109/TMTT.2002.807817

[8] Jun-Seok Park, Jun-Sik Yun and Dal Ahn, "A design of the novel coupled-line bandpass filter using defected ground structure with wide stopband performance," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 9, pp. 2037-2043, Sep 2002. doi: 10.1109/TMTT.2002.802313

[9] D. Kaddour *et al.*, "A compact and selective low-pass filter with reduced spurious responses, based on CPW tapered periodic structures," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 6, pp. 2367-2375, June 2006. doi: 10.1109/TMTT.2006.875290

[10] S. Parvez, N. Sakib and M. N. Mollah, "Demonstration on performance characteristics of microstrip low-pass filter for various substrates employing Tschebycheff distribution," *2015 International Conference on Advances in Electrical Engineering (ICAEE)*, Dhaka, Bangladesh, 2015, pp. 125-128. doi: 10.1109/ICAEE.2015.7506812

[11] S. Parvez, N. Sakib and M. N. Mollah, "Determination of appropriate substrate for microstrip low-pass filter employing square designed tschebycheff distribution," *2015 18th International Conference on Computer and Information Technology (ICCIT)*, Dhaka, 2015, pp. 554-559. doi: 10.1109/ICCITechn.2015.7488133

[12] C.A. Balanis, *Antenna Theory Analysis and Design*, 2 ed. New York: Wiley, 1997.

Temperature Dependent Structural Studies and Magnetic Characterization of Pure $BiFeO_3$ Nanoparticles Synthesized by Sol-gel Method

Abid Anjum Sifat,[1,*] Khandker Akif Aabrar,[1] Md. Khairul Bashar,[1] Bashir Ahmmad,[2] M. A. Basith,[3] and Md. Ziaur Rahman Khan[1]

[1]Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology
Dhaka-1000, Bangladesh
[2]Graduate School of Science and Engineering, Yamagata University
4-3-16 Jonan, Yonezawa 992-8510, Japan
[3]Department of Physics, Bangladesh University of Engineering and Technology
Dhaka-1000, Bangladesh
*sifatabid@yahoo.com

Abstract—**Pure $BiFeO_3$ nanoparticles were synthesized by a simple sol-gel method namely modified Pechini method. The nanoparticles were calcinated at four different temperatures (450^0C, 550^0C, 650^0C, 750^0C). Requirement for extra washing step for impurity phase reduction has been eliminated by elongation of the drying duration at oven from 12 hours to 28 hours along with intermediate grinding. FESEM and X-ray Diffraction (XRD) were performed which confirm that the samples calcinated at 450^0C and 550^0C show smaller particle size and greater correspondence. Hence XPS was performed on these two samples which showed lower oxygen vacancies for 550^0C. Finally the SQUID analysis has been carried out for the 550^0C sample. A higher value of remanent magnetization and coercive field at room temperature has been observed. Besides, an asymmetric shift in the field axis with unsaturated M-H curve may be attributed to the existence of exchange bias effect in the synthesized nanoparticles.**

Index Terms—**Multiferroics, sol-gel, nanoparticles, exchange bias effect.**

I. Introduction

The quest for multiferroic materials is actuated not only by the simultaneous existence of ferroelectricity and ferromagnetism but also by the possible coupling of their electric and magnetic orderings [1–3]. Such property allows mutual control of the electric polarization with a magnetic field or control of magnetization by an electric field. This magnetoelectric (ME) effect is prodigiously potential in developing novel memory and spintronic devices, spin valves, oscillators, filters, thin film capacitors and sensors [4–6].

Among the available multiferroic materials of type ABO_3, $BiFeO_3$ (BFO) having rhombohedrally distorted perovskite structure with lattice parameters of a = 5.571Å and c = 13.868Å at room temperature with high Curie temperature of T_C = 820–850^{0}C [1, 2], Neel temperature of T_N = 350–380^{0}C [3] has turned into a cynosure of many current studies. It has potential applications of magnetoelectric coupling at temperatures around room temperature [1, 5, 7, 8]. In BFO Bi^{3+} contributes to ferroelectricity whereas antiferromagnetism is attributed to Fe^{3+} ions. $BiFeO_3$ exhibits anti-ferromagnetic G-type spin conf guration along the $[111]_c$ or $[001]_h$ directions in its pseudocubic or rhombohedral structure. It has a cycloid spin structure with a periodicity of 62nm along the $[001]_h$ axis at room temperature which renders cancelation of the macroscopic magnetization inhibiting observation of the linear ME effect in the bulk form [9, 10]. The decrease in particle size has been proved to be effective in suppressing the spiral modulated spin structure and enhancing the magnetic moment of $BiFeO_3$ [7]. As the particle size decreases below 62nm, the magnetic response of $BiFeO_3$ greatly increases compared to the bulk material. So far, a variety of methods have been concocted for the synthesis of nanoparticles such as solid state reactions of oxide precursors at high temperature or by mechanical activation at lower temperatures [11], wet chemical methods under which comes solgel which includes metal complex [12], glycol reaction [13], pechini method [14] and polymeric precursor [15]etc.

In our present work, a short route has been proposed to the synthesis of BFO nanoparticle by sol-gel method which required no additional washing step for removal of impurities. The route consisted of modified pechini method accompanied by intermediate grinding and longer oven heating. Calcination was performed at four different temperatures. Morphological studies and structural characterization of the samples were performed by Field Emission Scanning Electron Microscope (FESEM) and X-ray Diffraction (XRD). Furthermore the oxidation states of the constituent atoms and the amount of oxygen vacancy were evaluated by X-ray Photoelectron Spectroscopy (XPS). Finally magnetization experiments were performed using Superconducting Quantum Interference Device Magnetometer (SQUID) to investigate the magnetic behavior of the synthesized nanoparticles which shows the existence of exchange bias in the sample with a biasing field of 202.9Oe at room temperature.

II. Experimental Details

In this investigation, BFO nanoparticles were synthesized by sol-gel method (modified pechini) using $Bi(NO_3)_3.5H_2O$ and $Fe(NO_3)_3.9H_2O$ as the initial components. Calculated amount of $Bi(NO_3)_3.5H_2O$ and $Fe(NO_3)_3.9H_2O$ were initially taken into a beaker with about 200-250 ml of 2-methoxyethanol as the solvent. Then the solution was stirred for about 4 hour until all the components got dissolved and clear solution was obtained. After that the pH of the solution was increased to around 5 by drop wise addition of NH_4OH. Afterwards acetic acid (CH_3COOH) and ethylene glycol

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

($OHCH_2CH_2OH$) in 1:1 molar ratio were added to the solution sequentially. After that the solution was heated at constant temperature at about 70^{0}C to 80^{0}C with continuous stirring for about 14 to 16 hours until all the solvents got evaporated and the solution turned thicker. Finally a thin gel like layer of dark brown color appeared. The obtained sample was dried in the oven for about 28 hours with two or three intermediate grinding. The intermediate grinding ensured proper esterification and uniform drying. The dried gel was then grinded into powder in the mortar. The obtained powder was then calcinated for 2 hours at four different temperatures 450^{0}C, 550^{0}C, 650^{0}C and 750^{0}C with the heating and cooling rate of the furnace, maintained at 3^{0}C/min. The calcinated samples were again grinded and used for various measurements and characterization.

For phase identification of the samples XRD patterns were collected. The micro-structures of the prepared samples were observed from FESEM images. To investigate the oxidation state of Bi and Fe and to observe the oxygen vacancy in the samples XPS measurements were performed. The magnetic properties of the prepared samples were determined through SQUID analysis with a maximum applied field of 30 KOe.

III. Results and Discussions

A. X-ray Diffraction Pattern Analysis

The XRD patterns of BFO nanoparticles with different drying duration in the oven are shown in Fig. 1. In one of our experiments, shorter drying duration about 12 hours in the oven without any intermediate grinding resulted in the red colored XRD pattern shown in Fig. 1. Here a shift in the base line and the insignificant peaks might indicate presence of defects and impurity. Whereas a longer duration in oven about 28 hours with intermediate grindings resulted in the black colored XRD patterns shown in Fig. 1 with precise peaks and seemingly no traces of extra peaks ensuring absence of any impurity phases.

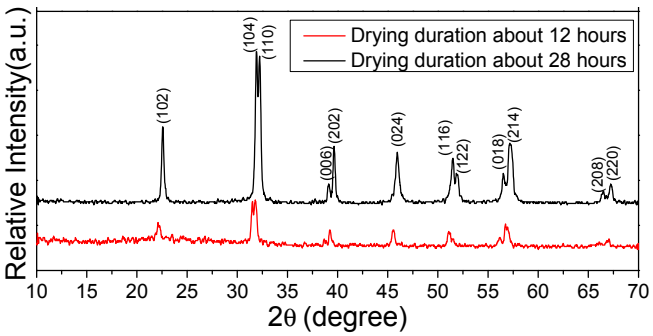


Fig. 1 X-ray diffraction pattern of $BiFeO_3$ nanoparticle prepared with drying duration about 12 hours (red curve) and 28 hours (black curve).

Several previous investigations [16-18] on $BiFeO_3$ nanoparticle synthesis reported generation of impurity phases which needed nitric acid leaching for proper removal. This resulted in coarser powders, poor reproducibility, and additional high temperature annealing process or addition of excess initial component (Bi) to compensate for the loss in thermal treatment [14]. As evident in Fig. 1 it was possible to obtain completely pure phase using acetic acid as the chelating agent and a longer drying duration with intermediate grinding. Thus no further treatment for impurity phase reduction was required.

The XRD patterns of our samples calcinated at 450^{0}C, 550^{0}C, 650^{0}C and 750^{0}C are shown in Fig. 2. Almost all patterns except the black curve were well matched with rhombohedrally distorted perovskite of $BiFeO_3$ with space group R3c as stated in JCPDS database (JCPDS file no. 86–1518). But the sample for $550°C$ showed the best alignment. No additional peaks of secondary impurities such as Fe_2O_3, $Bi_2Fe_4O_9$, Bi_2O_3 and $Bi_{25}FeO_{39}$ [19, 20] were detected.

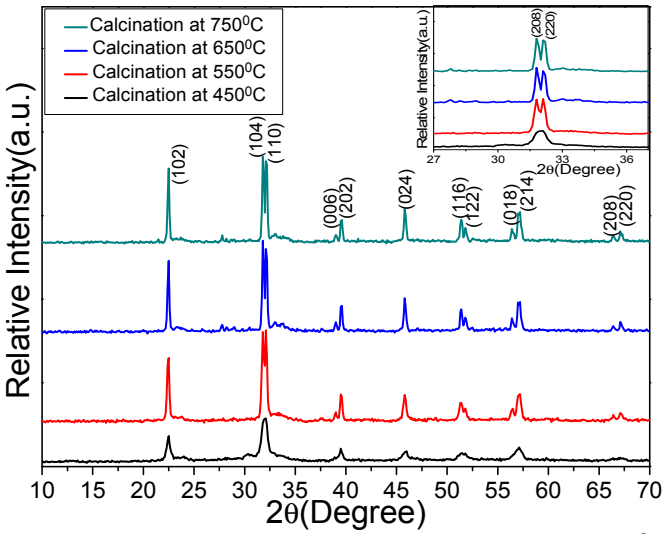


Fig. 2 X-ray diffraction pattern of $BiFeO_3$ nanoparticles calcinated at 450^{0}C, 550^{0}C, 650^{0}C and 750^{0}C. The inset shows the magnified XRD patterns of (104) and (110) reflection planes of the four samples.

From the magnified view of the peaks at $2\Theta = 32^0$ shown in the inset of Fig. 2 it was evident that as the temperature was decreased the reduced splitting of the peak revealed a tendency of structural transformation from rhombohedral to orthorhombic [21].

B. FESEM Micrograph Analysis

The FESEM images of the $BiFeO_3$ nanoparticles calcinated at 450^{0}C, 550^{0}C, 650^{0}C and 750^{0}C are shown in Fig. 3. From the figure it was clear that as the calcination temperature was increased the particle size increased too. Especially as the temperature crosses 550^{0}C, the particles got too larger which might be due to coarsening of particles as high temperature accelerates the process.

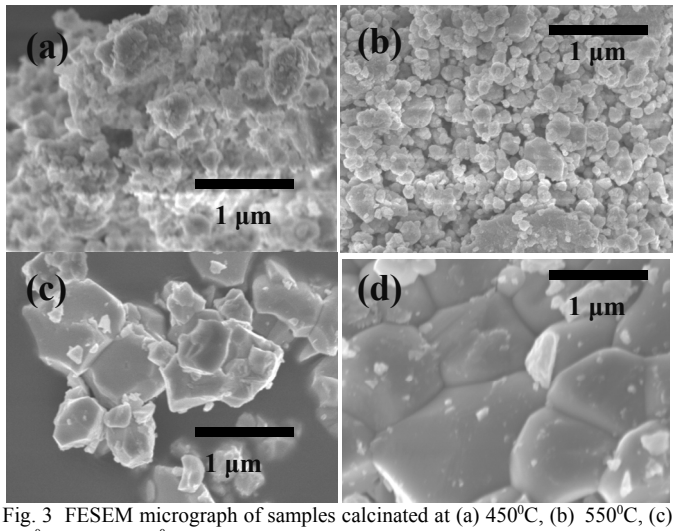


Fig. 3 FESEM micrograph of samples calcinated at (a) 450^{0}C, (b) 550^{0}C, (c) 650^{0}C and (d) 750^{0}C

Though the sample calcinated at 450^{0}C produced lower size particle, it was not in our required single rhomohedral phase as evident from the XRD pattern. The best particle size obtained for the sample calcinated at 550^{0}C which also gave required XRD pattern. From the highly magnified micrograph of 550^{0}C sample shown in Fig. 4 a good number of particles with sizes below 100nm has been observed.

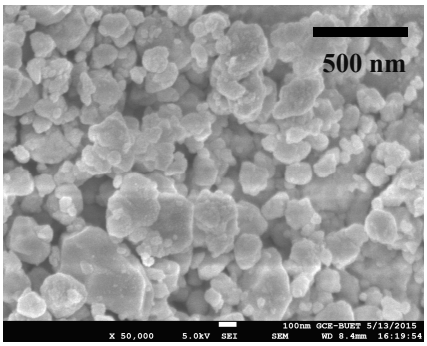


Fig. 4 Highly magnified FESEM micrograph of samples calcinated 550^{0}C

C. XPS (X-ray Photoelectron Spectroscopy) Data Analysis

The XRD patterns and FESEM micrographs motivated us to perform the XPS measurements on the samples calcinated at 450^{0}C and 550^{0}C. Fig. 5 shows the complete XPS pattern of the two samples.

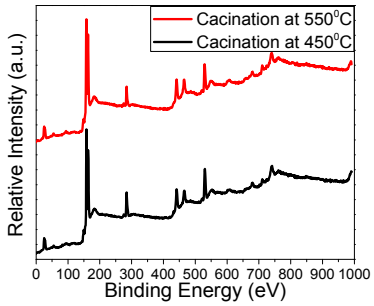


Fig. 5 The complete XPS spectra of $BiFeO_3$ calcinated at 450^{0}C & 550^{0}C

Fig. 6(a) shows the portion of XPS spectra due to Bi component. It showed a doublet located at 158.1eV and 163.4eV for the sample calcinated at 550^{0}C (Red curve) and a doublet at 158.6eV and 163.9eV for the sample calcinated at 450^{0}C (Black curve) corresponding to the core lines of $4f_{7/2}$ and $4f_{5/2}$ of Bi^{3+}. The doublet is a characteristic of the XPS spectra of any p, d and f orbitals arising due to spin-orbit interaction. Similar observations have been reported in [22].

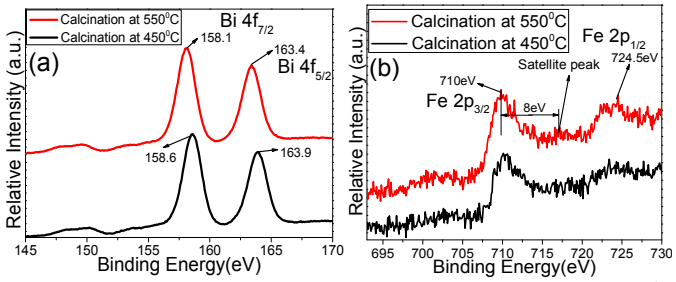


Fig. 6 XPS spectra of (a) Bi & (b) Fe elements of $BiFeO_3$ calcinated at 450^{0}C & 550^{0}C

Fig. 6(b) shows the Fe valance state of the samples. The red curve in Fig. 6(b) showed peaks at binding energies 710eV and 724.5eV which represent $2p_{3/2}$ and $2p_{1/2}$ peaks of Fe^{3+} for the sample calcinated at 550^{0}C. No distinct peak of Fe^{2+} was observed at 709.5eV for this sample. Presence of the satellite peak for Fe $2p_{3/2}$ at a difference of 8eV furthermore assured the sole existence of Fe^{3+} in this sample. Though the peak positions of $2p_{3/2}$ and $2p_{1/2}$ for the sample calcinated at 450^{0}C (Black curve) were same, the peaks seemed less precise and a bit broader than that obtained from the other sample. The peak broadening raised the possibility of Fe^{2+}. Moreover the satellite peak for Fe $2p_{3/2}$ at a difference of 8eV was merely visible for this sample.

The O 1s spectrum of the sample, calcinated at 450^{0}C, is shown on Fig. 7(a). It has been fitted to two symmetric Gaussian curves with peak positions 529.22eV and 530.93eV. Fig. 7(b) shows the spectrum for 550^{0}C which has also been fitted to two symmetric Gaussian curves with peaks at 528.82eV and 530.56eV. The peak for lower binding energy corresponds to the O 1s core spectrum of the BFO phase [23] while the higher binding energy peak can be attributed to the oxygen vacancy [24, 25] in the samples.

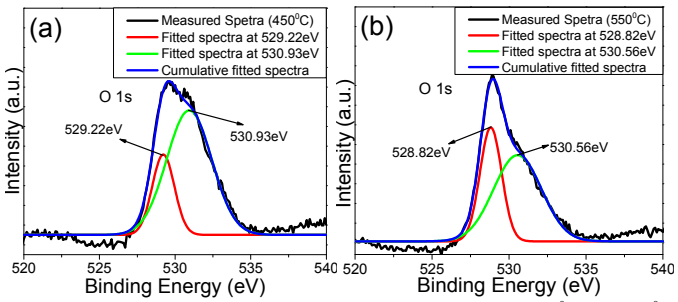


Fig. 7 XPS spectra of O element of $BiFeO_3$ calcinated at (a) 450^{0}C, (b) 550^{0}C

The area ratio of the two gaussian peaks (lower/higher binding energy) of the samples calcinated at 450^{0}C and 550^{0}C has been found to be 0.32 and 0.6. The higher area ratio of the sample, calcinated at 550^{0}C, compared to the other one indicates lower oxygen vacancy concentration in that sample.

D. Magnetic Characterization (SQUID Data Analysis)

Magnetization measurements were carried out using SQUID magnetometer for the sample calcinated at 550^{0}C due to its greater correspondence than the other samples. Fig. 8(a) shows the room temperature M-H loop of our synthesized $BiFeO_3$ nanoparicles. From the magnified view of the loop shown in Fig. 8(b) a non-zero remanent magnetization, M_r and coercive field, H_c were observed. The values of M_r and H_c of our synthesized sample were found to be 0.00726emu/g and 631.5Oe at room temperature which were noticeably larger than the values 0.0009emu/g and 110Oe found for bulk $BiFeO_3$ [26]. The bulk $BiFeO_3$ is antiferromagnetic and the rise of such remnant magnetization for decreased particle size can be modeled as antiferromagnetic core with ferromagnetic surface [27]. For smaller particles, there is disruption at the surface because of missing oxygen ions. Thus surface spin disorder is generated and number of uncompensated spins of Fe^{3+} ions increases giving enhanced magnetization.

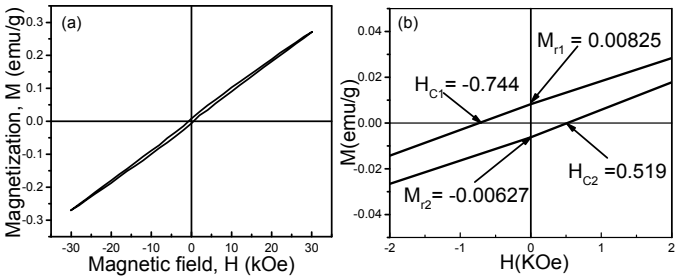


Fig. 8(a) The room temperature M-H loop of $BiFeO_3$ nanoparticles calcinated at 550^{0}C (b) The magnified view of the curve at the origin.

The most intriguing feature observed from Fig. 8(b) was the asymmetric behavior exhibiting shifts in both the field and magnetization axes. The unsaturated hysteresis loop with an applied magnetic field of up to 30kOe and a value of coercivity of 0.632kOe indicated the presence of both ferri/ferromagnetic and antiferromagnetic domains in these nanoparticles [28]. The exchange bias effect usually occurs in ferro and antiferromagnetic bilayers or multilayers in which the two coercive fields H_{c1} and H_{c2} of the M-H loop are not symmetric, and the centre of the loop shifts to either right or left [29]. The exchange bias along the field axis and the magnetization axis can be quantified as $H_{EB} = -(H_{c1}+H_{c2})/2$ where H_{c1} and H_{c2} are the left and right intercepts of the curve with the field axis [29] and $M_{EB} = (M_{r1}+M_{r2})/2$ where M_{r1} and M_{r2} are the positive and negative points of intersection with H=0 [26]. In our sample the observed shift in M_{EB} and H_{EB} has been found to be 0.000991emu/g and 202.9Oe which was significantly higher than the value reported in [30] (H_{EB} = 36Oe) for bulk $BiFeO_3$ ceramic. And such value of exchange bias is comparatively higher particularly at room temperature.

IV. Conclusions

Bismuth Ferrite nanoparticles were synthesised using modified pechini method. For producing this nanoparticle no washing step was needed after formation of gel. Elongation of the drying duration ensured pure phase of bismuth ferrite. Calcinations were done at four different temperatures and among them samples calcinated at 550^0C exhibited best particles considering the size of the particles and their crystalinity, proper oxidation states and lower oxygen vacancy. Significant amount of remnant magnetization and coersivity of 0.00726emu/g and 631.5Oe at room temperature were observed in the sample calcinated at 550^0C. An asymmetric shift along the field axis in unsaturated M-H curve ensured the presence of sufficient exchange bias of H_{EB} = 202.9Oe in the synthesized bismuth ferrite nanoparticles. Their improved structural and magnetic behaviour may be worthwhile in potential applications. The ferroelectric properties of the synthesized nanoparticles are being investigated.

Acknowledgment

We would like to thank Infrastructure Development Company Limited (IDCOL) for providing support under a research grant. We also sincerely acknowledge Department of GCE, BUET for their FESEM facilities.

References

[1] M. Fiebig, "Revival of the magnetoelectric effect," *Journal of Physics D: Applied Physics*, vol. 38, pp. R123-R152, April 2005.

[2] J. M. Caicedo, J. A. Zapata, M. E. Gómez, and P. Prieto, "Magnetoelectric coefficient in BiFeO3 compounds," *J. Appl. Phys.,* vol. 103, p. 07E306, April 2008.

[3] Sang-Wook Cheong and Maxim Mostovoy, "Multiferroics: a magnetic twist for ferroelectricity," *Nature Materials,* vol. 6, pp. 13-20, Jan. 2007.

[4] H Béa, M Gajek, M Bibes and A Barthélémy, "Spintronics with multiferroics," *J. Phys.: Condens. Matter,* vol. 20, p. 434221, 2008.

[5] Nicola A. Hill, "Why Are There so Few Magnetic Ferroelectrics?," *J. Phys. Chem. B*, 2000, vol. 104, pp 6694-6709, June 2000.

[6] W. Eerenstein, N. D. Mathur and J. F. Scott, "Multiferroic and magnetoelectric materials," *Nature,* vol. 442, pp. 759-765, Aug. 2006.

[7] R. Mazumder, P. Sujatha Devi, Dipten Bhattacharya, P. Choudhury, A. Sen and M. Raja, "Ferromagnetism in nanoscale $BiFeO_3$," *Appl. Phys. Lett.,* vol. 91, p. 062510, Aug. 2007.

[8] T.T. Carvalho and P.B. Tavares, "Synthesis and thermodynamic stability of multiferroic $BiFeO_3$," *Mater.Lett.*, vol. 62, pp. 3984–3986, Sep. 2008.

[9] Claude Ederer and Nicola A. Spaldin, "Weak ferromagnetism and magnetoelectric coupling in bismuth ferrite," *Phys.Rev.B*, vol. 71, p. 060401, Feb. 2005.

[10] I Sosnowska, T P Neumaier and E Steichele, "Spiral magnetic ordering in bismuth ferrite," *J. Phys. C: Solid State Phys.*, vol. 15, pp. 4835-4846, Aug. 1982.

[11] J. M. Xue, D. M. Wan, and J. Wang, "Functional ceramics of nanocrystallinity by mechanical activation", *Solid State Ionics*, vol. 151, pp. 403–412, Nov. 2002.

[12] K. Matsumuto, K. Yamaguchi, T.Fuji and A.Ueno, "Preparation of bismuth-substituted yttrium iron garnet powders by the citrate gel process," *J.Appl.Phys.,*, vol. 69, pp. 5918–5920, April 1991.

[13] K. Babooram and Z.-G. Ye, "New Soft Chemical Routes to Ferroelectric $SrBi_2Ta_2O_9$," *Chem.Mater.,* vol. 18, pp. 532–540, Dec. 2006.

[14] M.P.Pechini, "Method of preparing lead and alkaline earth titanates and niobates and coating method using the same to form a capacitor," U.S.PatentNo. 3 330 697, July11, 1967.

[15] M.A.Guelguen, M.H.Nguyen and W.M.Kriven, "Polymerized organic-irorganic synthesis of mixed oxides," *J.Am.Ceram.Soc.,* vol. 82, pp. 556–560, March 1999.

[16] Glenda Biasotto, Alexandre Z. Simões, César R. Foschini, Selma G. Antônio, Maria A. Zaghete and Jose A. Varela, "A novel synthesis of perovskite bismuth ferrite nanoparticles," *Processing and Application of Ceramics,* vol. 5, pp. 171-179, 2011

[17] Yongming Hu, Linfeng Fei, Yiling Zhang, Jikang Yuan and Yu Wang, Haoshuang Gu, "Synthesis of Bismuth Ferrite Nanoparticles via a Wet Chemical Route at Low Temperature," *Journal of Nanomaterials,* Vol. 2011, Article ID 797639, 2011

[18] Tahseen H Mubarak, Bruska Azhdar, Karim H Hassan and Chia H Kareem, "Effect of Temperature on Structural and Electrical Properties of Bismuth Ferrite Nanoparticles Prepared By Sol Gel Method," *International Journal of Innovative Research in Science, Engineering and Technology,* vol. 3, pp. 17034-17941, Oct. 2014

[19] M. I. Morozov, N. A. Lomanova and V. V. Gusarov, "Specific Features of $BiFeO_3$ Formation in a Mixture of Bismuth(III) and Iron(III) Oxides," *Russ. J. Gen. Chem.,* vol. 73, pp. 1676-1680, Nov. 2003.

[20] M. Valant, A. K. Axelsson, and N. Alford, "Peculiarities of a Solid-State Synthesis of Multiferroic Polycrystalline $BiFeO_3$," *Chem. Mater.,* vol. 19, pp. 5431–5436, Oct. 2007.

[21] Gurmeet Singh Lotey and N. K. Verma, "Structural, magnetic, and electrical properties of Gd-doped BiFeO3 nanoparticles with reduced particle size," *Journal of Nanoparticle Research*, vol 14, p. 742, Feb. 2012

[22] Rajasree Das, Tanushree Sarkar and K Mandal, "Multiferroic properties of Ba2+ and Gd3+ co-doped bismuth ferrite: magnetic, ferroelectric and impedance spectroscopic analysis," *J. Phys. D: Appl. Phys.,* vol. 45, p. 455002, Oct. 2012.

[23] Fang L, Liu J, Ju S, Zheng F, Dong W and Shen M, "Experimental and theoretical evidence of enhanced ferromagnetism in sonochemical synthesized BiFeO3 nanoparticles," *Appl. Phys. Lett.,* vol. 97, p. 242501, Dec. 2010.

[24] Y B Lin,Y M Yang, B Zhuang, S L Huang, L P Wu, Z G Huang, F M Zhang and Y W Du, "Ferromagnetism of Co-doped TiO2 films prepared by plasma enhanced chemical vapour deposition (PECVD) method," *J. Phys. D: Appl. Phys.*, vol. 4, p. 195007, Oct. 2008

[25] Rath C, Mohanty P, Pandey A C and Mishra N C, "Oxygen vacancy induced structural phase transformation in TiO2 nanoparticles," *J. Phys. D: Appl. Phys.*, vol. 42, p. 205101, Sep. 2009

[26] M. A. Basith, O. Kurni, M. S. Alam, B. L. Sinha and Bashir Ahmmad, "Room temperature dielectric and magnetic properties of Gd and Ti co-doped BiFeO3 ceramics," *Journal of Applied Physics,* vol. 115, p. 024102, Jan. 2014

[27] Fengzhen Huang, Zhijun Wang, Xiaomei Lu, Junting Zhang, Kangli Min, Weiwei Lin, Ruixia Ti, TingTing Xu, Ju He, Chen Yue and Jinsong Zhu, "Peculiar magnetism of $BiFeO_3$ nanoparticles with size approaching the period of the spiral spin structure," *Scientific Reports*, vol. 3, Article number: 2907, Oct. 2013

[28] M. A. Basith, F. A. Khan, Bashir Ahmmad, Shigeru Kubota, Fumihiko Hirose, D.-T. Ngo, Q.-H. Tran and K. Mølhave, "Tunable exchange bias effect in magnetic Bi0.9Gd0.1Fe0.9Ti0.1O3 nanoparticles at temperatures up to 250 K," *Journal of Applied Physics,* vol.118, p. 023901, July 2015

[29] Y. Guo, L. Shi, S. Zhou, J. Zhao, C. Wang, W. Liu and S. Wei, "Tunable exchange bias effect in Sr-doped double perovskite La_2NiMnO_6," *J. Phys. D: Appl. Phys.,* vol.46, p.175302, April 2013.

[30] R. Yao, C. Cao, C. Zheng and Q. Lei, "Lattice mismatch induced strained phase for magnetization, exchange bias and polarization in multiferroic BiFeO3," *RSC Adv.,* vol.3, pp. 24231-24236, Oct. 2013.

A High Efficiency Transformerless PV Grid-Connected Inverter with Leakage Current Suppression

Shakil Ahamed Khan,* Youguang Guo, and Jianguo Zhu

School of Electrical, Mechanical and Mechatronic Systems, University of Technology Sydney
81 Broadway, Ultimo, NSW 2007, Australia
*Shakil.A.Khan@student.uts.edu.au

Abstract—**This paper presents a new diode free freewheeling and common-mode voltage (CMV) clamping branches for single phase transformerless grid connected photovoltaic (PV) inverter for complete leakage current elimination and low conduction losses. In the past, various isolation techniques have been proposed for leakage current elimination in transformerless PV inverters. However, galvanic isolation only cannot completely eliminate leakage current due to that a resonant path is created by the switch junction capacitors, which also generate leakage current. The proposed freewheeling branch consists of four MOSFETs along with a clamping branch, which consists of two MOSFETs and a capacitor divider. The divider is connected to the DC side of the converter to keep constant CMV in the freewheeling path. As a result, the improved CMV clamping has been achieved for complete leakage current elimination. The unipolar sinusoidal pulse width modulation (SPWM) technique and modified HERIC topology with AC-decoupling for galvanic isolation is adopted for lower conduction losses. The proposed topology consists of only MOSFET in the freewheeling and clamping path which provides lower conduction losses compared with diode based topologies. The performances of the proposed topology in terms of common mode characteristics, leakage current, total harmonic distortion and conversion efficiency are analyzed and compared with H5, H6, HERIC and HBZBR topologies. The detail analyses are performed using MATLAB/Simulink and PSIM.**

***Index Terms*—inverters, leakage current suppression, common-mode voltage, photovoltaic systems.**

I. Introduction

Energy from renewable sources is becoming a common interest of research due to the rapid increase of energy demand and the exhaustion of global resources. Among renewable power sources, photovoltaic (PV) energy source is known as one great alternative all over the world due to the inexhaustible and pollution-free nature of solar power [1]. The integration of PV energy into the power grids is popular now. Therefore, efforts are being made to increase the percentage of energy sourced from PV energy into the grid [2]. However, what discourages PV system from expanding is the cost of the conversion systems due to expensive multi-winding transformer and filters in the grid connected inverter circuit. Therefore, it is important for a PV system to have a low cost and high efficiency grid connected inverter.

Photovoltaic inverter topologies can be classified into the transformerless inverter and the inverter with transformer. Generally, a high frequency and high efficiency transformer is used in the DC side or a low frequency bulky transformer is used in the AC side of the inverter to provide galvanic isolation for leakage current elimination and to avoid DC current injection into the grid. However, additional losses caused by the transformers in the inverter circuit deteriorate the overall system efficiency. Besides, the cost and size of the inverter increase due to use of the transformer. To overcome these problems, a research will be to design a transformerless grid-connected PV inverter. A direct path can be created between the PV module and the utility grid for a transformerless inverter due to nonisolation. As a result, the leakage current can flow from the PV system to the grid. In order to improve the safety and reliability of transformerless PV system, galvanic isolation is the main concern. A high leakage current is generated by the PV stray capacitance and the switch junction capacitance due to charging and discharging operation by the fluctuating common-mode voltage (CMV). According to VDE0126-1-1 standard, the leakage current flow from the PV system to the grid should be carefully handled [3-5].

In order to eliminate the leakage current of the transformerless grid connected inverter due to fluctuating common-mode voltage, a lot of in-depth research works have been conducted [3-10], where the new common mode clamping techniques and freewheeling branches have been proposed to generate constant common mode voltage and to separate the PV array from the grid in the freewheeling period. Several proposed topologies such as H5, H6, HERIC, oH5 and HBZVR can be divided into the AC bypass and the DC bypass based topology to disconnect the PV array and the utility grid in the freewheeling period. However, it is shown in the later section of this paper that the switch junction capacitances have unexpected influence on the leakage current generation. Based on the common-mode behavior of a transformerless inverter as explained in [6, 7], it is important to keep constant common mode voltage by using clamping branches instead of using only freewheeling branches to disconnect the PV module from the grid. Therefore, to completely eliminate the leakage current, common-mode clamping branches have been introduced in oH5, H6, HBZVR [8-11] topologies, where oH5 and H6 are DC bypass based topology. It can be seen that AC bypass based topologies bring lower conduction losses compared with the DC bypass based topologies due to the minimum switch count in the conduction path during power transfer mode. Furthermore, HBZVR topology cannot completely eliminate the leakage current [3].

In this paper, the leakage current generation mechanism in the single phase transformerless PV inverters is briefly analyzed, and then a modified HERIC-D topology is proposed. The proposed topology consists of an active

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

common mode clamping branch to keep constant common mode voltage in the whole switching period. In addition, the freewheeling path is modified in the AC side of the inverter to bring the benef t of the low-loss AC bypass based method.

The rest of this paper is organized as follows. Section II describes the principle of leakage current generation in a typical transformerless inverter. In section III, the proposed topology structure and operation principles are described in details. Section IV presents the simulation results. Finally, in section V the conclusions are presented to summarize the f ndings.

II. Leakage Current Generation Principle

A. Common Mode Behavior

The common mode characteristics of a full bridge transformerless inverter for symmetrical f lter inductor is explained in [12]. The typical single phase transformerless H5 voltage source inverter is illustrated in Fig. 1(a), which is used for defining common mode voltage and differential mode voltage and to explain the effect of switch junction capacitance for additional leakage current generation. According to the def nition in [12], the common-mode voltage V_{CMV} and differential-mode voltage V_{DM} are as follows.

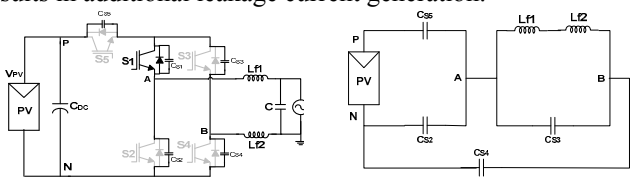
$$V_{CMV} = \frac{V_{AN} + V_{BN}}{2} \tag{1}$$

$$V_{DM} = V_{AN} - V_{BN} \tag{2}$$

where V_{AN} and V_{BN} are the voltages of terminals A and B to terminal N. The leakage current generation in a transformerless inverter depends on V_{AN}, V_{BN}, grid voltage, stray capacitance and f lter inductance [12].

B. Switch Junction Capacitance Effect

The switch junction capacitors have unexpected influence to the leakage current generation in the transformerless inverter circuits. In this section, the DC decoupled based transformerless H5 inverter is used to explain the leakage current generation due to the switch junction capacitance effect [6] as shown in Fig. 1(a). During the positive half cycle of the grid voltage, the switches S1, S4 and S5 are in on-state in the power transfer mode while the switches S2 and S3 are in off-state. The switches S4 and S5 are turned off in the next stage to go into the freewheeling mode. However, the antiparallel diode of the switch S3 does not conduct to change the state into the freewheeling mode. As a result, the junction capacitors of the switches S4 and S5 are charged, and the junction capacitors of the switches S2 and S3 are discharged. The charging and discharging operations during freewheeling stage and equivalent circuit are shown in Fig. 1(a) and Fig. 1(b) respectively. The voltage V_{AN} increases due to charging operation while the voltage V_{BN} decreases due to discharging operation. As a result, the common mode voltage V_{CMV} is not constant due to charging and discharging operation, which results in additional leakage current generation.



(a) (b)

Fig. 1 (a) Transient operation of the typical H5 topology, (b) Equivalent circuit [6].

In addition, during the freewheeling operation mode, the switches S1 and S3 are in on-state while all other switches are in off-state. In this mode, terminals A and B are on floating state, and further a resonant path is formed. As a result, additional leakage current is generated due to high frequency resonant phenomenon occurring in the circuit.

III. Structure and Operation Principle

A. Structure of the Proposed Topology

AC decoupling network based HERIC transformerless inverter circuits do not consist of complete clamping ability. Therefore, additional leakage current is generated due to a resonant path which is created by the switch junction capacitors like H5 topology. In order to avoid resonance circuit to be formed due to floating operation and to maintain constant common-mode voltage by effective clamping, an active clamping circuit has been added. The proposed circuit topology and gate drive signals are shown in Fig. 2 and Fig. 3 respectively. Two clamping switches (S_a and S_b) have been inserted between the midpoints of the two freewheeling branches (S5-S6 or S7-S8) and the midpoint of the series bus capacitors, respectively.

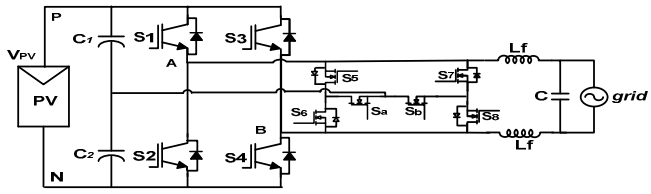


Fig. 2 Proposed topology.

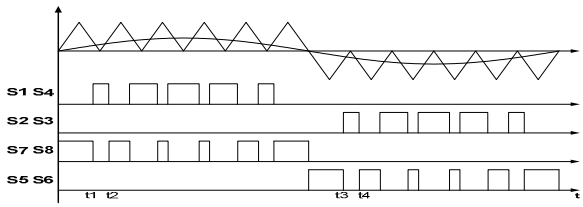


Fig. 3 Gate drive signals of the proposed topology.

B. Operation Principles and Analysis

Mode 1: Refer to Fig. 3 and Fig. 4(a). At t_1, the switches S_1 and S_4 are turned on and all other switches are off. At this stage, power flows from the PV to the grid. The current i_L increases linearly until t_2 and f ows through the switches S1, S4 and the filter inductor. The inductor current becomes

$$i_L(t) = \frac{V_{PV} - v_{grid}}{L}(t - t_1) \tag{3}$$

The common mode voltage (CMV) becomes

$$V_{CMV} = \frac{V_{AN} + V_{BN}}{2} = \frac{1}{2}(V_{DC} + 0) = \frac{V_{DC}}{2} \tag{4}$$

Mode 2: Refer to Fig. 3 and Fig. 4(b). At t_2, the main switches S_1 and S_4 are turned off. The freewheeling path switches S_7 and S_8 are turned on to create a freewheeling path during positive half cycle. The current i_L decreases and freewheels through the switch S8 the antiparallel diodes of the switches S_7 and the grid. The inductor current becomes

$$i_L(t) = \frac{-v_{grid}}{L}(t - t_2) \tag{5}$$

The common mode voltage becomes

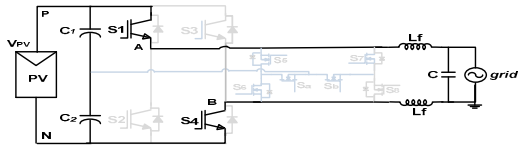
$$V_{CMV} = \frac{V_{AN} + V_{BN}}{2} = \frac{1}{2}\left(\frac{V_{DC}}{2} + \frac{V_{DC}}{2}\right) = \frac{V_{DC}}{2} \tag{6}$$

Mode 3: Refer to Fig. 3 and Fig. 4(c). In this stage, the switches S_2 and S_3 are turned on and all other switches are off. At this stage, power flows from the PV to the grid. The current i_L increases linearly until t_5 and f ows through the switches S_2, S_3 and the filter inductor. The inductor current becomes

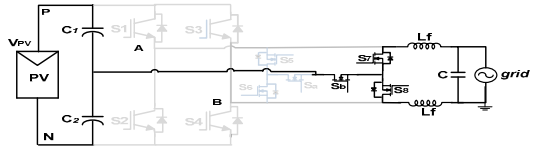
$$i_L(t) = \frac{V_{PV} - v_{grid}}{L}(t - t_3) \tag{7}$$

The common mode voltage becomes

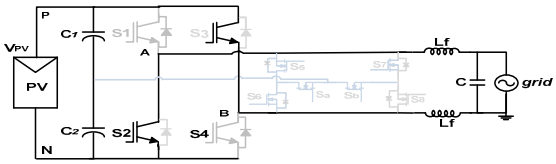
$$V_{CMV} = \frac{V_{AN} + V_{BN}}{2} = \frac{1}{2}(0 + V_{DC}) = \frac{V_{DC}}{2} \tag{8}$$



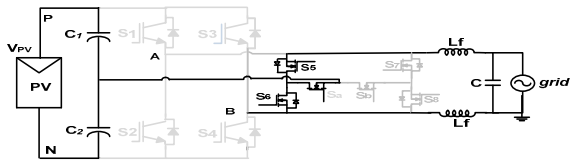
(a) Mode 1- conduction mode.



(b) Mode 2- freewheeling mode.



(c) Mode 3- conduction mode.



(d) Mode 4- freewheeling mode.

Fig. 4 Operation modes of the proposed topology.

Mode 4: Refer to Fig. 3 and Fig. 4(d). At this stage, the main switches S_2 and S_3 are turned off. The freewheeling path switches S_5 and S_6 are turned on to create a freewheeling path during negative half cycle. The current i_L decreases and freewheels through the switch S6, antiparallel diodes of the switches S5 and the grid. The inductor current becomes

$$i_L(t) = \frac{-v_{grid}}{L}(t - t_4) \tag{9}$$

The common mode voltage becomes

$$V_{CMV} = \frac{V_{AN} + V_{BN}}{2} = \frac{1}{2}\left(\frac{V_{DC}}{2} + \frac{V_{DC}}{2}\right) = \frac{V_{DC}}{2} \tag{10}$$

It is clear that the V_{CMV} is completely clamped to half of the bus voltage during the entire switching period by turning on the active clamping switches.

C. Operation Principle of the Improved Clamping Branch

The clamping switch S_b and the bidirectional switches S5-S6 are turned on when the voltage of freewheeling path drops. Therefore, the common mode voltage is clamped to half of the DC-link voltage through midpoint voltage of the series bus capacitors, which can successfully eliminate the further leakage current generation by resonant circuit effect. Similarly, by turning-on the gate signals of the clamping switch S_a and bidirectional switches S7-S8 when the voltage of freewheeling path increases, the additional leakage current generation can be eliminated. In other words, when the freewheeling voltage is higher than the half of the DC-link voltage, the switch S_a is turned on to clamp the freewheeling path voltage to $V_{DC}/2$. The common-mode current f ows through the clamping switch S_a. On the other hand, when the freewheeling path voltage is less than half of the DC-link voltage, the switch S_b is turned on to increase the freewheeling path voltage to $V_{DC}/2$. The common-mode current f ows through the clamping switch S_b. As a result, constant CMV is achieved in the freewheeling path during all switching modes and the leakage current is completely eliminated.

IV. Results

In order to compare the overall performance of the proposed topology with existing H5, H6, HERIC and HBZBR topologies, simulation studies have been carried out using MATLAB/Simulink and PSIM with the same parameters. Figs. 5-10 show the simulation results of the output voltage, leakage current and common-mode voltage waveforms for different topologies. Refer to Fig. 5(a), V_{CMV} is not constant and large oscillations are observed up to 350 V for H5 topology. As a result, high leakage current is found as shown in Fig. 5(b). Refer to Fig. 6(a), large oscillations are also observed in CMV for HERIC topology. Therefore, the leakage current is not effectively eliminated as shown in Fig. 6(b). The existing HBZBR topology utilizes common mode clamping technique. However, oscillatory V_{CMV} is observed with the magnitudes up to 500 V as shown in Fig. 7(a). As a result, the leakage current is not completely eliminated which is shown in Fig. 7(b). Refer to Fig. 8(a) and Fig. 8(b), constant V_{CMV} and reduced leakage current are observed for H6 topology. However, four switches conduct during power transfer mode, which reduces the inverter efficiency.

In contrast, the proposed HERIC based inverter topology can effectively eliminate the leakage current during the entire switching period by clamping the common-mode voltage to the midpoint of the series bus capacitors using active clamping technique. The proposed inverter output voltage, output current, V_{CMV} and leakage current are shown in Figs. 9(a) to 10(b) respectively. The V_{CMV} is completely clamped to half of the bus voltage during the entire switching period by turning on the active clamping switches. Therefore, the leakage is eliminated. In addition, the galvanic isolation of the proposed inverter topology is achieved on the AC side of the inverter, which reduces the conduction losses compared with the DC-based galvanic isolation technique due to the reduced switch count in conduction at any switching interval. The output current f ows through only two power switches in the freewheeling and power processing interval. The performance comparison among different topologies has been summarized in Table I for 2 kW output power to demonstrate the effectiveness of the proposed transformerless inverter topology. It is clear that the proposed topology gives the best overall performance.

V. Conclusions

A modified HERIC based topology with improved CMV clamping and freewheeling branches for transformerless grid connected inverter has been proposed in this paper. Besides, a comprehensive review, comparison and analysis of several

recently proposed single phase transformerless PV inverter topologies have been presented. The derived topology is an excellent candidate with improved conversion efficiency and minimized leakage current. The topology is obtained by adding six MOSFETS in the AC side of the converter which achieves lower conduction loss compared with other existing diode based topologies and DC-decoupling topologies. The performances of the derived topology in terms of common mode characteristics, leakage current, total harmonic distortion and efficiency are analyzed and compared with those of existing H5, H6, HERIC and HBZVR topologies. It can be concluded that the proposed topology provides the best performances and is suitable for transformerless single phase grid connected photovoltaic inverter.

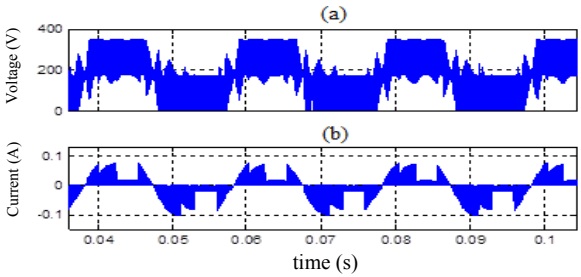


Fig. 5 H5 topology: (a) Common mode voltage, (b) Leakage current.

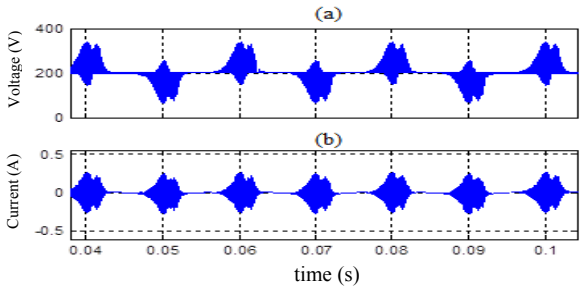


Fig. 6 HERIC topology: (a) Common mode voltage, (b) Leakage current.

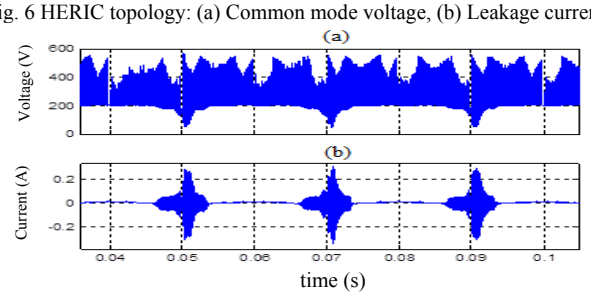


Fig. 7 HBZVR topology: (a) Common mode voltage, (b) Leakage current.

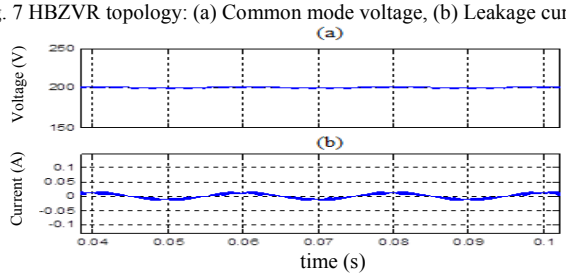


Fig. 8 H6 topology: (a) Common mode voltage, (b) Leakage current.

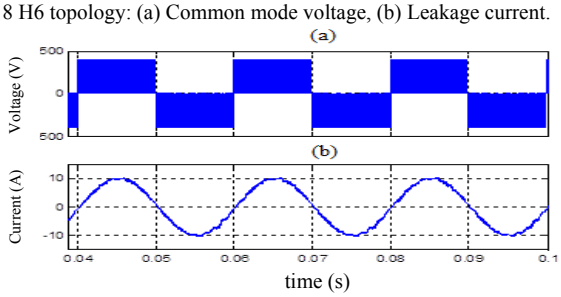


Fig. 9 Proposed topology: (a) Output voltage, (b) Output current.

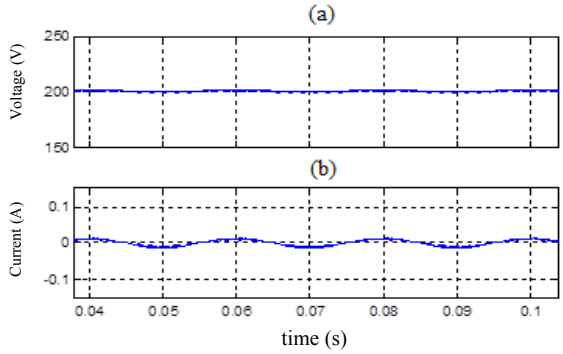


Fig. 10 Proposed topology: (a) Common mode voltage, (b) Leakage current.

TABLE I
PERFORMANCE COMPARISON OF TRANSFORMERLESS INVERTER TOPOLOGIES

Topologies	H5	H6	HERIC	HBZVR	Proposed
Leakage current (mA, RMS)	33	7.7	75	19	7.7
Efficiency (%)	98	97.5	98.2	98.1	99
THD (%)	3.87	5	3.2	2.8	3.5
Switch count during power transfer mode	3	4	2	2	2
Active switch losses (W)	22.8	27.4	17.13	19.6	16.58

REFERENCES

[1] M. R. Islam, Y. G. Guo, and J. G. Zhu, “A multilevel medium-voltage inverter for step-up-transformer-less grid connection of photovoltaic power plants,” *IEEE J. Photovoltaics*, vol. 4, no. 3, pp. 881-889, May 2014.

[2] M. R. Islam, Y. G. Guo, Z. W. Lin, and J. G. Zhu, “An amorphous alloy core medium frequency magnetic-link for medium voltage photovoltaic inverters,” *J. Applied Physics*, vol. 115, pp. 17E710-1-3, 2014.

[3] T. K. S. Freddy, N. A. Rahim, H. Wooi-Ping, and C. Hang Seng, "Comparison and analysis of single-phase transformerless grid-connected PV inverters," *IEEE Trans. Power Electron.,* vol. 29, pp. 5358-5369, Oct. 2014.

[4] N. Vazquez, M. Rosas, C. Hernandez, E. Vazquez, and F. Perez-Pinal, "A new common-mode transformerless photovoltaic inverter," *IEEE Trans. Ind. Electron.,* vol. 62, pp. 6381-6391, Oct. 2015.

[5] M. Islam, N. Afrin, and S. Mekhilef, "Efficient single phase transformerless inverter for grid-tied PVG system with reactive power control," *IEEE Trans. Sustainable Energy,* vol. 7, pp. 1205-1215, Jul. 2016.

[6] W. Li, Y. Gu, H. Luo, W. Cui, X. He, and C. Xia, "Topology review and derivation methodology of single-phase transformerless photovoltaic inverters for leakage current suppression," *IEEE Trans. Ind. Electron.,* vol. 62, pp. 4537-4551, Jul. 2015.

[7] H. F. Xiao, K. Lan, and L. Zhang, "A quasi-unipolar SPWM full-bridge transformerless PV grid-connected inverter with constant common-mode voltage," *IEEE Trans. Power Electron.,* vol. 30, pp. 3122-3132, Jun. 2015.

[8] H. Xiao, S. Xie, Y. Chen, and R. Huang, "An optimized transformerless photovoltaic grid-connected inverter," *IEEE Trans. Ind. Electron.,* vol. 58, pp. 1887-1895, Nov. 2011.

[9] B. N. Alajmi, K. H. Ahmed, G. P. Adam, and B. W. Williams, "Single-phase single-stage transformer less grid-connected PV system," *IEEE Trans. Power Electron.,* vol. 28, pp. 2664-2676, Jun. 2013.

[10] Y. Gu, W. Li, Y. Zhao, B. Yang, C. Li, and X. He, "Transformerless inverter with virtual DC bus concept for cost-effective grid-connected PV power systems," *IEEE Trans. Power Electron.,* vol. 28, pp. 793-805, Feb. 2013.

[11] T. Kerekes, R. Teodorescu, P. Rodriguez, G. Vazquez, and E. Aldabas, "A new high-efficiency single-phase transformerless PV inverter topology," *IEEE Trans. Ind. Electron.,* vol. 58, pp. 184-191, Jan. 2011.

[12] H. Xiao and S. Xie, "Leakage current analytical model and application in single-phase transformerless photovoltaic grid-connected inverter," *IEEE Trans. Electromagnetic Compatibility,* vol. 52, pp. 902-913, Nov. 2010.

Towards Optimal Shallow ANN for Recognizing Isolated Handwritten Bengali Numerals

Al Mehdi Saadat Chowdhury,[1] and M. Shahidur Rahman

Department of Computer Science and Engineering, Shahjalal University of Science and Technology
Sylhet, Bangladesh

[1]saadat_cse@yahoo.com

***Abstract*—This work attempts to find the most optimal setting for shallow artificial neural network (ANN) for Bengali digit dataset. Recognition of handwritten Bengali numerals has recently gained much interest among researchers due to significant performance gain found in the recognition of English numerals using artificial neural network. In this work, a new dataset of 70,000 samples were created first by taking handwriting of 1750 persons where 982 persons were male and the rests were female. These individual image samples are then converted to grayscale, normalized, inverted and pickled to complete the data preprocessing step. Later this dataset was recognized using several artificial neural network settings where the most optimal setting being found to be one hidden layer with rectified linear unit activation and one output layer with softmax activation. In this study, we have given enough evidence to support our choice of these activations. We proposed the optimal number of neurons that can be used in the hidden layer is 700; also the optimal batch size is 200. The maximum accuracy on the test set has been found to be 96.05% which is a good result due to the complexity of the recognition task.**

***Keywords*—Artificial Neural Network, Bengali Numerals, Handwritten Numeral Recognition, Rectified Linear Unit, Softmax Activation.**

I. Introduction

Current research on Artificial Intelligence has shifted from state space search algorithms to machine learning algorithms due to the enormous success found by applying the later in tasks like classification, inference and recognition. One of the major machine learning models that gained huge popularity over time is Artificial Neural Network (ANN) based models that tries to imitate the working mechanism of a biological neuron found in most living animals. Bengali is one of the most spoken languages of the world and is the first language of Bangladeshi citizens. Recognizing Bengali handwritten numerals is thus of much interest to a broad range of people. Earlier, researchers have achieved significant progress in recognition of English handwritten digits mostly applied on MNIST datasets where the past progress was summarized in [9] by LeCun et al. They found that, a fully connected multi-layer artificial neural network with one input layer of 400 inputs, one hidden layer of 300 neurons and one output layer of 10 neurons achieved an error rate of 1.6% on the test set of MNIST dataset which is significantly better than most other methods. Motivated by the success of handwritten English digit recognition, recognition tasks of other languages like Roman, Chinese and Arabic script [5, 6, 8] were also proposed, but recognition of Bengali numerals have always remain as a challenge for the researchers. This is due to the major similarity among Bengali digits.

Some of the previous research on recognizing handwritten Bengali digits was based on feature extraction and neural network [7], but complex feature extraction approaches introduce added complexity to the recognition task. In [7], authors worked with a feature set of 76 features which they divided into two groups, where the first group contains 24 shadow features and 16 octant centroid features and the second group contains 36 longest run features. Dividing the data into feature sets and then recognition based on them adds extra complexity to the task.

Bashar et al. [4] developed a recognition system based on windowing and histogram technique. In [3], authors used evolutionary artificial neural network based recognition. In [6], authors used Dempster-Shafer (DS) theory to combine the classification decision found from two MLP based classifiers using two different feature sets [2]. Again all these works add additional complexity in classification task.

In this work, we have used artificial neural network to recognize handwritten Bengali numerals. Firstly, we've created a new dataset containing Bengali digits and find most optimal parameters of the ANN model that we used for the recognition task. We've used from one to four hidden layers and one output layer as our network architecture and found that the shallowest network of all, the single hidden layer network outperforms all other architecture setups.

The rest of the paper is organized as follows. Section II explains how we prepared our dataset for the recognition task. Section III presents our proposed network architecture and different choices for the activation and the target of reducing overfitting of the network. The result and discussion section in section IV, justifies our choice and shows the result of the proposed method in classification of Bengali handwritten digits. Finally, section V concludes the paper.

II. Dataset Preparation

The famous MNIST dataset used for the recognition of English digits has 70,000 samples divided into training set, validation set and test set containing 50000, 10000 and 10000 digit images respectively. Image size is 28 by 28 pixels having a total of 784 pixels and pixel intensity is in grayscale form in the range of [0,1]. Each of these sets is further divided into two elements. The first element is a matrix with rows representing images and columns representing pixels of the image. The second element is an integer number representing the true digits that the image represents. Since MNIST dataset is hugely popular for English digit recognition task, a similar dataset in Bengali would be equally better. One major problem of the datasets found in literature is, either the size of the dataset is not adequate for recognition or they are not available for general use without prior permission. For example, the most common dataset used in Bengali handwritten recognition is the CMATERdb 3.1.1, which has only 6,000 samples in it. Another Bengali handwritten numeral dataset is the ISI dataset which contains 23,392 samples but the major problem associated with the dataset is the images it contains are not

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

normalized. All these problems motivate us to collect our own dataset and prepare them by making them compatible with the MNIST dataset. That way we think that the dataset will be more accessible to the researchers around.

A. Collections of Data

To create a new dataset, we've collected handwritten Bengali numerals of 70,000 samples from 1750 persons, each have written the digits 0-9, a single digit for four times to capture the within class variance of each individual writer. A sample of the handwritten Bengali digits and their corresponding English counterpart is given in Fig. 1:

Fig. 1: Bengali Handwritten digits and their English counterparts

From those 1750 persons, 982 persons were male and the rest were female. Most of the writers were university students while some of the others were school students, businessman, government employees etc. One major distinction made from MNIST to our dataset is, we've isolated each digit with a size of 32x32 pixels instead of 28x28. This choice was made due to get an even scaling on a power of two. Thus, if we want to reduce the size to 16x16, this 32x32 size will be a better match than 28x28 sizes.

B. Preprocessing of the Data

The goal for the dataset preparation is to make it compatible with the MNIST dataset. For this reason, we have followed below mentioned steps:

1) Grayscale Conversion, normalization and Inversion: As a starting step, the image is converted from RGB to grayscale and this conversion results the pixels to span in a range of [0,255] which is then normalized to [0,1]. Thus the black pixels intensity becomes 0, and white pixels intensity becomes 1. Furthermore, we revert the image such that white pixels are 0, thus we obtained sparsity in the image which will help to make the future calculations easier.

2) Converting image from 2D to 1D numpy array: Since, our images are of size 32x32, as we convert our images from 2D to 1D, the array size becomes 1024.

3) Shuffling images: All of our images were collected such that, all zeros be placed first, then all ones and so on. Combining the dataset in a similar manner may overfit our result, because the network will first be trained on all zeros and see no other digits, and then it will be trained on all ones seeing no other digits again. For this reason, the whole dataset is shuffled so that an uneven number of each digit is put on all three sets.

4) Pickling and Compressing: To collect and combine the whole dataset in a single place, we have used one of the python's most famous serializer mechanism known as pickling which gives us a single pickle file of size more than 2GB. Then compressing the pickled file with gzip compression reduces the file size to just over 55MB. This reduction is done for faster transfer of the data through physical network.

The whole dataset is now prepared for further processing. This whole dataset is further divided into three sets as training set, validation set and test set containing 50,000 samples, 10,000 samples and 10,000 samples respectively. Number of samples thus matched exactly to the MNIST dataset. Hence, this dataset can be used with any algorithms that was being tested with MNIST.

III. Bengali Handwritten Numerals Recognition

In this section, optimal parameters for recognition of Bengali handwritten numerals using artificial neural network (RBHN-ANN) is proposed.

A. RBHN-ANN Architecture

Fig. 2 shows one instance of the artificial neural network with two hidden layers and one output layer to recognize handwritten Bengali numerals.

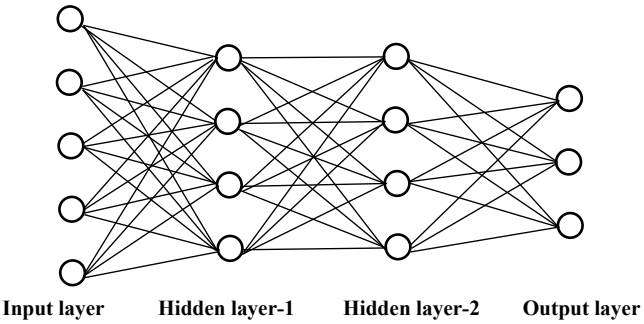


Fig. 2: Architecture of a two hidden layer neural network

The circles in all the layers except the input layer are artificial neurons that can take several inputs and produce a single output based on some computation given by the activation function. That single output is fed to all neurons in the next layer. Hence, our architecture is a fully connected architecture in all layers. Input layer circles are there to make completeness to the figure.

B. Activation Functions

Most common activation functions based on which the output of a neuron is calculated are linear activation, sigmoid activation, hyperbolic tangent activation, softmax activation, rectified linear unit activation etc. Given that, $X = x_0, x_1 ... x_{1024}$ are the inputs to the network, $W = W_{i,j}^k$ are the weights of the kth layer and $B = b_0, b_1 ... b_n$ are the bias of the neuron, the basic equations for these activations are listed in the following equations respectively.

$$z = \left(W^T X + B\right) \tag{1}$$

$$sigmoid(z) = \frac{1}{1+e^{-z}} \tag{2}$$

$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \tag{3}$$

$$softmax(x) = \frac{e^z}{\sum_j e^z} \tag{4}$$

$$relu(z) = \max(0, z) \tag{5}$$

C. Solution to the Overfitting Problem

Overfitting is a problem which occurs when the network try to fit every data of the training set. Normally, the data we get from the real world are noisy; hence, if the network predicts a hypothesis that passes through all data points in the training

set, we say that, our model will not generalize well to the test data then. For overcoming the overfitting problem, two techniques namely early stopping and regularization were employed in our work. They are described below.

1) Early Stopping: In this work, we have used a special type of early stopping mechanism that uses a strategy based on geometrically increasing the amount of patience when the validation cost is lower than the best validation cost times a specific improvement threshold.

2) Regularization: Regularization adds additional terms to the cost function that penalizes it when it is significantly larger or smaller, but remains unchanged if the cost is in the average region of the classification threshold. For this work, we've employed L_2 squared regularizer which is defined for $p=2$ as

$$\|\theta_p\| = \left(\sum_{j=0}^{|\theta|} |\theta_j|^p \right)^{\frac{1}{p}} \tag{6}$$

IV. Results and Discussions

In this section, we exhibit test results found from our proposed network architecture. The main target is to find the optimal open parameters for the recognition task.

A. Activation for the Output Layer

TABLE I
FINDING OPTIMAL OUTPUT LAYER ACTIVATION

No. of Hidden Neuron	Hidden Activation	Output Activation	Classification Accuracy on Test Set (Percentage) on n-Hidden Layers (HL)			
			1-HL	2-HL	3-HL	4-HL
50	Sigmoid	Sigmoid	44.82	26.17	20.0	15.11
	Tanh	Sigmoid	43.02	28.96	24.51	21.22
	Softmax	Sigmoid	11.63	10.18	10.18	10.18
	ReLU	Sigmoid	39.14	18.84	14.77	14.34
	Sigmoid	Tanh	9.81	9.18	9.81	9.81
	Tanh	Tanh	9.81	9.81	9.81	9.81
	Softmax	Tanh	9.81	9.81	9.81	9.81
	ReLU	Tanh	9.81	9.81	9.81	9.81
	Sigmoid	Softmax	**77.96**	**80.28**	**79.33**	**79.47**
	Tanh	Softmax	**89.74**	**93.56**	**93.3**	**93.11**
	Softmax	Softmax	**19.72**	**10.18**	**10.18**	**10.18**
	ReLU	Softmax	**82.84**	**94.07**	**93.01**	**92.8**
	Sigmoid	ReLU	9.81	9.81	9.81	9.81
	Tanh	ReLU	9.81	9.81	9.81	9.81
	Softmax	ReLU	9.81	9.81	9.81	9.81
	ReLU	ReLU	9.81	9.81	9.81	9.81

Finding the optimal activation function for any neural network is subject to the trial and error process. Hence, for our network architecture, we tried with a small number of hidden layers and observed performances by selecting various output layer activations. The findings are listed in Table I.

Table I suggests that, the best performance came when softmax activation is used as the output activation. That's why we test further by taking softmax activation as the activation for the output layer.

B. Activation for the Hidden Layer

Table II, along with Table I indicates that, the most prominent activation for hidden layer would be either tanh activation or rectified linear unit activation. Also our observation from Table II suggests that, instead of gaining performance by the increase of neurons, the tanh activation is slowly showing decreased performance while rectified linear unit is still showing good results. Hence, a natural choice for the hidden unit is the rectified linear unit activation.

TABLE II
FINDING OPTIMAL HIDDEN LAYER ACTIVATION

No. of Hidden Neurons	Hidden Activation	Output Activation	Classification Accuracy on Test Set (Percentage) on n-Hidden Layers (HL)			
			1-HL	2-HL	3-HL	4-HL
100	Sigmoid	Softmax	80.95	80.14	79.63	81.62
	Tanh	Softmax	82.26	83.97	94.54	94.16
	Softmax	Softmax	16.25	10.18	10.18	10.18
	ReLU	Softmax	**94.91**	94.76	93.97	92.75
200	Sigmoid	Softmax	77.36	79.15	79.59	80.86
	Tanh	Softmax	81.98	84.69	95.19	94.51
	Softmax	Softmax	11.56	10.18	10.18	10.18
	ReLU	Softmax	95.03	**95.23**	94.7	93.44
300	Sigmoid	Softmax	78.28	80.23	80.8	81.36
	Tanh	Softmax	81.83	83.2	95.08	94.96
	Softmax	Softmax	10.48	10.18	10.18	10.18
	ReLU	Softmax	95.13	**95.23**	94.72	93.56

C. Optimal Number of Hidden Layers and Neurons

One interesting observation found from Table III is that, neurons with even multiple of 100 works best for two hidden layer, odd multiple of 100 works best for one hidden layer. Another observation is that, after 700 neurons, notable performance gain is found only from one hidden layer network. And the observed best classification accuracy of 95.92% is obtained from one hidden layer with both 700 neurons and 1800 neurons. In this case, it is obvious that, choosing 700 hidden neurons rather than 1800 is the best decision, because more hidden neurons means more training time and more resource utilization. Hence, our final choice for the optimal number of hidden layer is one and optimal number of hidden neuron is 700.

However, we have still one unanswered question remains. The question is why adding more hidden layers such as three or four hidden layer network is not showing better performances, most obvious reason for this is the well-known vanishing gradient problem. Solution to the vanishing gradient problem is currently one of the major fields of neural network research.

D. Optimal Number of Batch Size

TABLE III
FINDING OPTIMAL NUMBER OF HIDDEN NEURONS

No. of Hidden Neurons	Classification Accuracy on Hidden layer activations (%) (with output layer activation fixed at softmax)				
	Activations	1-HL	2-HL	3-HL	4-HL
400	Tanh	82.07	82.74	95.23	95.05
	ReLu	95.18	**95.53**	94.86	94.07
500	Tanh	81.85	82.9	94.93	95.14
	ReLu	**95.54**	95.42	94.57	94.04
600	Tanh	82.0	82.94	95.04	94.92
	ReLu	95.48	**95.51**	94.92	94.18
700	Tanh	82.0	82.81	94.97	95.07
	ReLu	**95.92**	95.49	95.04	94.12
800	Tanh	81.88	82.45	95.2	95.1
	ReLu	95.25	**95.33**	94.96	94.26
900	Tanh	81.71	82.76	83.77	95.27
	ReLu	**95.79**	95.25	95.01	94.46
1000	Tanh	81.92	82.6	83.82	95.04
	ReLu	95.2	**95.48**	94.92	94.44
1200	Tanh	81.78	82.72	83.49	95.04
	ReLu	**95.55**	95.39	95.01	94.2
1400	Tanh	81.82	82.63	83.36	95.28
	ReLu	95.38	**95.41**	94.78	94.24
1600	Tanh	81.99	82.42	83.35	95.21
	ReLu	95.49	**95.56**	95.06	94.45
1800	Tanh	81.91	82.44	83.19	95.18
	ReLu	**95.92**	95.55	95.37	94.58
2000	Tanh	82.23	82.52	83.15	95.12
	ReLu	95.48	**95.52**	95.17	94.57

In training all these neural networks with one to four hidden layers and 50 to 2000 hidden neurons, we had fixed our batch size to 600. Thereafter, we have continued training our network with different batch sizes and listed the results in table IV. From table IV, finally we have found that the optimal number of batch size is 200 which give us a best classification accuracy of **96.05%**.

TABLE IV
FINDING OPTIMAL NUMBER OF HIDDEN NEURONS

Batch Size	Classification Accuracy (%)
50	95.97
100	95.97
200	**96.05**
300	95.82
400	95.99
500	95.42
600	95.92
700	95.4
800	84.56
900	95.42
1000	86.34

E. Comparison with Other Models

For evaluation, we have compared our artificial neural network based models with some well-established models whose results are listed in Table V. All these models are implemented using Pythons scikit-learn module, running on their generalized default settings.

It is apparent from Table V that, all other models fall short compared to the multi-layer neural network based model on Bengali handwritten digit classification.

V. CONCLUSION

In this work, we have created a new isolated Bengali handwritten digit dataset for recognizing them using artificial neural network based architecture. The main target of this research is to find the optimal parameter settings of this well-known architecture. We have tested from one to four hidden layers with sigmoid, softmax, tanh and rectified linear unit activation and found that the softmax activation for output unit, rectified linear unit activation for hidden unit, a single hidden layer with 700 neurons is the best setting for the recognition of Bengali handwritten digit dataset. The optimal number of batch size is 200. A best classification accuracy of 96.05% has been found in the process. Our future work encompasses a target to increase the classification accuracy and reduce training time.

TABLE V
FINDING OPTIMAL NUMBER OF HIDDEN NEURONS

Models	Classification Accuracy (%)
Logistic Regression	80.59
Decision Tree	72.58
K-Nearest Neighbors (K=3)	87.27
Nearest Centroids	69.47
Gaussian Naïve Bayes	64.82
Multinomial Naïve Bayes	72.7
Bernoulli Naïve Bayes	69.35
Support Vector Machine (SVM)	83.35
SVM with Linear Kernel	83.01
SVM with RBF Kernel	83.35
SVM with Sigmoid Kernel	10.18

ACKNOWLEDGMENT

We thank Asia Khanom, Tamanna Islam Chowdhury and Afaz Uddin from North East University Bangladesh for their enormous help in data collection and outcome listing.

REFERENCES

[1] G. E. Hinton, S. Osindero and Y. W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, pp. 1527-1554, 2006.

[2] M. A. H. Akhand, M. M. Rahman, P. C. Shill, S. Islam and M. M. H. Rahman, "Bangla handwritten numeral recognition using convolutional neural network," in *proceedings of 2nd Int'l Conf. on Electrical Engineering and Information & Communication Technology (ICEEICT)*, Jahangirnagar University, Dhaka, pp. 1-5, May 2015.

[3] M. M. R. Khan, S. M. A. Rahman and M. M. Alam, "Bangla handwritten digit recognition using evolutionary artificial neural network," in *proc. Of 7th International Conference of Computer and Information Technology (ICCIT)*, Dhaka, pp. 26-28, December 2004.

[4] M. R. Bashar, M. A. F. M. R. Hasan, M. A. Hossain and D. Das, "Handwritten bangla numerical digit recognition using histogram technique," *Asian Journal of Information Technology*, vol. 3, pp. 611-615, 2004.

[5] R. Plamondon and S. N. Srihari, "On-line and off-line handwritten recognition: A comprehensive survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 62-84, 2000.

[6] S. Basu, R. Sarkar, N. Das, M. Kundu, M. Nasipuri and D. K. Basu, *Handwritten bangla digit recognition using classifier combination through DS technique*, Lecture Notes in Computer Science, vol. 3776, pp. 236-241, 2005.

[7] S. Basu, N. Das, R. Sarkar, M. Kundu, M. Nasipuri and D. K. Basu, "An MLP based approach for recognition of handwritten bangla numerals," in *proc. Of 2nd Indian International Conference on Artificial Intelligence (IICAI)*, Pune, pp. 407-417, December 2005.

[8] U. Pal, C. B. B. Chaudhuri and A. Belaid, "A system for bangla handwritten numeral recognition," *IETE Journal of Research*, vol. 52, no. 1, pp. 27-34, 2006.

[9] Y. LeCun, et al., "Comparison of learning algorithms for handwritten digit recognition", in *F. Fogelman-Soulie, P. Gallinari (Eds.), Proceedings of the International Conference on Artificial Neural Networks*, Nanterre, France, 1995, pp. 53-60.

Traffic Sign Recognition Using SURF: Speeded Up Robust Feature Descriptor and Artificial Neural Network Classifier

Md. Zainal Abedin[1], Prashengit Dhar[1] ,Kaushik Deb[2]
[1]University of Science and Technology Chittagong(USTC),Chittagong,Bangladesh , [2]Chittagong University of Engineering andTechnology (CUET),Chittagong,Bangladesh
jakcse99@gmail.com, nixon.dhar@gmail.com

***Abstract*—Traffic Sign Recognition (TSR) system is a vital component of intelligent transport system. It plays an important role by enhancing the safety of the drivers, pedestrians and vehicles as traffic signs provide important information of the traffic environment of the road and assist the drivers to drive more safely and easily by guiding and warning. This paper represents road sign detection and recognition system based on speeded up robust feature (SURF) descriptor that is invariant to rotation, skew and occlusion of the sign and artificial neural network (ANN) classifier . The proposed system consists of three steps. In the first step the road scene image is segmented using color based thresholding algorithm followed by some post processing like morphological analysis and filtering using region properties to extract the road sign as a blob of region. Next, shape verification is carried out by using translation, scaling and rotation invariant features vector named Distance to Borders (DtBs) of the segmented blob .Finally the recognition of the traffic sign is implemented using ANN classifier with the training of SURF descriptor. The proposed system simulated on offline road scene images shows a high success rate in the recognition stage. The performance of the classifier is illustrated in terms of cross entropy, confusion matrix and receiver operating characteristic (ROC) curve.**

Keywords—Traffic Sign Recognition ; Speeded Up Robust Feature; Artificial Neural Network ; Segmentation; Distance to Borders ;Confusion matrix; Receiver Operating characteristic Curve ,Cross Entropy

I. INTRODUCTION

Traffic sign detection and recognition (TSR) is an important research topic that continuously keeps wider interest to the research in the field of intelligent transport system because of its application in the driver assistant system that helps to regulate the traffic, indicate the state of the road, guiding and warning drivers and pedestrians.

Due to the tremendous increase of road vehicles all over the world, the number of road accidents has also increased significantly. Among different causes of accidents, some major causes are ignorance of the road sign, occlusion of the road sign and distraction of the drivers. A driver assistant system like TSR helps the drivers and pedestrians to recognize and alert which help them to keep safe from road accidents. Traffic signs do not look same all over the world. Hence it is not possible to develop a universal TSR system. The TSR system depends on the policy and regulation of each country. There have been done a lot of researches to develop TSR system in the developed countries such as Europe, US. But, for Bangladeshi traffic signs a very few research works have been done. Hence we are strongly influenced to do some research in this regard in order to give a new dimension in the research of TSR for Bangladeshi road sign.

Most of the detection algorithm proposed in the literature used color information for segmentation of the image [1, 3, 4, 5, 6, 7, 8, 11]. For recognition feature vectors of segmented road sign is extracted. Recognition was implemented in [1] using SIFT, SURF and BRISK feature descriptor along with performance comparison. In [8] HOG feature vector was used to classify the road sign. In supervised learning paradigm the feature vector is fed to the classifier model and these models learn to discriminate the different road signs. Classifiers used for recognition of road sign include artificial neural networks (ANN) [7], support vector machine (SVM) [5, 9, 10, 11] and K-Nearest Neighbor (KNN) [1]. The purpose of this research is to develop a TSR system for Bangladeshi road sign based on the color information for segmentation, shape verification using distance to border vector and recognition by SURF descriptor and neural network classifier.

The organization of the rest of the paper is given as follows: Section II illustrates the proposed system using a flow chart. Section III represents the segmentation process. Then in section IV shape verification is explained. After this, the recognition along with performance analysis of the classifier is elucidated in section V. Finally, conclusion and future works are explicated in section V1.

II. SYSTEM OVERVIEW

In this paper we present a traffic sign recognition system (TSR) for the road sign of Bangladeshi traffic environment those have red rim and triangular shape categorized as warning signs. The flowchart of algorithm for the proposed system is depicted in Fig.1. The system consists of three main components: segmentation, shape verification and recognition. For segmentation the input image is normalized in the red channel and then this normalized image is segmented using the color based thresholding technique. After this, some morphological operations are applied to make fine the binary image. In the binary image, not all the regions are region of interest. Hence some sorts of filtering are applied using area and aspect ratio limit parameters after the connected component labeling of the binary image. Then, to reject the noisy shape, verification of the candidate's blob is carried by robust distance to border vectors. In the final stage, the recognition of the traffic sign is implemented using the SURF

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

descriptor and the ANN classifier with scaled conjugate gradient back propagation learning algorithm.

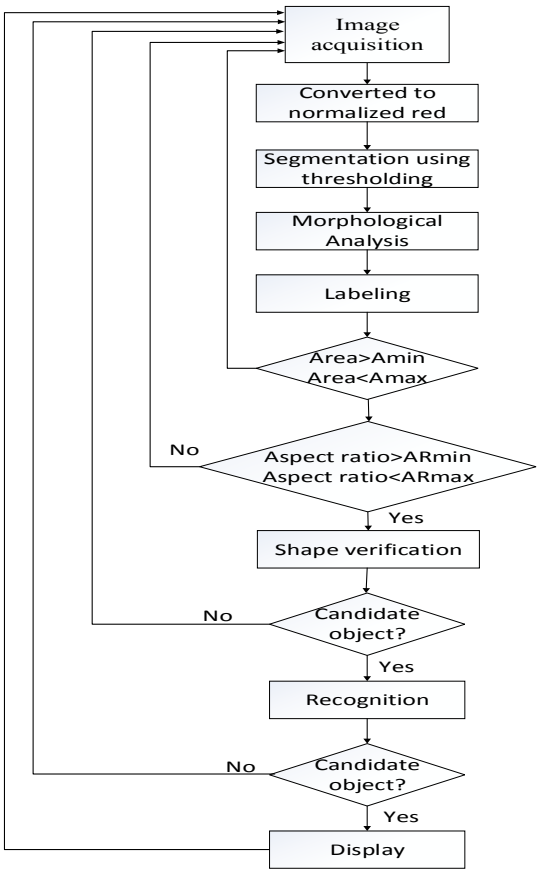


Fig.1. Algorithm for Proposed TSR system

III. SEGMENTATION

Detection of traffic sign is like finding the region of interest (ROI) that may contain road signs. In this system color based segmentation is used to extract possible ROI from the image. As we concentrate on only the red color sign, the border color rim of the sign is red which can be easily identified. As RGB consists of red, green and blue components, red component can be easily distinguished from an image. In our system, prior to apply color based thresholding the road scene image is converted into normalized red which is the ratio of the red channel to the sum of all channels. Fig.2 (b) shows the normalized red image of original image. Pixel values are higher for red and lower for others. After normalization in the red channel, the image is segmented using thresholding methods. The thresholding value is chosen by analyzing the histogram of the normalized image. Morphological operation with a squared structuring element is performed after color segmentation to fill small gaps and missing boundaries and to filter many spark noise which sometimes spreads on the image. Morphological closing operation is used to recover the gapes smaller than 5 consecutive pixels on the segmented objects. For spark noise such as single dots or small blobs, morphological filter is an effective method. Fig.2 (c) shows the binary image after morphological operations.

Then the connected components of the image are needed to label to find the region properties. These features can be used to filter the binary image to get the desired region of interest. In this system we use eight connected components for labeling. To get the desired region, geometrical features such as area and aspect ratio are calculated. Regions that satisfy the properties mentioned in Table1 are considered as ROI.The parameters used as filtering criteria are determined empirically from the image dataset. After the segmentation process, the sign area is extracted using the bounding box parameters and this segmented area is sent to the shape verification and recognition stage to derive the DtBs and SURF features vector respectively. Fig.2 depicts the complete segmentation process.

TABLE 1 PROPERTY USED FOR FILTERING

Parameter	Minimum	Maximum
Area	300	850
Width	35	53
Height	35	49

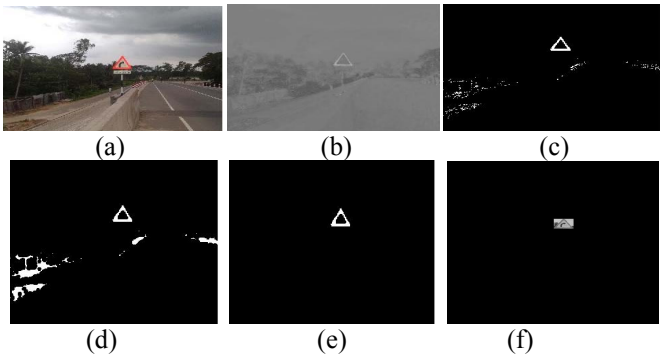


Fig.2. Steps in sign detection a) original image b) normalized image c) after thresholding d) after morphological operation and e) after filtering and f) ROI extracted using Bounding Box

IV. SHAPE VERIFICATION

The blobs which are extracted after the segmentation are classified and verified in this stage to discard noisy shape In order to classify the shape, in this paper we use the distant to border (DtBs) features of the candidate blob as introduced in [10]. DtBs are distances from the external edge of the blob to its bounding box. Fig.3 (a) shows these distances for a triangular shape where D1, D2, D3 and D4 are the left, right, upper and bottom DtBs respectively. Distance to Borders (DtBs) is a very efficient way to recognize the shape of road sign as it is invariant to translation, rotation and scaling of the blob [10]. Due to the use of four distance vector for each blob this method is robust to occlusions. Fig.4a shows the distance vectors for segmented sign area of Fig.3 (b). The blobs that are successfully tracked are processed for features extraction to recognize sign.

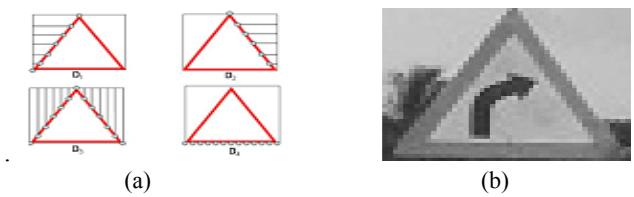


Fig.3.a) DtBs of triangular shape and b) ROI after segmentation

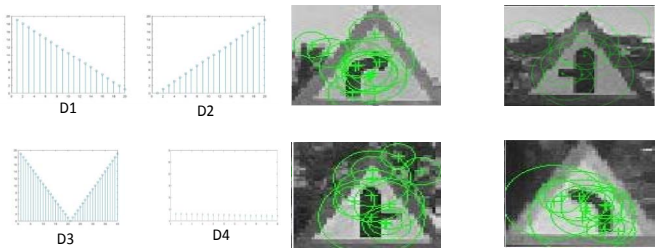


Fig.4. (a) DtBs vector of 3b (b) Interest Point for SURF

V. RECOGNITION

A. *Speeded up Robust Features (SURF):*

The most widely and accepted SIFT feature descriptor exhibits high computational cost. In 2006, H. Bay et al. [] proposed a speeded-up variant of SIFT which was termed as SURF (Speeded up Robust Features). SIFT computes a 128 dimensional feature descriptor of neighbors of each key point based on histogram of local gradients On the other hand SURF depends on sums of Haar wavelet components producing a 64 dimensional feature descriptor. Instead of difference of Gaussian SURF is approximated by laplacian of Gaussian with a box filter. Because of using integral images which allow a very fast computation of box type filters the computation time has reduced significantly in SURF. The detection of interest point is based on the determinant of Hessian matrix. Based on local maxima of determinant of the Hessian matrix at each location in the image over different scales the interest point is selected. For each key point an orientation is associated to allow rotation invariance. The orientation is calculated by computing the Haar wavelet response in a circular neighborhood of each key point. The wavelet responses are plotted as points with the strength of horizontal and vertical responses on the two axes. All responses within sliding window are calculated and the leading orientation is selected as the orientation of the key point. The SURF descriptors are calculated by considering a square box filter of size 20 around each key point oriented along its respective orientation. The region around the key point is divided into sub-regions of size 4 x 4 and for each block Haar wavelet response in horizontal and vertical directions is computed. Each sub-region is then represented by a four dimensional descriptor which includes the sum of wavelet responses (d_x and d_y) and the sum of the absolute values of d_x and d_y as summarized in Eq. 1 [1][2][3].

$$v = (\sum d_x , \sum d_y , \sum |d_x| , \sum |d_y|) \quad (1)$$

It is claimed that SURF is invariant to scale, rotation and skew of the sign [3].Fig.4b represents the location of key point in ROI for feature extraction.

B. *Dataset*

To train the classifier it is crucial to create a dataset on all possible environments to avoid misclassification. However gathering sufficient set of real data is difficult. For this we create syntactically road images from the real image by changing brightness, contrast, adding noise and blurring. The real images have been collected from the different high way roads of Bangladesh in different illumination conditions and different weather conditions. The dataset consists of four different road signs with 50 images per sign and hence a total of about 200 images. Some of the images are presented in Fig.5.

Fig.5. some examples of considered road signs

C. *Training And Classification*

For recognition we use artificial neural network which is extensively used for pattern classification task. To train the network SURF features are computed for every segmented region. The learning environment and parameters are presented in Table 2.

TABLE 2 LEARNING ENVIRONMENT AND PARAMETRES

Parameters	Value
Learning tool	Neural pattern recognition app of MATLAB 2O15A
Learning environment	Supervised
Learning algorithm	Scaled Conjugate Gradient Backpropagation algorithm
Transfer function in Hidden/Output layer	Sigmoid/ Softmax
Data Division	Random(70% for training, 15% for validation and 15% for testing)
Network architecture	Multilayer(64 x10x4) with one hidden layer
Performance	Confusion matrix ROC, Cross Entropy
No. of Class	Four
Size of Features vector(Input)	64 x 200

D. *Performances of the classifier*

In the field of machine learning a confusion matrix is a measurement that visualizes the performance of supervised learning algorithm. Each column of the matrix represents the instances in a predicted class while each row represents the instances in an actual class. The confusion matrix of the classifier is depicted in Fig.6.

A receiver operating characteristic (ROC) curve is a graphical plot that illustrates the performance of a classifier. The curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. The ROC curve of the classifier is depicted in Fig.7.

From the confusion matrix it is evident that the true positive rate in the all class is very satisfactory and the false positive rate is very low. In the test case of the classifier, true positive rate for class 1 is 85.7%, for class 2 and 4 it is 100% and 90% for class 3. Finally the combined true positive rate is 97.0% and false positive rate is 3.0%.

The area under the ROC curve is an indication of the performance of the classifier. The maximum area under the ROC curve is 1 which represents that true positive rate is 100%. It is observed from the ROC curve in Fig.7 that the areas under the ROC curves for all classes are very close to 1.Hence the performance of the classifier is quite satisfactory.

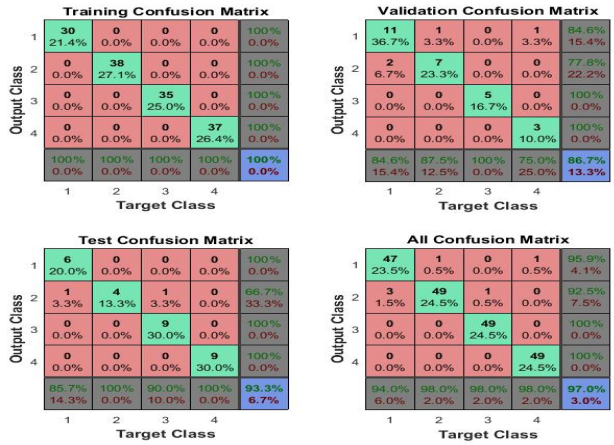


Fig.6. Confusion matrix of the classifier

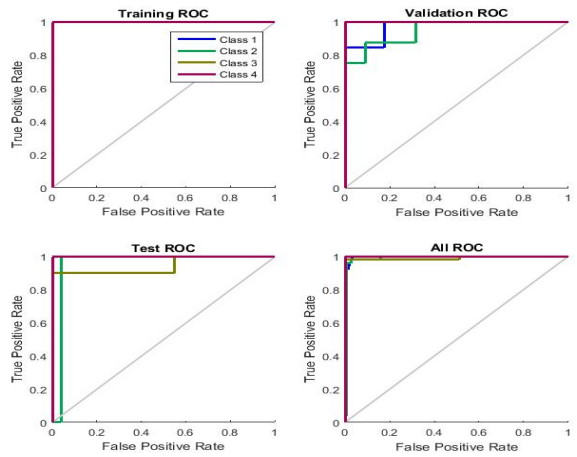


Fig. 7. ROC Curves of the Classifier

Cross entropy calculates a network performance on given targets and outputs. Minimizing cross entropy results in good classification. Fig.8 illustrates the cross entropy of the classifier.

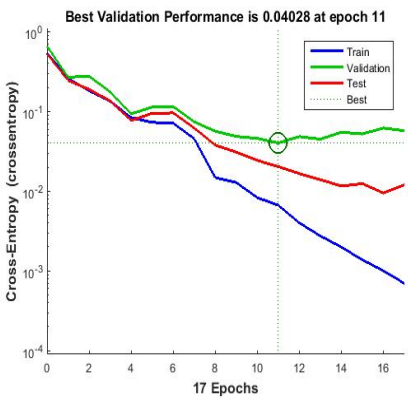


Fig. 8. Cross Entropy of the Classifier

VI. CONCLUSION AND FUTURE WORK

This paper illustrates a new detection and recognition approach of Bangladeshi road sign which is invariant to scaling, rotation and skewing of the sign due to robust SURF descriptor. The system is developed for red rim triangular road sign and simulated for four types of sign. Here color based segmentation is used along with filtering by morphological and geometrical properties. The shape verification is accomplished using robust distance to border vector. A precise road sign recognition system with low false positive rate is very crucial to contribute more safety and efficiency. In this regard, the experimental results represent that this system carries significant classification rate. From the confusion matrix it is evident that the overall correct classification rate of the ANN classifier is 97.0 % and the areas under the ROC curve are almost near 1 for all class.

In future work we want to take into account a number of challenges for further improvement of the proposed system. In this regard we are planning to address the following issues: 1) to recognize all types of road signs 2) to recognize array of road signs on the scene 3) to recognize signs when images are taken under night illumination and 4) to develop this system in real time.

REFERENCES

[1] Zumra Malik and Imran Siddiqi,"Detection and Recognition of Traffic Signs from Road Scene Image",12th IEEE International Conference on Frontiers of Information Technology,pp.330-335,2014 .

[2] Bay, H., A. Ess, T. Tuytelaars, and L. Van Gool. "SURF:Speeded Up Robust Features" ,*Computer Vision and Image Understanding (CVIU)*.Vol. 110, No. 3, pp. 346–359, 2008

[3] Yan HanKushal; Virupakshappa & Erdal Oruklu ,"Robust traffic sign recognition with feature extraction and k-NN classification methods",IEEE International Conference on Electro/Information Technology (EIT),2015 pp. 484 - 488

[4] Broggi, A., Cerri, P., Medici, P., Porta, P. P., & Ghisio G., "Real time road signs recognition"., IEEE Intelligent Vehicles Symposium, 2007, pp. 981-986.

[5] Maldonado-Bascon, S., Lafuente-Arroyo, S., Siegmann, P., GomezMoreno, H., & Acevedo-Rodriguez, F. J.,"Traffic sign recognition system for inventory purposes", In Intelligent Vehicles Symposium, 2008, pp. 590-595.

[6] Huang H., Chen, C., Jia, Y. & Tang, S., "Automatic detection and recognition of circular road sign", In Proc. of International Conference on Mechtronic and Embedded Systems and Applications, 2008, pp. 626-630.

[7] De La Escalera, A., Moreno, L. E., Salichs, M. A., & Armingol, J. M., "Road traffic sign detection and classification", IEEE Transactions on Industrial Electronics, 44(6), 848-859.

[8] Soumenn Chakraborty and Kaushik Deb,"Bangladeshi Road Sign Detection Based on YCbCr color model and DtBs Vector", 1st international Conference on Computer and Information Engineering, 26-27 November, 2015 .

[9] Jack Greenhalgh and Majid Mirmehdi,"Real-Time Detection and Recognitionof Road Traffic Signs" ,IEEE Transactions on Intelligent Transport Systems, VOL. 13, NO. 4, December 2012 .

[10] Saturnino Maldonado-Bascón , Sergio Lafuente-Arroyo, Pedro Gil-Jiménez ,Hilario Gómez-Moren and Francisco López-Ferreras," Road-Sign Detection and Recognition Based on Support Vector Machines" IEEE Transactions on Intelligent Transport Systems, , VOL. 8, NO. 2, June 2007 .

[11] C. G. Kiran, L. V. Prabhu, R. V. Abdu, and K. Rajeev, "Traffic sign detection and pattern recognition using support vector machine," inProc. ICAPR, 2009, pp. 87–90 .

A Memristor-based 6T1M Hybrid Memory Cell without State Drift during Successive Read

Mohammad Nazmus Sakib*, Rakibul Hassan and Satyen Biswas
Department of Electrical and Electronic Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh
*shuvo.sakib@gmail.com

Abstract—**This research proposes a new 6T1M memory structure, which is designed using 6 transistors and a single memristor. The proposed cell is capable of storing data in bidirectional fashion. This memory model is not only space efficient but also operates faster than other conventional structures. It does not require any refresh operation as it prevents state drift during successive read operation. Extensive simulation results by employing LTspice demonstrates the excellent performance and competency in terms of write time, read time and power dissipation of the proposed model .**

***Index Terms*— Memristance; Hybrid; Bidirectional; State Drift; Precharge.**

I. Introduction

CMOS based conventional memory technologies have reached its zenith to achieve higher density with the continuous constriction in nano scale regime [1-3]. Moore scaling will be difficult to meet as the end of scaling paradigm is imminent with the current shrinking rate of CMOS. Very recently, a highly capable despite of its infinitesimal size has been demonstrated which is the first physical realization of the elusive fourth fundamental circuit element, memristor. Because of the less power consumption along with the occupation of less die area, memristor has got the potential to supplant CMOS transistors in next generation computers [4-5]. Memristor has got some fascinating features like non volatility, nonlinearity at the boundaries, high packing density in crossbar arrays which have spurred extensive attention to the researchers to design memristor based nonvolatile memory arrays. Moreover, fabrication compatibility of memristor with CMOS technology boosted the hybrid approach of designing nonvolatile memory.

A novel 6T1M hybrid memory cell is proposed in this research. This model is the modified version of our previously proposed 7T1M [6] memory cell. The passive two terminal memristor is employed in this research to hold data in terms of resistance. The proposed circuit stores data bidirectionally and also does not require any refresh operation while successive read operation is performed. The main hindrance to design memristor based nonvolatile memory is the state drift phenomena in successive read which is referred to the change in memristor state from R_{OFF} to R_{ON} or vice versa. In the proposed cell, only one memristor is utilized to store data (0 or 1). So, there is a possibility to affect either a "0" or a "1" while retrieving data depending on the read pattern. If the read pattern is like the write 0 (1) pattern, stored 1 (0) can be affected after each read pattern and state drift phenomena can be occurred after some successive read operation. A new approach is provided to eradicate the state drift phenomena in this paper. The proposed circuit also outperforms some recently developed memristor based memories [7-11] in terms of die area, write and read times.

II. Background

The missing memristor was first postulated by Chua [12] in 1971 based on the circuit theoretical reasonings. Memristor is considered as the so called fourth fundamental circuit element because of its canonical nature [13] (i.e., missing static relationship between flux and electric charge). The missing relationship between flux and electric charge having memristance M of the memristor is $d\varphi=Mdq$, where φ and q are the flux and electric charge respectively. The resistance of the memristor which is the so called memristance does not change even if the supply voltage is removed. The physical model of memristor (see fig. 1) presented in [14] is composed of a two layer of TiO_2 thin film which is sandwiched between the top and bottom electrodes. The physical structure of the two electrodes is basically metal nanowires on platinum. State of the memristor can be changed according to the direction of current through the device or applied biasing (i.e., forward or reverse). R_{ON} and R_{OFF} state can be achieved by applying positive and negative voltage to the top electrode respectively. Oxygen vacancies drift from the doped (TiO_{2-x}) to the undoped (TiO_2) region or vice versa while the state of the memristor gets to the ON or OFF state respectively. The total resistance of the memristor is given by

$$M(t)=\left\{R_{ON}\frac{w(t)}{D}+R_{OFF}\left(1-\frac{w(t)}{D}\right)\right\} \quad (1)$$

Where R_{ON} and w(t) are the resistance and width of the completely doped region while R_{OFF} is the resistance of the

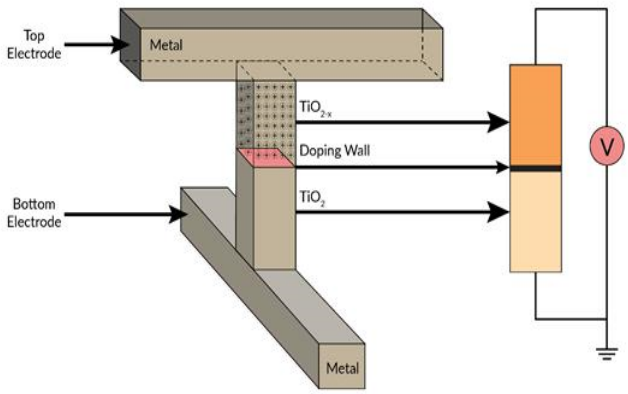


Fig. 1. Physical structure of memristor

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

completely undoped region. The parameter D denotes the thickness of TiO_2.

III. PROPOSED MEMORY CIRCUIT

The proposed bidirectional memory cell (see Fig. 2) which is hybrid in nature is composed of only six transistors for switching purpose and one memristor as a nonvolatile storage element. Among the six transistors used to design the bidirectional memory cell, two of them are PMOS transistors and the rest of them are NMOS transistors like the conventional 6T SRAM cell. But the operation of the proposed memory circuit is different from the conventional SRAM cell and does not need standby power to hold the binary data like the SRAM cell. In case of data extraction from one memristor, state drift phenomena is the main hindrance to design a stable hybrid memory cell. The bidirectional binary data storing of the proposed memory circuit is rapid and also prevents state drift for repetitive read.

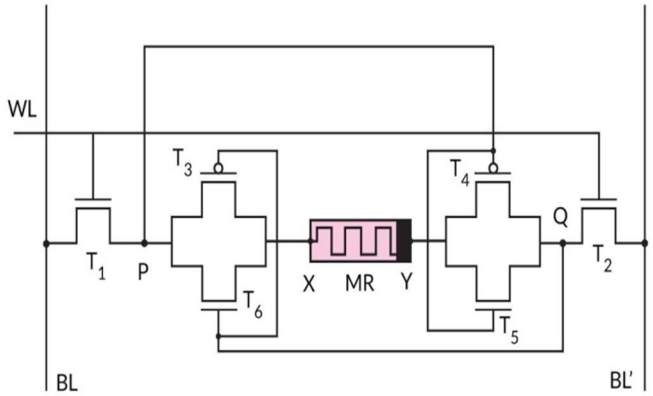


Fig. 2. Proposed 6T1M memory cell

A. Write Operation

The nonvolatile memristor gets the raw data to store through the bit line (BL) and inverse bit line (BL'). The word line (WL) selects the memory cell during data storing and extraction operation. The word line (WL) is set to high (VDD) to send the binary data to the memristor through the access transistors T_1 and T_2 for storing.

1. Write a "0"

When the bit line (BL) and inverse bit line (BL') are set to high and low respectively, current flows from node P to node Q that means from the undoped to the doped region of the memristor. Transistor T_3 and T_5 will be ON during this biasing. So, there will be a positive voltage from the undoped to the doped side of the memristor ensuring the full enhancement of the boundary of the undoped region over doped region. Thus, maximum memristance value will be achieved (i.e., R_{OFF} state) and a "0" will be written in the cell.

2. Write a "1"

The current flow during a write "1" operation is opposite to the one of writing a "0" operation as inverse biasing of writing a "0" needs to apply to get minimum memristance value. The memristor will be forward biased as the bit line (BL) and the inverse bit line (BL') are set to low and high respectively. Transistor T_4 and T_6 will be ON while storing a "1". The presence of negative voltage from the undoped to the doped side of the memristor will ensure the complete enhancement of the boundary of the doped region over the undoped region. Thus, minimum memristance value will be achieved (i.e., R_{ON} state) and a "1" will be stored in the cell.

Hence, bidirectional write operation is performed by the proposed circuit. The write time for the proposed memory circuit is determined by taking the time to change the memristance value from minimum (R_{ON} state) to maximum (R_{OFF} state) or vice versa.

B. READ OPERATION

Read operation starts by precharging the bit line (BL) and the inverse bit line (BL') to high (VDD) and low (GND) respectively. The biasing to extract binary data is identical to the one of writing a "0". This biasing to retrieve binary data eradicates the state drift problem during successive read operation. If the write "1" biasing would have been applied to retrieve the binary data, only "0" would get affected and R_{OFF} state would shift towards the R_{ON} state after every read pattern is applied. This is because there will be a large positive voltage to the doped region of the memristor which will enhance the boundary of the doped region after every read operation. But when write "0" biasing is applied to retrieve the data, a large positive voltage is present to the undoped region (i.e., X terminal) of the memristor which will not cause any state change. If a "1" is stored in the cell, the memristance can be changed while retrieving the "1" because of the opposite directional current flow through the memristor. But the memristance (R_{ON}) is constant as a result of the quick isolation from the bitline while retieving the data. Because of this quick isolation from the bit line (BL), there will not be enough voltage across the memristor to bias it to the R_{OFF} state. This isolation was done by adding a transistor in 7T1M cell [6] which can cause state drift during repetitive read. Detailed explanation of the read operation is carried out next.

As mentioned earlier, the bit line (BL) and the inverse bit line (BL') are precharged to VDD and GND respectively during read operation. So, the transistor T_3 and T_5 will be ON and the transistor T_4 and T_6 will be OFF while the read pattern is applied. Though, the PMOS transistor, T_3 gets OFF just after the starting of read operation and isolates the circuit from the bit line (BL) as the transistor T_6 is already OFF. If a "1" is stored in the memory cell, the voltage at node Q rapidly increases from zero to a large positive voltage as the memristance is lower. So, the source to gate voltage, Vsg of the PMOS transistor, T_3 becomes lower rapidly. The PMOS transistor, T_3 gets OFF as $Vsg < |Vt|$. Here, the threshold voltage $|Vt|$ is 0.16V. As a result of the isolation of the circuit from the bit line (BL), almost zero current flow through the memristor will be achieved. While retrieving a "0", same phenomena is noticed. But the time to have almost zero current through the memristor is higher than retrieving a "1". This is because of the larger memristance value which prevents the rapid increment of voltage at node Q. Fig. 3 depicts that the source to drain voltage (Vsg) of the PMOS transistor, T_3 gets lower as the read operation proceeds. It is also clearly depicting from Fig. 4 that the circuit gets isolated from the bit line (BL) during data retrieving operation. The stored data is detected by the disparate bit line voltage differences ($V_{BL\text{-}BL'}$) for "0" and "1".

The voltage difference across the memristor plays an important role in determining state change during repetitive read operation. Fig. 5 demonstrates the voltage difference across the memristor while extracting a "1" is much lower in comparison with the voltage difference across the memristor

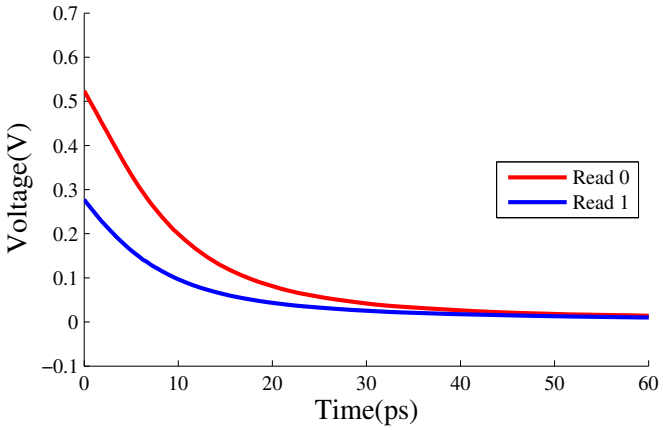


Fig. 3. Source to gate voltage of the transistor, T_3 during read operation

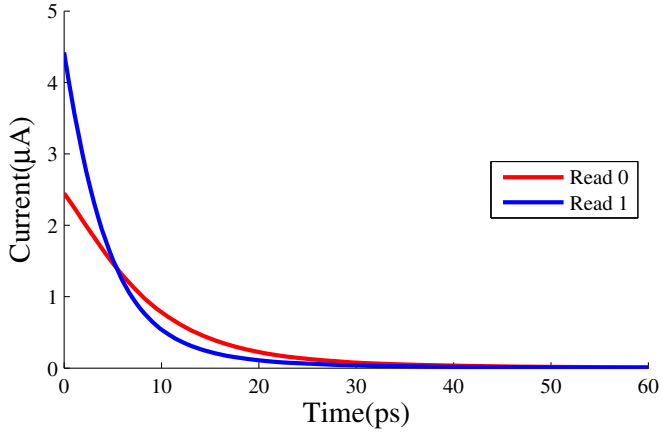


Fig. 4. Current through the memristor (undoped to the doped region of fig. 2) during read operation

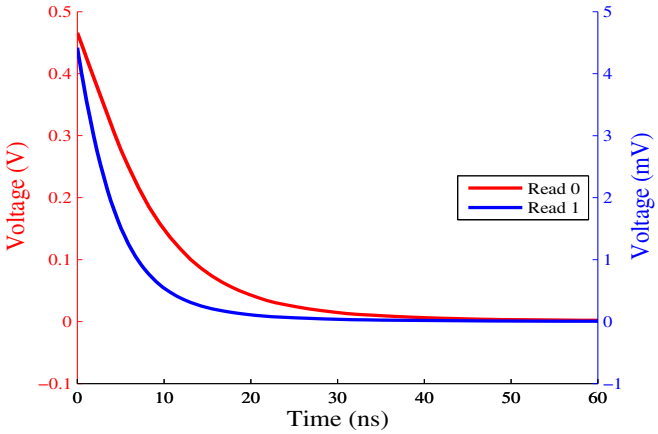
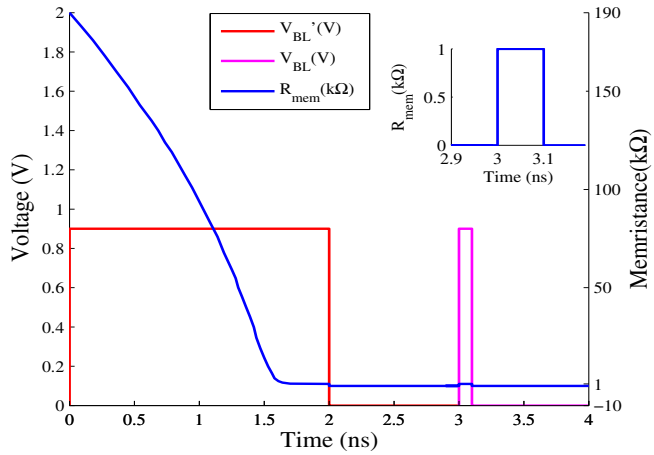


Fig. 5. Voltage across the memristor (V_x-V_y) during read operation

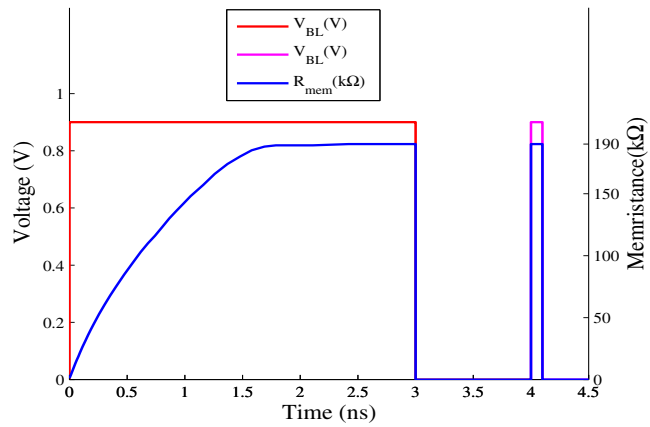
for "0". Moreover, the voltage drop across the memristor for a "1" extraction falls to almost zero abruptly after the starting of read operation. As mentioned earlier, zero will not get affected while the read pattern is identical to wrote "0" pattern. Despite having a high voltage drop across the memristor while retrieving a "0", the boundary of the undoped region cannot move further as it has already been enhanced to the highest level. Now, the negligible voltage drop across the memristor for a "1" extraction cannot change the memristance value. Thus, memristor state will not be perturbed ensuring the prevention of state drift phenomena.

Fig. 6 (a) & (b) delineates that change in memristance is absent for the proposed memory circuit by applying the read pattern hich is identical to write "0" pattern. In fig. 6 (a) the inverse bit line (BL') is high for 2 ns and a "1" is written in the cell as the memristor is biased to the R_{ON} state. From 2ns -3ns, both the bit line (BL) and the inverse bit line (BL') are set to low (i.e., no read or write pattern is given). From 3 ns-3.1 ns, read pattern is given to precharge the bitline to VDD. It is unequivocal from fig. 6(a) that no change in memristance is present while extracting a stored "1" which eradicates the state drift phenomena. The memristance of the memristor before and after the read pattern is applied while retrieving "1" is shown in the inset fig. 6(a). Fig .6 (b) shows the storing and retrieving a "0" in the proposed memory cell. It can also be seen from fig. 6 that zero standby power is required to hold the binary data as the oxygen vacancies get completely immobile when there is no pulse[12],[15]. This unique characteristic will cause a significant reduction of power dissipation in the proposed memory cell in comparison with the conventional SRAM cell and recently developed flash memories.

As mentioned earlier, the stored binary data is detected by the disparate bit line voltage difference for "0" and "1". Fig. 7 depicts the variation of bit line voltage difference for "0" and "1". The reason behind this dissimilar bit line voltage difference is the dissimilar memristance value for "0" and "1" (i.e., R_{OFF} and R_{ON}). If a "1" is already in the cell, the lower memristance value influences rapid voltage transfer from the bit line (BL) to the inverse bit line (BL'). On the other hand, sluggish transfer of voltage from the bit line to inverse bit line is noticed if a "0" is in the cell. From fig. 7, it is unequivocal



(a)



(b)

Fig. 6. Read and Write pattern and memristance for (a) "1" (b) "0"

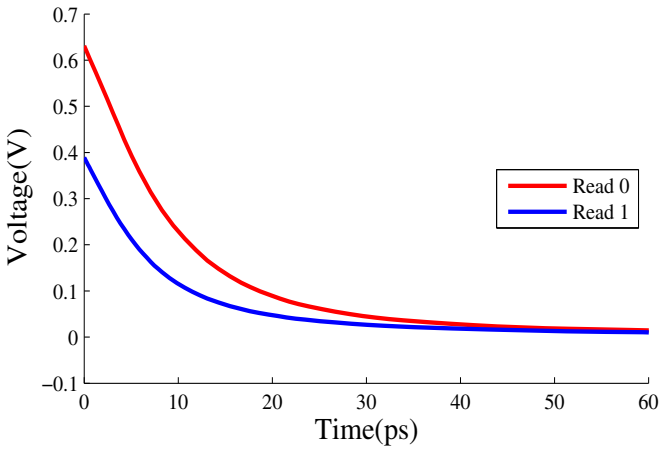


Fig. 7. Bitline voltage difference ($V_{BL-BL'}$) for retrieving "0" and "1"

that the bit line voltage difference for "0" is larger than "1" which is because of the sluggish abatement of the bit line voltage for a "0" extraction than a "1" extraction. The bit line voltage difference for retrieving "0" and "1" are 254 mV and 129 mV while the read time is 9 pS. The difference in voltage between the bit lines for "0" and "1" gets smaller if the read time is taken high.

IV. SIMULATION RESULTS AND COMPARISON

By employing LTspice as a simulation tool, extensive simulation experiments have been carried out to evaluate the competency of the proposed memory cell. 32 nm [16] feature size of CMOS technology has been employed by setting L_{eff} = 12.6 nm, T_{ox} = 1nm, VDD = 0.9 V and threshold voltage of PMOS and NMOS are -0.16 V and 0.16V respectively. Highly scalable memristor model of [17] having the R_{OFF}/R_{ON} ratio of 190 (i.e., R_{OFF} = 190 K and R_{ON} = 1K) and x_0 =0.4, a=50, p=10 are employed to exploit the characteristics of the proposed memory cell. Other parameters [18] used for simulations are: D= 1 nm, μ_v = 1e-7($m^2s^{-1}V^{-1}$).

TABLE I

COMPARISON OF THE PROPOSED CELL WITH RECENTLY DEVELOPD MEMRISTOR BASED MEMORY CELL

	Proposed Cell	Ambipolar based memory[7-8]	MCAM [10]
Write '0'	2.189 ns	219 ns	201 ns
Write '1'	1.7 ns	219 ns	201 ns
Read	9.00 ps	1.095 ns	12 ns
Supply Voltage	0.9 V	0.9 V	0.9 V

The proposed 6T1M memory cell does not only require refresh operation but also capable of rapid bidirectional storage operation. Recently developed memristor based memory cells such as Ambipolar based memory [8] requires large number of transistors to perform various operations and also refresh operation is required for repetitive read. The MCAM [10] cell requires search operation to match the stored data which causes high read time. Though the memory cell of [7] does not require any refresh operation, large die area is required to perform the bidirectional storage operation. On the other hand, if the threshold voltage of the ambipolar transistor in [7] is taken high, erroneous read operation can be occurred due to the belated switching and turning off the PMOS and ambipolar transistor respectively. The proposed circuit does not suffer the problem of erroneous read operation while varying the threshold voltage of the PMOS transistor. Though the memory cell of [10] does not require any refresh operation, higher supply voltage is needed to operate the cell because of the use of diode. This is the reason of higher dynamic power consumption for the 2M1D cell [11]. A comparison of the proposed memory cell and other recently developed memristor based memory cells [7-10] is shown in Table 1.

V. CONCLUSIONS

A new type of memristor based compact memory cell is proposed. This hybrid model used only 4 NMOS and 2 PMOS transistors along with a memristor and stores data from both directions. As the model does not need refresh during successive read operation, so less amount of power is required and the operating time becomes shorter than ambipolar based structure. The proposed model can be widely used due to its superior performance and linearity.

REFERENCES

[1] S. Borkar, T. Karnik, and V. De, "Design and reliability challenges in nanometer technologies," in Proc. 41st Annu. Design Autom. Conf., 2004, p. 75.

[2] C. T. Chuang, S. Mukhopadhyay, J. J. Kim, K. Kim, and R. Rao, "High-performance SRAM in nanoscale CMOS: Design challenges and techniques," in Proc. IEEE Int. Workshop MTDT, Dec. 2007, pp. 4–12.

[3] K. Kuhn, "Considerations for ultimate CMOS scaling," IEEE Trans. Electron Devices, vol. 59, no. 7, pp. 1813–1828, Jul. 2012.

[4] J. Borghetti *et al.*, "A hybrid nanomemristor/transistor logic circuit capable of self-programming," *Proc. Nat. Acad. Sci.*, vol. 106, no. 6, pp. 1699–1703, 2009.

[5] P. J. Kuekes, D. R. Stewart, and R. S. Williams, "The crossbar latch: Logic value storage, restoration, and inversion in crossbar circuits," *J. Appl. Phys.*, vol. 97, no. 3, pp. 034301-1–034301-5,Feb. 2005.

[6] M. N. Sakib, R. Hassan, S. Biswas, "Design a Memristor-based Hybrid Memory Cell having Faster Bidirectional Storage Operation" in 5th IEEE International Conference on Informatics, Electronics & Vision, Dhaka, pp.xx-xx , May 2016.

[7] P. Junsangsri, F. Lombardi, "A memristor-based memory cell with no refresh" Proceedings of the 14th IEEE International Conference on Nanotechnology Toronto, pp.947-950, 2014.

[8] P. Junsangsri, F. Lombardi,"Design of a Hybrid Memory Cell Using Memristance and Ambipolarity" IEEE Trans. Nanotechnology,Vol.12, Issue 1, pp.71-80, Jan 2013.

[9] P. Junsangsri, F. Lombardi, "A Memristor-Based Memory Cell using Ambipolar Operation" *Proc IEEE ICCD 2011*, pp. 148-153, Oct 10 2011, MA USA.

[10] K. Eshraghina, K.R. Cho, O. Kavehei, S.K Kang, D, Abbott, S.M.Steve Kang, "Memristor MOS Content Addressable Memory (MCAM): Hybrid Architecture for Future High Performance Search Engines" IEEE Transactions on VLSI Systems, vol. 19, no. 8, pp.1407-1417, Aug 2011.

[11] M. Elshamy, H. Mostafa, Y. Ghallab, S.M. Said, "A novel non-destructive read/write circuit for memristor-based memory arrays," *IEEE Transactions on VeryLarge Scale Integration,* vol. 23 no. 11, pp. 2648–2656, Dec. 2014.

[12] L. O. Chua, "Memristor—the missing circuit element," IEEE Trans. Circuit Theory, vol. CT-18, no. 5, pp. 507–519, Sep. 1971.

[13] L. O. Chua, C. A. Desoer, and E. S. Kuh, Linear and Nonlinear Circuits, McGraw-Hill, 1987.

[14] D. B. Strukov, G. S. Snider, D. R. Stewart, and R. S. Williams, "The missing memristor found," Nature, vol. 453, pp. 80–83, May 2008.

[15] L. O. Chua, "Device modeling via basic nonlinear circuit elements,"*IEEE Trans. Circuits Syst.*, vol. CAS-27, pp. 1014–1044, Nov. 1980.

[16] http://ptm.asu.edu/

[17] Z. Biolek, D. Biolek, and V. Biolkov, "Reliable Modeling of Ideal Generic Memristorsvia State-Space Transformation," Radio eng. J., vol 24, no. 2, p. 393, 2015.

[18] Y. Zhang, Y. Shen, X. Wang, Y. Guo, "A Novel Design for MemristorBased OR Gate," Circuits and Systems II: Express Briefs, IEEE Transactions on , vol.PP, no.99, pp.1,1, 2015

Bias dependent non-linear electrical characteristics of poly-silicon nanowire and assessment of biosensing application using liquid gate

Tanvir Alam Shifat, Rubiya Nila, Muktadir Imam Jan and M. M. A. Hakim*
East West University, Main Road, Jahurul Islam City, Aftabnagar, Dhaka-1212, Bangladesh.
*Email: dmmah@ewubd.ed

Abstract: We investigate the bias dependent modulation of the electrical characteristics of poly-silicon nanowire and perform a feasibility study of poly-silicon nanowire for biosensing application using liquid gate arrangement. Electrical characteristics of a 100 nm thick and 10 µm long p-type poly-silicon nanowire with a doping density of $10^{16}/cm^3$ is investigated. For positive drain voltage application nanowire's output characteristics exhibit a non-linear diode like behavior whereas for negative drain voltages output characteristics exhibit a perfect transistor behavior with saturation in the characteristics. While nanowire's output characteristics reveal quite a drastic change with drain bias polarity, the sub-threshold characteristics exhibit excellent behavior for both positive and negative drain bias application with a sub-threshold slope of 102.61 mv/decade implying that poly silicon nanowire is good candidate for bio sensing operation with an achievable sensitivity around 800%. It is also found that for positive drain voltages sub-threshold characteristics shifts with the change in drain bias providing the flexibility of a wide range of liquid gated voltages for physiological activity monitoring. For negative drain voltages sub-threshold characteristics does not shift thereby restricting the applicable liquid gate voltages for sensitive operation of biosensors. As such, this study reveals the applicable bias conditions of a 10 µm long 100 nm thick low doped p-type poly-silicon nanowire biosensors physiological process monitoring and the maximum achievable sensitivity for detection of biomarkers.

***Index Terms*—Nanowire, biosensor, liquid gating.**

I. INTRODUCTION

Over the last decade, silicon nanowires (NW) have emerged as important candidates for biochemical sensors due to its excellent sensitivity, real-time and label-free detection without expensive optical components [1]. This ultra-high sensitivity detection can be attributed to their smaller size and large surface to volume ratio, enabling local charge transfers to result in a current change due to a field effect when analytic molecules bind to a specific recognition molecule at the surface of the nanowire [1].

Plenty of works can be found in the literature employing Si-NWs for detecting biomarkers, individual bacteria or viruses. Cui et al. successfully fabricated silicon nanowire biosensors on p-type semiconductor where the typical output characteristics (I_{DS}-V_{DS}) characteristics were linear [1]. The ability of the fabricated biosensors was tested through pH response monitoring with or without surface functionalisation [1]. Chen et al. [2] fabricated p-type silicon nanowires using wet etch where silicon nanowires have height of 140 nm, width of 100 nm and a uniform doping concentration of $10^{17}/cm^3$. Measured I_{DS}-V_{DS} characteristics exhibited typically non-linear diode like behavior and these nanowires also demonstrated successful pH sensing with a sensitivity around 40mV/pH [2]. In a seminal paper Colinge et al reported that gated nanowires exhibit transistor like behavior with perfect saturation in output characteristics rather than diode like behavior and such a junctionless nanowire transistor has been claimed to have full CMOS functionality with ideal sub-threshold slope and extremely low leakage currents [3]. Although quite different electrical characteristics of silicon nanowires has been reported in the past, it has been utilized in bisosensor design without proper justification of applicable bias voltages which could have resulted in much better sensitivity and operating conditions of nanowire biosensors for physiological activity monitoring. It is worth mentioning that a promising set up of NW biosensor application is liquid gating (Fig. 1) where a fixed voltage V_{DS} is applied across the NW and the operating point of the bio-sensor can be set within the linear region of the sub-threshold regime for sensitive operation by providing required gate-bias in the buffer using Au or AgCl elctrode. There is no report available in the literature that is studying the effect of drain bias polarity and liquid gate voltage assisted perturbation of nanowire conduction path on the electrical characteristics of nanowire to find out the applicable bias window for a viable monitoring of physiological activity or biomarker detection.

In this report, we investigate the effects of drain bias polarity and liquid gate voltages on the electrical characteristics of a low doped ($10^{16}/cm^3$), 100 nm thick, 10 µm long poly-silicon nanowire to find out applicable bias window and extract its conductance charge with the change of the top gate voltages for ensuring sensitive operation of biosensor.

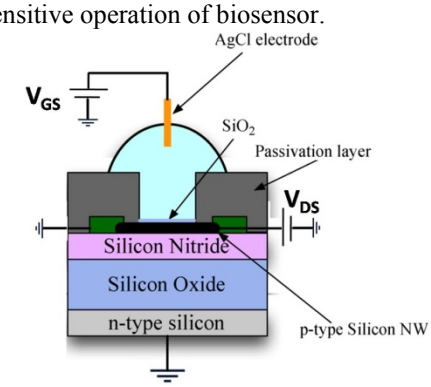


Fig. 1: Measurement setup of silicon nanowire biosensor using liquid gate.

I. SIMULATION METHODOLOGY

The investigation of polysilicon nanowire biosensors was done with the help of numerical commercial simulator SILVACO [4] installed in the VLSI lab of East West University. A p-type silicon nanowire with 10 µm length, 100 nm thick and a doping concentration of $10^{16}/cm^3$ was created on a 500 nm Silicon Nitride (Si_3N_4) which was deposited on a 500

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

nm buried Si layer. Liquid gate effect is simulated using a top gate electrode with work function set to 0. The gate oxide thickness was 2 nm. In the silicon nanowire, two heavily doped p-type regions (doping density of $10^{20}/cm^3$) in the two sides of the channel were implemented to ensure ohmic contacts on the source/drain regions. The 2D coupled Poisson's drift-diffusion solver was used to model carrier transport in Polysilicon NW. The Shockley-Read-Hall model for carrier emission and absorption was used to reflect the recombination phenomena within the device. Moreover, the bandgap narrowing(BGN) effects in the heavily doped source/drain regions were included by Slotboom and de Graaf's analytical BGN model. For accuracy, the polysilicon is modeled using a continuum of trap states consist of both donor-like and acceptor-like states distribution across the energy band gap. The defect states as a combination of exponentially decaying band Tail states and Gaussian distribution of mid gap states are used with parameters previously used in the literature for successful modeling of polysilicon thin film transistors [5].

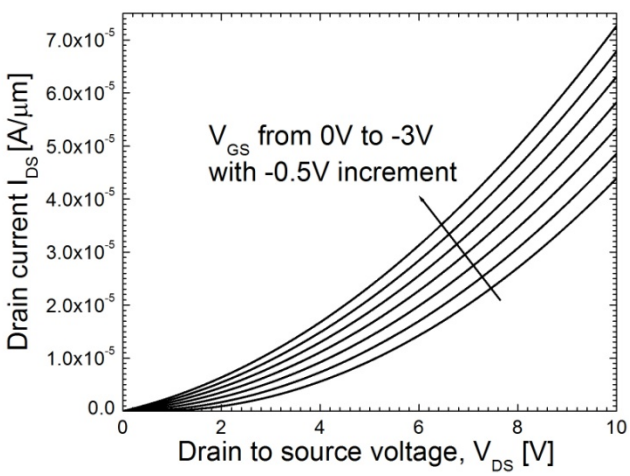


Fig. 2: Output characteristics (I_{DS} vs. V_{DS}) of polysilicon nanowire biosensor for different negative liquid gate voltages (V_{GS}) when the drain voltage (V_{DS}) is positive. The nanowire is 10 μm long, 100 nm thick and has a doping concentration of $10^{16}/cm^3$.

III. RESULT AND DISCUSSION

Figure 2 shows output characteristics (I_{DS} vs. V_{DS}) of 10 μm long, 100 nm thick polysilicon nanowires for positive drain (V_{DS}) voltages and for different negative liquid gate voltages (V_{GS}) while the doping density is $10^{16}/cm^3$. For positive drain voltages, the I_{DS} vs V_{DS} characteristics is inherently non-linear at $V_{Gs} = 0$ V with a very low level of conduction up to 1.5 V of V_{DS} value in linear scale. This characteristic resembles the electrical characteristics observed in p-type silicon nanowires processed by wet anisotropic plane dependent etching by Chen et al [2]. Application of negative liquid gate voltage (V_{GS}) results in the increase of I_{DS} values which increases from 9.48 μA/μm at V_{GS} = 0V to 23.5 μA/μm at V_{GS} = -3 V for a V_{DS} = 5 V. This can be explained by the modulation of nanowire conduction due to strong hole accumulation upon application of the negative V_{GS} voltages.

Figure 3 shows output characteristics (I_{DS} vs. V_{DS}) of 10 μm long, 100 nm thick polysilicon nanowires for positive drain (V_{DS}) voltages and for different positive liquid gate voltages (V_{GS}) while the doping density is $10^{16}/cm^3$. The I_{DS} vs V_{DS} characteristics is again found to be non-linear at different values of positive V_{Gs} values with no significant conduction up to certain V_{DS} values in linear scale. However, it is worth noting that with the increase of the positive liquid gate voltages (V_{GS}) the electrical characteristics becomes more non-linear and the typically less conductive characteristics of the nanowire increases. Application of positive liquid gate voltage (V_{GS}) results in the decrease of I_{DS} values which decreases from 9.48 μA/μm at V_{GS} = 0V to 0.96 μA/μm at V_{GS} = 3 V for a V_{DS} = 5 V. The decrease of the drive current and the increase of non-linearity in the characteristics with the increase of positive liquid gate voltages (V_{GS}) can be explained by the depletion of p-type nanowire upon application of positive V_{GS} values.

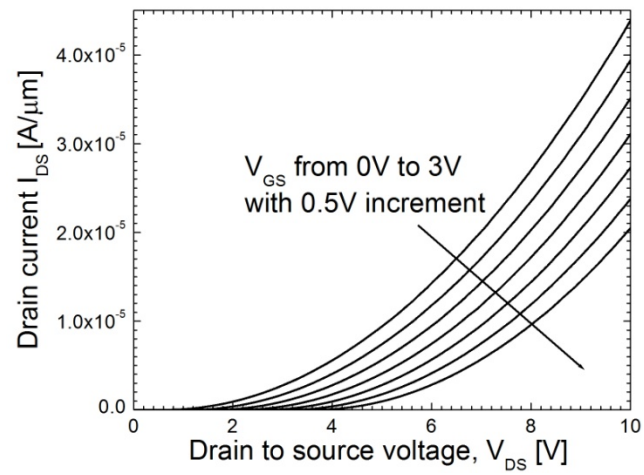


Fig. 3: Output characteristics (I_{DS} vs. V_{DS}) of polysilicon nanowire biosensor for different positive liquid gate voltages (V_{GS}) when the drain voltage (V_{DS}) is positive. The nanowire is 10 μm long, 100 nm thick and has a doping concentration of $10^{16}/cm^3$.

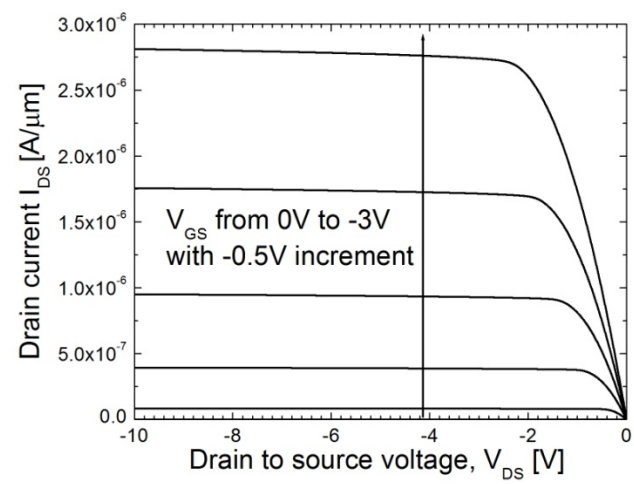


Fig. 4: Output characteristics (I_{DS} vs. V_{DS}) of polysilicon nanowire biosensor for different negative liquid gate voltages (V_{GS}) when the drain voltage (V_{DS}) is negative. The nanowire is 10 μm long, 100 nm thick and has a doping concentration of $10^{16}/cm^3$.

Figure 4 shows output characteristics (I_{DS} vs. V_{DS}) of 10 μm long, 100 nm thick polysilicon nanowires for different negative liquid gate voltages (V_{GS}) while the drain bias (V_{DS}) is negative. For negative V_{DS} values, a drastic change in the behavior of poly-silicon nanowire can be seen with nanowires exhibiting

excellent transistor characteristics with perfect saturation in behavior. However, I_{DS} current increase with the increase of the negative liquid gate voltage (V_{GS}) due to hole accumulation which increases from 1.64×10^{-13} A/µm at V_{GS} = 0V to 2.77×10^{-6} A/µm at V_{GS} = -3 V for a V_{DS} = -5 V. We have also investigated the output characteristics (I_{DS} vs. V_{DS}) of 10 µm long, 100 nm thick polysilicon nanowires for different positive liquid gate voltages (V_{GS}) while the drain bias (V_{DS}) is negative. In this condition the drive current is insignificant and the poly-silicon nanowire shows anomalous conduction characteristics due positive liquid gate voltage (V_{GS}) associated depletion of carriers.

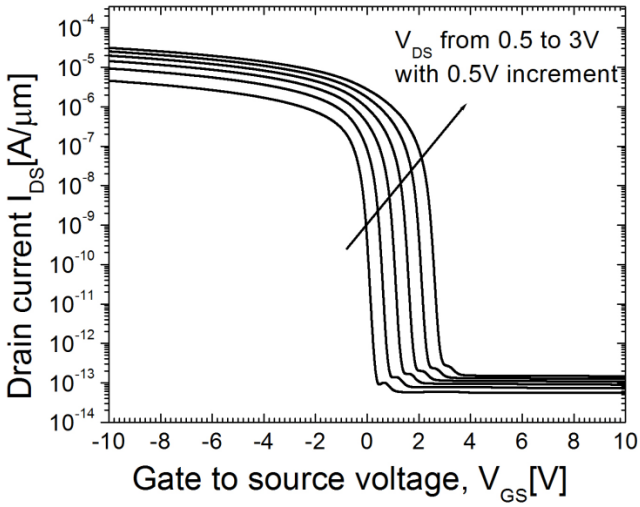


Fig. 5: Sub-threshold (I_{DS} vs. V_{GS}) of polysilicon nanowire biosensor when the drain voltage (V_{DS}) is positive. The nanowire is 10 µm long, 100 nm thick and has a doping concentration of 10^{16}/cm^3.

To further investigate the bias dependent variable characteristics of poly-silicon nanowire, figure 5 shows sub threshold characteristics (I_{DS} vs V_{GS}) of the p-type poly-silicon nanowire when drain voltage (V_{DS}) is positive. While the output characteristic of the nanowire does not exhibit conventional transistor behavior for positive V_{DS} application (no saturation in figure 2 & 3), the sub threshold characteristic shows quite a decent transitor like behavior. The exhibited sub-threshold slope of the 10 µm long, 100 nm thick poly-silicon nanowire with doping 10^{16}/cm^3 is around 102.61 mV/decade. Although this value of sub-threshold slope is much higher than the ideal value of 60 mV/decade, it is still suitable for biosensor operation considering the fact that gate charge modulation can be detected if operating point of the bio-sensor can be set within the linear region of the sub-threshold regime. However, it is interesting to note that the sub-threshold characteristics shifts towards positive V_{GS} values with the increase of V_{DS} voltage (1030 mV/V). This shift is known as drain induced barrier lowering (DIBL) in conventional CMOS technology and traditionally known to be disadvantageous for logic application. However, for biosensing application this DIBL somehow appears to be beneficiary. Physiological activity monitoring, biomarker detection and testing of drugs can be significantly accelerated by providing electric fields in the buffer whose polarity and strength depends on the analytes under consideration. Figure 6 shows that positive V_{DS} application is providing a unique platform to choose a wide range of liquid gated voltage (V_{GS}) values (i.e. polarity and strength) while the sensitive operation of poly-silicon nanowire biosensor can be maintained with the stringent condition for drain voltage (V_{DS}) and liquid gate voltage (V_{GS}) to ensure the operating point of the bio-sensor within the linear region of the sub-threshold regime.

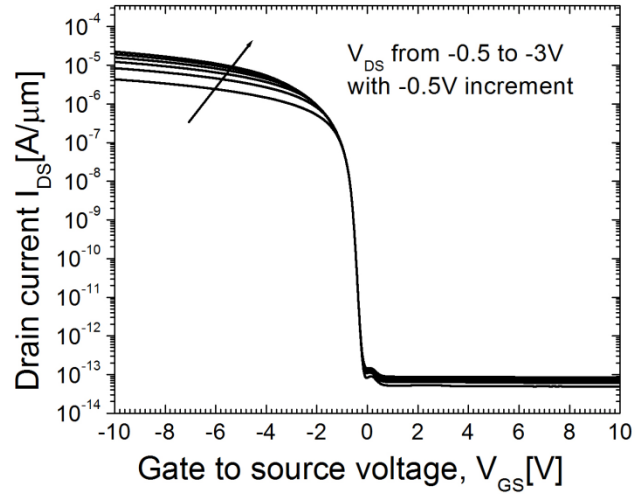


Fig. 6: Sub-threshold (I_{DS} vs. V_{GS}) of polysilicon nanowire biosensor when the drain voltage (V_{DS}) is negative. The nanowire is 10 µm long, 100 nm thick and has a doping concentration of 10^{16}/cm^3.

Figure 6 shows sub-threshold characteristics (I_{DS} vs V_{GS}) of the p-type poly-silicon nanowire when drain voltage (V_{DS}) negative. For negative V_{DS} application, the device exhibits a sub-threshold slope of 102.61 mV/decade which is similar to the sub-threshold slope with positive V_{DS} voltage implying that a same sensitivity of poly-silicon nanowires for biosensing application may be achieved using both positive and negative V_{DS} voltages by proper biasing scheme. However, for negative V_{DS} application, no shift in I_{DS} vs V_{GS} characteristic can be observed with negative V_{DS} voltages. This result shows that negative V_{DS} application will restrict applicable liquid gate voltage (V_{GS}) as sensitive operation of biosensor would automatically demand the operating point be set within the linear region of the sub-threshold regime. As can be seen from figure 6, a sensitive operation of the 10 µm long, 100 nm thick poly-silicon nanowire with doping 10^{16}/cm^3 may be set to a top liquid gate voltage (V_{GS}) around -1.25 V for all V_{DS} values. Such a fixed value of buffer voltage is obviously not suitable for all physiological monitoring, biomarker detection and testing of drugs.

To investigate the reasons behind the different sub-threshold characteristics (I_{DS} vs V_{GS}) of poly-silicon nanowires with positive and negative V_{DS} voltages, potential profile of nanowire with positive and negative V_{DS} voltages are examined (not shown). It is found that for positive V_{DS} voltages drain-body potential barrier is very low due to a forward biased p+p junction at drain side. As a result there is a significant reduction of the source-body potential barrier with increasing V_{DS} voltages thereby resulting in the shift of the sub-threshold characteristics (I_{DS} vs V_{GS}) with increasing V_{DS} voltages. For negative V_{DS} voltages drain-body potential barrier is very high (reverse biased p+p junction) and hence, there is an

insignificant modulation of the source-body potential barrier with increasing negative V_{DS} voltages and hence, no shift in the sub-threshold characteristics (I_{DS} vs V_{GS}) of poly-silicon nanowires are observed with increasing negative V_{DS} voltages. The strong drain-body junction depletion region is also responsible for the onset of saturation in the output characteristics (I_{DS} vs. V_{DS}) nanowires when drain bias (V_{DS}) is negative.

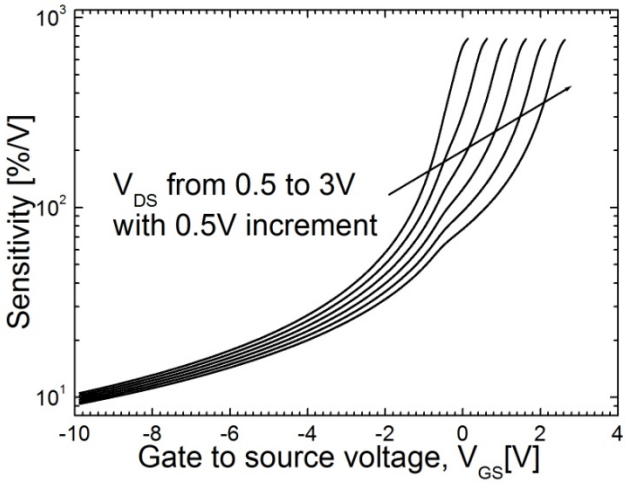


Fig. 7: Sensitivity as a function of liquid gate voltages (V_{GS}) for different values of positive V_{DS} applications. The nanowire is 10 µm long, 100 nm thick and has a doping concentration of $10^{16}/cm^3$.

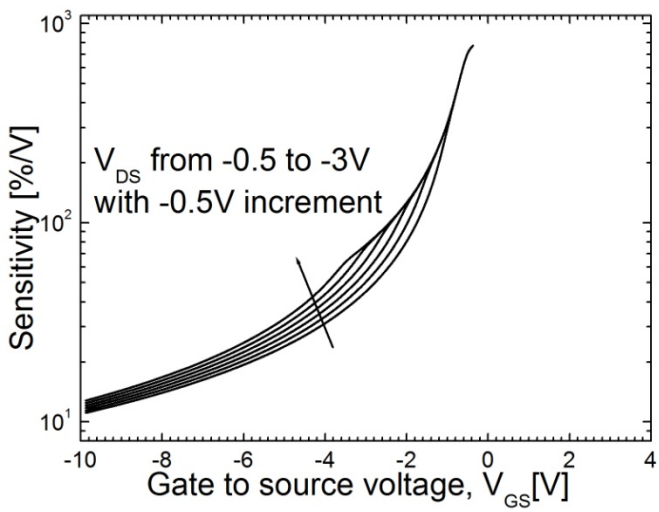


Fig. 8: Sensitivity as a function of liquid gate voltages (V_{GS}) for different values of negative V_{DS} applications. The nanowire is 10 µm long, 100 nm thick and has a doping concentration of $10^{16}/cm^3$.

Finally we extract the maximum achievable sensitivity of a biosensor using 10 µm length, 100 nm thick poly-silicon nanowire with a doping concentration of $10^{16}/cm^3$ and the plausible bias conditions to achieve this. Figure 7 shows sensitivity as a function of liquid gate voltages (V_{GS}) for different values of positive V_{DS} applications. The sensitivity is extracted by using a first order equation $(I_{GS2}-I_{GS1})/[I_{GS2}.(V_{GS1}-V_{GS2})]$ in percentage to show the possible change in the nanowire conductance when biomarkers will be immobilized on the surface of the nanowire. The sensitivity curve shifts towards positive V_{GS} values with increasing V_{DS} voltages and the sensitivity reduces with the increase of negative V_{GS} values for a fixed V_{DS}. However, it can be seen that for positive V_{DS} voltages a maximum sensitivity around 800% can be achieved if the operating point can be chosen at the starting of the linear section of the sub-threshold region. For positive V_{DS} applications this can be ensured by a (V_{GS}, V_{DS}) combinations of (0.125, 0.5); (0.625, 1.0); (1.125, 1.5); (1.625, 2.0); (2.125, 2.5) and (2.625, 3.0). An almost similar characteristic can be observed in the sensitivity trend for negative V_{DS} voltages (figure 8) except the fact that the sensitivity curve does not shift towards positive V_{GS} values with increasing negative V_{DS} voltages. In this case also a maximum 800% sensitivity can be achieved using 10 µm length, 100 nm thick poly-silicon nanowire with a doping concentration of $10^{16}/cm^3$. However, this can be achieved for a fixed liquid gate voltage V_{GS} = 0.125V even though different negative V_{DS} voltages are applied.

IV. CONCLUSIONS

We have investigated bias dependent non-linear electrical characteristics of poly-silicon nanowire to find out applicable bias window using liquid gate arrangement for ensuring sensitive operation of poly-silicon nanowire biosensors. Quite a different electrical characteristic has been observed for positive and negative drain voltage applications. Nanowires have exhibited a non-linear diode like behavior for positive drain voltages whereas for negative drain voltages nanowires have exhibited a perfect transistor behavior with saturation in the output characteristics. The sub-threshold characteristics of nanowires have shown excellent behavior for both positive and negative drain bias application implying that poly silicon nanowire is good candidate for bio sensing. It has also been found that positive drain voltage application provides the flexibility of a wide range of liquid gated voltages for physiological activity monitoring whereas negative drain voltage application restricts the applicable liquid gate voltages for sensitive operation of biosensors. As such, this study reveals the applicable bias conditions of a 10 µm long 100 nm thick low doped p-type poly-silicon nanowire biosensors physiological process monitoring and the maximum achievable sensitivity for detection of biomarkers.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support of the Center for Research and Training (CRT), East West University, under HEQEP sub-project Knowledge Transfer and Capacity Development of Academic Staff.

REFERENCES

[1] Y. Cui, Q. Wei, H. Park and C. M. Lieber, "Nanowire and Nano sensor for highly sensitive and selective detection of biological and Chemical species," *Science*,vol. 293, pp. 1289-1292, 2001.

[2] S. Chen, J. G. Bomer, W. G. Van der Wiel, E. T. Carlen and A. Van den Berg, "Top-Down Fabrication of Sub-30 nm Monocrystalline Silicon Nanowires Using Conventional Microfabrication," *ACS Nano*, vol. 3, pp. 3485-3492, 2009.

[3] J. P. Colinge Et al., "Nanowire transistors without junctions," *Nature Nanotecnology*, vol.5, pp. 225-229, 2010.

[4] Silvaco international, *Atlas User's Manual Device Simulation Software*, Silvao International Ltd., Santa Clara, Dec. 2002.

[5] P. Walker, H. Mizuta, S. Uno, Y. Furuta and D. Hasko, "Improved off-current and subthreshold slope in aggressively scaled poly-si TFTs with single grain boundary in the channel," IEEE Trans. On Electron Devices, vol. 51, pp. 212-219, 2004.

An investigation of the effects of doping and thickness on the electrical characteristics of polycrystalline silicon nanowire biosensors

Laila Parvin Poly, Md. Asrarul Haque and M. M. A. Hakim*
East West University, Main Road, Jahurul Islam City, Aftabnagar, Dhaka-1212, Bangladesh.
*Email: dmmah@ewubd.edu

Abstract: We study for the first time the effect of nanowire thicknesses and doping concentrations on the electrical characteristics of polycrystalline silicon nanowires to find out the proper combination of nanowire thickness and doping for sensitive operation of polycrystalline silicon nanowire biosensors. For nanowire thicknesses of 100 nm and 75 nm, a plausible sub-threshold slope around 100 mV/dec for a viable biosensor operation can only be achieved if doping concentration is $2\times10^{16}/cm^3$ or below. For a 50 nm nanowire thickness a relatively wide doping concentration range can be chosen for biosensor design while maintaining decent sub-threshold characteristics. In this thickness a doping up to $4\times10^{17}/cm^3$ with a sub-threshold slope around 100 mV/dec can be chosen. The widest range of doping concentrations can be chosen for 25 nm and 10 nm nanowire thicknesses with a maximum doping up to $10^{18}/cm^3$ while maintaining a promising sub-threshold slope around 95 mV/dec for a viable biosesnsor design using polycrystalline silicon nanowires. As such this research reveals the possible combinations of nanowire thickness and doping to ensure the sensitive operation of polycrystalline silicon nanowire biosensors.

***Index Terms*—Nanowire, biosensor, polysilicon.**

I. INTRODUCTION

Over the last decade, silicon nanowires have emerged as a promising candidate for biochemical sensors due to its excellent sensitivity, real-time and label-free detection without expensive optical components [1]. Due to their small size and large surface-to-volume ratio the depletion and accumulation of charge-carriers produced by specific binding of biological macromolecule on the surface affects the entire cross sectional conduction pathway and hence, NW devices give extraordinary sensitivity when compared to other transducers reported in the literature [2]. Although this advantage of nanowires has been exploited to sense ions, proteins, DNA and viruses, some outstanding issues remain for mass commercialization. These are reproducible integration issues, cost of fabrication, and notably the appropriate choice of thickness, length and doping of nanowires for ensuring sensitive operation.

The silicon nanowire nanofabrication toolbox consists of techniques that can be broadly classified as either bottom-up or top-down. The bottom-up approaches usually employ metal-catalytic growth for the preparation of NWs [1, 3] which is simple in principle; however, suitable methods for accurate NW positioning are lacking, and electrical contact formation is problematic, making it difficult to construct functional device arrays. Top-down approaches overcome these shortcomings and several groups used nano-patterning techniques such as deep-UV [4] lithography and electron beam lithography [5] to fabricate silicon nanowires on silicon-on-insulator (SOI) substrate. This has the advantage of CMOS compatibility, but a serious disadvantage is the high cost associated with these advanced lithography techniques and expensive SOI wafers. Recently a low cost, top-down approach to nanowire fabrication has been reported that uses thin film technology and spacer etch technique [6-7]. This approach is particularly attractive because it produces nanowires with nano-scale dimensions using mature lithography in combination with standard deposition and spacer etch techniques that are widely available in industry. More importantly, this process can be used for glass or plastic substrates suitable for the realization of low cost disposable diagnostic kits. However, in this approach defined nanowires are usually amorphous and/ or polysilicon depending on the deposition and annealing conditions and hence, nanowire material usually composed of grain boundaries and defects which may also affect its electrical characteristics eventually affecting biosensor performance realized in these nanowires. So far there have been very few works on polysilicon nanowire biosensors and there is no report available showing inherent thickness and doping dependency of the electrical characteristics of polycrystalline silicon nanowire biosensors.

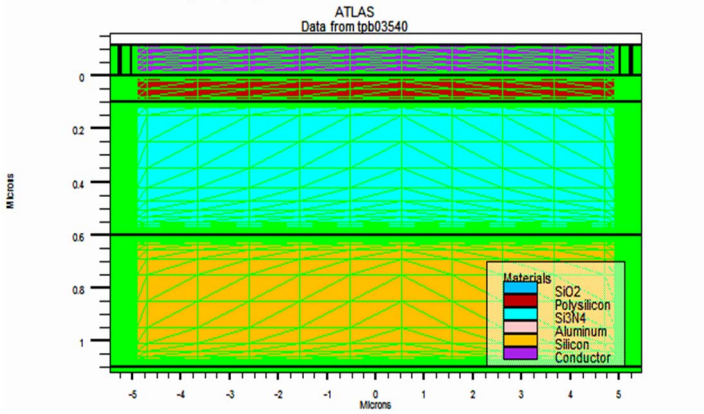


Figure 1: Cross-sectional view of polycrystalline silicon nanowire biosensor showing the mesh density.

In this report, we study for the first time the effect of nanowire thickness and doping concentration on the electrical characteristics of polycrystalline silicon nanowires to find out the proper combination of nanowire thickness and doping to ensure the sensitive operation of polycrystalline silicon nanowire biosensors. The electrical characteristics of a p-type polysilicon nanowire having a nanowire length of 10 µm is studied at various thicknesses and doping concentrations. 10 nm, 25 nm, 50 nm, 75 nm and 100 nm thick nanowires with doping density varying from $10^{16}/cm^3$ to $10^{18}/cm^3$ are investigated to find out the best combination of doping concentrations and nanowire thicknesses for polycrystalline silicon nanowire biosensor fabrication.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

II. SIMULATION METHODOLOGY

The investigation of polysilicon nanowire biosensors was done with the help of numerical commercial simulator SILVACO [8] installed in the VLSI lab of East West University. A p-type silicon nanowire with 10 µm length was created on a 500 nm Silicon Nitride (Si_3N_4) which was deposited on a 500 nm buried Si layer (Fig.1). An Al (Aluminium) electrode was used as top gate material and its work function was set to 0. The gate oxide thickness was 2 nm. In the silicon nanowire, two heavily doped regions (doping density of $10^{20}/cm^3$) in the two sides of the channel were implemented to ensure ohmic contacts on the source/drain regions. The heavily doped source/drain regions and the nanowire were p-type. Simulations carried out for nanowire thicknesses of 10nm, 25nm, 50nm, 75nm and 100nm while the doping of the nanowire was varied from 10^{16} /cm^3 to 10^{18} /cm^3. The 2D coupled Poisson's drift-diffusion solver was used to model carrier transport in Polysilicon NW. The Shockley-Read-Hall model for carrier emission and absorption was used to reflect the recombination phenomena within the device. Moreover, the bandgap narrowing(BGN) effects in the heavily doped source/drain regions were included by Slotboom and de Graaf's analytical BGN model. For accuracy, the polysilicon is modeled using a continuum of trap states consist of both donor-like and acceptor-like states distribution across the energy band gap. The defect states as a combination of exponentially decaying band Tail states and Gaussian distribution of mid gap states are used with parameters previously used in the literature for successful modeling of polysilicon thin film transistors [9].

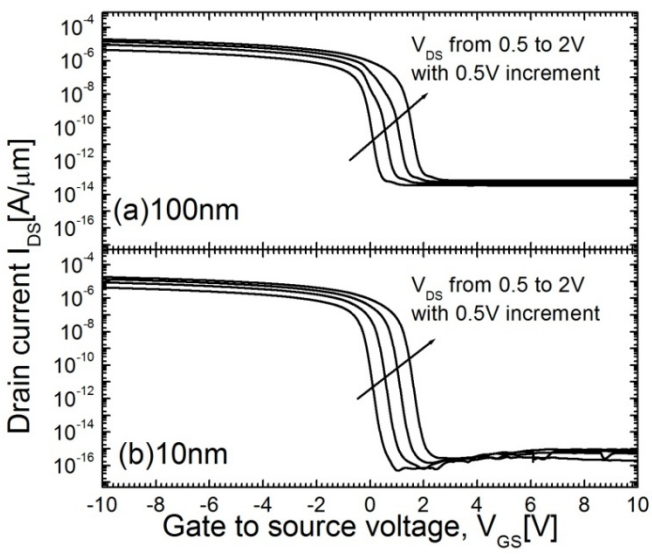


Figure 2: Transfer characteristics of p-type polysilicon nanowires with $10^{16}/cm^3$ doping concentration at different thicknesses; (a) 100nm and (b) 10nm.

III. RESULTS& DISCUSSION

Figure 2 shows transfer characteristics (I_{DS} vs V_{GS}) of p-type Si-NWs for different nanowire thicknesses. The length of the nanowire is 10 µm and the doping density is $10^{16}/cm^3$. For the naowire thickness of 100 nm in figure 2(a), the sub-threshold slope is found to be 102.61 mV/decade and the DIBL is found to be around 1030 mV/V. Although the value of sub-threshold slope is much higher than the ideal value of 60 mV/decade, it is still suitable for biosensor operation considering the fact that gate charge modulation can be detected if operating point of the bio-sensor can be set within the linear region of the sub-threshold regime. For the nanowire thickness of 10 nm in figure 2(b), the sub-threshold slope and DIBL are found to be 93 mV/decade and 980 mV/V. The sub-threshold characteristic at this thickness is much better than that of the 100 nm thickness and the sub-threshold slope is also close to the ideal value of 60 mV/V. Hence, 10 nm thick 10 µm long nanowire with a doping of $10^{16}/cm^3$ is a promising candidate for biosensor application as setting the operating point within the linear region of sub-threshold region would ensure an excellent response to the gate charge modulation upon immobilization of analytes. These results imply that 100 nm to 10 nm thick polysilicon nanowires can be used as biosensor with a doping density of $10^{16}/cm^3$.

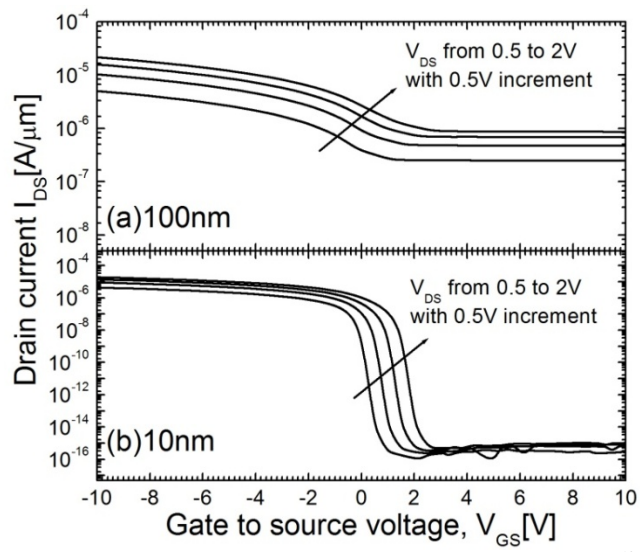


Figure 3: Transfer characteristics of p-type polysilicon nanowires with $10^{18}/cm^3$ doping concentration at different thicknesses; (a) 100nm and (b) 10nm.

Figure 3 shows transfer characteristics (I_{DS} vs V_{GS}) of p-type Si-NWs for different nanowire thicknesses while the doping density is $10^{18}/cm^3$. For the nanowire thickness of 100 nm in figure 3(a), the characteristics is severely degraded with a sub-threshold slope around 5246 mV/decade and a DIBL around 3220 mV/V. These drastically inferior sub-threshold characteristics render 100 nm thick polysilicon nanowire with a doping density of $10^{18}/cm^3$ unsuitable for a viable biosensor operation. However, for the nanowire thickness of 10 nm in figure 3(b), the sub-threshold characteristics are drastically changed. While sub-threshold characteristics of 100 nm thick nanowire with a doping of $10^{18}/cm^3$ is not promising for biosensor, the 10 nm thick polysilicon nanowire at this doping concentration exhibits a sub-threshold slope of 104 mV/decade and a DIBL around 1030 mV/V, which is quite comparable to the $10^{16}/cm^3$ doped nanowires and obviously promising for biosensor design. This result show that a heavy doping around $10^{18}/cm^3$ can be used in biosensor design if nanowire thickness is scaled down to 10 nm or below. Low doped NW is traditionally used for biosensor design due to its extreme sensitivity which may allow single molecule analyte detection for early disease diagnosis. However, extreme sensitivity of NW also makes it vulnerable to noise as due to small size and large surface-to-volume ratio minor perturbations at surface

may deplete or accumulate the tiny volume of NW. As a result, to improve the signal to noise ratio in NW biosensor design doping concentration may be required to be increased. It is apparent from figure. 2 and 3 that the doping concentration of nanowire cannot be increased arbitrarily. For viable biosensor performance nanowire doping is inherently related to its thickness and if a design requires a doping around $10^{18}/cm^3$, then the thickness of the nanowire should be scaled down to 10 nm or below.

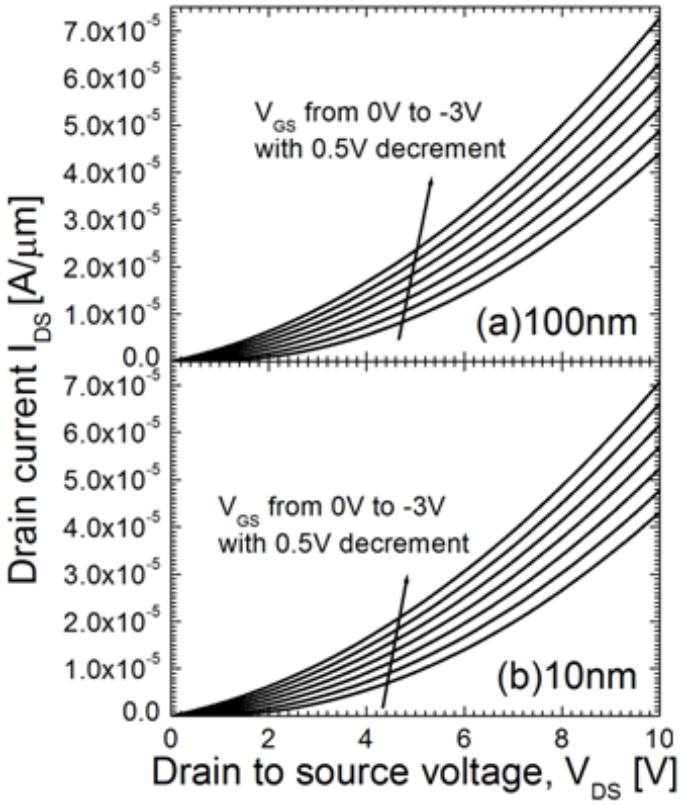


Figure 4: Output characteristics of p-type polysilicon nanowires with $10^{16}/cm^3$ doping concentration at different thicknesses; (a) 100nm and (b) 10nm.

Figure 4 shows output characteristics (I_{DS} vs. V_{DS}) of 10 μm long polysilicon nanowires for different nanowire thicknesses while the doping density is $10^{16}/cm^3$. For the 100 nm thick nanowire in figure 4(a), I_{DS} vs V_{DS} characteristics is inherently non-linear at $V_{Gs} = 0$ V with a very low level of conduction up to 1.5 V of V_{DS} value in linear scale. This characteristic resembles the electrical characteristics observed in p-type silicon nanowires processed by wet anisotropic plane dependent etching by Chen et al [10]. Application of negative $V_{GS,}$ results in the increase of I_{DS} values which increases from 9.48 μA/μm at $V_{GS} = 0$V to 23.5 μA/μm at $V_{GS} = -3$ V for a $V_{DS} = 5$ V (60% change). This can be explained by the modulation of nanowire conduction due to strong hole accumulation upon application of the negative V_{GS} voltages. A similar characteristic can be observed for 10 nm thick nanowire (figure 4(b)). However, drive current of 10 nm thick nanowire is slightly lower than 100 nm thick nanowire and a similar change in the drive current can be observed for 10 nm thick polysilicon nanowire for a V_{GS} change from 0 to -3 V at $V_{DS} = 5$ V (60% change). This result imply that for a doping density of $10^{16}/cm^3$, nanowire thickness scaling has mild effect on the electrical characteristics of polyslicon nanowires. This result also agrees with the result of figure 2 that is showing a non-phenomenological change of nanowires sub-threshlod characteristics while thickness has been scaled from 100 nm to 10 nm at a doping density of $10^{16}/cm^3$. As a result, it can be decided that thickness scaling of nanowires at the doping density of $10^{16}/cm^3$ beyond 100 nm has no obvious benefit for biosensor operation and any nanowire thickness ranging from 100 nm to 10 nm can be chosen for biosensor design at this doping density.

Figure 5 shows output characteristics (I_{DS} vs. V_{DS}) of 10 μm long polysilicon nanowires for different nanowire thicknesses while the doping density is $10^{18}/cm^3$. In general, the I_{DS} vs. V_{DS} characteristics at $10^{18}/cm^3$ is similar to $10^{16}/cm^3$ doping density but the drive current is significantly increased. However, for the 100 nm thick nanowire (figure 5(a)), the drive current change for a V_{GS} change from 0 to -3 V at $V_{DS} = 5$ V is reduced to a value of 49% at a doping density of $10^{18}/cm^3$ implying a less gate control on polysilicon nanowire at elevated doping level for 100 nm nanowire thickness. For the 10 nm thick nanowire at a doping density of $10^{18}/cm^3$ (figure 5(b)), the drive current is slightly lower than 100 nm nanowire thickness. However, the drive current change for a V_{GS} change from 0 to -3 V at $V_{DS} = 5$ V is around 60% which is much higher than 100 nm nanowire thickness at a doping density of $10^{18}/cm^3$ and similar to the drive current change found for $10^{16}/cm^3$doping density. This result agrees with figure 3 and implies that for a doping around $10^{18}/cm^3$, the thickness of the nanowire should be scaled down to 10 nm or below for a viable biosensor operation.

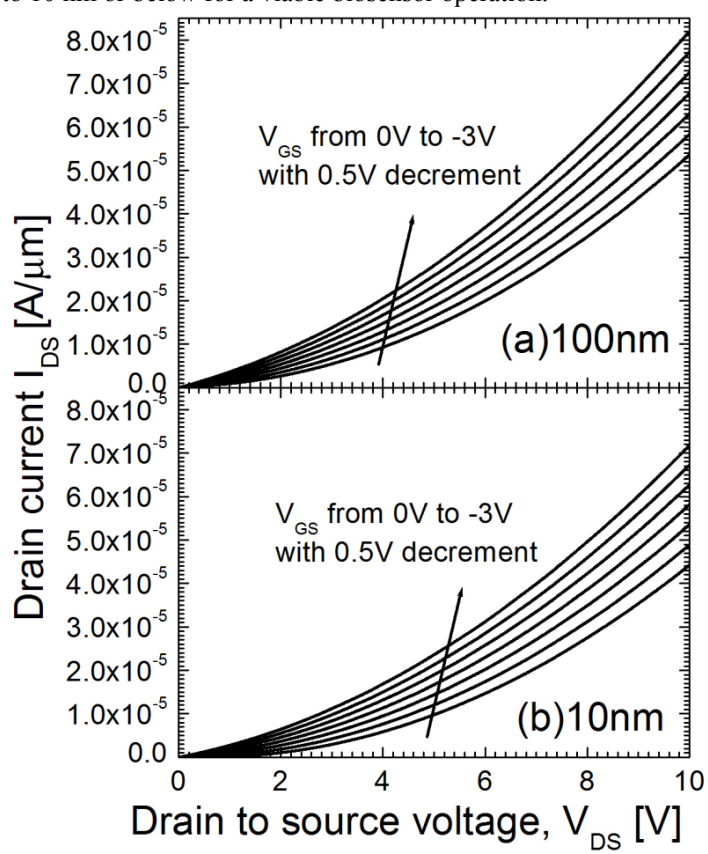


Figure 5: Output characteristics of p-type polysilicon nanowires with $10^{18}/cm^3$ doping concentration at different thicknesses; (a) 100nm and (b) 10nm.

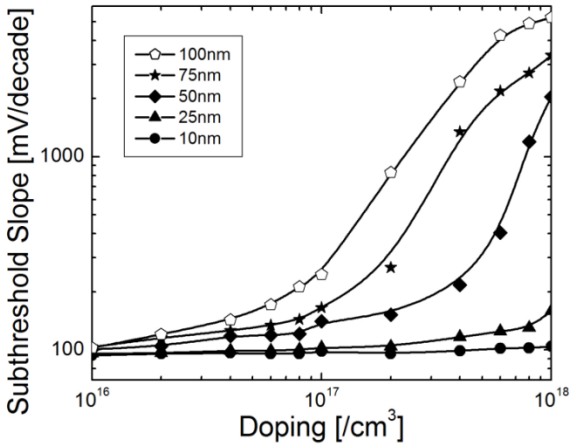


Figure 6: Sub-threshold slopes as a function doping concentrations for different thicknesses of polysilicon nanowires.

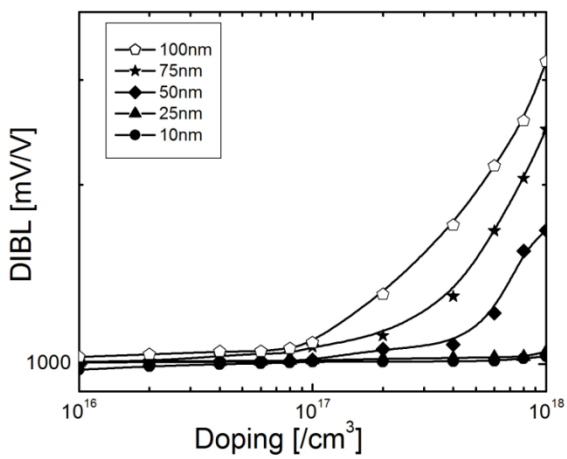


Figure 7: Drain induced barrier lowering (DIBL) as a function doping concentrations for different thicknesses of polysilicon nanowires.

To elucidate the proper combination of nanowire thickness and doping to ensure the sensitive operation of polycrystalline silicon nanowire biosensors, figure. 6 summarizes the extracted values of sub-threshold slopes as a function doping concentrations for different thicknesses of polysilicon nanowires. For nanowire thicknesses of 100 nm and 75 nm, a severely degraded sub-threshold slope can be observed when doping concentrations are above $10^{17}/cm^3$. The polysilicon nanowires exhibit sub-threshold slopes of 5246 mV/dec and 3344 mV/dec respectively for 100 nm and 75 nm nanowire thicknesses at a doping concentration of $10^{18}/cm^3$. A plausible sub-threshold slope around 100 mV/dec for a viable biosensor operation at these thicknesses can only be achieved if doping concentration is $2\times10^{16}/cm^3$ or below. For a 50 nm nanowire thickness a relatively wide doping concentration range can be chosen for biosensor design while maintaining decent sub-threshold characteristics. In this thickness a doping up to $4\times10^{17}/cm^3$ with a sub-threshold slope around 100 mV/dec can be chosen. The widest range of doping concentrations can be chosen for 25 nm and 10 nm nanowire thicknesses with a maximum doping up to $10^{18}/cm^3$ while maintaining a promising sub-threshold slope around 95 mV/dec for a viable biosesnsor design using polycrystalline silicon nanowires. Finally figure 7 summarizes the extracted values of drain induced barrier lowering (DIBL) as a function doping concentrations for different thicknesses of polysilicon nanowires. The DIBL trend of polysilicon nanowires is similar to the observed sub-threshold slopes at different doping concentrations and nanowire thicknesses (figure 6) and showing the plausible combinations of doping concentrations and nanowire thicknesses for polycrystalline silicon nanowire biosensor fabrication.

IV. CONCLUSIONS

We have investigated the effect of nanowire thickness and doping concentration on the electrical characteristics of polycrystalline silicon nanowire biosensors. For nanowire thicknesses of 100 nm and 75 nm, a plausible sub-threshold slope around 100 mV/dec for a viable biosensor operation can only be achieved if doping concentration is $2\times10^{16}/cm^3$ or below. For a 50 nm nanowire thickness a relatively wide doping concentration range with a maximum doping up to $4\times10^{17}/cm^3$ can be chosen for biosensor design while maintaining decent sub-threshold characteristics. The widest range of doping concentrations can be chosen for 25 nm and 10 nm nanowire thicknesses with a maximum doping up to $10^{18}/cm^3$ while maintaining a promising sub-threshold slope around 95 mV/dec for a viable biosesnsor design using polycrystalline silicon nanowires. As such this research reveals the possible combinations of nanowire thickness and doping to ensure the sensitive operation of polycrystalline silicon nanowire biosensors.

ACKNOWLEDGEMENT

The authors would like to acknowledge the support of the Center for Research and Training (CRT), East West University, under HEQEP sub-project Knowledge Transfer and Capacity Development of Academic Staff.

REFERENCES

[1] Y. Cui, Q. Wei, H. Park and C. M. Lieber, “Nanowire and Nano sensor for highly sensitive and selective detection of biological and Chemical species,” *Science*,vol. 293, pp. 1289-1292, 2001.

[2] F. Patolsky, G. Zheng and C.M. Lieber, “Nanowire-Based Biosensors,” *Anal. Chern*. Vol. 78, pp. 4260-4269, 2006.

[3] M. C. McAlpine, R. S. Friedman, S. Jin, K. H. Lin, W. U. Wang and C. M. Lieber, “Highperformance nanowire electronics and photonics on glass and plastic substrates,”*Nano Lett.*, vol. 3, pp. 1531-1533, 2003.

[4] Z. Q. Gao, A. Agarwal, A. D. Trigg, N. Singh, C. Fang, C. H. Tung, Y. Fan, K. D. Buddharaju and J. M. Kong, “Silicon Nanowire Arrays for Ultrasensitive Label-Free Detection of DNA,” *Analytical Chemistry*, vol. 79, pp. 3291-3297, 2007.

[5] I. Park, Z. Li, A. P Pisano and R. S. Williams, “Top-down fabricated silicon nanowire sensors for real-time chemical detection,”*Nanotech.*,vol. 21, pp. 1-9, 2010.

[6] M. M. A. Hakim, M. Lombardini, K.Sun, F.Giustiniano, P.L.Roach, D.E. Davies, P.H.Howarth, M. R. R.de Planque, H. Morgan, P. Ashburn, “Thin film polycrystalline silicon nanowire biosensors,” *Nano Letters*, vol. 12, pp. 1868-1872, 2012.

[7] H. C. Lin, M. H. Lee, C. J. Su, T. Y. Huang, C. C. Lee and Y. S. Yang, “A simple and low-cost method to fabricate TFTs with poly-si nanowire channel,” *IEEE Electron Device Lett*., vol. 26, pp. 643-645, 2005.

[8] Silvaco international, *Atlas User’s Manual Device Simulation Software*, Silvao International Ltd., Santa Clara, Dec. 2002.

[9] P. Walker, H. Mizuta, S. Uno, Y. Furuta and D. Hasko, “Improved off-current and subthreshold slope in aggressively scaled poly-si TFTs with single grain boundary in the channel,” IEEE Trans. On Electron Devices, vol. 51, pp. 212-219, 2004.

[10] S. Chen, J. G. Bomer, W. G. Van der Wiel, E. T. Carlen and A. Van den Berg, “Top-Down Fabrication of Sub-30 nm Monocrystalline Silicon Nanowires Using Conventional Microfabrication,” *ACS Nano*, vol. 3, pp. 3485-3492, 2009.

Use of Single Bin Fourier Transform Algorithm for High Speed Tone Detection and Parallel Processing

Shahnewaz Karim Sakib[1], Farah Ferdaus[2], and Md. Saifur Rahman[3],
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka-1205, Bangladesh
Email: shahnewaz.karim.sakib@gmail.com[1], farahferdaus13@gmail.com[2], saifur@eee.buet.ac.bd[3]

***Abstract*— Fourier transform is the classical tool for performing spectral analysis of a short-duration signal and fast Fourier transform (FFT) is the widely used algorithm for achieving the same with a high degree of computational efficiency. Recently a *"single bin Fourier transform algorithm"* has been proposed by the present authors, which is designed to compute the spectral amplitude and phase of a single frequency present in the time signal to be analyzed. The traditional detection scheme involves sampling the analog signal first using the Nyquist's sampling theorem, apply the FFT on the sampled signal and then identify the desired frequency from the FFT output. However, in case of tone detection, only the presence of a single frequency is to be tested. The computation of the whole FFT output dataset is superfluous in this case. In this paper, we have used the single bin Fourier transform algorithm for tone detection, which makes the detection scheme less complex and computationally more efficient than the traditional FFT algorithm. Moreover, the method of parallel processing can be applied to make the system more efficient in detecting the existence of more than one frequency component present in the incoming signal. The main objective of the proposed scheme is to improve the detection speed of a single frequency, or a narrow band of frequencies, present in a signal.**

***Index Terms*—FFT, Single Bin Fourier Transform, Tone Detection, Parallel Processing.**

I. Introduction

The discrete Fourier transform (DFT) converts a finite set of equally spaced samples of a function into the set of coefficients of a finite combination of complex sinusoids, ordered by their frequencies, that has the same number of sample values. Mathematically, the equations of DFT can be written as [1]:

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{kn} \quad \dots\dots\dots (1)$$

where, W_N= Phase Factor = $e^{-j2\pi/N}$

FFT performs the computation of the above DFT in a faster manner. It is the most popular tool in the field of engineering, which has numerous applications ranging from the estimation of power spectra to the faster computation of convolution algorithms, even to the multiple sequence alignment in molecular biology [2]-[4].

Among the various implementations of FFT, radix-2 algorithm, proposed by James W. Cooley and John W. Tukey, is the most prevalent one [5]. FFT considers a range of frequency from a corresponding time function and computes the spectra for all the frequency components present in the signal. However, there are certain practical applications, such as tone detection, dual tone multi frequency signaling (DTMF), radar, electronic warfare or sonar system, where the existence of a single frequency or the presence of a narrow band of frequencies is required. The traditional FFT algorithm analyzes the magnitude and phase information of all the frequency components present in the signal (as dictated by the time window, T, and sampling frequency f_s). Isolating a particular frequency from the FFT output may be further time-consuming and tedious.

In the past, many algorithms have been proposed for single tone detection. In reference [6], various methods for single tone detection are evaluated and compared, but like the FFT algorithm, they consider all the components over the frequency domain thoroughly. The Goertzel Algorithm [7] is the most common method used for single tone detection. However, the Goertzel filter is marginally stable and vulnerable to numerical error accumulation when computed using low precision arithmetic and long input sequence [8]. Moreover, an impulse during the sampling period can cause the power detected by Goertzel filter to be significant and may result in possible false detection.

Hence, a method needs to be developed for the faster detection of a single frequency or a narrow band of frequencies. Fortunately, an algorithm, known as ***"single bin Fourier transform algorithm"***, has been proposed recently by the present authors [9], which serves the said purpose. In this paper the developed *"single bin Fourier transform algorithm"* has been used to improve the detection speed of a single frequency or a narrow-band of frequencies present in the signal. This algorithm requires much less computation than the traditional FFT algorithm. Furthermore, the main performance criterion, the detection speed, has been evaluated to show the superior performance of the proposed algorithm. The rest of the paper is organized as follows: Section II describes the difficulty of the existing FFT algorithm in tone detection. Section III presents the possible applications of the *"single bin Fourier transform algorithm"*. Section IV demonstrates the results and compares the performance with those of the FFT and other algorithms. Finally, Section V concludes the outcome of the research.

II. Difficulty Of Using The Existing FFT Algorithm In Tone Detection

Numerous FFT algorithms have been developed over the years. However, most FFT algorithms employ the symmetry properties of the complex exponential, which is the keystone of DFT. The Radix-2 algorithm of Cooley and Tukey [5] is a divide and conquer algorithm that recursively breaks down a DFT of any composite size into smaller DFTs, along with

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

multiplications by the complex roots of unity, traditionally called the "twiddle factors".

In the traditional FFT algorithms, the discrete time signal is generated using sampling and the corresponding sequence of FFT output is generated by using those samples as input. The samples, corresponding to the positive frequencies, are placed first and then, the sequence of samples, pertaining to the negative frequencies, are integrated to it.

The main theme of FFT is to convert signals into samples depending upon the sampling frequency and analyze all the samples to give the FFT output corresponding to a fixed range of frequencies. This method can be cumbersome in some cases, where the sample value of only a small fraction of the dataset is required. In this case, computing the value of all the samples is redundant. Moreover, if only one sample value is required, then this desired sample must be isolated out from all the samples of dataset and this isolation process will require additional delay in the overall process of detection or analysis.

The Goertzel Algorithm [7], for single tone detection, operates by generating individual DFT coefficients using a simple recursive filter and a second order digital resonator. Many algorithms have been proposed in the past for single tone detection. However, like the FFT algorithm, they consider computing all the frequency components in the range. There is no provision for computing a single frequency or a particular range of frequencies in those algorithms.

III. Possible Applications Of The Single Bin Fourier Transform Algorithm

In this section, use of the "single bin Fourier transform algorithm" for tone detection and parallel processing are presented as possible applications. It should be noted here that a single sample in frequency domain is known as a "frequency bin".

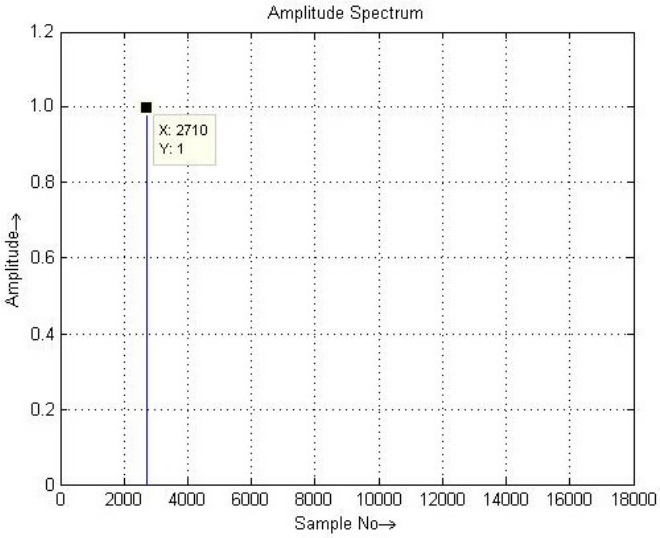
A. Use of Single Bin Fourier Transform Algorithm for Tone Detection

In order to remove redundancy in the computation of the amount of unwanted FFT samples, an effective method is to calculate the twiddle factor for the desired output sample only. To perform this operation, we need to evaluate the nature of the power of phase factor, *kn*, as the phase factor is constant for each sample for a given signal.

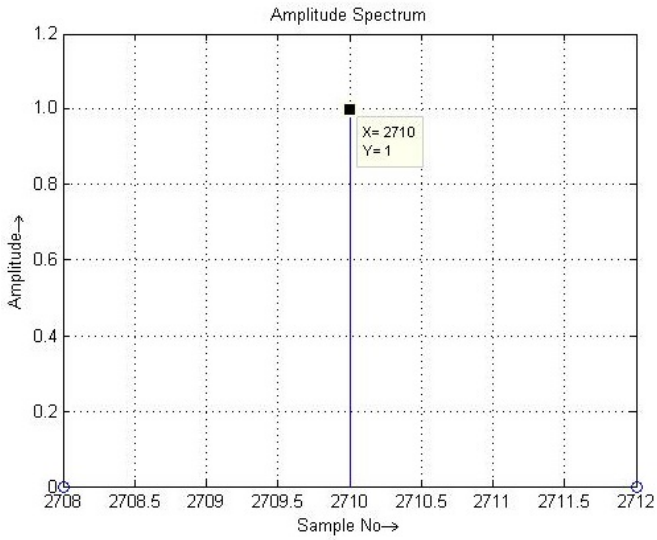
If the length of FFT is N, then we can generate a descending series from N to 1 and multiply it with the desired sample number, k. Finally, each term is divided by N and the remainder of each is considered to be the value of *kn*.

Therefore, the value of *kn* for each sample is created and the power of phase factor is raised to this sequence to get the twiddle factor corresponding to that sample only. Finally, the signal is multiplied with this twiddle factor and then summed to get the FFT value of the corresponding sample. Therefore, the whole algorithm requires only one adder and one multiplier. The detail of the "single bin Fourier transform algorithm" is included in reference [9].

Fig. 1(a) depicts the frequency domain sample output using the traditional FFT on the time signal and isolating the desired frequency bin, and Fig. 1(b), on the other hand, shows the result of the single bin Fourier transform directly.



(a)



(b)

Fig. 1: Tone Detection: (a) Amplitude spectrum using the traditional FFT followed by isolation, and (b) Amplitude spectrum obtained directly using the single bin Fourier transform algorithm.

As evident from Fig. 1(b), the detection of a single frequency from the input signal using the single bin Fourier transform algorithm is exactly the same as detected by the traditional FFT. However, the beauty of the single bin Fourier transform algorithm is that it improves the detection speed considerably, making the system computationally more efficient. This is further discussed in Section IV.

B. Use of Single Bin Fourier Transform Algorithm for Parallel Processing

Parallel processing scheme using the single bin Fourier transform algorithm can be applied in detecting multiple frequency components present in the input signal. A mode of operation in which a process is split into parts, and is executed concurrently is known as parallel processing. In Signal processing parallel checking of multiple frequencies can be performed.

Decisions based on the presence of certain frequencies can be performed simultaneously, and then combining them in the right order, we get the desired output; hence the process reduces computational time. This will make the spectral analysis process or the detection scheme more efficient. Such a parallel processing scheme is shown in Fig. 2.

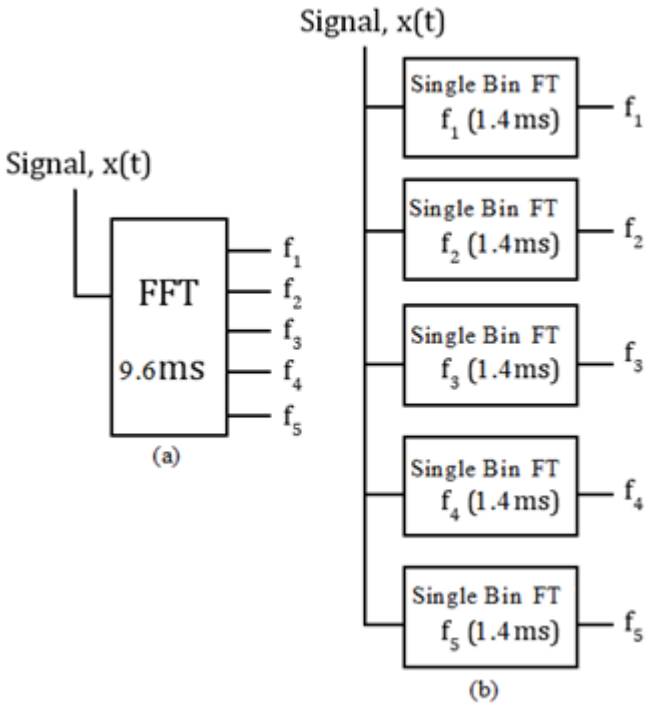


Fig. 2:(a) Traditional FFT block showing 5 output samples in frequency domain, (b) 5*"single bin Fourier transform algorithm"* blocks for the parallel processing scheme.

Fig. 2(b) shows 5 single bin Fourier transform algorithm blocks to detect five separate single frequencies. The corresponding traditional FFT algorithm block is shown in Fig. 2(a). Let us assume that the received signal, *x(t)*, of time window T (the data window time, when sampled), has some frequency components. Suppose we are interested to check if some specific frequencies like, $f_1=\Delta f$, $f_2=2\Delta f$, $f_3=3\Delta f$, $f_4=4\Delta f$ and $f_5=5\Delta f$ are present in the received signal. (Here, Δf is the frequency in Hz, which represents the spectral resolution of the traditional FFT algorithm and is given by $\Delta f = 1/T$). Therefore, we shall pass the received signal, *x(t)*, into the five blocks, where each block is designed to detect the corresponding single frequency. This whole process can be detected far quickly using the proposed method. The traditional FFT algorithm takes 9.6ms on the average, to detect each of them, whereas the proposed parallel processing scheme takes 1.4*ms* to detect. Thus, the computational efficiency is apparent from the example, as the desired information can be extracted in a much faster time (refer to Table I in Section IV).

This scheme can be used to perform spectral analysis of, or enable the detection of a narrow band of frequencies present in the incoming/received signal. It may also be used in place of the traditional FFT algorithm. For example, if the traditional FFT performs analysis of a frequency range of 0, Δf, $2\Delta f$, … … $(N/2-1)\Delta f$ Hz, the same scheme can be implemented to include $N/2$ no. of parallel "single bin Fourier transform algorithm" blocks (similar to those shown in Fig. 2(b)) to perform the same spectral analysis. As each single bin Fourier transform block now requires much less time than that of the traditional FFT, the whole process speeds up to a great extent.

The *"single bin Fourier transform algorithm"* can primarily be used for fast detection of the presence or absence of a signal of desired frequency in an incoming or received signal.

Efficient detection of a tone of particular frequency is easily achievable by using this method. It can also be used to find the energy of the signal of a certain frequency. Besides, other necessary spectral information, such as the dominant frequency, distortion, bandwidth etc. can easily be examined by applying this algorithm either in single mode or in parallel processing mode.

Again, making proper choice of the parallel processing scheme will enable us to dynamically apply the technique to a multi-tone detection system that can be used for paging and alerting. This algorithm will make sure of the proper detection of a tone sent by another user and can also be used for detection of pitch.

IV. MATLAB SIMULATION AND RESULTS

The simulation results, presented in this paper, are performed using the MATLAB [10] programming environment. A signal has been generated and passed through the system as defined by the Single bin Fourier transform algorithm. The outcome has been monitored for examining various performance criteria. The results are prepared with a Core-i3 processor, having a clock frequency of 1.7GHz and a fixed RAM of 4GB.

It has been found that there is no difference between the output amplitudes, obtained from the single bin Fourier transform and the traditional FFT followed by isolation for the desired frequency bin.

Table I reports the computational performance between the traditional approach in computing FFT and isolating the desired frequency, and the *"single bin Fourier transform algorithm"* Five different samples have been considered for the analysis to assess the performance between the two methods (Here, T=2 s, f_s=5,290 Hz and N=16,384 were used).

TABLE I. PERFORMANCE COMPARISON BETWEEN THE TRADITIONAL FFT FOLLOWED BY ISOLATION AND THE SINGLE BIN FOURIER TRANSFORM

Trial No.	**Computation Time** (*ms*)	
	Performing FFT and isolating the desired frequency bin	***Single bin Fourier transform***
01	9.709	1.634
02	9.576	1.391
03	9.412	1.416
04	9.396	1.358
05	9.413	1.411

From Table I, it can be noted that the proposed single bin Fourier transform scheme reduces the computation time substantially and makes the detection scheme almost seven times faster.

Table II indicates the comparison of computation time among various existing FFT algorithms [11] and the proposed approach for various FFT order. The advantage of the single bin Fourier transform algorithm for the detection scheme is visible here as the computation time for each case is lower in the single bin Fourier transform than that computed from the traditional FFT.

TABLE II. COMPARISON OF COMPUTATION TIME FOR VARIOUS ALGORITHMS

Algorithm	FFT Order		
	1024	***4096***	***16384***
Computation Time	*(μS)*	*(μS)*	*(μS)*
RAD2 [10]	1960	6800	30500
RAD4 [10]	1800	6940	29000
SRFFT [10]	660	3700	17260
FHT [10]	560	3240	14020
QFT [10]	1020	5460	27760
DITF [10]	1780	8780	40200
Single bin Fourier transform	486	1004	2026

Finally, Table III specifies the memory usage (in Kilobytes) for each of the existing algorithm and for the proposed method. The memory required for our method is also less than those of the existing FFT algorithms.

TABLE III. MEMORY USAGE IN COMPUTING 1024 POINT FFT

Algorithm	Memory Usage (*Kilobytes*)
RAD2	98.2
RAD4	94.2
SRFFT	94.6
FHT	90.2
QFT	101.9
DITF	96.4
Single bin Fourier transform	86.3

V. CONCLUSION

FFT is the traditional tool for signal analysis and it may be used for tone detection. However, single bin Fourier transform algorithm surpasses it in signal detection, especially in detecting a single frequency tone and a narrow band of frequencies present in a signal by using the parallel processing scheme proposed in this paper. The advantage provided by the reduced computation time can be versatile and various possible applications conform to this statement. Now a days, high speed processors and computers have become available. If employed in parallel processing mode, it can also outperform the traditional FFT in computing the spectral components of a block of time samples of a signal (N-sample points), by employing the parallel processing scheme of the single bin Fourier transform algorithm as proposed in Fig. 2(b).

Therefore, the proposed scheme is more preferable compared to the established methods for some specific applications, and the associated improvement makes it more desirable for the transmission and detection of communications signal.

VI. REFERENCES

[1] K. Katoh, K. Misawa, K. Kuma, T. Miyata, "MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform", Nucleic Acids Research, 2002.

[2] J. G. Proakis, D. G. Manolakis, "Digital Signal Processing- Principles, Algorithms and Applications", 3rd Ed., 1996, Prentice Hall International Inc., Upper Saddle River, NJ, USA.

[3] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms", IEEE Transactions on Audio and Electroacoustics, Vol. 5, Issue. 2, pp. 70-73, Jun 1967.

[4] H. J. Nussbaumer, "Fast Fourier Transform and Convolution Algorithms", 2012.

[5] J. W. Cooley, J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series", Mathematics of Computation, 1965.

[6] Y. T. Chan, Q. Ma, H. C. So, and R. Inkol, "Evaluation of Various FFT Methods for Single Tone Detection and Frequency Estimation", IEEE Canadian Conference on Electrical and Computer Engineering, Vol. , May 1997, pp. 211-214.

[7] R. Beck, A. G. Dempster, and I. Kale, "Finite Precision Goertzel Filters Used for Single Tone Detection", IEEE Trans. On Circuits and Systems – II: Analog and Digital Signal Processing, Vol. 48, No. 6, June 2001, pp. 691-700

[8] W. M. Gentlemen, "An error analysis of Goertzel's (Watt's) method for computing Fourier coefficients", The Computer Journal 12(2), February 1969.

[9] M. S. Rahman, S. K. Sakib and F. Ferdaus, "Analyses of Fourier transform, FFT and Single Bin Fourier Transform Algorithms for High Speed Spectral Analysis and Computation", July 2016, (Unpublished).

[10] MATLAB [Online]. Available: http://www.mathworks.com/products/matlab

[11] M. Soni, P. Kunthe, "A General Comparison of FFT algorithms", 9th National Conference, Pioneer Journal for IT and Management, March 2011.

Part Definition and Selection for Part-Based Speed Invariant Gait Recognition

Md. Rokanujjaman[1,*], Md. Altab Hossain[1,3], Md. Rezaul Islam[2], AKM Akhtar Hossain[1], and Alexander Ferworn[3]

[1]Department of Computer Science and Engineering, University of Rajshahi
Rajshahi 6205, Bangladesh

[2]Department of Applied Physics and Electronic Engineering, University of Rajshahi
Rajshahi 6205, Bangladesh

[3]Department of Computer Science, Ryerson University
Toronto, Ontario, Canada

[*]rokon_cstru@yahoo.com

Abstract— **Walking speed change is one of the most common cofactor that affects the gait signature. The intra-class variations increases by the changes of walking speed and the training data set fails to model the variations. As a consequence, the performance of the gait recognition systems is degraded. Naturally, human reconstruct the model by considering the unaffected body parts to recognize a person when some parts are affect due to the walking speed transition. In this paper, we propose a general feature set based method to recover from the speed transition by using the less affected body parts. The OU-ISIR Treadmill Database A is used to show the effectiveness of the proposed method.**

Index Terms— **Gait recognition, parts definition, parts selection, speed invariant, intra-class variation.**

I. Introduction

Gait of a person is derived from his/her physical build and body weight. However, the gait is affected by many cofactors including clothing, shoe type, heel height, walking speed, walking surface, illness, emotional state, and environment. Human gait is one of the behavioural biometrics that has several unique characteristics: (i) it can be recognized even at low image resolution [7], [13], (ii) it can be captured without the understanding of a person, and (iii) it is difficult to spoof. Considering these characteristics, applications of the gait identification are ranging from video-based wide-area surveillance [16], [18] to criminal investigation (e.g., authenticating a perpetrator at a crime scene).

Gait recognition approaches are generally divided into two broad categories: (i) Model-based and (ii) Model-free techniques.

The model-based approaches are used a priori knowledge of the human gait. More specifically, those methods extract the gait features such as shapes and motions by fitting the human model to input images. These gait features contain kinematics of leg motion by Fourier analysis [22], static shape parameters and gait period with an articulated body model [19], and joint angles with an articulated body model [18].

These methods have limited efficiency because of the high computational costs based on complex matching and searching. They also often suffer by model fitting errors. In fact, the study [20] reports that high quality gait image sequences are required to achieve a high accuracy of model fitting.

On the other hand, model free methods [12] operates directly on a video sequence involving individuals walking and characterize the manner of walking based upon the prediction or estimation of features related to position, velocity, shape, texture and colour. Model-free methods [3], [21] are comparatively insensitive to the quality of gait silhouettes and have the benefit of less computational costs and complexity compared to model-based approaches. These approaches show better performance in comparison with model-based methods. However, significant change in appearance due to the effect of different cofactors makes the model free gait recognition task much more complicated.

The human gait is composed of different body parts. The effects of different cofactors do not change all body parts. It may affect some parts of whole gait whereas other parts that are useful for gait identification remained unchanged.

With regard to the different part-based gait representation techniques, there is still a major issue to define the effective body components that influence the gait recognition under the effect of different cofactors. Although a variety of part-based techniques [6], [10], [13], [14] have been proposed, the body parts are defined manually in all the studies. Part-based techniques [11], [12] define more effective and less effective body parts based on the experimental result.

There are two important ways to improve the gait recognition accuracy in case of different cofactors: (i) Gait features should be represented with the most discriminating information, and (ii) For selecting the appropriate body parts, more affected and less affected body parts should be considered.

Out of the different cofactors walking speed variation is the most challenging because the of affected dynamic body parts of the gait due to the speed changes. Recent study [2] shows the walking speed invariant identification has two main categories: (i) speed transformation based, and (ii) invariant feature based.

The first category of methods are based on the transformation of features from one speed to other speed before classification. In [4] Tanwongsuwan et. al. developed a stride normalization procedure to model walking gaits across different speeds. Tsuji et. al. [5] proposed a method of gait silhouette transformation from one speed to another by factorization based speed transformation model using SVD.

The second type of methods are based on the extraction of speed invariant feature from gait template to minimize intra class variations. Procrustes Shape Analysis (PSA) descriptors based method [2], [7] have developed to handle large changes of walking speed. In [7], Kusakunniran et. al. proposed Higher-order derivative Shape Configuration (HSC) to extract speed invariant gait features from PSA descriptors. They also extended the HSC based on Differential Composition Model

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

(DCM) [2] to compute and assign weights to different body parts. DCM shows significant performance improvements than HSC considering large speed changes. A classifier ensemble framework based on Random Subspace Method (RSM) [8] is proposed by Guan et. al. to solve the cross speed gait recognition problem. In RSM, each experiment is executed ten times and calculate the mean CCR for the randomness of their method. Although they achieved better performance, the method does not guarantee the best accuracy [9] due to the randomness property of the technique. The time complexity of the RMS technique is definitely increased due to multi-time execution of each experiment.

In this paper, we propose a general feature based method for parts definition and selection to solve the speed invariant gait recognition problem by discarding the affected body parts. The main contribution of the proposed system is to define the body parts and construct a look up table to select the most effective parts from different speed combinations of the gait and discard the redundant part to minimize the intra-class variations.

II. GAIT FEATURE REPRESENTATION AND WALKING SPEED ESTIMATION

The OU-ISIR gait database was recorded at 60 f/s and the image size is 640 × 480 pixel. To represent the gait features, most of the model-free gait recognition approaches use silhouettes. Silhouettes are extracted from each frame of a sequence by background subtraction. In pre-processing stage, each background-subtracted silhouette is registered to obtain the spatio-temporal gait silhouette volume (GSV) [16]. Then, the silhouettes are normalized into a fixed size of 128 × 88 pixel and calculate the normalized autocorrelation of a GSV to detect the gait cycle N_{gait}. In this paper, we used 3 gait cycles for each sequence.

A. Feature Representation

The average silhouettes over a complete gait cycle is called Gait Energy Image (GEI) [15]. The example of GEI is shown in Fig. 1. Most of the appearance-based approaches [11], [12], [15] used GEI as gait features. GEI represents both the static and dynamic areas of the gait. Given the pre-processed binary gait silhouettes $B(x, y, n)$ at time n in a sequence, the GEI is defined as,

$$GEI(x, y) = \frac{1}{N}\sum_{n=1}^{N} B(x, y, n) \qquad (1)$$

where, N is the number of frames in a gait cycle, n is the frame number and x and y values in the 2D image coordinates.

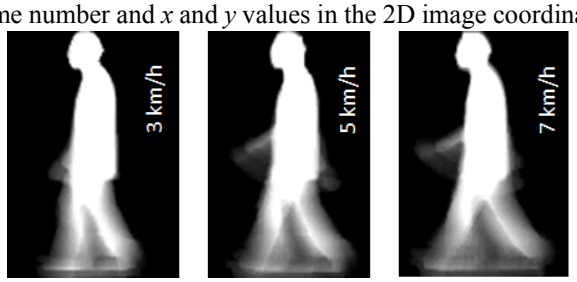


Fig. 1 Gait Energy Image (GEI).

B. Gait Speed Estimation

It is possible to automatically estimate test dataset walking speed of each gait sequence from training dataset by using classification. Usually, walking speed is measured in a unit of km/h. Each gait cycle contains a number of frames depending on walking speeds. Table 1 shows the number of average frames per gait cycle estimated under the various walking speeds based on training dataset. The number of frames decreases with walking speed change from lower to higher.

TABLE I
NUMBER OF FRAMES PER GAIT CYCLE FOR DIFFERENT SPEEDS.

Speed (km/h)	Average no. of Frames per gait cycle	Speed (km/h)	Average no. of Frames per gait cycle
2	88	5	63
3	75	6	59
4	69	7	55

III. MATCHING MEASURE

The Euclidean distances between feature vectors for each gallery and probe for the sub-sequences with N_{gait} frames are used as the matching measure.

Let $\{x_j^p\}$ (j=1 to N_{gait}) and $\{x_k^g\}$ (j=1 to N_{gait}) be the subsequence's for the probe $\boldsymbol{x}^p$ and the gallery $\boldsymbol{x}^g$ respectively. The Euclidean distance between a probe and a gallery $d^{subs}(x^p, x^g)$, is the matching measure for the sub-sequences.

The median value of the minimum distances of the combinations of sub-sequences of each probe and gallery is defined as:

$$d^{\min}(x^p, x^g) = \min_j[\min_k\{d^{subs}(x_j^p, x_k^g)\}] \qquad (2)$$

IV. PARTS DEFINITION AND SELECTION

For dividing the body horizontally into different parts, we study the contribution of small segment in terms of identification rate in the training dataset. The smallest segment in this case is a row as a pixel has no effect for horizontal part selection.

A. Part Definition

We measure the recognition rate in row wise starting from top row of GEI feature and then merging a next lower single row in each step.

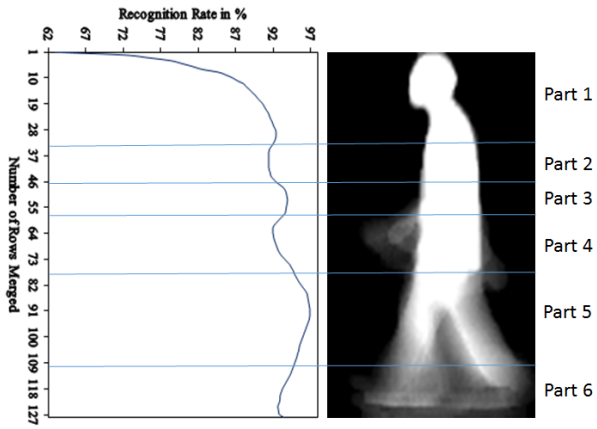


Fig. 2 Parts definition based on local maxima and local minima.

The process continues from top to bottom. Stepwise recognition rate is shown in Fig. 2. It clearly shows sometimes a row has positive effect and sometimes has negative effect. Form this indication; we can define the body parts based on the positive or negative contributions of consecutives rows.

TABLE II
DIFFERENT COMBINATION OF PARTS (1-6) WITH CROSS SPEED RECOGNITION RATE WITHIN BRACKET (%).

Probe / Gallery	2km/h	3 km/h	4 km/h	5 km/h	6 km/h	7 km/h
2 km/h	1, 2, 3, 4, 5, 6 (100%) 1, 2, 3 (100%)	1, 2, 3, 5, 6 (100%) 1, 2 (100%)	1, 2, 3 (94.4%) 1, 2 (94.4%)	1 (94.4%)	1 (83.3%)	1 (72.2%)
3 km/h	1, 2, 3, 5 (100%) 1, 2, 3 (100%) 1, 2 (100%)	1, 2, 3, 4, 5, 6 (100%) 1, 2 (100%)	1, 2, 3, 4, 5, 6 (94.4%) 1, 2, 3 (94.4%) 1, 3, 6 (94.4%) 1, 2 (94.4%)	1, 2, 5 (83.3%) 1, 2 (83.3%)	1 (83.3%)	1 (72.2%)
4 km/h	1, 2, 3, 6 (94.4%) 1, 2, 3 (94.4%) 1, 2 (94.4%)	1, 3, 6 (83.3%) 1, 3 (83.3%)	1, 3, 4, 5, 6 (94.4%) 1, 3, 4, 5 (94.4%) 1, 3, 4 (94.4%) 1, 3 (94.4%)	1, 2, 3 (94.4%)	1 (77.8%)	1 (77.8)
5 km/h	1, 2 (88.9%)	1, 2, 6 (83.3%) 1, 2 (83.3%)	1, 2, 3 (100%)	1, 3, 4, 5, 6 (100%) 1, 3, 4, 5 (100%) 1, 3, 4 (100%)	1 (88.9%) 1, 6 (88.9%)	1 (88.9%)
6 km/h	1 (88.9%)	1, 2 (94.4%)	1, 2 (72.2%) 1 (72.2%)	1 (100%)	1, 2, 3, 4, 5, 6 (100%) 1, 2, 3, 4, 5 (100%) 1, 2, 3, 4 (100%)	1, 6 (94.4%) 1 (94.4%)
7 km/h	1 (72.2%)	1, 2, 6 (55.6%) 1, 2 (55.6%) 1 (55.6%)	1 (77.8%)	1 (88.9%)	1, 3, 4 (94.4%) 1, 3 (94.4%) 1, 4 (94.4%) 1(94.4%)	1, 3, 4 (94.4%) 1, 3 (94.4%) 1, 4 (94.4%) 1 (94.4%)

TABLE III
PARTS (1-6) SELECTION LOOKUP TABLE FOR CROSS SPEED GAIT RECOGNITION.

Probe / Gallery	2 km/h	3 km/h	4 km/h	5 km/h	6 km/h	7 km/h
2 km/h	1, 2, 3, 4, 5, 6	1, 2, 3, 5	1, 2, 3	1	1	1
3 km/h	1, 2, 3, 5	1, 2, 3, 4, 5, 6	1, 3, 6	1, 2	1	1
4 km/h	1, 2, 3	1, 3, 6	1, 3, 4, 5, 6	1, 2, 3	1	1
5 km/h	1	1, 2	1, 2, 3	1, 3, 4, 5, 6	1	1
6 km/h	1	1	1	1	1, 2, 3, 4, 5, 6	1
7 km/h	1	1	1	1	1	1, 3, 4

We follow the curve from top row (minima here) until local maxima is found. After getting one local maxima, we search for next local minima, we find out mean between the first local maxima and second local minima and divide the body here. We then search for another local maxima and take the mean value of the second minima and divide the body. Similarly, by searching maxima-minima or minima-maxima and calculate their mean and divide the body there. If the last minima is the bottom, there is no division of the body. Following the procedure, human body is divided into six parts (Fig. 2).

B. Part Selection

As described in section IV (A), the body is divided into six parts based on the positive or negative contributions of consecutives rows the body. This body parts are selected for cross speed gait recognition based on training set (Table 3). We measure recognition rate for cross speed gait by combining different body parts of the training set. We observed that for some combination the recognition rate increased and for some combination the recognition rate decreased. The best combinations is shown in Table 2 with recognition rate in bracket.

Form Table 2, it is observed that for same gallery probe combinations more than one combinations shows same recognition rate. In addition, for cross speed shows that different combination of parts is effective. Therefore, we can select only the common parts as effective for the cross speed gait recognition task (Table 3).

V. EXPERIMENTS

For part definition and selection experiments, the OU-ISIR Treadmill Gait Database A [17] is used which includes a large range of walking speed. The database contains 34 subjects (26 males and 8 females), and speed variation ranges from 2 km/h to 7 km/h at 1 km/h interval. According to [5] nine subjects are selected for training process and remaining 25 subjects are used to evaluate gait recognition performance. Two video sequences were recorded for each subject from each speed. We used a training dataset containing 108 sequences of 9 subjects of different speed from the gait database.

For performance evaluation, a gallery set of 25 subjects excluding the 9 training subjects is used from the database. The probe set includes 300 sequences for these 25 subjects with 2 km/h to 7 km/h walking speed.

TABLE IV
CROSS SPEED RECOGNITION RATE OF THE PROPOSED METHOD (%).

Probe / Gallery	2 km/h	3 km/h	4 km/h	5 km/h	6 km/h	7 km/h
2 km/h	100	92	94	88	84	74
3 km/h	92	100	88	92	84	74
4 km/h	78	100	100	98	94	88
5 km/h	86	100	96	100	86	84
6 km/h	74	80	86	94	100	90
7 km/h	76	72	82	88	94	100

The rank 1 recognition rate for the proposed method is shown in Table 4.

TABLE V
CROSS SPEED RECOGNITION RATE WITHOUT DCM METHOD [2] (%).

Probe / Gallery	2 km/h	3 km/h	4 km/h	5 km/h	6 km/h	7 km/h
2 km/h	100	96	84	72	72	72
3 km/h	100	100	96	80	76	60
4 km/h	76	96	96	92	92	80
5 km/h	76	76	96	96	100	96
6 km/h	64	68	80	96	100	96
7 km/h	56	68	80	96	100	100

VI. RESULT AND DISCUSSION

The cross speed gait recognition rate in rank-1 is reported in Table 4 and compared in Table 5. The proposed method shows better performance with respect to cross speed gait recognition without DCM [2]. It also shows that recognition rate span from 72% to 100% of the proposed method, however recognition rate without DCM [2] span from 56% to 100%.

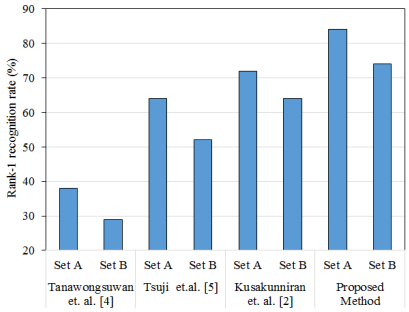


Fig. 3 (a) Comparison with other methods with large speed changes.

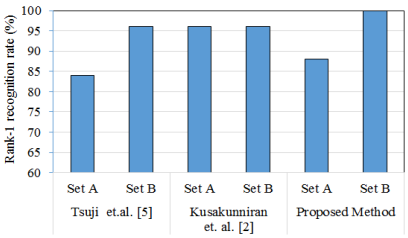


Fig. 3 (b) Comparison with other methods with small speed changes.

The proposed method is compared with Kusakunniran et. al. [2] without DCM, Tanwongsuwan et. al. [4], and Tsuji et. al. [5]. The results are shown in Fig. 3(a) and Fig. 3(b). Here, dataset of slower walkers is used in Set A as the gallery and vice versa in Set B. In Fig. 3(a), large speed change Tanwongsuwan et. al. [4] match between 2.5 km/h and 5.8 km/h, and Tsuji et. al. [5], without DCM [2] and the proposed method match between 2 km/h and 6 km/h. In Fig. 3(a), the proposed method shows much better results. In Fig. 3(b), small speed change Tsuji et. al. [5], Kusakunniran et. al. [2] without DCM and the proposed method match between 3 km/h and 4 km/h and achieved recognition rate 100% for Set B. Overall, the performance of the proposed method is better than other methods.

VII. CONCLUSION

We propose a method of part definition and selection for human gait recognition to cope with the intra-class variations due to the large changes of walking speed. For the cross speed gait recognition, the identification rate is higher when the speed changes is smaller, but the recognition rate decreases with the increases of walking speed difference which is reported in Table 4 and Table 5. The dynamic features are greatly dependent on speed changes while the static features are less dependent (Table 2). A future works is the extension of the proposed method to improve the feature extraction technique for the selected parts.

REFERENCES

[1] Ariyanto G, Nixon M, *Marionette mass-spring model for 3D gait biometrics,* In: Proceedings of the 5th IAPR international conference on biometrics. New Delhi, pp 354–359, 2012.

[2] W. Kusakunniran, Q. Wu, J. Zhang, and H. Li, *Gait recognition across various walking speeds using higher order shape configuration based on a differential composition model*, IEEE Transactions on Systems, Man, and Cybernetics, PartB: Cybernetics, 42(6):1654–1668, 2012.

[3] Bashir K, Xiang T, Gong S, *Gait recognition without subject cooperation,* Pattern Recog. Lett 31:2052–2060, 2010.

[4] R. Tanawongsuwan, A. Bobick, *Modelling the effects of walking speed on appearance-based gait recognition*, In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, volume 2, pages 783–790, 2004.

[5] A. Tsuji, Y. Makihara, Y. Yagi, *Silhouette transformation based on walking speed for gait identification*, In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pages 717–722, 2010.

[6] Boulgouris N, Chi Z, *Human gait recognition based on matching of body components,* Pattern Recogn 40:1763–1770, 2007.

[7] W. Kusakunniran, Q. Wu, J. Zhang, H. Li, *Speed invariant gait recognition based on procrustes shape analysis using higher-order shape configuration*, In Proceedings of IEEE International Conference on Image Processing, pages 545–548, 2011.

[8] Guan Y, Li C-T, *A Robust Speed-Invariant Gait Recognition System for Walker and Runner Identification*, International Conference on Biometrics, Madrid, pages 1-8, 2013.

[9] Kuncheva LI, Roli F, Marcialis G, Shipp C, *Complexity of data subsets generated by random subspace method: an experimental investigation, multiple classifier systems*, In: Proceedings of 2nd international workshop on multiple classifier systems, vol 2096. pp 349–358, 2001.

[10] Hossain M, Makihara Y, Wang J, Yagi Y, *Clothing invariant gait identification using part-based clothing categorization and adaptive weight control*, Pattern Recog. 43:2281–2291, 2010.

[11] M Rokanujjaman, MS Islam, MA Hossain, MR Islam, Y Makihara, Y Yagi, *Effective part-based gait identification using frequency-domain gait entropy features*, Multimedia Tools and Applications 74 (9), 3099-3120, 2015.

[12] Rokanujjaman M, Hossain M, Islam M, *Effective part selection for part-based gait identification*, 7th international conference on electrical and computer engineering. BUET, pp 17–19, 2012.

[13] Lee L, Grimson W, *Gait analysis for recognition and classification*, In: Proceedings of the fifth IEEE conference on face and gesture recognition. Washington DC, pp 148–155, 2002.

[14] Li X, Maybank S, Yan S, Tao D, Xu D, *Gait components and their application to gender recognition*, IEEE Trans Syst Man Cybern 38:145–155, 2008.

[15] Liu Z, Sarkar S, *Simplest representation yet for gait recognition: averaged silhouette*, In: Proceedings of international conference on pattern recognition, Cambridge, pp 211–214, 2004.

[16] Makihara Y, Sagawa R, Mukaigawa Y, Echigo T, Yagi Y, *Gait recognition using a view transformation model in the frequency domain*,. In: Proceedings of the ninth European conference on computer vision. Graz, pp 151–163, 2006.

[17] The OU-ISIR gait database, http://www.am.sanken.osaka-u.ac.jp/BiometricDB/GaitTM.html, 2010.

[18] Urtasun R, Fua P, *3D tracking for gait characterization and recognition*, 6th IEEE international conference on automatic face and gesture recognition. Seoul, pp 17–22, 2004.

[19] Wagg D, Nixon M, *On automated model-based extraction and analysis of gait*, 6th IEEE international conference on automatic face and gesture recognition. Seoul, pp 11–16, 2004.

[20] Wang J, She M, Nahavandi S, Kouzani A, *A review of vision-based gait recognition methods for human identification*, International conference on digital image computing techniques and applications, Sydney, pp 320–327, 2010.

[21] Wang L, Tan T, Ning H, Hu W, *Silhouette analysis based gait recognition for human identification*, IEEE Trans Pattern Anal Mach Intell 25:pp 1505–1518, 2003.

[22] Yam C, Nixon M, Carter J, Automated person recognition by walking and running via model based approaches, Pattern Recogn 37:pp 1057–1072, 2004.

Numerical Study on ITO Surface Modification Technique To Improve Single Emissive Layer Based Organic Light Emitting Diode Performance

Riku Chowdhury,[1,*] Md. Rashedul Haq,[1] Md. Sarwar Uddin Chowdhury,[2] Sharmin Afrose,[2] and Sukanta Paul [3]

[1]Electrical & Electronic Engineering Department, University of Science and Technology Chittagong
Foy's Lake, Chittagong, Bangladesh

[2]Electrical & Electronic Engineering Department, Chittagong University of Engineering and Technology
Raojan, Pahartoli, Chittagong, Bangladesh

[3]Computer Science and Engineering Department, University of Science and Technology Chittagong
Foy's Lake, Chittagong, Bangladesh

[*]eeeriku@gmail.com

***Abstract*— To improve the performance and operational life time of an organic light emitting diode (OLED), a single uniformly mixed-host (MH) emissive layer (EML) based OLED structure with modified anode surface has been proposed in this work. Through the numerical simulation, the feasibility of the proposed modified work function of ITO based device structure has been justified from the electrical characteristics and efficiency curves. By the ITO surface modification technique, the interfacial energy barrier of ITO/EML is reduced enough to increase hole injection rate into the EML. Therefore, the carriers become more balanced which raise the recombination rate in the EML. From the calculated current balance factor from the simulation results, it is confirmed that due to ITO surface modification, the proposed MH-OLED able to give better performance in terms of brightness, efficiency and longer operating time over unmodified ITO based MH-OLED.**

Keywords*— *indium tin oxide (ITO) surface modification; hole injection rate; charge transport; carrier recombination; current balance factor.

I. INTRODUCTION

For next generation display technology, it is being considered that organic light emitting diodes (OLEDs) is the most promising because of it's simple fabrication process, wide viewing angle, swift switching time (<1μs), broad operating temperature range (-46 °C to 70 °C) and lower turn-on voltage [1]. It was in 1987 when Tang and Vanslyke first successfully introduced the fabrication of this revolutionary OLED consisting of bi-layer organic materials [2]. In consequent years, numerous research work have been carried out to bring improvement in device performance over the conventional bi-layer OLED via modification of organic materials, layer architecture and composition [3-4]. To ensure improved performance and reliability in the conventional heterojunction (HJ) OLED, a mixed-host (MH) structure had been introduced where hole transport material (HTM) and electron transport material (ETM) mixed uniformly as the host material in the single emissive layer (EML) [3-4]. By using these mentioned transport materials the interface of hole transport layer (HTL) and electron transport layer (ETL) can be eliminated which enhances operation lifetime about 6 times as compared to the conventional HJ-OLED [4].

Ongoing OLEDs research articles are being focusing on charge balance or confinement in the EML for better recombination by improving the electron injection rate through the cathode [5]. The functionality of OLEDs is controlled by energy barrier at interface and recently an innovative research on surface work function of metal oxide electrodes corroborates its importance [6]. Having better electrical conductivity as well as excellent optical transmission properties, tin-doped indium oxide (ITO) has been well developed in industry as electrode which is widely used in opto-electronics devices ranging from solar cells to OLEDs [3-6]. But the potential barrier between ITO and the organic layer severely limits the efficiency of hole-injection. Recently, there has been reported in varied research articles and quotes a breakthrough in boosting the ITO surface work function by O_2 and Cl_2 plasma treatment [7], post annealing treatments, adding buffer layer [6] and UV-ozone treatments [8] . Simply by soaking ITO in different concentrations of Cesium Fluoride (CsF) solution and subsequent UV-ozone treatment to modify the ITO surface work function has been demonstrated by H.-W. Lu et al [9] qualitatively analyses the performance of conventional HJ structure.

Motivating from this simple OLED performance improvement technique, a single EML based uniform MH-OLED structure which consist of modified ITO has been proposed in this work. The physical structure of the proposed OLED is: ITO (modified with 5 wt% and 10 wt% CsF solutions) / NPB : Alq_3 (5:2) (100 nm) /LiF/Al. The experimental feasibility of this proposed MH-OLED has been justified through the numerical simulation.

Though, systematic study of OLED device performance enables to explain some effects (J-L-V characteristic and efficiency) qualitatively, but except numerical simulation explanation of these effects becomes abstruse [10]. Numerical simulation is more viable to get any analytical results directly

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

from the OLEDs where complicated working process and several interfaces are involved. Therefore, this numerical simulation work at first led to get the electrical characteristics curves (current density- luminance - voltage (J-L-V) and efficiency curves) to ensure the performance improvement of proposed MH device over unmodified ITO based device. Finally, the details internal device physics like, charge injection through anode/EML interface, transport and recombination along the EML those are impossible to measure or explain from the experimental results, have investigated to explain the effects of ITO surface modification on device performance improvement and increase operation life time.

II. Simulation Process

A. Electrical Model

To analysis the effect of ITO work function modification on the underlying physics and performance of MH-OLED, drift-diffusion electrical model has been implemented using a numerical simulator called 'SETFOS' (Semiconducting Thin Film Optics Simulation Software). Poisson equation and continuity equation as presented in "(1)," and "(2)," correspondingly considered in that electrical model to solve the electric field distribution and time evaluation of the system.

$$\frac{dE}{dx} = \frac{q}{\varepsilon\varepsilon_0}(p-n) \tag{1}$$

$$\frac{dn}{dt} = \frac{dJ_n}{qdx} - R - T \tag{2}$$

where, E is the electric field inside the OLED, q the elementary charge, ε the dielectric permittivity of the organic materials, ε_0 the dielectric constant, p and n are the hole density and the electron density respectively, J_n the electron current density, R the recombination rate and T the trapping rate for free carriers. The carrier density distribution can be obtained by solving the drift and diffusion equation based on the Langevin recombination (R) process those are calculated using "(3)," and "(4)," respectively. The conventional electric filed dependent 'Poole-Frenkel mobility (PFM)' model as expressed in "(5)," applied for carrier mobility calculation in the device. The current balance factor (CBF) (b) in "(6)," used to measure the device quantum efficiency [10] which is defined as the recombination current with the total current density.

$$J_n = qp\mu_n E + qD_n \frac{dn}{dx} \tag{3}$$

$$R = \frac{q}{\varepsilon\varepsilon_0}(\mu_p + \mu_n) \tag{4}$$

$$\mu(E) = \mu_0 \exp(\gamma * \sqrt{E}) \tag{5}$$

$$b = \frac{\int_0^L qR(x)dx}{J_T} \tag{6}$$

Here, μ_p and μ_n are the hole and electron mobility respectively, D_n the diffusion constant, $\mu(E)$ is the charge carrier mobility under electric field (E), μ_0 is the mobility under zero electric field, γ is the Poole-Frenkel field dependent factor, J_T is the total current density and L is the total device thickness.

B. MH-OLEDs Structure and Input parameters for simulation

Three uniform MH-OLED structures, namely A, B and C are presented in "Fig. 1," where NPB (HTM) and Alq_3 (ETM) mixed with a uniform ratio of 5:2 (NPB:Alq_3). For OLED simulations, parameters have to be defined for setting the energy level and mobilities of the materials. The parameters setting in EML of UM-OLEDs is the main problem due to the mixed transport materials. It is reported [11] that HOMO and LUMO in mixed-host EML may not be changed for HTM and ETM, respectively. Following the experimental observations [12], the mobilities in the EML are taken to be power-dependent of the concentration of mixture, therefore Equations (7) and (8) are used to calculate these mobilities:

$$\mu_{mix,P} = \mu_{HTM,P}^{C_{HTM}} * \mu_{ETM,P}^{(1-C_{HTM})} \tag{7}$$

$$\mu_{mix,n} = \mu_{HTM,n}^{C_{HTM}} * \mu_{ETM,n}^{(1-C_{HTM})} \tag{8}$$

where $\mu_{mix,P(n)}$ is the hole (or electron) mobility in mixed-host layer, $\mu_{HTM,P(n)}$ the hole (or electron) mobility of HTM, and $\mu_{ETM,P(n)}$ is the hole (or electron) mobility of ETM. C_{HTM} is the concentration of hole transport material in the EML. The used work function of ITO before surface modification is 4.78 eV which is increased to 5.01 eV and 5.11 eV after modification through CsF solutions of 5 wt% and 10 wt% respectively [9]. The performances of MH-OLEDs are simulated and analyzed from the numerical results after defining the energy level and mobilities of the mixed host layer. The energy levels and calculated mobilities of the mixed-host materials in the EML using the mobility of pure ETM and HTM are listed in Table 1.

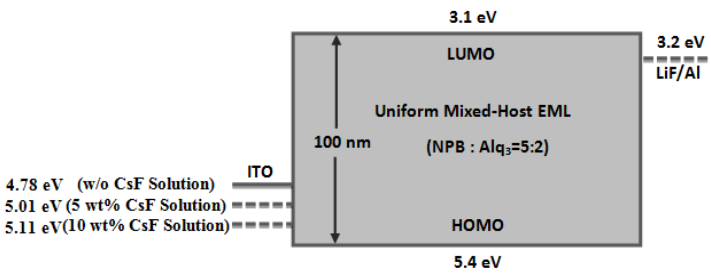


Fig. 1. MH-OLEDs based on W/ and W/O Modified ITO

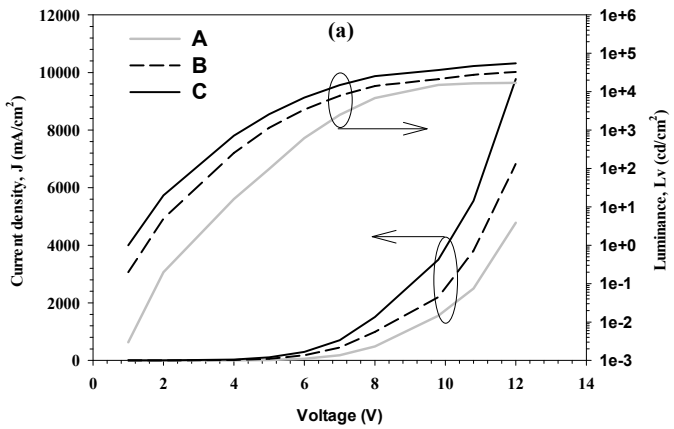
TABLE I. INPUT PARAMETERS FOR MH-OLEDS SIMULATION.

Parameters	NPB [12]	Alq3 [12]	Uniform MH-OLEDs		
			A	B	C
HUMO (eV)	-5.4	-5.8	-5.4	-5.4	-5.4
LUMO (eV)	-2.4	-3.1	-3.1	-3.1	-3.1
μ_{p_0} (cm^2/ Vs)	2.6×10^{-4}	1.4×10^{-6}	1.1×10^{-5}	1.9×10^{-5}	3.2×10^{-5}
γ_{p_0} (cm$^{1/2}$/ V$^{1/2}$)	1.3×10^{-3}	2×10^{-4}	1.4×10^{-3}	1.4×10^{-3}	1.4×10^{-3}
μ_{n_0} (cm^2/ Vs)	1.4×10^{-6}	2×10^{-8}	1.1×10^{-7}	1.6×10^{-7}	2.5×10^{-7}
γ_{n_0} (cm$^{1/2}$/ V$^{1/2}$)	2.8×10^{-4}	9.9×10^{-3}	1.4×10^{-3}	1.4×10^{-3}	1.4×10^{-3}
LiF-Al (eV)			- 3.2 [4]		

III. RESULTS AND DISCUSSIONS

A. J-L-V Characteristics and Efficiencies

Figure 2 shows the current density-voltage-luminance (J-V-L) characteristic curve of unmodified ITO based MH device A and two MH-OLEDs: B and C those are based on modified ITO. Because of ITO surface modification, the energy barrier for hole injection at ITO/EML interface is reduced from 0.62 eV for device A to 0.39 eV and 0.29 eV for device B and C respectively. Therefore, the hole injection rate increased enormously in case of device B and C over the reference device A. Hence, under a given operating voltage, the current densities of device B and C are higher than the device A. The luminescence of device B and C shows the same trend as that of current density with the given operating voltage, which indicates that high luminescence from a device can be obtained with high current density. The turn-on voltage (the voltage as measurable luminescence of 1 cd/m2) decreased from 2.7 V for reference device A to 1.4 V and 1.0 V for device B and C respectively.



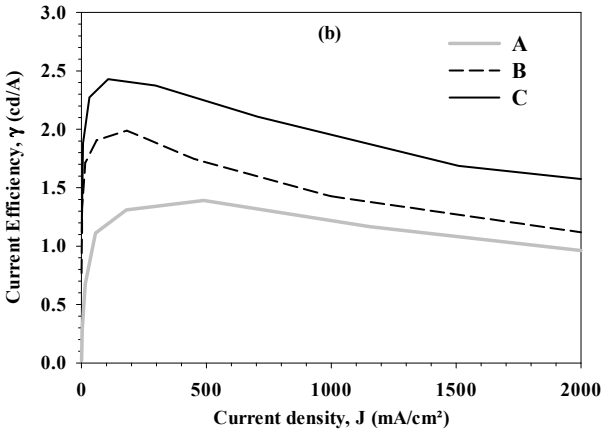


Fig. 2. Characteristics curves of MH-OLEDs in terms of (a) Current density-Luminance-Voltage and (b) Electroluminescence Efficiency vs Current density curves

Moreover, the EL efficiency-current density curve as shown in above Fig. 2(b), express that modified device B and C have the improved values of EL efficiency, γ of 1.98 cd/A and 2.43 cd/A respectively over the reference device A which has EL efficiency of 1.41 cd/A. These significant performance improvements of proposed MH-OLEDs based modified anode by surface modification can be visualized and explained form the simulated charge injection, carrier density and recombination profiles.

B. Effects of ITO surface modication on carrier distribution, recombination and balancing in the emissive layer

From the carrier concentration profile as shown in Fig. 3 (a), it is observed that due to reduction of hole injection barrier at ITO/EML, the hole density at ITO/EML interface for device B and C improved around 1 and 1.5 orders respectively. As a result, the total current density due to hole in Fig. 2 (a) is much higher in the EML in case of device B and C than the device A. In modified devices (B and C), the higher charge density (Fig. 3(a)) at both ITO/EML and EML/LiF-Al interfaces generate higher electric field with respect to device A which influence the carrier recombination rate. Hence, modified device gives the higher recombination rate in each electrode/EML interface which is shown in Fig. 3(b). Due to balance of charge density, the carriers recombination rate become almost stable in middle of the EML for all MH-OLEDs but modified devices keeps the peak magnitude of carrier recombination rate over reference device. However, this enormous improvement of hole injection in the proposed MH devices B and C enhanced the current balancing as calculated from the current balance factor (CBF) which is plotted with respect to operating voltage in Fig. 3(c). As a high value for CBF can be directly correlated to better OLED performance and stability, therefore the computed characteristic of CBFs indicates that the performance of all MH-OLEDs followed as device C> device B > device A.

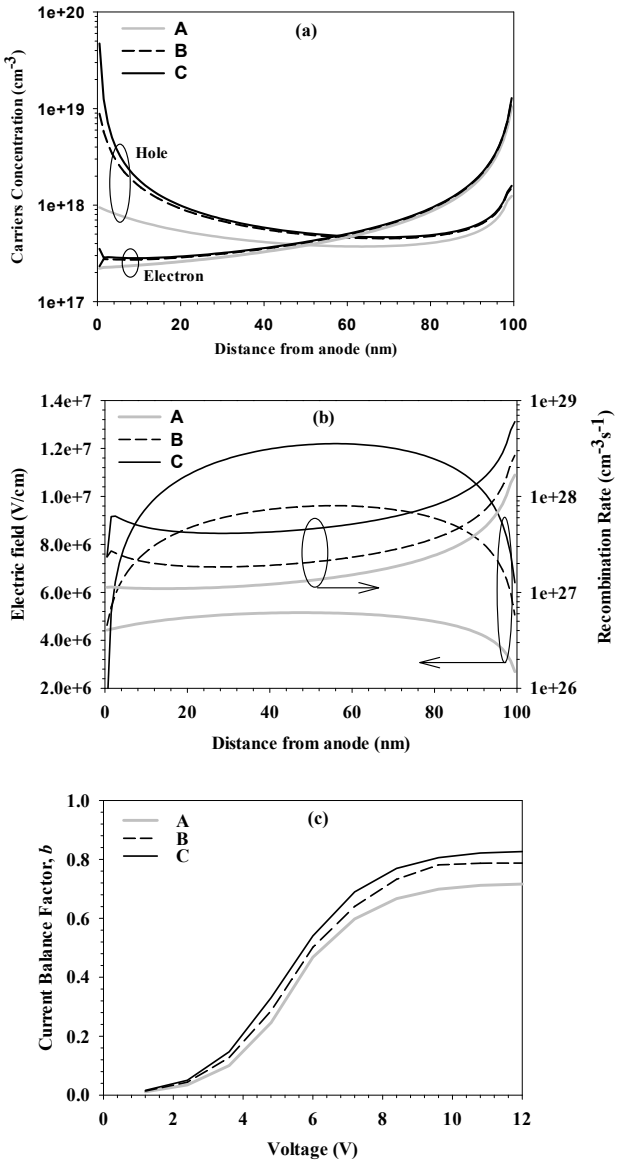


Fig. 3. (a) Carriers density profile at 10V, (b) Electric field and Carrier recombination zone at 10V and (c) Current balance factor at all operating voltage for untreated and treated ITO based MH-OLEDs.

C. Discussions

Due to the ITO surface modification technique, the energy barrier at anode/emissive layer interface reduced by 0.23 and 0.33 eV respectively from the original value (0.62 eV) which ultimately boost up the hole injection rate through that interfacial zone. Therefore, electroluminescence efficiencies of treated ITO based devices are improved by 1.41 and 1.73 times. About 1.1 times and 1.2 times improvements (with 5 wt% and 10 wt% CsF solution, correspondingly) in current balance factor over normal condition also confirmed the improvement of device external quantum efficiency which is related to device life time.

IV. Conclusions

The feasibility of the experimental process of proposed ITO surface modification technique based uniform MH-OLED performance has been analyzed from the 'SETFOS' based numerical simulation results. By ITO surface modification, the energy barrier for hole injection reduced by around 2.37 times (from 0.69 eV for device A to 0.29 eV for device C). Therefore, EL. efficiency is raised around 1.73 times (device C) which correspondingly improved the value of current balance factor is also improved from 0.71 to 0.83. Indeed, the quantitative results of modified ITO based uniform MH-OLEDs performances those are obtained from the numerical simulation results, have proved the performance improvement and longer operation life time of this proposed uniform MH-OLED structure as well as experimental feasibility.

References

[1] N. C. Erickson and R. J. Holmes, "Highly efficient, single-layer organic light-emitting devices based on a graded-composition emissive layer," *Appl. Phys. Lett.*, vol. 97, no. 8, p. 083308, 2010.

[2] C. W. Tang and S. a. VanSlyke, "Organic electroluminescent diodes," *Appl. Phys. Lett.*, vol. 51, no. 12, p. 913, 1987.

[3] D. Ma, C. S. Lee, S. T. Lee, and L. S. Hung, "Improved efficiency by a graded emissive region in organic light-emitting diodes," *Appl. Phys. Lett.*, vol. 80, no. 19, p. 3641, 2002.

[4] C.-H. Hsiao, Y.-H. Chen, T.-C. Lin, C.-C. Hsiao, and J.-H. Lee, "Recombination zone in mixed-host organic light-emitting devices," *Appl. Phys. Lett.*, vol. 89, no. 16, p. 163511, 2006.

[5] M. Q. Wang, and I. G. Hill, "Fluorinated alkyl phosphonic acid SAMs replace PEDOT:PSS in polymer semiconductor devices," *Organic Electonics*, vol. 13, no. 3, p. 498-505, 2012.

[6] H. Wang, K. P. Klubek, and C. W. Tang, "Current efficiency in organic light-emitting diodes with a hole-injection layer," *Appl. Phys. Lett.,* vol. 93, no. 9, p. 093306, 2008.

[7] X. A. Cao and Y. Q. Zhang, "Performance enhancement of organic light-emitting diodes by chlorine plasma treatment of indium tin oxide," *Appl. Phys. Lett.*, vol. 100, no. 18, p. 183304, May 2012.

[8] Baldo, M. A., Thompson, M. E. & Forrest, S. R., "High-efficiency fluorescent organic light-emitting devices using a phosphorescent sensitizer, " *Nature*, vol. 403, p. 750–753, 2000.

[9] H.-W. Lu, C.-W. Huang, P.-C. Kao, S.-Y. Chu, and Y.-D. Juang, "Effects of ITO Electrode Modification Using CsF Solution on Performance of Organic Light-Emitting Diodes," *ECS J. Solid State Sci. Technol.*, vol. 4, no. 3, p. R54–R59, Jan. 2015.

[10] E. Knapp and B. Ruhstaller, "Numerical analysis of steady-state and transient charge transport in organic semiconductor devices," *Opt. Quantum Electron.*, vol. 42, no. 11–13, p. 667–677, Feb. 2011

[11] Hui Wang, Kevin P. Klubek, C. W. Tang, "Current Efficiency in Organic Lightemitting Diodes with a Hole-injection Layer", *Appl. Phys. Lett.*, vol. 93, p. 093306, 2008.

[12] S.-W. Liu, J.-H. Lee, C.-C. Lee, C.-T. Chen, and J.-K. Wang, "Charge carrier mobility of mixed-layer organic light-emitting diodes," *Appl. Phys. Lett.*, vol. 91, no. 14, p. 142106, 2007.

A Highly Birefringent Slotted-Core THz Fiber

Sohel Rana[1], Md Saiful Islam[2], Jakeya Sultana[3], KM Samaun Reza[4], Md Ashfaq Uddin[5], Mohammad Faisal[6], Mohammad Rakibul Islam[7] and Raonaqul Islam[8]

[1,8] Electrical and Electronic Engineering , Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh.
[2-5,7]Electrical and Electronic Engineering, Islamic University of Technology, Gazipur, Bangladesh.
[6]Electrical and Electronic Engineering, Bangladesh University of Engineering & Technology, Dhaka, Bangladesh.
Corresponding Author: dsmsrana@gmail.com

***Abstract*— We present a slotted porous-core microstructure polymer fiber for polarization maintaining (PM) guidance of terahertz (THz) waves. The numerical analysis is performed using a full vector finite element method (FEM). Our proposed structure contains a compact hexagonal lattice in the cladding and rectangular slotted air-holes inside the core. Geometries inside the core have been designed to introduce asymmetry which results in a high modal birefringence of ~0.08. Moreover, majority of the mode power is propagated in air medium to exhibit a very low effective material loss of ~0.08 cm^{-1}. Some other modal characteristics of the fiber are also discussed rigorously, which include confinement loss, mode power fraction, dispersion properties etc.**

***Index Terms*— Far-infrared or terahertz; Photonic Crystal Fibers; Birefringence; Microstructured fibers; Polymers**

I. Introduction

IN recent times, there has been a growing interest towards the spectral range of THz waves (10^{11}-10^{13} Hz).This is due to the fact that it serves numerous scientific and technological progresses in significant fields such as pulsed imaging [1], sensing [2],time-domain spectroscopy [3],and biomedical engineering [4] etc. Devices like THz couplers are also gathering attentions [5].Unfortunately, due to strong dielectric and conductive losses in this frequency region, propagating THz waves through waveguides for long distance communications has turned into an enormous technological challenge. It is assumed that dry air exhibits nearly no material absorption loss in this band. Yet free space transmission of THz radiation is not beneficial due to the unavoidable losses introduced during coupling, transporting and managing THz beam [6]. Previously, Plastic [7], metallic ribbon [8] waveguides, dielectric tubes [9] and polymer PCFs [10], [11] were demonstrated for THz wave guidance. Most recently, researchers' attention is turning towards porous-core PCFs due to their low absorption loss and high modal power confinement in dry air [12-13].Highly birefringent porous fibers also demonstrate beneficial functions in the THz regime since they have potential applications in THz polarized filters and sensors. Basically, high birefringence can be achieved by either introducing asymmetry to the fiber-core [14-15] or breaking the symmetry of the fiber cladding [16]. Following these ways, some remarkable works have been reported to date. For example, Atakaramians *et al*. designed and fabricated a porous fiber [17], [18] with rectangular air-holes that showed a high birefringence of ~0.012 at 0.65 THz. Chen *et al.* proposed a squeezed lattice elliptical-hole THz fiber [19] with a birefringence of the order 10^{-2}. A plastic based PCF [20] showed high birefringence of ~0.021, but the reported propagation loss was very high. In 2013, Chen *et al.* offered a porous fiber with rotated elliptical air-holes [6]. It showed a high birefringence of ~0.045 and a very high effective absorption loss. Most recently, Islam *et al.* proposed a slotted core PCF [21] with circular lattice cladding that exhibited a high birefringence of ~0.075 with a low transmission loss. Nevertheless, all these works mentioned in the literature used either solid air-cladding [10, 17-19] or complex lattice holey cladding [21]. These factors cause high absorption loss, coupling with external environment and fabrication difficulties resulting in an overall degradation of the performance of the fiber.

In this paper, we report on a slotted-core PCF incorporating a hexagonal cladding. The proposed fiber exhibits simultaneously ultra-high birefringence and low effective absorption loss with a better confinement factor. Slotted fibers with air-cladding [17], [18] and circular lattice cladding were reported earlier [21]. But to the best of our knowledge, slotted-core fiber with a hexagonal cladding has never been discussed in THz researches before.

II. Design Methodology

Figure 1 shows the geometry of the proposed rectangular slotted fiber. The slotted core is surrounded by a regular hexagonal lattice structure comprising five rings of air-holes. The Proposed fiber consists of 9 rectangular slots in the core and 5 rings of air-holes in the cladding. We choose different lengths for the slots which are related to the core diameter (D_{core}). D_{core} indicates the diameter of the core from the centre of the core to the edge of each air hole of first air-holes ring shown in Fig. 1.The five slots in the middle have a length of $0.91D_{core}$, the two outermost slots have $0.35D_{core}$, and the rest of the slots have $0.9D_{core}$of lengths. Throughout our simulations, width of the slots has been varied while lengths of the slots were kept fixed. Variations of the width of the slots mainly affect the porosity. Porosity is the measure of the void (i.e. "empty") spaces in a material, and is a fraction of the volume of voids over the total volume. In our design, the ratio of area of the total number of slots to the area of the core defines the porosity. The center to center vertical distance between two adjacent slots is called the core pitch, indicated by Λ_c. It is somewhat possible to obtain a higher birefringence by further extending the length of the slots. However, considering the ease of fabrication, we always maintain a minimum distance between the edge of each slot and the edge of each air-hole of the first ring of cladding. The size of the air-holes in the cladding is kept the same. The symbol Λ denotes the distance between two adjacent air-holes on the same ring and that between two adjacent rings in the outer lattice.

We choose TOPAS cyclic-olefin copolymer (COC) as the background material for our proposed fiber for its excellent inherent characteristics like constant refractive index (n = 1.53) between 0.1–2 THz [12], very low bulk material absorption loss (0.2 cm^{-1} at f = 1 THz) [11] insensitivity to humidity [22], the fact that it is good for biosensing [23] and can be made from different grades for high glass transition temperature [24].

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

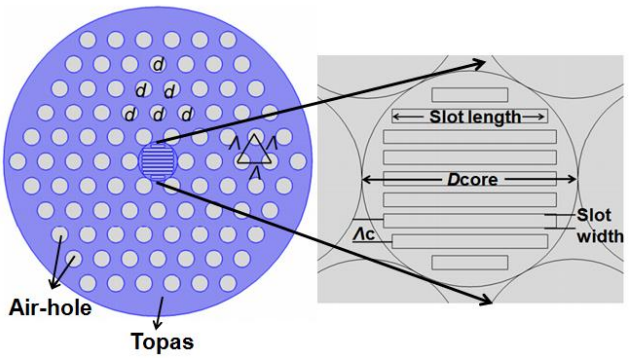


Fig. 1.Cross section of the proposed fiber and the enlarged view of its core.

III. Results and Discussions

To calculate the guiding properties of the proposed fiber, we used the finite element method (FEM) based commercially available *COMSOL* software. An artificial absorbing layer, known as the perfectly matched layer (PML) is applied outside our proposed structure to limit the computational domain and provide anti-reflectivity. In our design, the thickness of the PML region is 9% of the total radius of the fiber. Throughout our simulations, D_{core} = 350 µm and frequency f =1.1 THz are considered as the optimal parameters. Fig. 2 depicts the power flow distribution of the proposed fiber for the two orthogonal polarization components at different porosities. One can resolve from Fig. 2 that the mode field is well confined.

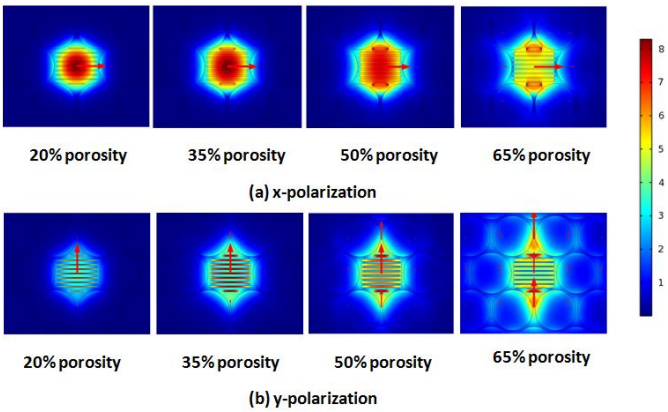


Fig. 2.Electric field distribution of the normalized power flow for the proposed fiber at different porosities.

At the beginning, we calculate the birefringence expressed as [6]

$$B = |n_x - n_y|, \qquad (1)$$

Where n_x and n_y are the refractive indices of two orthogonal components of the polarization maintaining wave. Birefringence as a function of D_{core} at different porosities is shown in Fig. 3. It is observed from Fig. 3 that the birefringence is reduced when porosity is increased. The reason is that at a higher porosity, the mode field is delocalized from the core and spread into the cladding. As a consequence, less power is controlled by the slots and birefringence is reduced. We also observed that the birefringence remains nearly constant with respect to D_{core} for porosities above 35%. Please note that the proposed fiber exhibits an ultra-high birefringence of ~0.08 for D_{core} = 350 µm, 35% porosity at 1.1 THz, which is obviously better than the previously reported works [6, 17-21].Next, the behaviors of the effective refractive index and the birefringence with respect to frequency are shown in Fig. 4.One would note an interesting observation that the birefringence increases with frequency for 0.7<f< 1.2 THz and then decreases for f> 1.2 THz. To understand this mechanism, we focus on the mode index property of the slotted fiber. First, in case of x-polarization mode, each of the wave fronts encounters either a single slot or solid material along the orthogonal polarization direction. So the effective mode index for an x-polarized wave remains high. On the other hand, in case of y-polarization mode, each of the waves faces every single slot while passing through the fiber. Therefore, the effective refractive index is very much lowered for y-polarized waves (see Fig. 4). Generally, mode index of a fiber increases with frequency for non-absorbing materials whereas for absorbing materials, it reduces when frequency is increased. Since the absorption co-efficient of the proposed fiber is negligible below 1.2 THz and starts to increase at f> 1.2 THz due to the increased sensitivity of scattering behavior, the increment rate of the mode index with respect to frequency is reduced for both the polarization modes at f> 1.2 THz resulting in lower birefringence. Nevertheless, it can be observed in Fig. 4 that the proposed fiber shows an ultra-high birefringence (~0.08) within 1-1.35 THz which is wider than the previously reported works [6, 17-21].

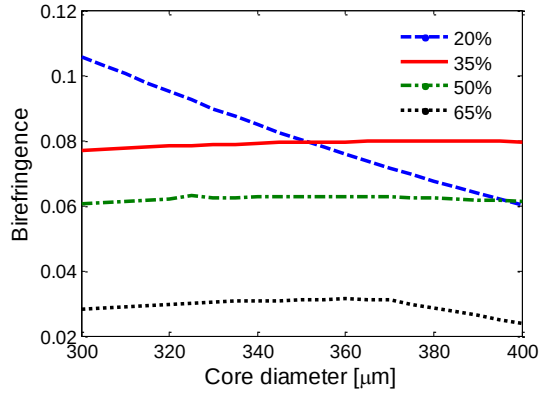


Fig. 3.Birefringence versus D_{core} for different porosities at 1.1 THz.

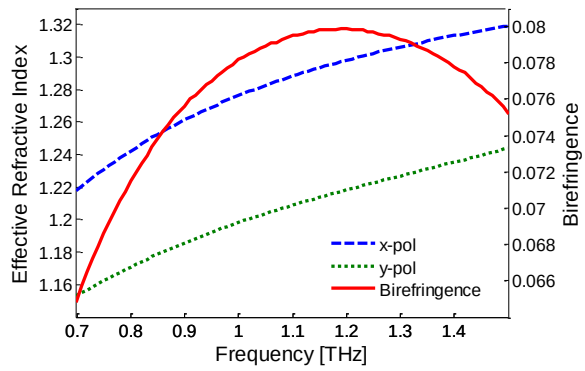


Fig. 4. Effective mode index (left vertical axis) and the birefringence (right vertical axis) versus frequency at 35% porosity, D_{core} = 350 µm.

We also investigated the material absorption loss. The effective material loss (EML) is quantified by [13]

$$\alpha_{eff} = \sqrt{\frac{\varepsilon_0}{\mu_0}}\left(\frac{\int_{mat} n_{mat}|E|^2\alpha_{mat}\,dA}{|\int_{all} S_z\,dA|}\right), \qquad (2)$$

Where ε_o and μ_o are the permittivity and permeability of vacuum, $n_{mat}$is the refractive index of the material used, E is the modal electric field, $\alpha_{mat}$is the bulk material absorption loss and $S_z$is the z-component of the Poynting vector, $S_z = \frac{1}{2}(\vec{E}\times\vec{H^*}).\hat{z}$. Integration in the numerator of the Eq. (2) is performed only over the solid material region [13]. Figure 5 reveals the EML property of the proposed waveguide with respect to the fiber geometry. It is observed from Fig. 5 that a higher porosity assists in reducing the EML because higher porosity ensures less material inside the core region. We found that the fiber exhibits EML of ~0.08 cm^{-1} for y-polarization mode and

0.135 cm^{-1} for *x*-polarization mode at 35% porosity, D_{core} = 350 µm and 1 THz, which is very much comparable to the previous works on related field [6, 17-21].The *y*-polarized waves show lower loss than the *x*-polarized waves because the index contrast for *y*-polarization mode is much lower than the *x*-polarization mode as described earlier (see Fig. 4). That means the *y*-polarized wave encounters less solid material than the *x*-polarized wave while propagating through the fiber which consequently reduces the EML. Next, Fig. 6 shows how the EML responds to the variation of frequency for the proposed fiber. It is observed that the EML increases with frequency and the increment rate is independent of the polarization mode.

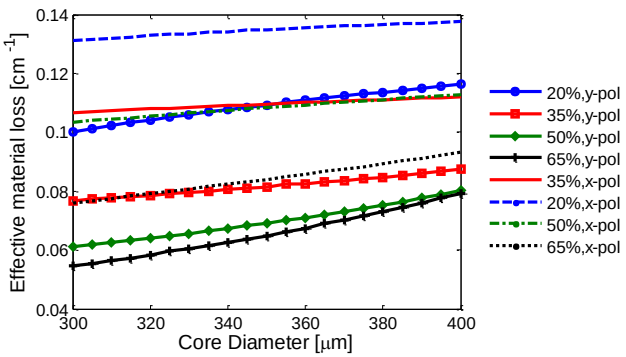


Fig. 5.Effective material loss as a function of core-diameter for different porosities at 1.1 THz.

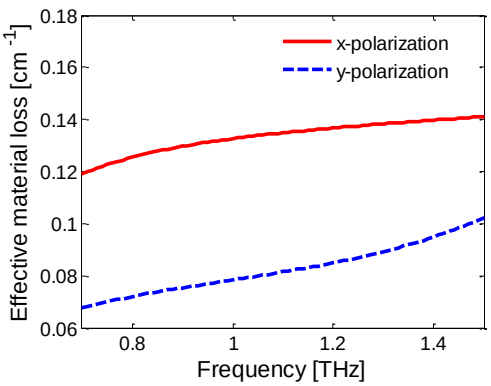


Fig. 6. Effective material loss as a function of frequency at D_{core} = 350 µm and 35% porosity.

The confinement loss occurs due to the finite extent of the periodic cladding and is obtained from the imaginary part of the complex refractive index, n_{eff}. Confinement loss is given by [21],

$$\alpha_{CL} = 8.686 k_0 \mathrm{Im}(n_{eff}) \qquad (3)$$

Where $k_0 = \frac{2\pi}{\lambda}$ and Im(n_{eff}) is the imaginary part of the refractive index of the guided mode. Figure 7 shows the confinement loss as a function of the frequency. One can deduce from Fig. 7 that for both the polarization modes, confinement loss is very low at f = 1.1 THz and D_{core} = 350 µm. Thus it can be neglected as compared to the obtained effective material loss.

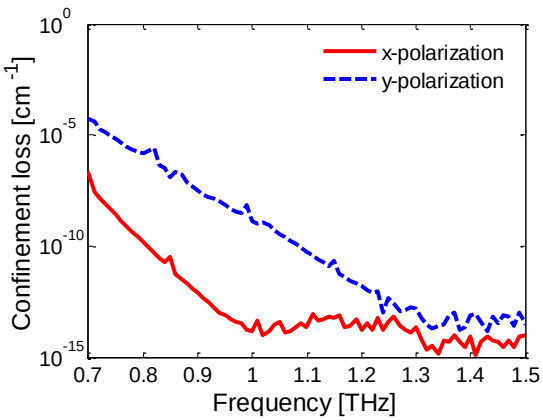


Fig. 7.Confinement loss versus frequency for D_{core} = 350µm and 35% porosity.

The power flow distribution in different regions of a fiber is called the power fraction, which is estimated by [6]

$$\eta' = \frac{\int_X S_z dA}{\int_{all} S_z dA}, \qquad (4)$$

Where *X* represents the integral region which one interests such as core region, cladding region, solid material region etc. In our proposed design, all of the criteria are thoroughly considered. Behaviors of mode power fraction is shown in Fig. 8, where it is observed that a moderate amount, about 41% and 31% of the total power for *y* and *x* polarization modes respectively are transmitted through the slotted air-holes at D_{core} = 350 µm and 35% porosity. It is possible to increase power fraction in air-slots by increasing the porosity. However, that would delocalize most of the mode power in the cladding region to incur additional leakage loss.

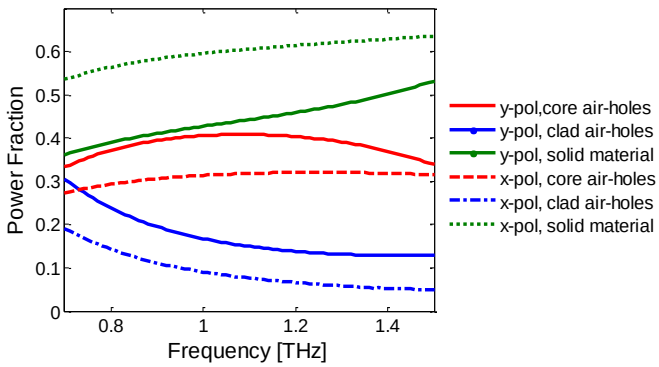


Fig. 8. Power fraction versus frequency at D_{core} = 350 µm and 35% porosity.

Also, we analyze the dispersion of the proposed fiber. Since refractive index of TOPAS is constant in between 0.1-2 THz [12], our proposed fiber shows only the waveguide dispersion. The dispersion parameter is calculated by [25]

$$\beta_2 = \frac{2}{c}\frac{dn_{eff}}{d\omega} + \frac{\omega}{c}\frac{d^2 n_{eff}}{d\omega^2}, \qquad (5)$$

where angular center frequency $\omega = 2\pi f$ and c is the velocity of electromagnetic waves. We have shown the dispersion characteristics in Fig. 9, where it is observed that a positively flat dispersion is obtained within the frequency range of 1.16-1.5 THz at optimal parameters. We found that the variation of the dispersion valueis 0.85±0.05 ps/THz/cm for *y*-polarization mode and 0.85±0.35 ps/THz/cm for x-polarization mode. Flat dispersion is essential in THz wave guiding because it allows transmitting several optical signals at the same time with nearly equal pulse spreading.

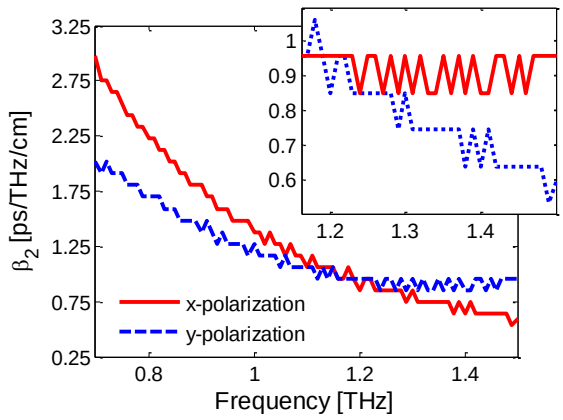


Fig. 9.Dispersion parameter versus frequency at D_{core} = 350 µm and 35% porosity.

Finally, we study the practical realization of the proposed PCF. It is easy to rely on the ongoing fabrication technology due to its excellent progresses in PCF fabrication. However, because of the rectangular slots, the fabrication of the proposed porous fiber is not possible by stacking or drilling preforms. It has been shown earlier that non-circular air holes in microstructure fibers made of polymer and soft glasses are feasible to be fabricated by a well-established extrusion technique [26].In 2009, an asymmetric porous fiber with rectangular slots was experimentally realized by Atakaramians *et al.* using the extrusion method [17]. They utilized a two-step process to fabricate the polymer fibers. We expect that it is possible to fabricate the proposed fiber using the verified methods described in Refs. [17, 26].

IV. Conclusion

A slotted-core PCF is designed and characterized for polarization maintaining THz wave guidance. The key aspects of the structure are its ultra-high birefringence (~0.08), very low effective material loss (~0.08cm^{-1}) and the regular hexagonal cladding which makes the fiber more practical. Furthermore, sufficient amount of mode power is transmitted through the core which results in very low leakage loss. In near future, we anticipate of this robust and compact structure to open a new window in the field of polarization maintaining THz technology.

References

[1] Q. Chen *et al.*,"Nearfield terahertz imaging with a dynamic aperture," *Opt. Lett.*, vol. 25, no. 11, pp. 1122-1124, 2000.

[2] R. H. Jacobsen *et al.*, "Chemical recognition of gases and gas mixtures with terahertz waves," *Opt. Lett.*, vol. 21, no. 24, pp. 2011- 2013, 1996.

[3] J. Q. Zhang *et al.*, "Waveguide terahertz time-domain spectroscopy of nanometer water layers," *Opt. Lett.*, vol. 29, no. 14, pp. 1617-1619, 2004.

[4] R. M. Woodward *et al.*, "Terahertz pulsed imaging of skin cancer in the time and frequency domain," *J. Biol. Phys.*, vol. 29, no.2, pp. 257–261, 2003.

[5] H. Bao *et al.*, "Design and optimization of mechanically downdoped terahertz fiber directional couplers," *Opt. Express*, vol. 22, no. 8, pp. 9486-9497, 2014.

[6] N. Chen *et al.*, "High-birefringence, low-loss porous fiber for single-mode terahertz- wave guidance," *Appl. Opt.*, vol. 52, no. 21, pp. 5297-5302, 2013.

[7] R. Mendis *et al.*, "Plastic Ribbon THz Waveguides," *J. Appl. Phys.*, vol. 88, pp.4449-4451, 2000.

[8] K.Wang *et al.*, "Metal wires for terahertz wave guiding," *Nature*, vol. 432, pp. 376–379, 2004.

[9] H. Bao *et al.*, "Dielectric tube waveguides with absorptive cladding for broadband, low-dispersion and low loss THz guiding," *Sci. Rep.*, vol. 5, pp. 1-9, 2015.

[10] M. Goto *et al.*,"Teflon photonic crystal fiber as terahertz waveguide," *Jpn. J. Appl. Phys, Part 2.*, vol. 43, pp. 317-319, 2004.

[11] K. Nielsen *et al.*, "Bendable, low-loss topas fibers for the terahertz frequency range," *Opt. Express*, vol. 17, no. 10, pp. 8592–8601, 2009.

[12] H. Bao *et al.*, "Fabrication and characterization of porous-core honeycomb bandgap THz fibers," *Opt. Express*, vol. 20, no. 28, pp. 29507-29517, 2012.

[13] K. Nielsen *et al.*, "Porous-core honeycomb bandgap THz fiber," *Opt. Lett.*, vol. 36, no.5, pp. 666 –668, 2011.

[14] S. E. Kim *et al.*, "Elliptical defected core photonic crystal fiber with high birefringence and negative flattened dispersion,"*Opt. Express*, vol.20, no. 2, pp.1385–1391, 2012.

[15] D. R. Chen *et al.*, "Highly birefringent elliptical-hole photonic crystal fibers with double defects," *J. Lightwave Technol.*, vol. 25, no. 9, pp.2700–2705, 2007.

[16] Y. Yue *et al.*, "Highly birefringent elliptical-hole photonic crystal fiber with squeezed hexagonal lattice," *Opt. Lett.*, vol. 32, no. 5, pp.469–471, 2007.

[17] S. Atakaramians *et al.*, "THz porous fibers: design, fabrication and experimental characterization," *Opt. Express*, vol. 17, no.16, pp.14053–14062, 2009.

[18] S. Atakaramians *et al.*, "Low loss,low dispersion and highly birefringent terahertz porous fibers," *Opt. Commun.*, vol. 282, pp.36–38, 2009.

[19] H. B. Chen *et al.*, "Squeezed lattice elliptical hole terahertz fiber with high birefringence," Appl. *Opt.*, vol. 48, no.20, pp. 3943-3947, 2009.

[20] M. Cho *et al.*, "Highly birefringent terahertz polarization maintaining plastic photonic crystal fibers," *Opt. Express*, vol. 16, no.1, pp.7-12, 2008.

[21] R. Islam, *et al.*, "Extremely High-birefringent Asymmetric Slotted-core Photonic Crystal Fiber in THz Regime," *IEEE Photonics Technol. Lett.*, vol. 27, no.21, pp. 2222-2225, 2015.

[22] W. Yuan *et al.*, "Humidityinsensitive TOPAS polymer fiber Bragggrating sensor," *Opt. Express*, vol. 19, no.20, pp.19731-19739, 2011.

[23] G. Emiliyanov *et al.*, "Selective serial multi-antibody biosensing withTOPAS microstructured polymer optical fibers," *Sensors*, vol. 13, no. 3, pp.3242-3251, 2013.

[24] C. Markos *et al.*, "High-Tg TOPAS microstructured polymer optical fiber for fiber Bragg grating strain sensing at 110 degrees," *Opt. Express*, vol. 21, no. 4, pp.4758-4765, 2013.

[25] J. Liang *et al.*, "Broadband, low-loss, dispersion flattened porous-core photonic bandgap fiber for terahertz (THz)-wave propagation," *Opt. Commun.*, vol. 295, pp. 257–261, 2013.

[26] H. Ebendorff-Heidepriem *et al.*, "Extrusion of complex preforms for microstructured optical fibers,"*Opt. Express*, vol. 15, no. 23, pp. 15086-15092, 2007.

Security in Multiple Keyhole MIMO Channels

R. Sultana, M. Z. I. Sarkar and M. S. Hossain
Department of Electrical and Electronic Engineering
Rajshahi University of Engineering and Technology
Rajshahi-6204, Bangladesh.
Email: dola.ruet@gmail.com

***Abstract*—A new closed form expression for the ergodic secrecy capacity with multi-keyhole Multiple-Input Multiple-Output (MIMO) channel has been derived in this paper. Assuming independent and identically distributed (i.i.d.) Rayleigh flat-fading between transmitting and receiving antennas and equal power allocation to each of the transmit antennas, the ergodic secrecy capacity is expressed as a finite sum of weighted Meiger G-functions for ease of calculation. The effect of the number of keyhole on the ergodic secrecy capacity as well as on the receiving antenna is also analyzed in this paper.**

***Index Terms*—Secrecy ergodic capacity, MIMO channel, multiple keyhole, Meiger G-function**

I. Introduction

Due to use of multiple antenna at the transmitting and receiving terminals, MIMO wireless communication system has received a great deal of attention in comparison with the single antenna system. On the other hand, keyhole effect can occur in the case when there exists a large separation between transmitters and receivers. The multi-keyhole channel model can establish a relationship between the keyhole and the Rayleigh fading MIMO channels [1].

Recently, a good number of papers appeared in the literature on the ergodic capacity calculation in multi-keyhole wireless MIMO channels. In [2] authors derived closed-form analytical expressions for the ergodic capacity and outage probability for the Nakagami m fading keyhole MIMO channels. The analytical results were verified via Monte-Carlo simulation. The closed-form expression for asymptotic pairwise error probability of space-time codes was derived in [3] to show the effect of diversity gain on the asymptotic pairwise error probability. The analytical closed-form expressions of the eigen-mode optimal power, rate adaption policy and sub-optimal eigen-mode channel for the capacity were derived in [4] considering double-correlated rank-deficient keyhole to investigate the effect of keyhole on the MIMO channels with instantaneous channel state information (CSI) at the transmitter and receiver. In [5], authors derived exact closed-form expressions for the ergodic mutual information for spatially uncorrelated multi-keyhole MIMO channels using non-asymptotic random matrix theory and observed that the capacity of multi-keyhole channels is worse than that of rich-scattering MIMO channels. In [6], the closed-form expressions for the Gallager random coding and expurgated error exponents, ergodic capacity and cut off rate had been derived based on Meiger G-function to investigate the effect of keyholes on the error exponents and the cut off rate. In order to show the effect of number of antennas and the number of keyholes on the decoding error probability, authors derived the closed-form expressions for the ergodic capacity in [7].

In this paper, we derive the closed-form analytical expression for the ergodic secrecy capacity of multi-keyhole MIMO channel in the presence of multiple eavesdropper and verify its validity via Monte-Carlo simulation.

The rest of the paper is organized as follows. The system model is discussed in Section II. In Section III, we derive the closed-form expression for the ergodic secrecy capacity of multi-keyhole MIMO channel. The numerical results are provided in Section IV. Finally, Section V describes the concluding remarks of this work.

II. System Model

We consider a confidential communication scenario through multi-keyhole MIMO channels as shown in Fig.1 in which a transmitter sends a common stream of information to a receiver in the presence of an eavesdropper. The transmitter, receiver and eavesdropper are equipped with n_T , n_R and n_E antennas, respectively. The channel coefficient between transmitter and receiver is $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$, and that between transmitter and eavesdropper is $\mathbf{G} \in \mathbb{C}^{n_E \times n_T}$. The total transmit power

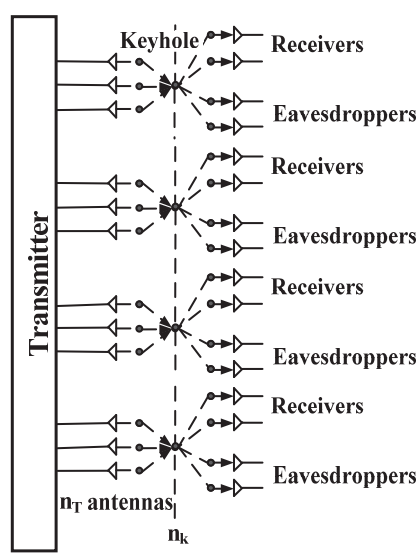


Fig. 1. System model.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

is denoted by P_T. Let $\mathbf{x} \sim \widetilde{\mathcal{N}}_{n_T}\left(0, \frac{P_T}{n_T}\mathbf{I}_{n_T}\right)$ denotes the transmitted signal. The received signals at the receiver and eavesdropper are respectively given by

$$\mathbf{y}_r = \mathbf{H}\mathbf{x} + \mathbf{z}_r, \tag{1}$$
$$\mathbf{y}_e = \mathbf{G}\mathbf{x} + \mathbf{z}_e, \tag{2}$$

where $\mathbf{z}_r \sim \mathcal{CN}\left(0, N_r\mathbf{I}_{n_R}\right)$ and $\mathbf{z}_e \sim \mathcal{CN}\left(0, N_e\mathbf{I}_{n_E}\right)$ represents the Gaussian noises imposed on the receiver and eavesdropper, respectively.

III. Ergodic Secrecy capacity of Multi-keyhole MIMO channel

For a MIMO channel in which the transmitter and receiver are equipped with n_T and n_R antennas, respectively, if the channel matrix is denoted by $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$, then the capacity with only receiver CSI can be expressed as

$$C = \log_2 \det\left[(\mathbf{I}_{n_R} + \frac{\gamma}{n_T}\mathbf{H}\mathbf{H}^\dagger)\right], \tag{3}$$

where γ denotes the received SNR. For a multi-keyhole MIMO channel with n_k keyholes, where the transmitter and receiver are equipped with n_T and n_R antennas, respectively, the channel matrix is modeled as

$$\mathbf{H} = \sum_{k=1}^{n_k} a_k \mathbf{h}_{r,k}\mathbf{h}_{t,k}^\dagger = \mathbf{H}_r\mathbf{A}\mathbf{H}_t^\dagger,$$

where $\mathbf{H}_r = [\mathbf{h}_{r,1}, \ldots, \mathbf{h}_{r,n_k}]$, $\mathbf{H}_t = [\mathbf{h}_{t,1}, \ldots, \mathbf{h}_{t,n_k}]$, $\mathbf{A} = \text{diag}(a_1, \ldots, a_{n_k})$. Here, a_k is the complex gain for the kth keyhole. $\mathbf{H}_r \sim \mathcal{CN}_{n_R,n_k}(\mathbf{0}_{n_R\times n_k}, \mathbf{I}\bigotimes\mathbf{I})$ and $\mathbf{H}_t \sim \mathcal{CN}_{n_T,n_k}(\mathbf{0}_{n_T\times n_k}, \mathbf{I}\bigotimes\mathbf{I})$ are mutually independent matrices. Defining $\mathbf{D} \triangleq \mathbf{A}\mathbf{A}^\dagger$, we have trace$(\mathbf{D}) = 1$. It is assumed that the channel is normalized i.e. $\mathbb{E}[\text{trace}\left(\mathbf{H}\mathbf{H}^\dagger\right)] = n_R n_T$, and therefore the capacity at the receiver for multi-keyhole MIMO channel is given by

$$C_r = \log_2 \det\left[\mathbf{I}_{n_R} + \frac{\gamma_1}{n_T}\mathbf{H}_t^\dagger\mathbf{A}^\dagger\mathbf{H}_r^\dagger\mathbf{H}_r\mathbf{A}\mathbf{H}_t\right], \tag{4}$$

where $\gamma_1 = \frac{P_1}{N_r}$ denotes SNR at the receiver. Similarly, capacity at the eavesdropper with multi-keyhole MIMO channel is given by

$$C_e = \log_2 \det\left[\mathbf{I}_{n_E} + \frac{\gamma_2}{n_T}\mathbf{G}_t^\dagger\mathbf{B}^\dagger\mathbf{G}_r^\dagger\mathbf{G}_r\mathbf{B}\mathbf{G}_t\right], \tag{5}$$

where $\gamma_2 = \frac{P_1}{N_e}$ denotes SNR at the eavesdropper, $\mathbf{G}_r = [\mathbf{g}_{r,1}, \ldots, \mathbf{g}_{r,n_k}]$, $\mathbf{G}_t = [\mathbf{g}_{t,1}, \ldots, \mathbf{g}_{t,n_k}]$, $\mathbf{B} = \text{diag}(b_1, \ldots, b_{n_k})$. Here, b_k is the complex gain for the kth keyhole. $\mathbf{G}_r \sim \mathcal{CN}_{n_E,n_k}(\mathbf{0}_{n_E\times n_k}, \mathbf{I}\bigotimes\mathbf{I})$ and $\mathbf{G}_t \sim \mathcal{CN}_{n_T,n_k}(\mathbf{0}_{n_T\times n_k}, \mathbf{I}\bigotimes\mathbf{I})$ are mutually independent matrices. Defining $\mathbf{F} \triangleq \mathbf{B}\mathbf{B}^\dagger$, we have trace$(\mathbf{F}) = 1$. It is assumed that the channel is normalized i.e. $\mathbb{E}[\text{trace}(\mathbf{G}\mathbf{G}^\dagger)] = n_E n_T$, and therefore the secrecy capacity at the receiver is given by

$$C_s = \log_2 \frac{\det\left(\mathbf{I}_{n_R} + \frac{\gamma_1}{n_T}\mathbf{H}_t^\dagger\mathbf{A}^\dagger\mathbf{H}_r^\dagger\mathbf{H}_r\mathbf{A}\mathbf{H}_t\right)}{\det\left(\mathbf{I}_{n_E} + \frac{\gamma_2}{n_T}\mathbf{G}_t^\dagger\mathbf{B}^\dagger\mathbf{G}_r^\dagger\mathbf{G}_r\mathbf{B}\mathbf{G}_t\right)}. \tag{6}$$

The ergodic secrecy capacity at the receiver for the multi-keyhole MIMO channel is given by

$$\begin{aligned}\langle C_s\rangle &= \mathbb{E}\left\{\log_2 \frac{\det\left[\left(\mathbf{I}_{n_R} + \frac{\gamma_1}{n_T}\mathbf{H}_t^\dagger\mathbf{A}^\dagger\mathbf{H}_r^\dagger\mathbf{H}_r\mathbf{A}\mathbf{H}_t\right)\right]}{\det\left[\left(\mathbf{I}_{n_E} + \frac{\gamma_2}{n_T}\mathbf{G}_t^\dagger\mathbf{B}^\dagger\mathbf{G}_r^\dagger\mathbf{G}_r\mathbf{B}\mathbf{G}_t\right)\right]}\right\} \\ &= p_1\mathbb{E}\left\{\log_2\left(1 + \frac{\gamma_1}{n_T}\lambda_1\right)\right\} - p_2\mathbb{E}\left\{\log_2\left(1 + \frac{\gamma_2}{n_T}\lambda_2\right)\right\} \\ &= p_1\mathcal{I}_1 - p_2\mathcal{I}_2,\end{aligned} \tag{7}$$

where

$$\begin{aligned}\mathcal{I}_1 &= \mathbb{E}\left[\log_2\left(1 + \frac{\gamma_1}{n_T}\lambda_1\right)\right] \\ &= \int_0^\infty \log_2\left(1 + \frac{1}{n_T}\lambda_1\right) f_{\lambda_1}(\lambda_1)\, d\lambda_1\end{aligned} \tag{8}$$

and

$$\begin{aligned}\mathcal{I}_2 &= \mathbb{E}\left[\log_2\left(1 + \frac{\gamma_2}{n_T}\lambda_2\right)\right] \\ &= \int_0^\infty \log_2\left(1 + \frac{1}{n_T}\lambda_2\right) f_{\lambda_2}(\lambda_2) d\lambda_2,\end{aligned} \tag{9}$$

where λ_1 and λ_2 denote the unordered eigenvalues of random matrices $\mathbf{H}_t^\dagger\mathbf{A}^\dagger\mathbf{H}_r^\dagger\mathbf{H}_r\mathbf{A}\mathbf{H}_t$ and $\mathbf{G}_t^\dagger\mathbf{B}^\dagger\mathbf{G}_r^\dagger\mathbf{G}_r\mathbf{B}\mathbf{G}_t$, respectively. $f_{\lambda_1}(\lambda_1)$ and $f_{\lambda_2}(\lambda_2)$ denote the marginal PDF of eigenvalues λ_1 and λ_2, respectively. $p_1 = \min(n_1, n_k)$, $n_1 = \max(n_T, n_R)$, $p_2 = \min(n_2, n_k)$ and $n_2 = \max(n_T, n_E)$. From [8], the marginal PDF of eigenvalue λ_1 is given by

$$\begin{aligned}f_{\lambda_1}(\lambda_1) &= \frac{1}{p_1\prod_{i<j}^{n_k}(d_i - d_j)}\sum_{i=1}^{n_k}\sum_{j=n_k-p_1+1}^{n_k}\mathcal{D}_{i,j} \\ &\times \frac{2d_i^{n_k-1-\frac{m_1+n_1}{2}}\lambda_1^{\frac{m_1+n_1}{2}-n_k+j-1}K_{n_1-m_1}\left(2\sqrt{\frac{\lambda_1}{d_i}}\right)}{\Gamma(n_1-n_k+j)\Gamma(m_1-n_k+j)} \\ &= \frac{1}{p_1\prod_{i<j}^{n_k}(d_i - d_j)}\sum_{i=1}^{n_k}\sum_{j=n_k-p_1+1}^{n_k}\lambda_1^{\frac{m_1+n_1}{2}-n_k+j-1} \\ &\times K_{n_1-m_1}\left(2\sqrt{\frac{\lambda_1}{d_i}}\right)\psi,\end{aligned} \tag{10}$$

where $m_1 = \min(n_T, n_R)$, $\psi = \frac{\mathcal{D}_{i,j}2d_i^{n_k-1-\frac{m_1+n_1}{2}}}{\Gamma(n_1-n_k+j)\Gamma(m_1-n_k+j)}$ and $d_1 \leq d_2 \leq \ldots \leq d_{n_k}$ denote the non zero eigenvalues of $\mathbf{D} \triangleq \mathbf{A}\mathbf{A}^\dagger$ and $\mathcal{D}_{i,j}$ is the (i,j)th cofactor of the matrix Ψ whose (l,k)th entry equals $[\Psi]_{l,k} = d_l^{k-1}$ for $1 \leq l,\ k \leq n_k$. Similarly, from [8] the marginal PDF of eigenvalue λ_2 is given

by

$$f_{\lambda_2}(\lambda_2) = \frac{1}{p_2 \prod_{i<j}^{n_k}(f_i - f_j)} \sum_{i=1}^{n_k} \sum_{j=n_k-p_2+1}^{n_k} \mathcal{F}_{i,j} \times \frac{2 d_i^{n_k-1-\frac{m_2+n_2}{2}} \lambda_1^{\frac{m_1+n_1}{2}-n_k+j-1} K_{n_2-m_2}\left(2\sqrt{\frac{\lambda_2}{f_i}}\right)}{\Gamma(n_2-n_k+j)\Gamma(m_2-n_k+j)}$$
$$= \frac{1}{p_2 \prod_{i<j}^{n_k}(f_i - f_j)} \sum_{i=1}^{n_k} \sum_{j=n_k-p_2+1}^{n_k} \lambda_1^{\frac{m_2+n_2}{2}-n_k+j-1} \times K_{n_2-m_2}\left(2\sqrt{\frac{\lambda_1}{d_i}}\right)\varphi, \quad (11)$$

where $m_2 = \min(n_T, n_E)$, $\varphi = \frac{\mathcal{F}_{i,j} 2 f_i^{n_k-1-\frac{m_2+n_2}{2}}}{\Gamma(n_2-n_k+j)\Gamma(m_2-n_k+j)}$. Here, $f_1 \leq f_2 \leq \ldots \leq f_{n_k}$ denote the non zero eigenvalues of $\mathbf{F} \triangleq \mathbf{B}\mathbf{B}^\dagger$ and $\mathcal{F}_{i,j}$ is the (i,j)th cofactor of the matrix ξ whose (l,k)th entry equals $[\xi]_{l,k} = f_l^{k-1}$ for $1 \leq l$, $k \leq n_k$. Substituting the value of $f_{\lambda_1}(\lambda_1)$ in equation (8) and performing integration, $\mathcal{I}_1$ can be written as

$$\mathcal{I}_1 = \frac{1}{p_1 \prod_{i<j}^{n_k}(d_i - d_j)} \sum_{i=1}^{n_k} \sum_{j=n_k-p_1+1}^{n_k} \psi \times \int_0^\infty \log_2\left(1 + \frac{1}{n_T}\lambda_1\right) \lambda_1^{\frac{m_1+n_1}{2}-n_k+j-1} \times K_{n_1-m_1}\left(2\sqrt{\frac{\lambda_1}{d_i}}\right) d\lambda_1. \quad (12)$$

Using the following identity of [9, eq. (8.4.6,5)]

$$\ln(1+x) = \frac{1}{\log_e 2} G_{2,2}^{1,2}\left[x|_{1,0}^{1,1}\right] \quad (13)$$

we have

$$\mathcal{I}_1 = \frac{1}{p_1 \prod_{i<j}^{n_k}(d_i - d_j)} \sum_{i=1}^{n_k} \sum_{j=n_k-p_1+1}^{n_k} \psi \times \int_0^\infty \frac{1}{\log_e 2} G_{2,2}^{1,2}\left[\frac{1}{n_T}\lambda_1|_{1,0}^{1,1}\right] \lambda_1^{\frac{m_1+n_1}{2}-n_k+j+1} \times K_{n_1-m_1}\left(2\sqrt{\frac{\lambda_1}{d_i}}\right) d\lambda_1.$$

Now using the following identity of [10, eq. (7.821.3)]

$$\int_0^\infty x^{-\rho} K_v(2\sqrt{x}) G_{pq}^{mn}\left(ax|_{b_1,\ldots,b_q}^{a_1,\ldots,a_p}\right) dx = \frac{1}{2} G_{p+2,q}^{m,n+2}\left(\infty|_{b_1,\ldots,b_q}^{\rho-\frac{1}{2}v,\rho+\frac{1}{2}v}, a_1, \ldots, a_p\right) \quad (14)$$

the closed-form expression for the $\mathcal{I}_1$ can be derived as

$$\mathcal{I}_1 = \frac{1}{\log_e 2 p_1 \prod_{i<j}^{n_k}(d_i - d_j)} \sum_{i=1}^{n_k} \sum_{j=n_k-p_1+1}^{n_k} \psi \times G_{4,2}^{1,4}\left[\frac{\lambda_1}{n_T} d_i|_{1,0}^{n_k-j+1-n_1, n_k-j+1-m_1,1,1}\right]. \quad (15)$$

Substituting the value of $f_{\lambda_2}(\lambda_2)$ in equation (9) and performing integration, $\mathcal{I}_2$ can be written as

$$\mathcal{I}_2 = \frac{1}{p_2 \prod_{i<j}^{n_k}(f_i - f_j)} \sum_{i=1}^{n_k} \sum_{j=n_k-p_2+1}^{n_k} \varphi \times \int_0^\infty \log_2\left(1 + \frac{1}{n_T}\lambda_2\right) \lambda_2^{\frac{m_2+n_2}{2}-n_k+j-1} \times K_{n_2-m_2}\left(2\sqrt{\frac{\lambda_2}{f_i}}\right) d\lambda_2. \quad (16)$$

Using the identity of equations (13) and (14), the closed-form expression for the $\mathcal{I}_2$ can be derived as

$$\mathcal{I}_2 = \frac{1}{\log_e 2 p_2 \prod_{i<j}^{n_k}(f_i - f_j)} \sum_{i=1}^{n_k} \sum_{j=n_k-p_2+1}^{n_k} \varphi \times G_{4,2}^{1,4}\left[\frac{\lambda_2}{n_T} f_i|_{1,0}^{n_k-j+1-n_2, n_k-j+1-m_2,1,1}\right]. \quad (17)$$

Finally, substituting equations (15) and (17) into (7), the closed-form analytical expression for the ergodic secrecy capacity for the multi-keyhole MIMO channel shown in equation (18) at the bottom of the next page.

IV. Numerical Results

In this section, we investigate the effect of number of keyholes on the ergodic secrecy capacity based on the analytical expressions derived in section III. Fig.2 depicts the analytical results of ergodic secrecy capacity versus SNR for the selected number of keyholes. Since the keyhole is assumed to ideally reradiate the captured energy, like the transmit and receive scatterers, hence the ergodic secrecy capacity increases with the number of keyholes.

Fig.3 describes the effect of transmitting, receiving and eavesdropper's antenna on the ergodic secrecy capacity as a function of SNR. We see that the ergodic secrecy capacity increases with the number of transmitting and receiving antennas and decreases with the number of eavesdropper's antenna.

In Fig.4, the ergodic secrecy capacity is plotted as a function of the number of keyhole for selected values of the number of receiving antennas. We see that the ergodic secrecy capacity increases exponentially for a certain number of keyhole. After that there is no effect of keyhole on the ergodic secrecy capacity for a particular receiving antenna. In this case, the

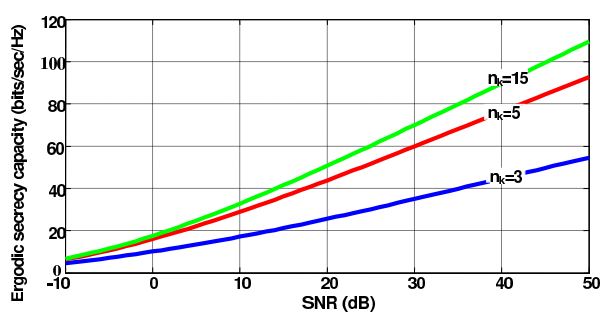


Fig. 2. The ergodic secrecy capacity versus SNR for selected values of keyhole.

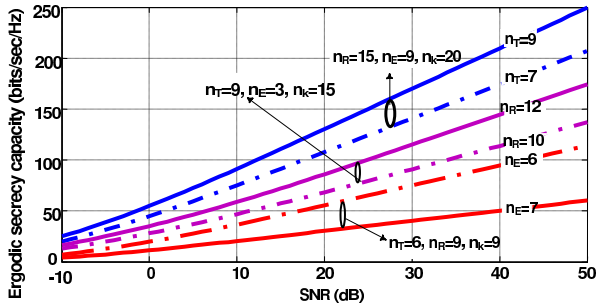


Fig. 3. The effect of transmitter's, receiver's and eavesdropper's antenna on the ergodic secrecy capacity as a function of SNR.

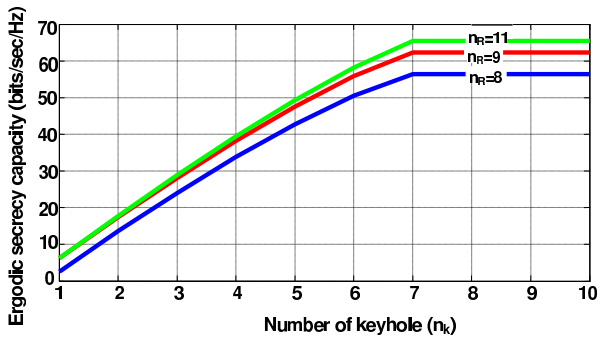


Fig. 4. Impact of the number of keyhole on the receiving antenna.

ergodic secrecy capacity increases only with the number of receiving antennas.

V. Conclusion

In order to investigate the effect of number of keyholes in multi-keyhole MIMO channel, a closed form expression of the ergodic secrecy capacity has been derived in this paper. Based on our observation and from numerical results, we may conclude that the ergodic secrecy capacity increases with the number of keyholes. It is also observed that the ergodic secrecy capacity depends on the number of keyhole upto a certain number of keyholes for a specific number of receiving antenna and after that it depends only on the number of receiving antenna.

VI. Acknowledgement

This work was supported by the Research Support and Publication Division of University Grant Commission, Bangladesh.

References

[1] H. Shin, *Capacity and Error Exponents for Multiple-Input Multiple-Output Wireless channels*, Seul National University, August 2004.

[2] A. Muller and J. Speidel, "Ergodic capacity and information outage probability of MIMO Nakagami -m keyhole channels with general branch parameters," *WCNC*, 2007.

[3] A. Nezampour, A. Nasri, and R. Schober, "Asymptonic analysis of space-time codes in generalized keyhole fading channels," *IEEE Wireless Communication and Networking Conference*, pp. 1–6, April 2010.

[4] A. Maaref and S. Aissa, "Impact of spatial fading correlation and keyhole on the capacity of MIMO systems with transmitter and receiver CSI," *IEEE Transactions on Wireless Communications*, vol. 7, no. 8, pp. 3218–3229, August 2008.

[5] C. Zhong, S. Jin, K.-K. Wong, and M. R. McKay, "Ergodic mutual information analysis for multi-keyhole MIMO channels," *IEEE Transactions on Wireless Communications*, vol. 10, no. 6, pp. 1754–1763, 20 June 2011.

[6] J. Xue, T. Ratnarajah, and C. Zhong, "Error exponents for multi-keyhole MIMO channels," *Problems of Information transmission*, vol. 51, no. 1, pp. 1–19, January 2015.

[7] J. Xue, M. Z. I. Sarkar, and T. Ratnarajah, "Error exponents for Rayleigh fading multi-keyhole MIMO channels," *International Conference on Communications*, pp. 3176–3180, June 2013.

[8] C. Zhong, *Capacity and Performance Analysis of Advance Multiple Antenna Communication System*, University College London, London, United Kingdom, March 2010.

[9] A. P. Prudnikov, Y. A. Brychkov, and O. I. Marichev, *Integrals and Series*, 3rd ed. Gordon and Breach Science Publisher, 1990.

[10] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. San Diego, CA: Academic, 2000.

$$\langle C_s \rangle = \frac{1}{\log_e 2 \prod_{i<j}^{n_k}(d_i - d_j)} \sum_{i=1}^{n_k} \sum_{j=n_k-p_1+1}^{n_k} \psi G_{4,2}^{1,4}\left[\frac{\lambda_1}{n_T} d_i|_{1,0}^{n_k-j+1-n_1, n_k-j+1-m_1, 1, 1}\right] - \frac{1}{\log_e 2 \prod_{i<j}^{n_k}(f_i - f_j)} \sum_{i=1}^{n_k} \sum_{j=n_k-p_2+1}^{n_k} \varphi G_{4,2}^{1,4}\left[\frac{\lambda_2}{n_T} f_i|_{1,0}^{n_k-j+1-n_2, n_k-j+1-m_2, 1, 1}\right]. \quad (18)$$

A Novel Low Buffered Optimized Solid State Drive Controller

Ratul Das[1,*], Baishakhi Rani Biswas[1], ABM Harun-ur Rashid[1], MA Abedin[2]
[1]Dept. of EEE, Bangladesh University of Engineering and Technology
Dhaka, Bangladesh
[2]Dept. of EEE, Dhaka University of Engineering and Technology
Gazipur, Bangladesh
*ratuldas@live.com

Abstract- **In this paper, a novel algorithm to reduce the intermediate buffer size of a Solid State Drive (SSD) controller has been proposed. The algorithm implements compression and encryption to reduce the stored data and secure it. The cyclic operation of compression and encryption reduces the intermediate buffer size substantially compared to serial operation. The algorithm has been optimized and tested for different standards of Xtensa processor and the best suitable standard has been reported in terms of processor cycles.**

***Index Terms*- LZ77, AES-128, processor cycles, minimum clock rate, optimization.**

I. Introduction

Solid State Drives (SSD) are used extensively in mobile and personal computing [1]. It is becoming popular because of its light-weight, low cost and low power consumption [2]. The performance and lifetime of a SSD can be improved using data compression by reducing the actual amount of data written in the SSD [3]. For sensitive data, there are a handful of cipher algorithms [4] to ensure data protection.

Typically the compression and encryption process are implemented separately where intermediate buffer size would have been as large as the compressed data size. In this work, lossless data compression algorithm LZ77 and standard encryption algorithm AES-128 have been implemented together in cyclic order (e.g. Fig. 1) to control a SSD by Xtensa processor [5] which uses intermediate buffer size only of 18 bytes in both forward and backward path. Implementation of compression/decompression and encryption/decryption in an Xtensa processor for the purpose of SSD controlling is the main objective of this work.

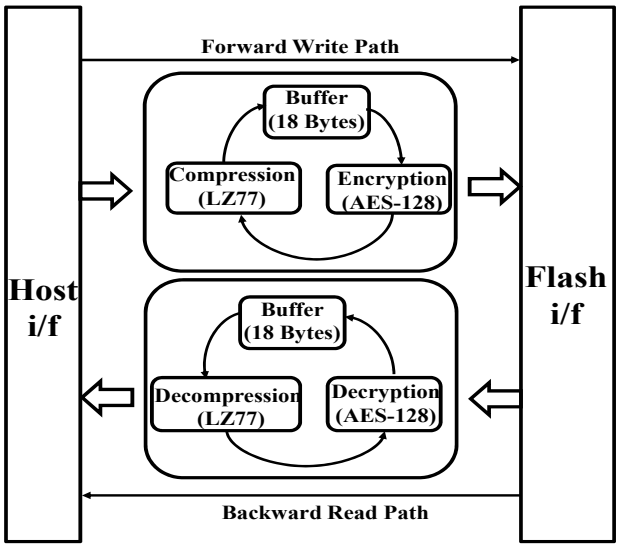


Fig. 1: Proposed SSD controller

Standard AES-128 algorithm optimization and effect of search buffer size variation in LZ77 on compression ratio has been discussed in section II and III respectively. Optimization of intermediate buffer sizes has been discussed in section IV. Algorithm optimization for Xtensa processor with different built in standards of Xtensa processor have been discussed in section V. Section VI discusses about the processor performance with offset length variation in LZ77 algorithm. The results have been reported in section VII.

II. Optimization of AES-128 algorithm

AES-128 is an encryption algorithm [6, 7] which uses a password or key to encrypt data. During the decryption process, the same password or key is required. The key is processed by key expansion which is a common procedure required in both encryption and decryption. Encryption consists of four subparts - subbyte, shift row, mixcolumn and add round key which are needed to be implemented sequentially. Decryption process requires the inverse of the subparts in the reverse order.

However, subbyte & shift row are commutable and mixcolumn & add round key are also commutable in the encryption process. Same situation holds true for inverse of subbyte & shift row and mixcolumn & add round key during the decryption process. Some of mentioned commutative characteristics of encryption and decryption steps results in scopes for merging loops which decreased the cycle numbers of forward write path and backward read path notably.

There are some constant values required during key expansion, subbyte, mix column and the inverse of subbyte & mixcolumn operation which can be provided directly without calculation. Although this will increase the size of the code but reduce the required processor cycles. As, this work concentrates on optimizing the SSD controlling with Xtensa processor, this is acceptable because the increase of the code size is compatible with most of the built in standards in Xtensa processor.

III. Effect of window size variation of LZ77 algorithm on compression ratio

LZ77 is an established algorithm in the field of data compression [8]. There have been a lot of researches for optimizing the algorithm and a few algorithm have been proposed based on LZ77 for data compression [9, 10]. In most cases, there is a trade-off between the compression ratio and computation complexity.

LZ77 is called a sliding window based algorithm. Sliding window consists of search buffer and look ahead buffer. Search buffer contains a portion of the sequence that has been encoded recently and the look ahead buffer contains the next portion [11].

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

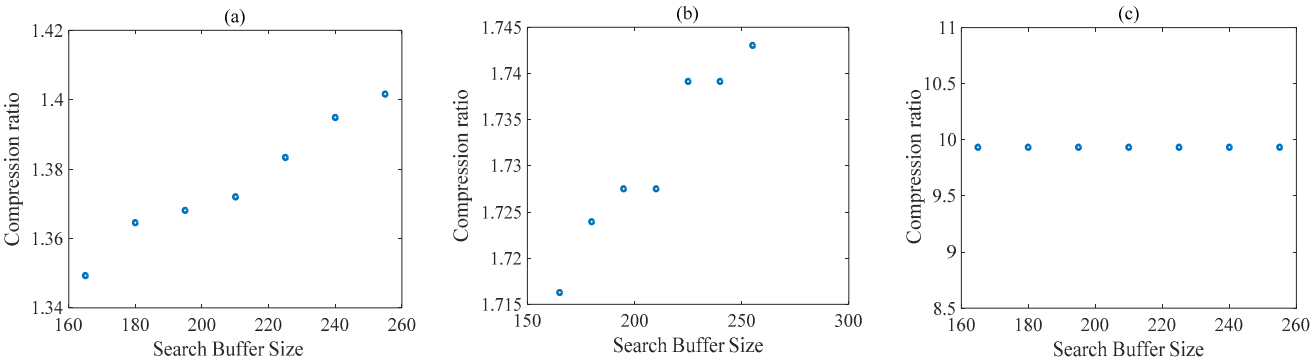


Fig. 2: Compression ratio vs search buffer size for (a) data 1, (b) data 2 and (c) data 3

The algorithm searches the sliding window for the longest match with the beginning of the look-ahead buffer. This searching requires a lot of processor cycles. Processor cycles and compression ratio are dependent on the length of search buffer and look-ahead buffer. Search buffer and look-ahead buffer length has been taken equal in this work. Highest search buffer length was selected as 255 which can be stored within 1 byte.

The compression ratio is defined as the ratio of uncompressed data and compressed data and depends on search buffer length [12]. But, it is the consistency of the data which mainly determines the compression ratio.

Three hex sample data of approximately 32 Kbytes have been used to observe the effect of search buffer size variation on compression ratio and the results are shown in Fig. 2(a), 2(b) and 2(c). In Fig. 2(c) the compression ratio is constant (approximately 10) over the variation of search buffer size whether in Fig. 2(a) and 2(b), the compression ratio varies slightly around 1.38 and 1.73 respectively for same search buffer sizes.

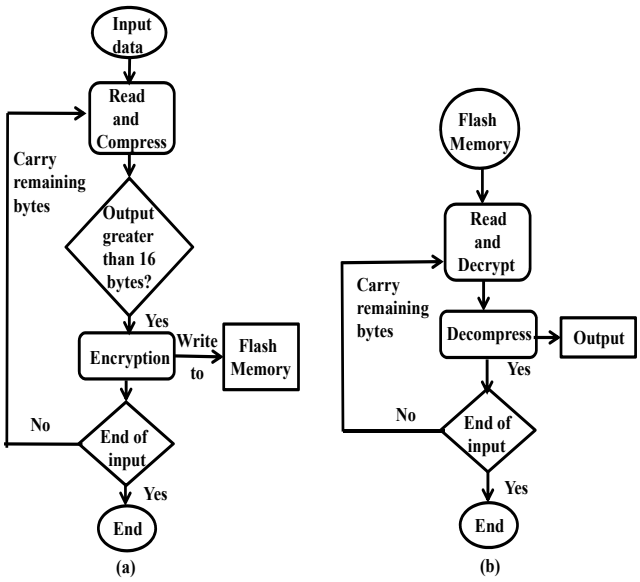


Fig. 3: Flow Chart of (a) Forward Write Path algorithm and (b) Backward Read Path algorithm

IV. Reduction of Intermediate Buffer Size

If serial operation of compression and encryption would have done, buffer size will be totally dependent on compression ratio. Such as if a sample data size is 32 Kbytes and compression ratio is 2 that means the compressed data size is 16 Kbytes and intermediate buffer size is also 16 Kbytes. This is also true for the reverse process. Cyclic operation in both forward and backward path reduces the intermediate buffer size to only 18 bytes.

LZ77 compression algorithm produces output of 3 bytes every time. On the other hand, the encryption algorithm can only encrypt 16 bytes of data at a time. So, when compression procedure produces accumulated data greater than or equal 16bytes, the first 16 bytes are encrypted, stored in the flash memory and the remaining bytes are carried for the next encryption cycle. By applying this idea, the size of intermediate buffer can be reduced substantially (18 bytes in this work).

While reading the encrypted data back and decompressing it for use, the same technique has been applied. The flow chart of forward write and backward read path algorithm are shown in Fig. 3.

V. Optimization Approach in Xtensa Processor

At first, the combined algorithm has been implemented and optimized for a data of size 1Kbytes on Xtensa processor. But computation cycles were too high which can be seen from the bar charts of Fig. 4(a)-1 and 4(a)-2. By applying optimization techniques provided by Xtensa C and C++ Compiler [13], minimum cycles for execution of the algorithm and the best built in standard has been ascertained.

1) Automatic Feedback

Use of automatic feedback improves run-time performance [13] significantly. It enables compiler to automatically decide which functions to compile for speed and which functions to compile for size. For the conditional tests in this work, use of automatic profiling feedback drastically reduced the cycle numbers.

2) Interprocedural Analysis

Interprocedural analysis [13] has been introduced for code optimization. Most other compiler optimizations focus on individual functions whereas it focuses on the whole program. This is a technique provided with the compiler optimization configuration.

3) SIMD Vectorization

This technique [13] performed a slight reduction in the execution cycles although the reduction was not as great as the previous two techniques.

Xtensa C and C++ compiler has 4 optimization level (O0, O1, O2 and O3) [12]. Simulating the algorithm with different optimization level & options and also with different Xtensa processor configurations, the best processor and compiler settings have been determined.

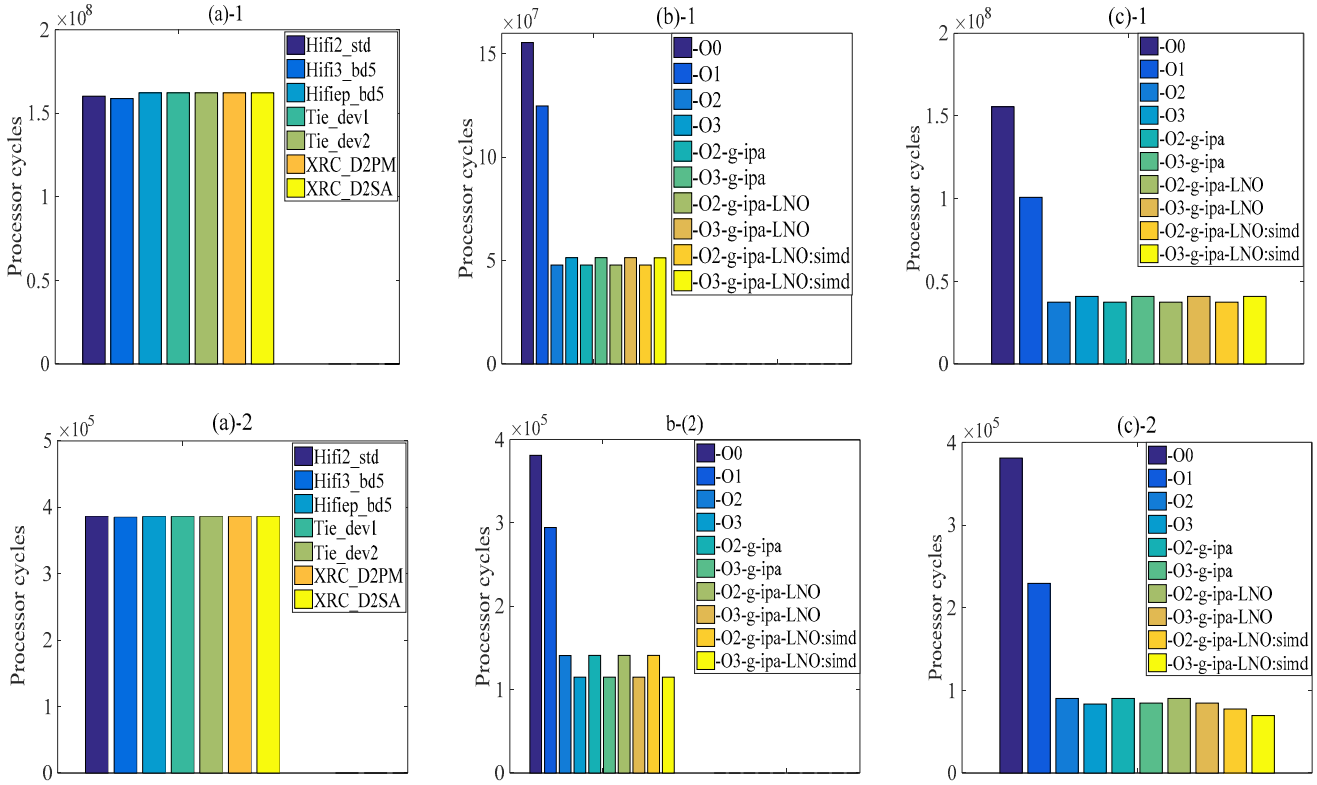


Fig. 4: (a) Required cycles for different built in standards of Xtensa processor for (1) Forward path (2) Backward path
(b) Required cycles for various optimization options with 'Hifi2_std' for (1) Forward path (2) Backward path
(c) Required cycles for various optimization options with 'XRC_D2PM' for (1) Forward path (2) Backward path

Every built in standard of Xtensa processor (Hifi2_std, Hifi3_bd5, Hifiep_bd5, Tie_dev1, Tie_dev2, XRC_D2PM and XRC_D2SA) [12] was compiled for optimization but 'XRC_D2PM' has provided the best performance whereas '-02 –g –ipa –LNO:simd' and '-O3 –g –ipa –LNO:simd' compiler configurations are required for the write and read path respectively. Here required processing cycles for Hifi2_std (e.g. Fig. 4(b)) and XRC_D2PM (e.g. Fig. 4(c)) has been shown in bar charts. Here, '-g' stands for the command to the compiler to include debug information, '-ipa' stands for interprocedural analysis and '-LNO:simd' stands for SIMD vectorization.

VI. Effect of Sliding Window Length Variation on Processor Performance

The effect of search buffer variation on write and read path cycles are shown in TABLE I. It can be observed that reduction in search buffer size drastically reduced the forward write path cycle which was expected. On the other hand, reverse read path cycles increased slightly. As compression ratio decreases with the reduction of the search buffer size, the increased compressed data size results in the increase of cycles in the backward read path. However, significant reduction in the search buffer size does not change the compression ratio and backward read path cycles much, but it can drastically reduce the forward path cycle.

Another important parameter for evaluating processor performance is minimum clock rate required to implement the algorithm. The minimum clock rate depends on input data rate as well as the designed algorithm. The processor cycles calculated for search buffer size (255 in this work) have been used to calculate the cycles required for per byte of data and used for determination of minimum clock rate at different input data rate and the results are shown in Fig. 5(a) and 5(b). The write path requires a high clock rate because of LZ77 algorithm's search of processed data. On the other hand, the read path requires a very low clock rate.

TABLE I
EFFECT OF BUFFER SIZE ON WRITE AND READ PATH CYCLES

Data No	Search Buffer size			Search Buffer size		
	165	195	255	165	195	255
	Forward write path cycles ($\times 10^9$)			Backward read path cycles ($\times 10^6$)		
Data1	2.35	3.25	5.41	5.58	5.5	5.38
Data2	1.88	2.62	4.44	4.36	4.34	4.3
Data3	0.442	0.632	1.17	0.86	0.86	0.86

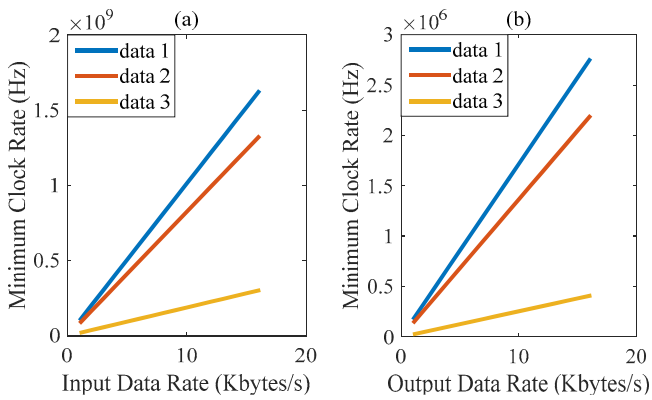


Fig. 5: (a) Minimum Clock Rate vs Input Data Rate for forward write path and (b) Minimum Clock Rate vs Output Data Rate for backward read path with 'XRC_D2PM'

VII. RESULTS

The results of the work are given in TABLE II. From the table it can be seen that the forward write path requires a high clock rate at a normal input data rate. This is because of the search buffer window length of the LZ77 algorithm. As, LZ77 and AES-28 algorithms are combined, summation of their minimum clock rate should be considered as the minimum clock rate. However, in the read path, the decryption cycles have not been shown separately because the backward read path minimum clock rate is already low. Taking search buffer window size of 80 instead of 195 resulted in more realistic minimum clock rate for forward path. However, reduction in search buffer length results in slightly lower compression ratio.

TABLE II
RESULTS OF THE WORK

Processor Configuration		XRC_D2PM	
Compiler Configuration for write path		-O2 –g –ipa –LNO:simd	
Compiler Configuration for read path		-O3 –g –ipa –LNO:simd	
Search buffer size in LZ77		195	80
Write path cycles	Data 1	3.24×10^9	6.11×10^8
	Data 2	2.62×10^9	5.85×10^8
	Data 3	6.32×10^8	8.29×10^7
Encryption Cycles	Data 1	4.51×10^6	5.03×10^6
	Data 2	4.85×10^6	4.72×10^6
	Data 3	6.53×10^5	6.98×10^5
Intermediate buffer size in write path		18 bytes	18 bytes
Read path cycles	Data 1	5.51×10^6	6.13×10^6
	Data 2	4.34×10^6	5.76×10^6
	Data 3	8.56×10^6	9.10×10^6
Intermediate buffer size in read path		18 bytes	18 bytes
Compression Ratio	Data 1	1.3682	1.228
	Data 2	1.7275	1.295
	Data 3	9.9318	9.28
Total Code size with compiler configuration '-O2 –g –ipa –LNO:simd' (bytes)		95858	95858
User Code size with compiler configuration '-O2 –g –ipa –LNO:simd' (bytes)		8384	8384
Minimum clock rate for write path at input data rate 8kbytes/s(MHz)	Data 1	810.365979	152.434809
	Data 2	660.022652	147.316773
	Data 3	150.522852	19.771034
Minimum clock rate for encryption at input data rate 8 Kbytes/s (MHz)	Data 1	1.12604	1.254323
	Data 2	1.092310	1.189419
	Data 3	0.155462	0.166209
Minimum clock rate required for read path at output data rate 8 Kbytes/s (MHz)	Data 1	1.374053	1.52926
	Data 2	1.092310	1.437526
	Data 3	0.203945	0.216852

VIII. CONCLUSION

This whole work can be categorized mainly in two stages. In the first stage, intermediate buffer size and processing cycles are minimized by introducing cyclic operation and loop minimization. Cyclic operations of compression and encryption in the forward path and decryption and decompression in the backward path have resulted in very low intermediate buffer size of only 18 bytes. The effect of search buffer size on compression ratio, processor cycles and minimum clock rate have also been discussed in details.

In the second stage, the solid state drive controller has been optimized on Xtensa processor by selecting the best built in standard and using some special techniques and commands. The code size, compression ratio and minimum required clock rate have been reported for different data samples.

However, the clock rate can be optimized more if parallel operation of compression and encryption in write path and decryption & decompression in read path are used. Moreover, other lossless compression algorithms might have resulted better processor performances [14].

ACKNOWLEDGEMENT

The authors would like to thank Cadence Design Systems, India for providing the software and license required for this work.

REFERENCES

[1] A. Kosuge, J. Hashiba, T. Kawajiri, S. Hasegawa, T.Shidei,Tadahiro Kuroda and K. Takeuchi, "An Inductively Powered Wireless Solid-State Drive System With Merged Error Correction of High-Speed Wireless Data Links and NAND Flash Memories," *IEE Journal of Solid State Circuits*, vol. 51, pp. 1041-1050, Apr. 2016.

[2] (2014) S. Yasarapu, "The second wave of the SSD revolution", Flash Memory Summit. [Online]. Available:http://www.f ashmemorysummit.com/English/Collaterals/ Proceedings/2014/20140805_C11_Yasarapu.pdf.

[3] S. Lee, J. Park, K. Fleming, Arvind and J. Kim, "Improving performance and lifetime of solid-state drives using hardware-accelerated compression," *IEEE Transactions on Consumer Electronics*, vol. 57, pp. 1732–1739, Nov. 2011.

[4] Code.org, "The Internet: Encryption & public keys," in YouTube, YouTube, 2015. [Online]. Available: https://www.youtube.com/watch?v=ZghMPWGXexs.

[5] (2014) Xtensa customizable processors - Cadence IP. [Online]. Available: http://ip.cadence.com/ipportfolio/tensilica-ip/xtensa-customizable

[6] United States National Institute of Standards and Technology (NIST), "Announcing the Advanced Encryption Standard (AES)," *Federal Information Processing Standards Publication 197,* Nov. 26, 2001.

[7] E. Zabala, "Rinjadel Cipher". [Online]. Available: http://www.formaestudio.com/rijndaelinspector/archivos/Rijndael_ Animation_v4_eng.swf.

[8] J. Ziv and A. Lempel, "A universal algorithm for sequential data compression," *IEEE Transactions on Information Theory*, vol. 23, pp. 337–343, May 1977.

[9] LZ77 and LZ78. [Online], Available: https://en.wikipedia.org/wiki/LZ77_and_LZ78

[10] "Category: Lossless Data Compression Algorithms," [Online], Available: https://en.wikipedia.org/wiki/Category: Lossless_ compression_algorithms

[11] H. Kumawat and J. Chaudhury, "Optimization of LZ77 Data Compression Algorithm," *International Journal of Computer Engineering and Technology,* vol. 4, pp. 42-48, Sept. 2013.

[12] [Online], Available: http://ethw.org/History_of_Lossless_ Data_ Compression_Algorithms

[13] Xtensa C and C++ Compiler User Guide, collected from Cadence Design Systems, India.

[14] H. Yuanfu and W. Xunsen, "The methods of improving the compression ratio of the LZ77 Family data compression algorithms," *Signal Processing 1996, 3rd international Conference,* vol. 1, pp. 698-701.

Spiral Photonic Crystal Fiber for Gas Sensing Application

Md. Ibadul Islam[1], Maksuda Khatun[1], Shuvo sen[1], Kawsar Ahmed[1] and Sayed Asaduzzaman[1,2,*]
[1]Department of Information and Communication Technology (ICT), Mawlana Bhashani Science and Technology University (MBSTU), Santosh, Tangail-1902, Bangladesh.
[2]Department of Software Engineering, Daffodil International University, Dhaka, Bangladesh
*sayed.asaduzzaman.bd@ieee.org

***Abstract*—This paper proposes, the design and characterization of a spiral shape photonic crystal fiber for gas sensing application. The core region of the proposed PCF is oriented with elliptical air holes in a hexagonal shape and the cladding is oriented in a spiral manner. The proposed PCF not only shows high sensitivity but also shows high birefringence as well. The geometrical parameters were altered to fix the optimized structure. By utilizing finite element method (FEM) the numerical analysis of the proposed PCF is performed. Our proposed PCF shows the high sensitivity of 42.27% and high birefringence of 1.94×10^{-2} at λ=1.33μm which is the absorption line of toxic gasses like (methane and hydrogen fluoride).**

***Keywords*— Birefrngence; Sensitivity; Finite element method; Evanescent field; Gas sensor; Spiral Photonic Crystal fiber**

I. Introduction

Photonic Crystal Fiber (PCFs) is currently a topic that greatly impacts not only in optical communication but also other modern technologies. It overcomes many drawbacks of conventional optical fiber. In one realization, photonic crystal fiber is a light guidance single-material fiber technology which consists of a more-or-less random and periodic array [1-2] of air holes that remains along the entire fiber, while a region of pure silica embedded within the fiber. In recent years, the PCF attracts the researchers due to fascinating structures [3], unique and exception features [4] enable to observe the guiding properties like dispersion [5], birefringence [6] and relative compatibility with optic telecommunications technology. PCF technology performs a large number of potential applications which have demonstrated sensing of toxic gasses [7], chemicals [8], biomedical [9] and environmental applications.

Various guiding properties of PCF can be achieved by changing of air hole shape, size, and their positions (geometrical parameters). PCF sensing technique is an interaction between passing light and analyte at the core which varies due to frequency, phase, wavelength, intensity and the polarization state of light [10]. Better guiding properties have already been achieved by different geometric shape lattice structures like Circular [11], Octagonal [12] and Decagonal [13]. However, various forms of polymer fibers have already existed for efficient transmissions such as plastic fiber [14], bragg fiber [15], hollow core fiber [16], and sub wavelength porous fiber [17]. The porous photonic crystal fibers attract the attention due to its features such as highly birefringence, low confinement loss, very small dispersion, low material absorption loss and flexibility in fabrication [18, 19]. In [7] a simple circular structure of PCF based gas sensors improved sensitivity 13.94% compares to [20] and article [21] improved sensitivity 20.10% which proved the merits of the modified structure. Four non-circular holes around a hollow high index GeO_2–doped silica based PCF was proposed in [22] that shows high sensitivity and low confinement loss. An asymmetry in porous-core spiral photonic crystal fiber can be obtained an ultrahigh birefringence of 0.0483[23]. In this article, we have proposed a spiral photonic crystal fiber where the core is formed with hexagonal shaped elliptical holes. The proposed PCF shows high sensitivity and high birefringence.

II. Geometrical Structure of the Proposed S-PCF

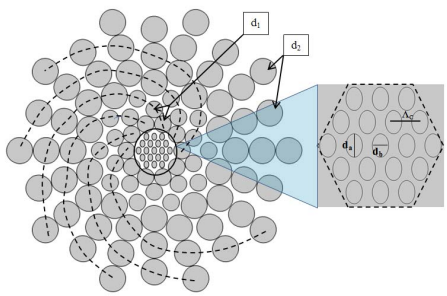


Fig.1. Cross sectional view of the proposed S-PCF

Fig. 1 shows the cross sectional view of the proposed spiral Photonic Crystal Fiber with the magnification of core region. The proposed PCF contains a cluster of elliptical air holes in the core region which are formed into a hexagonal shape and cladding is spiral formed. Spiral topology has attracted increased attention due to their excellent modal confinement properties compared to conventional hexagonal PCF [24]. Spiral symmetry consists of ten spiral arms, where each arm consists of nine circular air holes. The center to center distance between the two adjacent holes is called the pitch. The simplicity of the design demonstrated by $r_0 = \Lambda$ where r_0 is a distance of the first air hole in each spiral ring and Λ is the pitch of the lattice. The second air hole of each ring is placed from the center $r_1 = r_0 + (0.48 * \Lambda)$ distance. The proposed structure of PCF which contains the nth air hole in each spiral ring is placed from the center at a distance $r_n = r_{n-1} + (0.48 * \Lambda)$ with an angular displacement, $\theta_n = (n*360^o)/(2*N)$ where N is the number of circular rings. The air-hole diameter of the first spiral ring is d_1 and expands the diameter of second and third circular air-holes in the cladding are d_2. In the core region, elliptical air holes are arranged by the optimized diameters of the major axis is d_a and the minor axis d_b. The pitch between adjacent elliptical air holes in the core region is

Λ_c. Pure silica is used as the background material. The refractive index of silica changes with the wavelength according to the sellmier equation. Throughout the analysis, the simulation of selecting the core air-filling ratio offers maximum birefringence which will be fulfilled by the proposed structure.

III. Synopsis of Numerical Method

To simulate the optical properties of this proposed S-PCF, a finite element method (FEM) is used for solving Maxwell's equation. As a result, it provides a good approximation to prove the reliability of this S-PCF structure. We know that due to the finite number of air holes in the cladding, light can leak among the air holes which are called the leakage loss or confinement loss. So, to compute the leakage loss, an effective boundary condition is needed that produces no reflection at the boundary. In this case, Perfectly Matched Layers (PMLs) are the most efficient absorption boundary condition for this purpose.

The confinement loss (dB/m) can be calculated by the following equation:

$$L_c = 8.686 \times K_0 \, \mathrm{Im}\,[n_{eff}] \text{ (dB/m)} \quad (1)$$

Where, $K_0 = 2\pi/\lambda$; λ wavelength of light and $\mathrm{Im}[n_{eff}]$ is the imaginary part of the effective refractive index.

The difference between refractive index of x-polarization and y-polarization is called the birefringence which can be calculated as:

$$B = |\, n_x - n_y \,| \quad (2)$$

Where n_x and n_y are refractive index of x-polarization and y-polarization respectively.

On the other-hand, Beer-Lambert law is used to evaluate the absorbance of evanescent field by the gas samples. The relationship between gas concentration and optical intensity is given by the following equation:

$$I(\lambda) = I_0(\lambda) \exp[-r\alpha_m l_c] \quad (3)$$

where I and I_0 are the output light intensities respectively, α_m is the absorption coefficient, l is the length of the S-PCF used for gas detection, c is the gas concentration, and r is a relative sensitivity coefficient defined as:

$$r = \frac{n_r}{Re[n_{eff}]} f \quad (4)$$

Where n_r are refractive index of gas, $\mathrm{Re}[n_{eff}]$ is the real part of the effective mode index and f is the fraction of the total power located in the holes that can be expressed as:

$$f = \frac{\int_{\text{sample}} \mathrm{Re}(E_x H_y - E_y H_x)\, dx\, dy}{\int_{\text{total}} \mathrm{Re}(E_x H_y - E_y H_x)\, dx\, dy} \times 100 \quad (5)$$

Where E_x, E_y, H_x and H_y are the transverse electric and magnetic fields of the guided mode. These transverse fields and the effective index of mode n_{eff} can be calculated by solving Maxwell's equation by the process of FEM method.

In this research, the main focus is on the methane (CH4) and hydrogen fluoride (HF) gasses to have the absorption line at the wavelength of 1.33µm.We have tried to improve the relative sensitivity and confinement loss at this wavelength. Furthermore, for the other wavelength our proposed S-PCF has shown the better results.

IV. Numerical Results and Discussion

Figure 2 depicts the fundamental mode field distribution for the proposed S-PCF at the operating wavelength of 1.33 µm. From this figure, it is observed that the Gaussian distribution of the fields is uniform as well as the fundamental modes are strongly bound by the hexagonal core region of elliptical air holes. It also illustrates that the sample air holes close to the core region strongly interacts with the evanescent field that causes the sensitivity very high.

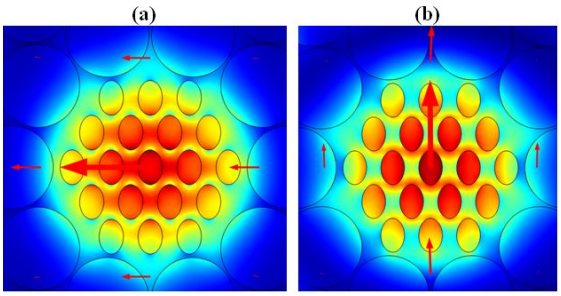


Fig. 2 Modal intensity distribution of proposed S-PCF at the wavelength of 1.33µm

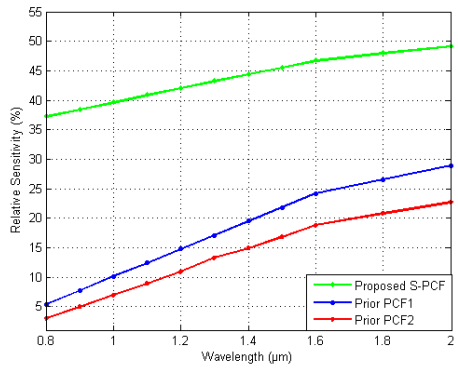


Fig. 3 Relative sensitivity versus wavelength for the proposed S-PCF for the optimum design parameters: Λ =1.6µm; d_a/d_b =0.65; d_1= 0.96µm and d_2= 1.5µm

The relative sensitivity versus wavelength curve for the optimum design of the proposed structure and prior structure has been shown in Fig. 3. The optimized values of the proposed structure are Λ =1.6µm, d_a/d_b =0.65, d_1= 0.96µm, and d_2= 1.5µm which was fixed by varying the parameters and investigating the behavior of guiding properties with it. It can be seen that 42.27% relative sensitivity can be obtained at the operating wavelength of 1.33µm. From above observations, it can be seen that the proposed S-PCF has a higher relative sensitivity compared to prior PCFs.

Figure 4 shows that if we vary the diameter of cladding d_1 kept fixed other parameters (d_a/d_b =0.65, d_1= 0.96, 0.90 and 0.86μm and d_2= 1.5μm), the relative sensitivity will be changed. With the variation of d_1 from up to down of optimized value, the relative sensitivity decreases. By the reduction of diameter d_1, the evanescent field is penetrated into cladding causing the fraction of the power of holes alleviates and leads to decrease the absorption power by core region which results in relative sensitivity decreases. By this way we optimized d_1.

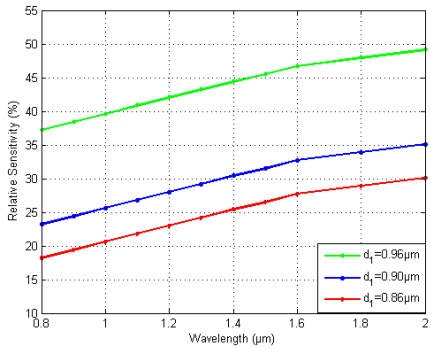


Fig. 4 Relative sensitivity versus wavelength for the proposed S-PCF for the parameters: d_1=0.96, 0.90 and 0.86μm and the other parameters were constant.

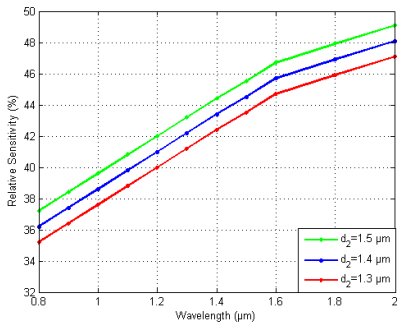


Fig. 5 Relative sensitivity versus wavelength of the proposed S-PCF for the parameters Λ=1.6μm; d_1=0.96μm; d_2=1.5, 1.4 and 1.3μm

Figure 5 also illustrates with varying the diameter of cladding d_2 kept fixed other parameters (d_a/d_b=0.65, d_1= 0.96μm and d_2=1.5, 1.4, 1.3μm), the relative sensitivity will be unstable. With the variation of d_2 from up to down of optimized value, the relative sensitivity decreases. By the reduction of the diameter d_2, the evanescent field is penetrated into cladding causing the fraction of the power of holes alleviates and leads to decrease the absorption power by core region which results in relative sensitivity decreases. From this analysis, we have optimized d_2.

Figure 6 depicts the relative sensitivity cure with increasing the ratio of minor and major axis from optimized value kept fixed other parameters (Λ=1.6μm, d_1= 0.96μm, and d_2=1.5μm) of ellipses of the core which leads to alleviating the absorption power into core causes decreases relative sensitivity.

Figure 7 exhibits the impact of changing global parameters of ±1%, and ±2%, on the relative sensitivity which leads to promoting the relative sensitivity with increase the parameters +1%, +2% and alleviate the sensitivity by the reduction of parameters -1% and -2%.

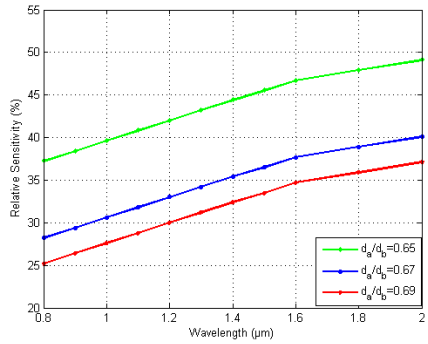


Fig. 6 Relative sensitivity versus wavelength for the proposed S-PCF for the parameters: Λ=1.6μm; d_a/d_b =0.65, 0.67 and 0.69; d_1= 0.96μm and d_2= 1.5μm.

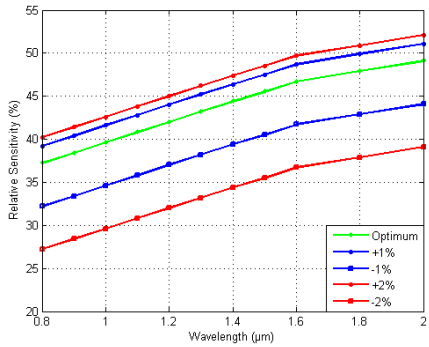


Fig. 7 The relative sensitivity versus wavelength of proposed S-PCF for optimum parameters and effect of changing in global parameters.

Figure 8 describes the impact of changing global parameters of ±1% and ±2%, on birefringence which governs dramatically to increase and decrease the birefringence with respect to wavelength. It also shows the effective refractive indices for the orthogonal axis of the optimum proposed S-PCF design with their respective fundamental electric field properties at a different wavelength. It is depicted that the effective refractive index of the fast-axis direction is lower than that of the slow-axis direction, and the difference exhibits a high birefringence.

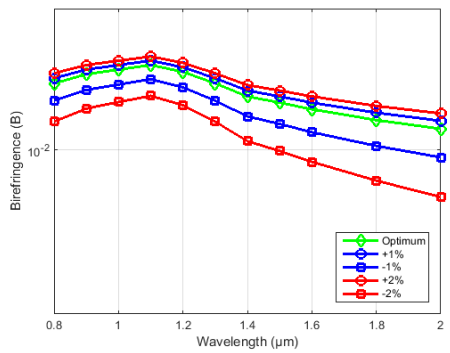


Fig. 8 The birefringence versus wavelength of proposed S-PCF for optimum parameters and effect of changing in global parameters.

TABLE I. COMAPRISION OF SIMUTATED RESULT AND STRUCTURE SHAPE AMONG PROPOSED PCF AND PRIOR PCFS AT λ=1.33 µm.

PCFs	Sen. (r%)	Birefringence	Structural Shape	
			Core	Cladding
Prior PCF_1 [25]	*13.3%*	-	*circular*	*Hexagonal*
Prior PCF_2 [22]	*16.88%*	-	*Defeated core*	*Hexagonal*
Proposed PCF	*42.27%*	1.94×10^{-2}	*Elliptical holes in hexagonal*	*Spiral*

Table 1 demonstrates the comparison between relative sensitivity, birefringence, structural shape of the proposed S-PCF with other prior designs in which prior PCF_1 and PCF_2 provide 13.3% and 16.88% sensitivity respectively and the proposed S-PCF provides 42.27% sensitivity and higher birefringence at a wavelength of 1.33µm. In the prior structure, birefringence properties were not analyzed.

Finally, the fabrication process is one of the fundamental issues in PCFs. In the proposed S-PCF three types of different air-hole are considered and shown in Fig. 1. However, the proposed structure can be fabricated by currently developed technology for technological advancement in the fabrication of PCFs [26]. Bise et al [27]. Provided a sol-gel technique to fabricate the PCFs with all structures, and they provide the freedom to adjust air-hole size, shape, and spacing. In such situation, for the proposed S-PCF the sol-gel casting method gives design flexibility that will be essential. After drawing the relative sensitivity curve to the expected level in the way just illustrated, we then verify the accuracy of the property of the proposed design. A standard fiber draw, ±1% variations in fiber global diameter may occur during the fabrication process.

V. Conclusion

A spiral photonic crystal fiber is illustrated that includes a core with elliptical air holes. It has been demonstrated that an asymmetry of air holes in the S-PCF core tends to high relative sensitivity (42.27%) and very high birefringence (0.01727) at a wavelength of 1.33µm. It is expected that the application of the proposed structure is more practical compared to prior PCFs and provides high relative sensitivity for colorless gas sensing.

Acknowledgment

The authors are grateful to those who participated in this research work.

References

[1] P. Russell, "Photonic crystal fibers," Science , vol. 299, pp. 358–362, 2003.

[2] J. C. Knight, T. A. Birks, P. S. J. Russell, and D. M. Atkin, "All-silica single-mode optical fiber with photonic crystal cladding," *Optics letters*, vol. 21, no. 19, pp. 1547–1549, 1996.

[3] F. Zolla, G. Renversez, A. Nicolet, B. Kuhlmey, S. Guenneau, D. Felbacq, A. Argyros, and S. Leon-Saval, *Foundations of photonic crystal fibres*. World Scientific, 2005.

[4] H. Ebendorff-Heidepriem, P. Petropoulos, S. Asimakis, V. Finazzi, R. Moore, K. Frampton, F. Koizumi, D. Richardson, and T. Monro, "Bismuth glass holey fibers with high nonlinearity," *Optics Express*, vol. 12, no. 21, pp. 5082–5087, 2004.

[5] F. Begum, Y. Namihira, S. A. Razzak, S. Kaijage, N. H. Hai, T. Kinjo, K. Miyagi, and N. Zou, "Design and analysis of novel highly nonlinear photonic crystal fibers with ultra-flattened chromatic dispersion," *Optics communications*, vol. 282, no. 7, pp. 1416–1421, 2009.

[6] M. S. Habib, M. S. Habib, S. A. Razzak, and M. A. Hossain, "Proposal for highly birefringent broadband dispersion compensating octagonal photonic crystal fiber," *Optical Fiber Technology*, vol. 19, no. 5, pp. 461–467, 2013.

[7] M. Morshed, S. Asaduzzaman, M. Faizul Huq Arif, and K. Ahmed, "Proposal of simple gas sensor based on micro structure optical fiber," in *Electrical Engineering and Information Communication Technology (ICEEICT), 2015 International Conference on*, 2015, pp. 1–5.

[8] S. Asaduzzaman, M. Kawsar Ahmed, F. H. Arif, and M. Morshed, "Application of Microarray-core Based Modified Photonic Crystal Fiber in Chemical Sensing."

[9] J. Jensen, P. Hoiby, G. Emiliyanov, O. Bang, L. Pedersen, and A. Bjarklev, "Selective detection of antibodies in microstructured polymer optical fibers," *Optics Express*, vol. 13, no. 15, pp. 5883–5889, 2005.

[10] B. Q. Wu, Y. Lu, C. J. Hao, L. C. Duan, N. N. Luan, Z. Q. Zhao, and J. Q. Yao, "Hollow-Core Photonic Crystal Fiber Based on C2H2 and NH3 Gas Sensor," in *Applied Mechanics and Materials*, 2013, vol. 411, pp. 1577–1580.

[11] S. Asaduzzaman, K. Ahmed, T. Bhuiyan, and T. Farah, "Hybrid photonic crystal fiber in chemical sensing," *SpringerPlus*, vol. 5, no. 1, pp. 1–11, 2016.

[12] H. Ademgil, "Highly sensitive octagonal photonic crystal fiber based sensor," *Optik-International Journal for Light and Electron Optics*, vol. 125, no. 20, pp. 6274–6278, 2014.

[13] S. A. Razzak, Y. Namihira, M. Abdul, F. Begum, and S. Kaijage, "Guiding properties of a decagonal photonic crystal fiber", *J. Microw. Optoelectron*, vol. 6, pp. 44–49, 2007.

[14] H. Han, H. Park, M. Cho, and J. Kim, "Terahertz pulse propagation in a plastic photonic crystal fiber," *Applied Physics Letters*, vol. 80, no. 15, pp. 2634–2636, 2002.

[15] H. Bao, K. Nielsen, H. K. Rasmussen, P. U. Jepsen, and O. Bang, "Fabrication and characterization of porous-core honeycomb bandgap THz fibers," *Optics express*, vol. 20, no. 28, pp. 29507–29517, 2012.

[16] X. Jiang, D. Chen, and G. Hu, "Suspended hollow core fiber for terahertz wave guiding," *Applied optics*, vol. 52, no. 4, pp. 770–774, 2013.

[17] M. I. Hasan, S. A. Razzak, G. K. M. Hasanuzzaman, and M. S. Habib, "Ultra-low material loss and dispersion flattened fiber for THz transmission," *IEEE Photonics Technology Letters*, vol. 26, no. 23, pp. 2372–2375, 2014.

[18] A. Hassani, A. Dupuis, and M. Skorobogatiy, "Low loss porous terahertz fibers containing multiple subwavelength holes," *Applied Physics Letters*, vol. 92, no. 7, p. 71101, 2008.

[19] J. Lei and Y. Jian-Quan, "Single Mode Condition and Power Fraction of Air-Cladding Total Refractive Guided Porous Polymer Terahertz Fibers," *Chinese Physics Letters*, vol. 28, no. 8, p. 84202, 2011.

[20] J. Park, S. Lee, S. Kim, and K. Oh, "Enhancement of chemical sensing capability in a photonic crystal fiber with a hollow high index ring defect at the center" *Optics express*, vol. 19, no. 3, pp. 1921–1929, 2011.

[21] M. Morshed, M. F. H. Arif, S. Asaduzzaman, and K. Ahmed, "DESIGN AND CHARACTERIZATION OF PHOTONIC CRYSTAL FIBER FOR SENSING APPLICATIONS" *European Scientific Journal*, vol. 11, no. 12, pp. 228-235, 2015.

[22] M. Morshed, M. I. Hasan, and S. A. Razzak, "Enhancement of the sensitivity of gas sensor based on microstructure optical fiber," *Photonic Sensors*, vol. 5, no. 4, pp. 312–320, 2015.

[23] M. R. Hasan, M. S. Anower, M. A. Islam, and S. M. A. Razzak, "Polarization-maintaining low-loss porous-core spiral photonic crystal fiber for terahertz wave guidance," *Applied Optics*, vol. 55, no. 15, pp. 4145–4152, 2016.

[24] A. Agrawal, N. Kejalakshmy, M. Uthman, B. M. A. Rahman, A. Kumar, and K. T. V. Grattan, "Ultra low bending loss equiangular spiral photonic crystal fibers in the terahertz regime," *AIP Advances*, vol. 2, no. 2, p. 22140, 2012.

[25] S. Olyaee, A. Naraghi, and V. Ahmadi, "High sensitivity evanescent-field gas sensor based on modified photonic crystal fiber for gas condensate and air pollution monitoring," *Optik-International Journal for Light and Electron Optics*, vol. 125, no. 1, pp. 596–600, 2014.

[26] K. M. Kiang, K. Frampton, T. M. Monro, R. Moore, J. Tucknott, D. W. Hewak, D. J. Richardson, and H. N. Rutt, "Extruded single-mode non-silica glass holey optical fibres," *Electronics Letters*, vol. 38, no. 12, pp. 546–547, 2002.

[27] R. T. Bise and D. J. Trevor, "Sol-gel derived microstructured fiber: Fabrication and characterization," in *Optical Fiber Communications Conf.(OFC)*, 2005, vol. 3.

Ultra-flattened Broadband Dispersion Compensating Photonic Crystal Fiber with Ultra-low Confinement Loss

Animesh Bala Ani[1], and Mohammad Faisal[2]

Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology, Dhaka 1205, Bangladesh
[1]animesh.ani@live.com, [2]mdfaisal@eee.buet.ac.bd

Abstract— **A simple index-guiding circular-lattice photonic crystal fiber is presented in this paper for compensation of dispersion with extremely low confinement loss. Full vector finite element method (FEM) with an anisotropic perfectly matched absorbing layer boundary condition is used to investigate the guiding properties. The designed dispersion compensating photonic crystal fiber (DC-DCF) shows that it is possible to obtain an ultra-flattened average dispersion of −124.0 ps/(nm.km) with a dispersion variation of ± 0.1252 ps/(nm.km) in the wavelength range of 1350 to 1700 nm (around E+S+C+L+U telecommunication band) at very low confinement loss less than 10^{-6} dB/km by using a modest number of design parameters and very simple cladding design. Furthermore, fiber nonlinearities and single-mode operation are also checked for the entire band.**

Index Terms— **Chromatic Dispersion, Confinement Loss, DCF, Nonlinearity, PCF**

I. Introduction

Photonic crystal fiber (PCF) has been a subject of extensive research all over the world due to the unusual guiding property of light through an arrayed distribution of micro-structured holes in a suitable background material, such as silica [1]. Based on the mechanisms of light guiding, PCFs have been widely categorized into two groups- index guiding PCFs and photonic band gap (PBG) guiding PCFs. Both kinds of fibers have important application areas which include dispersion compensation, supercontinuum generation, wavelength division multiplexing based communications, coherent communication, pressure sensing, CO_2 laser beam transmission etc. PCFs offer a great design flexibility providing better controls over the optical properties like dispersion, nonlinearity, birefringence, effective mode area, numerical aperture etc.

Optical fiber transmission system based on standard single-mode fiber having dispersion coefficient in the range of 15 to 20 ps/(nm.km) at 1550 nm has huge accumulated dispersion which is compensated for by DCF with negative dispersion coefficient. Conventional DCF has high losses, high dispersion slope and low effective area. Unlike DCF, the chromatic dispersion profile of a PCF can be easily controlled by varying the air hole diameter, the hole-to-hole spacing, and distribution of holes etc. Controllability of chromatic dispersion in PCFs is a very important problem for practical applications in optical communication systems, dispersion compensation, and nonlinear optics. So far, various PCFs with remarkable dispersion properties have been reported in recent years, for example, recently a complex design of equiangular spiral PCF with moderate negative dispersion is suggested for residual dispersion compensation [2]. DCF with a very large negative dispersion coefficient at 1550 nm has been proposed [3], however, such a high value is obtained only for a very narrow band which is not suitable for wavelength division multiplexed (WDM) networks. Using dual core structure, a very high value of dispersion has also been achieved for a narrow band [4]. The authors in [5] have used a hybrid structure of DC-PCF with wavelength tunable facility to compensate for dispersion in WDM network, however, the proposed structure is complex and has high confinement loss. Another hybrid PCF with high nonlinearity and birefringence has been proposed as DCF by Hasan et al. [6] whose dispersion slope is much higher with complicated structure that is a fabrication challenge. A defected core decagonal DCF has been proposed [7] with flattened dispersion with dispersion variation of 7 ps/(nm.km) which suffers from high confinement loss and defected core contributed scattering losses. A porous core PCF has been used for residual dispersion compensation over a wide range of E+S+L+C+U wavelength bands [8], however, the authors did not tell about confinement loss which is presumed to be high for porous fiber. The authors also did not mention about single-mode operation. Broadband DC-PCF has also been suggested [9-12] which has offered higher dispersion slope and/or confinement losses etc.

In this paper we propose a simple structure of a pure silica based index guiding PCF that gives an adequate negative dispersion of −124.0±0.1252 ps/(nm.km) in the wavelength range of 1350 to 1700 nm which is ultra-flattened in that range with a low confinement loss of less than 10^{-6} dB/km at the wavelength of 1550 nm. The calculated V-parameter at that wavelength range indicates that our proposed DC-PCF only associates with fundamental mode ensuring the fiber being immune to intermodal dispersion.

II. Fiber Structure

The design of the proposed fiber is kept as simple as possible so that fabrication becomes easier and the production of the fiber remains cost effective. Fig. 1 depicts the transverse cross-section of the proposed PCF. Material of the studied PCF is taken to be silica without any doping at the core region. The cladding is formed by nine circular rings of air holes keeping 6 air holes in the 1st ring, 13 air holes in the

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

2nd ring, 20 air holes in the 3rd ring, 26 air holes in the 4th ring, 34 air holes in the 5th ring, 40 air holes in the 6th ring, 46 air holes in the 7th ring, 52 air holes in the 8th ring and 58 air holes in the 9th ring as arranged in Fig 1. Spacing between holes of adjacent rings or pitch, Λ is adjusted along with air hole diameters of each ring to find the optimized result of ultra-flattened dispersion and ultra-low confinement loss.

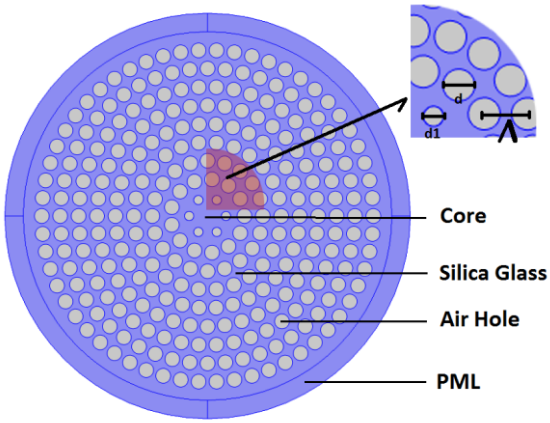


Fig. 1 Transverse cross-section of the proposed DC-PCF structure.

III. Theoretical Background

An efficient full vector model solver COMSOL Multiphysics 5.0 based on Finite Element Method (FEM) is used for analysing various properties of proposed PCF. In order to reduce the simulation window, anisotropic Perfectly Matched Layers (PML) are positioned outside the outermost ring as absorbing boundary. By solving eigenvalue problem drawn from Maxwell's equations using FEM, modal effective refractive indices n_{eff} are obtained. Once we obtain modal effective indices, we can calculate other required parameters like chromatic dispersion, confinement loss, and V-parameter using following suitable equations from their definitions.

The chromatic dispersion $D(\lambda)$ of PCFs can be easily calculated from the real part of the obtained effective refractive indices using following equation [13]

$$D(\lambda) = -\frac{\lambda}{c}\frac{d^2\,\mathrm{Re}[n_{eff}]}{d\lambda^2}\ , \tag{1}$$

where λ is operating wavelength and c is velocity of light in vacuum. After calculating dispersion, we can also calculate dispersion slope with following equation

$$S(\lambda) = \frac{dD(\lambda)}{d\lambda}\ . \tag{2}$$

The confinement loss L_c can be obtained from the imaginary part of n_{eff} as follows [13]

$$L_c = \frac{20\times10^6}{\ln(10)}k_0\,\mathrm{Im}[n_{eff}] \tag{3}$$

where $k_0 = 2\pi/\lambda$ is the wave number in free space. The effective mode area is a quantity based on modal field distribution. It was originally introduced as a measure of nonlinearities. Indeed, a low effective mode area gives a high density of power needed for nonlinear processes and a high effective mode area gives a low density of power that ignores nonlinear effects. A useful definition for effective mode area of the optical fiber is [14],

$$A_{eff} = \frac{\left(\iint |E|^2 dxdy\right)^2}{\iint |E|^4 dxdy} \tag{4}$$

where E is the amplitude of electric field. The nonlinearity coefficient (γ) of a fiber can be calculated from

$$\gamma = \frac{2\pi n_2}{\lambda A_{eff}} \tag{6}$$

where n_2 is the nonlinear index coefficient in the nonlinear part of the refractive index. The nonlinearity of a fiber can be controlled by n_2 and/or A_{eff}.

As the group velocity of light propagating in an optical communication fiber depends not only on the optical frequency, but also on the propagation mode involved, it is required that intermodal dispersion should not be occurred, requiring single modeness must be met which can be verified with the help of the effective V-parameter within the telecommunication band. V_{eff} for a circular-lattice PCF can be calculated by using the following formula [15]

$$V_{eff} = \frac{2\pi\Lambda}{\lambda}\sqrt{n_{core}^2 - n_{eff}^2} \tag{4}$$

where n_{core} can be calculated using the Sellmeier equation for silica. The condition for higher order mode cutoff is that V-parameter at any particular wavelength must be higher than $V_{\text{eff}} = \pi$ [15]. So it is required that V-parameter must be lower than π at the proposed wavelength range for the fiber operating only at fundamental mode.

IV. Results and Discussions

Electric field distribution of the proposed PCF at 1550nm wavelength is presented in Fig. 2. It depicts the confinement of field at the core region where red portion indicates the highest field intensity. Distance between holes of adjacent rings or pitch and diameters of air holes are varied to find the optimized ultra-flattened dispersion curve. Number of rings of air holes is varied to find the ultra-low confinement loss considering the structure remains immune to mechanical stress.

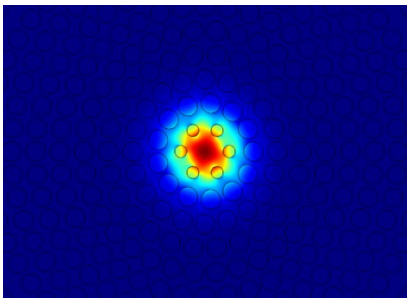


Fig. 2 Electric field distribution of the proposed DC-PCF structure at 1550nm wavelength.

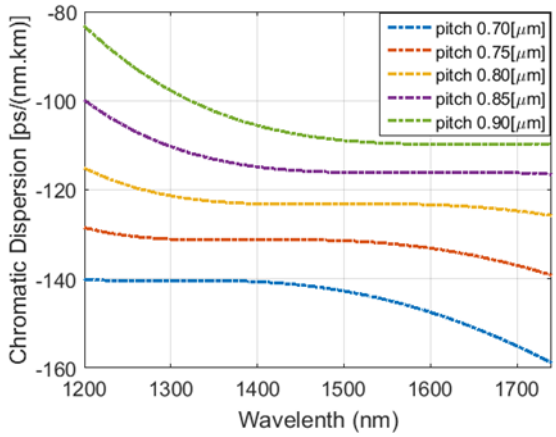


Fig. 3 Chromatic dispersion of the proposed DC-PCF versus wavelength for different pitch.

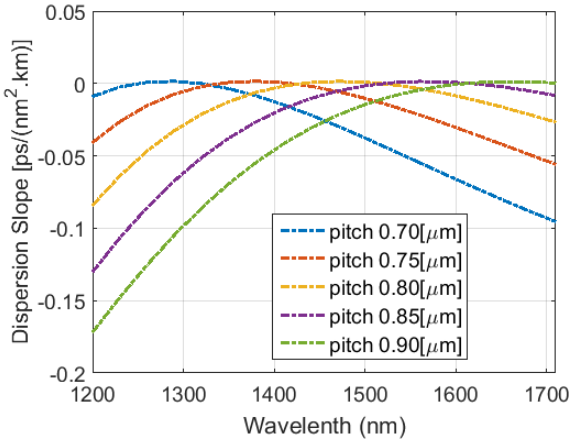


Fig. 4 Dispersion slope of the proposed PCF versus wavelength with pitch variation.

Fig. 3 and Fig. 4 demonstrate the chromatic dispersion and dispersion slope curve of the proposed PCF for different pitch. Fig. 3 depicts that chromatic dispersion is flattened from 1350 to 1700 nm wavelength with an average value of −124 ps/(nm.km) dispersion for pitch, $\Lambda = 0.8$ µm. For the same optimum pitch, very low dispersion slope around that wavelength range is achieved as revealed in Fig. 4.

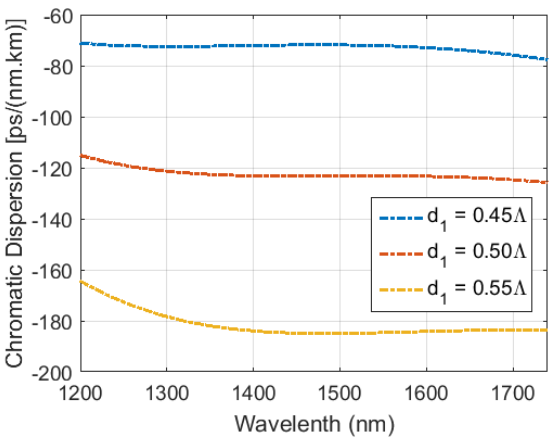


Fig. 5 Chromatic dispersion of the proposed DC-PCF versus wavelength for different diameter of holes in the 1st ring.

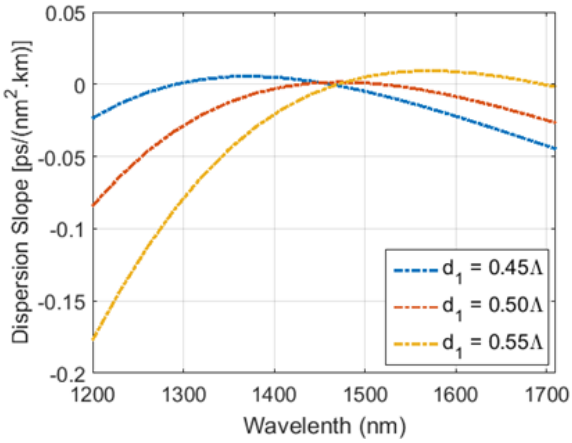


Fig. 6 Dispersion slope of the proposed PCF versus wavelength varying diameter of holes in 1st ring.

In Fig. 5 and Fig. 6, diameter of air holes in 1st ring, d_1 is varied keeping the pitch fixed at 0.8 µm; chromatic dispersion and dispersion slope is plotted against wavelength. In both figures, wide flattened dispersion and dispersion slope are achieved when $d_1 = 0.50\Lambda$ around the said wavelength range. However, the figures also indicate that more average negative dispersion could be achieved for higher value of d_1 at the cost of wavelength band.

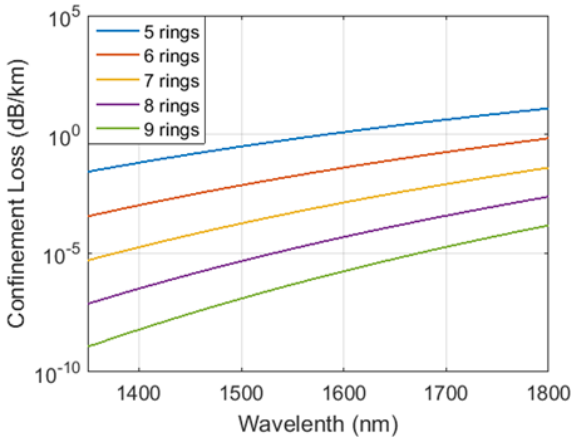


Fig. 7 Confinement loss of the proposed DC-PCF structure versus wavelength for different number of rings of air hole.

From Fig. 7, it is found that our required dispersion profile can be obtained only by using 5 rings of air holes, but confinement loss becomes considerably large. It is also found that dispersion profile remains almost unchanged adding extra rings of air holes around. So this technique can be applied to decrease the confinement loss which is quite comprehensible from the figure. It should be noted that increasing the diameter of cladding by adding more rings of air holes could make the fiber vulnerable, so 9 rings are chosen to find an optimized result. At this case, confinement loss at 1550 nm wavelength becomes 10^{-6} dB/km, which is ultra-low for dispersion compensating fibers.

Fig. 8 presents the nonlinearity of the proposed DC-PCF. We see, nonlinearity is moderate over the wavelength range of interest which could be suitable for dispersion compensation in long-haul optical fiber transmission. By increasing pitch,

nonlinearity could be further reduced. However, we consider $\Lambda = 0.8$ µm as optimum for favorable results of dispersion and confinement loss.

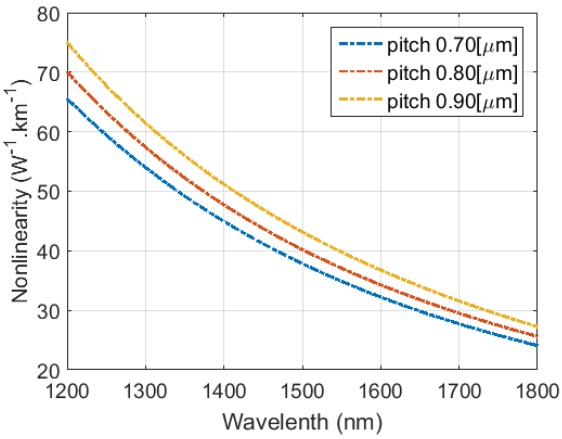


Fig. 8 Nonlinearity of DC-PCF against wavelength.

Fig. 9 depicts the V-parameters of the proposed fiber at different wavelengths for different pitch. Fiber being assisted with single mode greatly depends on core radius which is pretty small in our case, can be confirmed with the value of V-parameter being less than π for circular-lattice PCFs. From this figure, it is clear that V-parameter is less than π around the total wavelength range of 1200 to 1800 nm which confirms our fiber being a single mode fiber operating at fundamental mode only.

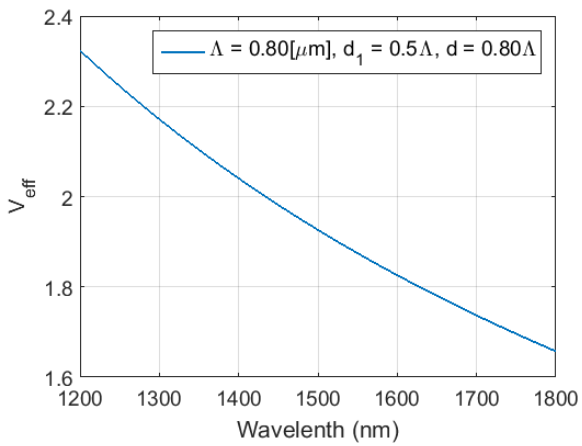


Fig. 9 V-parameter of the proposed DC- PCF structure versus wavelength.

V. Conclusions

In this paper, we have demonstrated a simple but outstanding design of PCF to flatten the chromatic dispersion curve with a significant negative value over E+S+L+C+U bands. We have obtained average dispersion of −124.0 ps/(nm.km) with a very low dispersion variation of ± 0.1252 ps/(nm.km) in the wavelength range of 1350 to 1700 nm. Dispersion slope is also flat in the range. More average negative dispersion could be achieved at the expense of operating band. More negative dispersion is available in the literature, however, the operating band is not so large and dispersion slope is not so flat over a huge band. The proposed PCF also shows confinement loss lower than 10^{-6} dB/km in that wavelength range. Moreover, comparatively few parameters are required to design the structure which makes the design very easy to fabricate. Considering all these, the proposed PCF is much suitable for dispersion compensation in WDM networks.

Future target could be attaining much more negative dispersion coefficient with low dispersion slope over the E+S+L+C+U bands.

References

[1] J. Knight, T. Birks, P. Russell and D. Atkin, "All-silica single-mode optical fiber with photonic crystal cladding: errata", Optics Letters, vol. 22, no. 7, p. 484, 1997.

[2] M.I. Hasan, R.R. Mahmud, M. Morshed and M.R. Hasan, "Ultra-flattened negative dispersion for residual dispersion compensation using soft glass equiangular spiral photonic crystal fiber", J. Modern Optics, vol. 63, no. 17, pp. 1681-1687, 2016.

[3] J.-M. Hsu, W.-H. Zheng, C.-L. Lee, J.-S. Horng, "Theoretical investigation of a dispersion compensating photonic crystal fiber with ultra-high dispersion coefficient and extremely low confinement loss." Photonics and Nanostructures, vol. 16, pp. 1–8, 2015.

[4] G. Prabhakar, A. Peer, V. Rastogi, and A. Kumar, "Large-effective area dispersion-compensating fiber design based on dual-core microstructure," Appl. Optics, vol. 52, no. 19, pp. 4505-4509, July 2013.

[5] J.-M. Hsu, C.-W. Yao and J.-Z. Chen, "Wavelength-tunable dispersion compensating photonic crystal fibers suitable for conventional/coarse wavelength division multiplexing systems," J. Lightwave Technol., vol. 33, no. 11, pp. 2240-2245, 2015.

[6] M.I. Hasan, M. Selim habib, M. Samiul Habib, and S. M. A. Razzak, "Highly nonlinear and highly birefringent dispersion compensating photonic crystal fiber," Opt. Fiber Technol., vol. 20, pp. 32-38, 2014.

[7] M. A. Islam, R. Ahmad, M. S. Ali, and K. M. Nasim, "Proposal for highly residual dispersion compensating defected core decagonal photonic crystal fiber over S+C+L+U wavelength band," Opt. Engineering, vol. 53, no. 7, paper 076106-1-5, July 2014.

[8] D.C. Tee, M. H. A. Bakar, N. Tamchek, and F. R. M. Adikan, "Photonic crystal fiber in photonic crystal fiber for residual dispersion compensation over E+S+C+L+U wavelength band," IEEE Photon. Journal, vol. 5, no. 3, June 2013.

[9] M. Selim Habib, M. Samiul Habib, S.M.A. Razzak, M. A. Hoaasin, "Proposal for highly birefringent broadband dispersion compensating octagonal photonic crystal fiber," Opt. Fiber Technol., vol. 19, pp. 461-467, 2013.

[10] M. M. Haque, M. S. Rahman, M. S. Habib, and S. M. A. Razzak, "Design and characterization of single mode circular photonic crystal fiber for broadband dispersion compensation," Optik, vol. 125, no. 11, pp. 2608-2611, 2014.

[11] F. Begum, Y. Namihira, S.M.A. Razzak, S. Kaijage, N.H. Hai, T. Kinjo, K. Miyagi, and N. Zou, "Novel broadband dispersion compensating photonic crystal fibers: applications in high-speed transmission systems," Opt. & Laser Technol., vol. 41, pp. 379-686, 2009.

[12] S. Yang, Y. Zhang, L. He, and S. Xie, "Broadband dispersion-compensating photonic crystal fiber," Opt. Lett., vol. 31, no. 19, 2006.

[13] K. Saitoh and M. Koshiba, "Leakage loss and group velocity dispersion in air-core photonic bandgap fibers", Opt. Express, vol. 11, no. 23, p. 3100, 2003.

[14] K. Saitoh, M. Koshiba, "Numerical modeling of photonic crystal fibers," J. Lightwave Technol., vol. 23, no. 11, pp. 3580–3590, 2005.

[15] N. Mortensen, J. Folkenberg, M. Nielsen and K. Hansen, "Modal cutoff and the V parameter in photonic crystal fibers", Optics Letters, vol. 28, no. 20, p. 1879, 2003.

Data Adaptive Filtering Approach to Improve the Classification Accuracy of Motor Imagery for BCI

Sanjoy Kumar Saha[1,*] and Md. Sujan Ali[2, 3]

[1]Department of Computer Science and Engineering, Begum Rokeya University
Rangpur, Bangladesh

[2]Signal Processing and Computational Neuroscience Lab, Department of Computer Science and Engineering
University of Rajshahi, Rajshahi, Bangladesh

[3]Department of Computer Science and Engineering, Jatiya Kabi Kazi Nazrul Islam University
Mymensingh, Bangladesh

[*]csesanjoy@gmail.com

Abstract—**This paper presents a data adaptive pre-filtering method to enhance the motor imagery (MI) classification for brain computer interface (BCI) implementation electroencephalography (EEG). The filtering is implemented using multivariate empirical mode decomposition (MEMD). The MEMD decomposes multivariate EEG signal into a finite set of time varying basis functions called intrinsic mode functions (IMFs). Several IMFs are suppressed to improve the intelligibility of the EEG for better classification. The features from pre-filtered EEG are extracted by using common spatial pattern (CSP) which is a spatial filtering approach. The two class MI classification is performed by support vector machine (SVM). The proposed method is evaluated by BCI competition III of dataset IVa. Its performance is higher compared to that of the recently developed algorithm.**

Index Terms—**Brain computer interfaces (BCI), electroencephalogram (EEG), multivariate empirical mode decomposition (MEMD), common spatial pattern (CSP), support vector machine (SVM).**

I. Introduction

A Brain-Computer Interfaces (BCI) which provides direct communication between human and computer based on human neurophysiological signals without depending on any muscle or peripheral nervous system activities [1]. Recently BCI based researches have been drown attention in the field of brain science, neural engineering and rehabilitation [1]. In particular, motor imaginary (MI) based BCI is the imagination of a motor action without any real limbs movements, has practical significance in the case of lock in and spinal cord injured patients to communicate with their family, doctor or others.A class of electroencephalogram (EEG)-based BCIs which rely on MI of users are of particular interest to the BCI community since this type of BCI has a relatively robust communication performance, detectable on the scalp non-invasively with affordable equipment[2].

Different research shows that spatial distribution of EEG signals is different for different imagined movements [3]. For successful BCI implementation, it is necessary to identify the features of EEG signals and then convert them into control commands. It has been shown in [4] that motor imagery could cause the energy of alpha (7-12Hz) and beta (13-30Hz) rhythms increased or decreased, which has been known as event-related synchronization (ERS) or desynchronization (ERD). Nevertheless, the truth is the brain does not produce tones with fixed frequencies and additionally EEG come with noise due to several causes like bad electrode location, electrode impedances, etc [5]. Although many methods have been reported for classification but a MI-based BCI system is not satisfactory due to the lack of classification accuracy. To tackle this problem Long Zhang et al. in [6] proposed different methods, where the EEG signal is decomposed into IMF using MEMD, and classified by CSP. Among different methods in [6] no single method showed better result over the entire subjects. The goal of the study is to improve the classification accuracy compared with reported methods in Long Zhang et al. in [6].

The proposed method focuses on the study of features extraction using data adaptive methods with common spatialpattern (CSP).EEG signals give an unclear image of brain activity because of the volume conduction. For this reason, CSP is used for discriminating between different classes of EEG activities based on their spatial configuration across electrodes.Common pattern (CSP) has been widely used in spatial filtering because of its efficiency and to improve BCI's performance [7] [8]. Empirical mode decomposition (EMD) is a fully data adaptive technique for analyzing nonlinear and non stationary signals [9]. The specialty of the method is that it decomposes the signal into intrinsic mode functions (IMFs) and the signal is reconstructed simply by summing up the bases with a negligible error. Recently it is reported in [10], its mode mixing is evident between intrinsic mode functions (IMFs) or a single mode is "leaked" into two IMFs. Bivariate and trivariate EMD (BEMD and TEMD) are also extension of EMD whereas performance depends on selecting variables in the case of multivariate data. By considering this problem, N. Rehman et al. [10] proposed multivariate EMD, is extension of standard EMD for analyzing multivariate data. The novelty of the method is that it reduces the mode mixing problem of the basis functions by taking signal projections along different directions in multi-dimensional spaces. The performance of CSP can be improved by decomposing EEG signal into IMFs and summing up optimal IMFs between 7 to 30 Hz due to smoothing CSP spatial filter with meaningful information. In our study, MEMD based common spatial pattern (MEMD-CSP) is presented to improve the performance of MI-BCI. The IMFs obtain using MEMD omitted based on examination. First IMF and last two IMFs are omitted because of, first IMF contains dominant frequency greater than 30 Hz and last two IMFs are residue.

After reconstructing signal, feature vector is obtained by CSP is classified using support vector machine (SVM). The

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

SVM utilizes a nonlinear map to translation the row training data into a higher dimension. Inside this new dimension, it seeks out for the linear optimal separating hyper-plane MI data from two classes can always be separated via a hyper-plane [11].

II. MULTIVARIATE EMPIRICAL MODE DECOMPOSITION (MEMD)

The Empirical Mode Decomposition (EMD) algorithm is a fully data-driven method designed for multi-scale decomposition and time-frequency analysis of real-world signals. Despite empirical mode decomposition (EMD) becoming a standard for time-frequency analysis of nonlinear and non-stationary signals, but it necessary to extend EMD for direct multichannel EEG data analysis. In EMD IMFs comes from different channel having a different number of IMFs, means they do not have the same frequency. To solve above problem multivariate EMD proposed by Rehman and Mandic [10] that a generic extension. Basic EMD find local mean by computing the average of upper and lower envelops. Local mean of n-dimensional in MEMD method signals cannot be computed directly. Multiple n-dimensional envelopes are generated by projecting the signal along different directions in n-variant spaces [10]; those projections are then averaged to acquire the local mean. This step is complex to perform due to the lack of formal definition of maxima and minima in *n*-dimensional domains in general EMD. The sampling based on low discrepancy Hammersley sequence is used to generate projections of input signal in [10]. Once the projections along different directions in multidimensional spaces are obtained, their extrema are interpolated via cubic spline interpolation to obtain multiple-signal envelopes. Thus, obtained envelopes are then averaged to obtain the local mean of the multivariate signal. The following algorithm proposed in [10] is employed here to decompose signal $x(t)$into a set of IMF components.

I. Generate the point-set-based on the Hammersley sequence for sampling on an *(n-1)*-sphere.
II. Calculate a projection, denoted by $\{P^{\theta_k}(t)\}_{t=1}^{T}$(here,$\{\theta_1, \theta_2, \dots, \theta_{n-1}\}$be a set of the *(n-1)* angular coordinates), of the input signal $\{x(t)\}_{t=1}^{T}$, along the direction vector X^{θ_k}, for all *k* (the whole set of direction vectors), giving $\{P^{\theta_k}(t)\}_{k=1}^{K}$ as the set of projections.
III. Find the time instants $\{t_i^{\theta_k}\}_{k=1}^{K}$corresponding to the maxima of the set of projected signals$\{P^{\theta_k}(t)\}_{k=1}^{K}$.
IV. Interpolate$[t_i^{\theta_k}, s(t_i^{\theta_k})]$, for all values of *k*, to obtain multivariate envelope curves $\{e^{\theta_k}(t)\}_{k=1}^{k}$.
V. For a set of *K* direction vectors, calculate the mean $\mu(t)$ of the envelope curves as

$$\mu(t) = \frac{1}{k}\sum_{k=1}^{K} e^{\theta_k}(t).$$

VI. Extract the "detail"$\mathrm{d(t)}$using $\mathrm{d(t) = X(t) - \mu(t)}$. If the "detail"$\mathrm{d(t)}$ fulfills the stoppage criterion for a multivariate IMF, apply the above procedure to $\mathrm{X(t) - \mu(t)}$otherwise apply it to$\mathrm{d(t)}$.

Basically, MEMD most significantly solve uniqueness problem associated with single channel EMD when processing multichannel data such as EEG signal.

III. COMMON SPATIAL PATTERNS

Firstly, H. Ramoser used common spatial patterns (CSP) technique to classify multi-channel EEG during imagined hand movements [12].The main idea is to use a linear transform to project the multi-channel EEG data into low dimensional spatial subspace with a projection matrix of which each row consist of weights for channel. This transformation can maximize the variance of two class signal. CSP method is based on the simultaneous diagonalization of the covariance matrices of both classes.

Details of the algorithm are described as follows with the example of classifying EEG during hand and foot movements. Here,$X_H$andX_F denote EEG matrices during single trail hand and foot imagination with dimensions N×T, where N is the number of channels and T is the number of samples per channel. The normalized spatial covariance of the EEG can be calculated from:

$$R_H = \frac{X_H X_H^T}{trace(X_H X_H^T)} R_F = \frac{X_F X_F^T}{trace(X_F X_F^T)} \quad (1)$$

$\mathrm{X^T}$is the transpose of X and $\mathrm{trace(X)}$computes the sum of the diagonal elements of X. The averaged normalized covariance $\overline{\mathrm{R_H}}$ and $\overline{\mathrm{R_F}}$are calculated by averaging over all the trails of each group. The composite spatial covariance can be factorized as

$$R = \overline{R_H} + \overline{R_F} = U_0 \sum U_0^T \quad (2)$$

where$U_0$the matrix of eigenvectors and $\sum$ is the diagonal matrix of eigenvalues. The whitening transformation matrix

$$P = \sum^{-1/2} U_0^T \quad (3)$$

Transforms the average covariance matrices as

$$S_H = P\overline{R_H}P^T \; S_H = P\overline{R_H}P^T \quad (4)$$

S_H and S_F share common eigenvectors and the sum of corresponding eigenvalues for the two matrices will always be the identity matrix,

$$S_H = U\sum{}_H U^T \; S_F = U\sum{}_F U^T \quad \sum{}_H + \sum{}_F = I \quad (5)$$

The eigenvectors with the largest eigenvalues for S_H have the smallest eigenvalues for S_F and vice versa. The transformation of whitened EEG onto the eigenvectors corresponding to the largest eigenvalues in $\sum_{\mathrm{H}}$ and $\sum_{\mathrm{F}}$ is optimal for separating variance in two signal matrices. The projection matrix $W$is denoted as

$$W = U^T P \quad (6)$$

With the projection matrix W, the original EEG can be transformed into uncorrelated components

$$Z = WX \quad (7)$$

Zcan be seen as EEG source components including common and specific components of different tasks. The original EEG Xcan be reconstructed by

$$X = W^{-1} Z \quad (8)$$

where$W^{-1}$is the inverse matrix of Wand the columns of W^{-1} are spatial patterns which can be considered as EEG source distribution vectors. The first and last columns of W^{-1} are the most important spatial patterns that explain the largest variance of one task and the smallest variance of the other.

IV. DATA AND METHOD

A. Data Acquisition

The EEG data were collected from the BCI Competition III [13] dataset IVa. This data set was recorded from five healthy subjects (labeled aa, al, av, aw, ay). Each subject is performed right hand (class 1) and right foot (class 2) motor imaginary task. The subjects sat in a comfortable chair with arms resting on armrests. The data is recoded from 118 electrodes

according to the international 10/20 system with sampling frequency 100 Hz using BrainAmp MR plus amplifiers and an Ag/AgCl electrode cap. There were 280 trails for each subject, where 140 trails for each task. In each trail, visual cues are displayed and subjects imagined one of the two possible tasks for 3.5 seconds.

B. Proposed Algorithm

1) Channel selection: It is known that, center parts of the brain are most active during motor imageries [14]. In this experiment, we selected 13 channels "FC3," "FC4," "Cz," "C1," "C2," "C3," "C4," "C5," "C6," "T7," "T8," "CP3," and "CP4" according to the 10–20 system. We can call channel for short 34, 38, 54, 53, 55, 52, 56, 51, 57, 50, 58, 70, and 74.
2) Basic filtering: Filtering data can help remove high-frequency artifacts and low frequency floats. In this experiment, we applied a fifth-order Butterworth band-pass filter at first to each channel, which the pass-bands for each subject were 7–30 Hz. Therefore, we choose the alpha (7–12 Hz) and beta (13–30 Hz) bands, which are more effective during motor imagery task.
3) Filtering with MEMD: To decompose multi-channel EEG signal into a set of IMFs, we used MEMD. Now, we have to select IMFs, which will increase, our accuracy of classification. It is observed that first IMF carries noise and component that have frequency higher than 30Hz which carries the maximum energy among all frequencies found in the spectrum and last two IMFs are residue. Therefore, we reject first IMF and last two IMFs after decomposing the original signal $x(t)$ using MEMD. Now each channel has the same number of IMF index$c_j(t)$. In our new method, we remove first IMF and last two IMFs, thus a new signal $\hat{x}(t)$is reconstructed by

$$\hat{x}(t) = \sum_{j=2}^{\Omega-2} c_j(t)$$

where$\Omega$is the total number of IMF, which is same in each channel and *t,* represents time. Now, reconstructed signal $\hat{x}(t)$ carry more useful information than the original signal.
4) Feature Extraction & Classification: There have many techniques for feature extraction, among them we select common spatial filter (CSP) for this work. Output of CSP is a projection matrix W, the original EEG X can be transformed into uncorrelated components Z.

$$Z = WX$$

We choose an optimal subset of features, to improve classification performance. We choose the first row of Z as a feature. For classification, we used support vector machines with radial basis function (rbf) kernels and rbf sigma (σ)=10, through 10-fold cross-validation.

V. Results And Discussion

In existing method [6], EMD and MEMD are used to analyze motor imagery response and CSP is used as a classifier. In another method [6] laplacian derivativeis applied before EMD, MEMD or CSP by referencing an electrode to the mean of its four nearest neighboring electrodes. But all those methods have subject's dependency. Our proposed method increased averaged classification accuracy along with five subject and reduced subject dependency and relatively stable than method at Long Zhang et al. Our processing framework serves to design more reliable MI-BCls, which is of great interest for speeding up communication in real-life applications.

Table I shows that classification accuracy obtained by using proposed method. From TABLE I, it is clear that proposed algorithm achieves a better classification performance both individually and overall to subjects. The aim of motor imagery classification experiments is to identify IMFs which are bearing significant relevant data and that is achievable using MEMD for time-frequency analysis. Significant improvements for two class motor imagery dataset indicate that, we successfully extract the alpha and beta rhythms of interest. Cross validation is a model evaluation technique to estimating how well the model learned from some training data and what is going to perform on unseen data. In our experiments we used k-fold cross validation, which works by dividing the training set into k parts of equal size, then using (k-1) parts as a training set and check the classification rate on the one remaining part. The performance of proposed method is evaluated with SVM classifier through a 10-fold cross validation method.

TABLE II
Performance comparisons using proposed method

Method	Classification accuracy rate (%)					
	aa	al	av	aw	ay	Avg
Proposed method	**82.0**	**92.2**	**69.4**	**83.8**	**89**	**83.3**
CSP1 [6]	71.3	88.4	48.6	89.9	79.9	75.6
CSP2 [6]	65.3	90.2	63.7	80.3	87.3	77.4
EMD-CSP [6]	68.4	89.6	64.1	82.5	86.9	78.3
MEMD1-CSP [6]	68.8	90.0	68.8	76.3	87.5	78.3
MEMD2-CSP [6]	60.3	82.9	55.3	60.7	74.0	66.4
CSP1 with laplacian derivative [6]	70.7	78.3	48.7	88.5	72.6	71.8
CSP2 with laplacian derivative [6]	63.7	87.7	64.0	79.1	86.3	76.2
EMD-CSP with laplacian derivative [6]	62.8	89.5	64.4	79.6	85.9	76.4
MEMD1-CSP with laplacian derivative [6]	63.8	95.0	63.8	71.2	88.1	76.4
MEMD2-CSP with laplacian derivative [6]	55.3	79.7	54.7	70.0	75.6	67.1

VI. Conclusions

In this experiment, we focus on realization of a brain computer interface for the paralyzed people to control assistive environmental devices. For this purpose, motor imagery tasks are classified using different signal processing methods. For filtering frequencies which are not related with motor imagery by basic band-pass filter and multivariate empirical mode decomposition (MEMD) to enhance the EEG signal. Unlike the standard single channel EMD algorithm, MEMD allowed a more stable estimation of the time-varying frequency response from multichannel EEG and preserve the cross cannel interdependence. CSP is used as discriminating tool that helps to differentiate different mental status. In this experiment, we use support vector machine (SVM) as a classifier with a non-linear rbf kernels. This study developed a promising method based on MEMD-CSP aid with SVM that shows better performance in MI based classification.

ACKNOWLEDGMENT

This work is supported by ICT fellowship-2015-16, information and communication technology division, govt. of the people's republic of Bangladesh.

REFERENCES

[1] M. Grosse-Wentrup, C. Liefhold, K. Gramann, and M. Buss, *Beam forming in noninvasive brain-computer interfaces*, IEEE Trans.Biomed. Eng., vol. 56, no. 4, pp. 1209–1219, Apr. 2009.

[2] G. Pfurtscheller, C. Brunner, A. Schlogl, and F. Lopes da Silva, "*Mu rhythm (de) synchronization and EEG single-trial classifacation of different motor imagery tasks,"*Neuroimage, vol. 31, no. 1, pp. 153-159, 2006.

[3] Le Quoc Thang, ChivalaiTemiyasathit, Increase performance of four-class classification for Motor-Imagery based Brain-Computer Interface.

[4] BAI Xiaoping, WANG Xiangzhou, ZHENG Shuhua, YU Mingxin, *The Offline Feature Extraction of Four-class Motor Imagery EEG Based on ICA and Wavelet-CSP*, Proceedings of the 33rd Chinese Control ConferenceJuly 28-30, 2014, Nanjing, China.

[5] M Agustina Garcés Correa and Eric LaciarLeber. Noise removal from eeg signals in polisomnographic records applying adaptive filters in cascade. 2011.

[6] Long Zhang, Cheng Zhang, Hiroshi Higashi, Jianting Cao, Toshihisa Tanaka, *Common Spatial Pattern Using Multivariate EMD for EEG classification*,APSIPA ASC 2011 Xi'an.

[7] J. M¨uller-Gerking, G. Pfurtscheller, and H. Flyvbjerg, Designing *optimal spatial filters for single-trial EEG classification in a movement task*, Clinical Neurophysiology, vol. 110, pp. 787–798, 1999.

[8] H. Ramoser, J. M¨uller-Gerking, and G. Pfurtscheller, Optimal spatial filtering of single trial EEG during imagined hand movement, IEEETrans. Rehabil. Eng., vol. 8, no. 4, pp. 441–446, 2000.

[9] Cheolsoo Park, David Looney, Naveed urRehman, AlirezaAhrabian, and Danilo P. Mandic*, Classification of Motor Imagery BCI Using Multivariate Empirical Mode Decomposition*, IEEE Trans. on Neural Systems andRehabilitation Eng, vo. 21, no. 1, January 2013.

[10] N. Rehman and D. P. Mandic, *Multivariate empirical mode decomposition*,Proc. R. Soc. A, vol. 466, no. 2117, pp. 1291–1302, 2010.

[11] H. Drucker, C. J. C. Burges, L. Kaufman, A. J. Smola, and V. Vapnik*,Support vector regression machines*, M. C. Mozer, M. I. Jordan, and T.Petsche, Eds. Cambridge, MA: MIT Press, 1997.

[12] H. Ramoser,J. M.Gerking and G. pfurtscheller, *Optimal spatial filtering of single trail EEG during imagined hand movements*, IEEE Trans. Rehab. Eng.,vol. 8, no. 4, pp.441-446,2000.

[13] BCI Competition III [Online]. Available: http://www.bbci.de/competition/iii

[14] C. Neuper, G. R. M¨uller-Putz, R. Scherer, and G. Pfurtscheller, *Motor imagery and EEG-based control of spelling devices and neuroprostheses*,Progress in Brain Research, vol. 159, pp. 393–409, 2006.

Porous Core Photonic Crystal Fiber based Chemical Sensor

Bikash Kumar Paul, Md. Shadidul Islam, Sawrab Chowdhury, Sayed Asaduzzaman* and Kawsar Ahmed

Department of Information and Communication Technology (ICT),
Mawlana Bhashani Science and Technology University (MBSTU)
Tangail-1902,Bangladesh
*sayed.asaduzzaman.bd@ieee.org

Abstract—**In this paper, a porous core photonic crystal fiber based chemical sensor has been proposed. The cladding of the proposed PCF is hybrid where inner three layers are hexagonal and the outer two layers are circular. The proposed PCF shows higher relative sensitivity and low confinement loss simultaneously. For the investigation of the proposed photonic crystal fiber a full vectorial finite element method (FEM) was used. Geometric parameters like diameters of the holes and pitch are varied to optimize the Proposed PCF. The proposed PCF can be used for detecting the lower index chemical like ethanol which is industrially valuable. Moreover, the proposed PCF can sense the chemicals in a wide range of wavelength from 0.8 µm to 1.9 µm.**

Index Terms— **Relative sensitivity; Confinement loss; lower indexed chemical sensor; Finite element method; Crystal fiber.**

I. Introduction

Recently, many outstanding inventions and discoveries are done by the "age of technology". Technology always opening different era, in different application areas. In telecommunication sector optical fiber brings a revolutionary change which ensures us faster communication of information. Microstructure optical fibers (MOFs) are fibers which consist of a continuously cylindrical air hole running along the whole length of the fiber. Photonic crystal fiber can be categories into two groups; one is a photonic bandgap (PBG) and another one high index guiding fibers.

For the purpose of optical fiber communication information is converted into light and that light propagates through the medium. Optical fiber communication follows well known mechanism named Total Internal Reflection (TIR) and indexed guiding mechanism for propagating the light. Optical fiber technology is well known as a communication medium due to its superb properties like fastest information transmission wide bandwidth etc. Nevertheless, it is not bounded as a communication area. It also plays a vital role in the area of sensing. It is used as chemical sensing [1, 2] and bio-sensing [3, 4] too. As a sensor, photonic crystal fiber contributes a remarkable contribution in the area of bio-sensing which sense HIV viral load measurement [5] and cancer cell detection [6]. PCF is also convenient for temperature and pressure sensing [7]. High sensitivity, flexibility, robustness, tiny size, cost effectively altogether provided by optical fiber based sensor.

Silica glasses are widely used as a background material for photonic crystal fiber, but recently polymer is also used to manufacture these [8].For geometric structural diversity we found different PCF design architecture. It consist a number of geometric parameters and some of them are liable for changing the guiding properties [9] .But for excellent sensing application photonic crystal fiber must two fundamental parameters which are sensitivity and confinement loss reported in previous literature [10]. Diverse geometric compositions are found like circular, honey comb cladding [11], elliptical [12], hexagonal [13], octagonal [14], and decagonal [15] and many more.

Researchers in the area of optical fiber, are relentlessly working to improve the performance of optical fiber till now. In 2013 Park et al. [16] proposed a photonic crystal fiber which enhancing chemical sensing capability. In the same year Zhang et al. [17] proposed a PCF for gas sensing applications. It represents a high (about 93%) air filling ratio in the core region. As a result of the high air filling ratio in the core region confinement losses are minimized and relative sensitivity increases. In 2014 Zhixin Tan et al. [18] proposed a PCF by filling liquid in a hollow core which improving the sensitivity of fiber.

In this paper, we propose an optimized structured PCF which is superior than the PCF proposed by Sayed Asaduzzaman et al. [10] We investigated on our proposed and gets better sensitivity when we use two layer porous core than three layers .The core is filled with the chemical (Ethanol whose refractive index = 1.354). There is no hindrance to fabricate our proposed PCF due to its structural simplicity and superb modern technology. Such types of fiber can be easily fabricated by sol-gel technique.

II. Geometrical Structure of the Proposed PCF

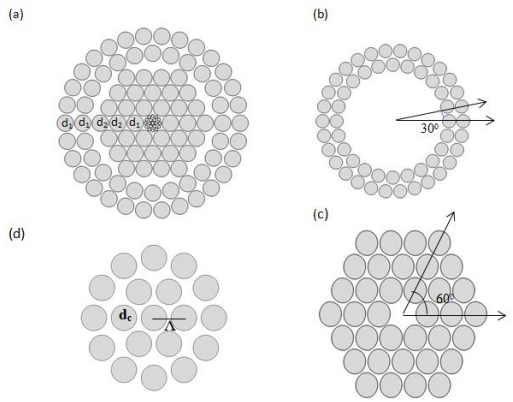


Fig. 1 (a) Transverse cross section of the proposed PCF structure (b) Outermost two layers (c) Innermost three layers (d) Porous core.

Figure 1(a) represents geometries of the transverse cross-section of the proposed PCF structure having five air-hole rings in the cladding region. Among the cladding layers of proposed PCF has two circular and three hexagonal air-hole

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

rings. Figure 1(b) and 1(c) shows the orientation of the cladding of hexagonal shape and circular shape with sector angles. The angular movement of two adjacent air holes of circular ring is 30º and hexagonal ring is 60º. The diameter of the two outermost circular rings and diameter of innermost ring is specified by the symbol d_1 whereas d_2 is used as the hexagonal second and third air-hole rings. Figure 1(d) shows the geometries structure of porous core which contains two air-hole rings. The distance between air-holes on the adjacent ring of the porous core is called the pitch which is denoted by Λ.

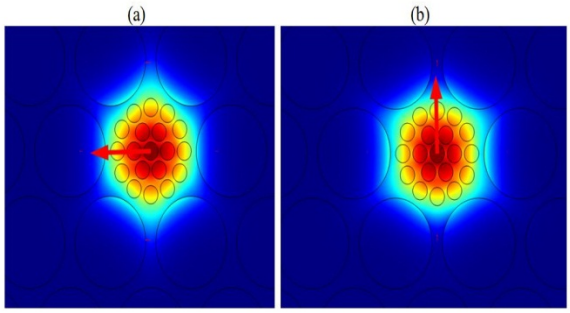


Fig. 2 Mode field distribution of the fundamental mode for (a) x polarization (b) y polarization.

III. Synopsis of Numerical Method

The Finite Element Method (FEM) has been used for electromagnetic simulation of the proposed chemical sensor PCF structure and solving Maxwell's equation due to its proven reliability. Moreover anisotropic perfectly match layer (APML) is employed here. The overall numerical simulation procedure was finished by FEM based software COMSOL Multiphysics version 4.2.

Confinement loss and relative sensitivity are the most important characteristics of the photonic crystal fiber.

The confinement loss L_c is determined through the imaginary part of effective mode index n_{eff} as follows

$$L_c = 8.686 k_0 \, I_m[n_{eff}] \qquad (1)$$

Where, $Im[n_{eff}]$ is the imaginary part of the effective mode index reported in [1,2,10]. The Coefficient r is defined as relative sensitivity it measures the interaction between light and the material to be sensed and it can be calculated through the following equation (2).

$$r = \frac{n_r}{n_{eff}} f \qquad (2)$$

Where, n_r represents the refractive index of sensed material within the air holes and n_{eff} is modal effective index. Here, f is the percentage ratio of the air hole power and the total power carried by that mode and expressed as

$$f = \frac{\int_{sample} Re(E_x H_y - E_y H_x)\, dx\, dy}{\int_{total} Re(E_x H_y - E_y H_x)\, dx\, dy} \times 100 \quad (3)$$

Where, E_x and H_x are transverse electric field and magnetic field, E_y and H_y are longitudinal electric field and magnetic field. The mode field pattern E_x, H_x, E_y, H_y are calculated by solving Maxwell's equation using finite element method (FEM). Refractive index of the background material (silica) finds out by the Sellmeiers equation.

IV. Result and Discussion

In this section, two major guiding properties of the proposed Photonic Crystal Fiber like the relative sensitivity and confinement loss are calculated according to the wavelengths. The Geometrical properties like as inner layer air hole diameters, outer layer air hole diameters, pitches of core and cladding are analysed to fix the optimized structure.

The fundamental mode for x-polarization and y-polarization of the proposed PCF has been shown in the Figure 2 (a) and (b). Here it is incontestable that the electrical field distribution is tightly confined at the core.

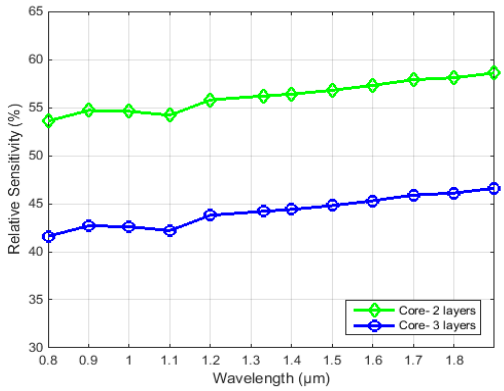


Fig. 3 Comparison of the relative sensitivity versus wavelength (μm) of the proposed PCF with 2 and 3 layers porous core for the optimized parameters.

Figure 3 illustrates the effect of 2-layers and 3-layers porous core on relative sensitivity of the proposed PCF. From the figure it is clear that for 2-layers porous core shows higher sensitivity than the 3-layers porous core. The relative sensitivity increases a little bit with the wavelength but a distinction occurs at wavelength 1.1μm.

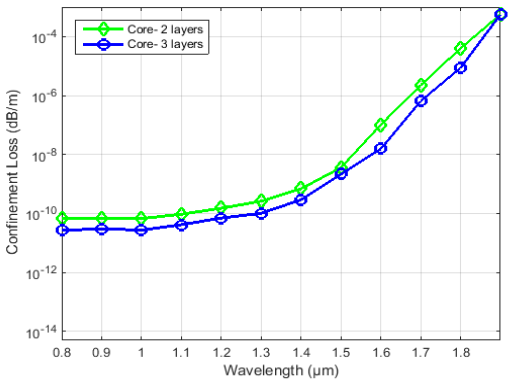


Fig. 4 Comparison of the Confinement loss versus wavelength of the proposed-PCF for 2 and 3 layers porous core when optimized parameters were constant.

From figure 4, it is obvious that proposed-PCF with 2-layers porous core show a little high Confinement than the 3-layers porous core. This figure also shows that Confinement loss is almost same at the wavelength 1μm and gradually increased at the wavelength 1.3 μm when it dramatically increases from the wavelength 1.4μm. The proposed-PCF for 3-layers porous core shows lower Confinement loss (dB/m) than 2-layers porous core, but our aim is to increase relative sensitivity. Since proposed-PCF for 2-layers porous core shows higher relative sensitivity, proposed-PCF for 2-layers porous core is chosen.

Figure 5, clearly illustrates the effect of various cladding diameter d_1. Relative sensitivity is high for d_1 = 1.50μm than

d_1 = 1.45µm and 1.40 µm. This figure also shows that relative sensitivity is decreased at the wavelength 1.1 µm then it's gradually increased. So d_1 = 1.50µm. is chosen as the optimal value.

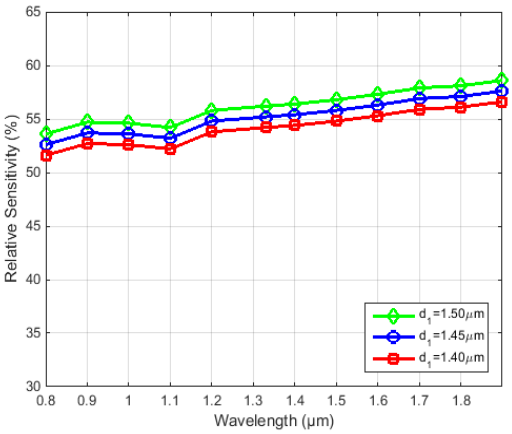


Fig. 5 Comparison of the relative sensitivity versus wavelength of the proposed 2-layers PCF for d_1= 1.50µm, 1.45µm and 1.40µm when other parameters were kept constant.

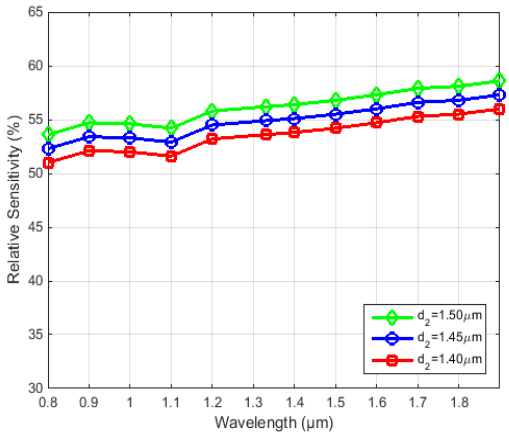


Fig. 6 Comparison of the relative sensitivity versus wavelength of the proposed 2-layers PCF for d_2= 1.50 µm, 1.45µm and 1.40µm when other parameters were kept constant.

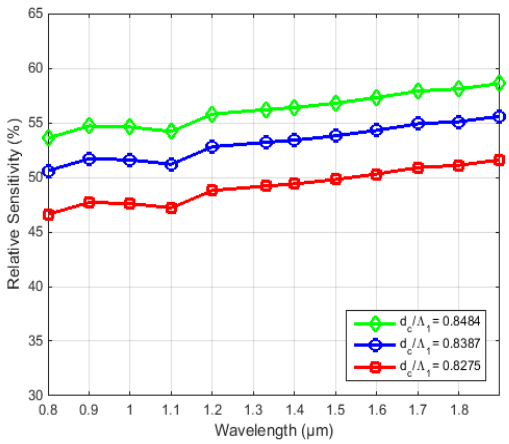


Fig. 7 Comparison of the relative sensitivity (%) versus wavelength (µm) of the proposed-PCF with 2-layers porous core for different values of air filling ratio (d_c/Λ) of core when other parameters are kept constant.

Figure 6 visualizes that the various diameter of cladding shows different relative sensitivity. For different values of d_2 = 1.50µm, 1.45µm and 1.40µm the proposed PCF shows relative sensitivity of 57.03%, 54.88% and 52.67% respectively. So of d_2 =1.50µm is chosen as the optimal value of inner circular air hole at two middle Hexagonal layers in cladding.

The effect of varying air filling ratio of core (d_c/Λ_1) on relative sensitivity has been shown in Figure 7. When d_c/Λ_1= 0.8484, the proposed-PCF shows higher relative sensitivity of 57.03%. Whereas relative sensitivity are 53.77% and 48.60% for the air filling ratio d_c/Λ_1 = 0.8387 and d_c/Λ_1= 0.8275. All air filling ratios are calculated at the wavelength 1.33 µm. So here proposed air filling ratio d_c/Λ_1= 0.8484 for better relative sensitivity.

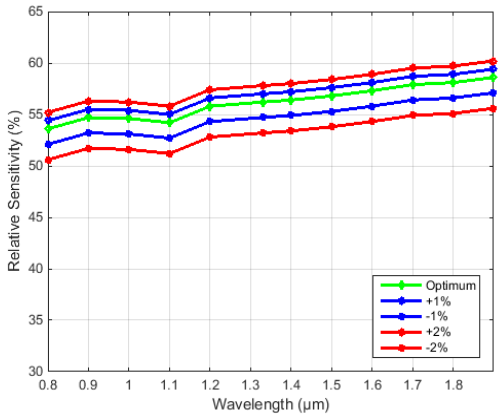


Fig. 8 Comparison of the relative sensitivity (%) versus wavelength (µm) and changing the parameters (Λ, d_1, Λ_1, and d_c) +1%, -1%, +2%, -2% with optimum value (Λ =2.4 µm, d_1=1.75 µm, Λ_1=0.7 µm, and d_c = 0.45µm).

During the fabrication process ±1% to ±2% global diameters of the holes may be varied. We have investigated the changing behaviour of sensitivity on global parameters change in Figure 8. Filling the core with the analytes is a challenging criterion. By the modern technology we can fill the core with the chemical to be sensed [1, 10]. Besides our proposed PCF is hybrid cladding with porous core. The proposed PCF can be fabricated by the modern fabrication techniques [19].

Table I. shows the comparison of proposed PCF and prior PCFs. From the figure it can be demonstrated that the proposed PCF not only shows higher sensitivity for Ethanol but also shows lower confinement loss compared to prior PCF.

TABLE I. COMAPRISION OF SIMUTATED RESULT AND STRUCTURE SHAPE AMONG PROPOSED PCF AND PRIOR PCFS AT λ=1.33 µm.

PCFs	**Sen. (r%)**	**Con. Loss (dB/m)**	**No. of Ring**	**Structural Shape**	
				Core	**Cladding**
Prior PCF_1 [2]	*31.26%*	*8.75×10^{-9}*	*3*	*Elliptical*	*Hybrid*
Prior PCF_2 [10]	*40.32%*	*7.55×10^{-7}*	*2*	*Elliptical*	*circular*
Proposed PCF	*57.02%*	*3.71×10^{-10}*	*5*	*Porous*	*Hybrid*

V. CONCLUSION

We have investigated relative sensitivity and confinement loss for proposed porous core photonic crystal fiber. The investigation is done by varying the wavelength. We use ethanol as an analytes, none the less it can redact different chemical like (Methanol, Butanol, Propanol, Pentanol etc). We have changed various parameters of the proposed PCF and measure confinement loss and relative sensitivity. After

varying such parameters we found an optimized design which provided better sensitivity and reduced confinement loss. Our proposed PCFs sensitivity is better than previous proposed PCF [10].

ACKNOWLEDGMENT

The authors are grateful to those who participated in this research work.

REFERENCES

[1] S. Asaduzzaman, K. Ahmed, T. Bhuiyan, and T. Farah, “Hybrid photonic crystal fiber in chemical sensing,” *SpringerPlus*, vol. 5, no. 1, pp. 1–11, 2016.

[2] S. Asaduzzaman, M. Kawsar Ahmed, F. H. Arif, and M. Morshed, “Application of Microarray-core Based Modified Photonic Crystal Fiber in Chemical Sensing.”, *International Conference on Electrical & Electronic Engineering (ICEEE)*, pp. 41-44. IEEE, 2015.

[3] X. Yu, D. Yong, and Y. Zhang, “Photonic Crystal Fiber-Based Biosensors,” in *Handbook of Photonics for Biomedical Engineering*, A. H.-P. Ho, D. Kim, and M. G. Somekh, Eds. Springer Netherlands, 2013, pp. 1–22.

[4] L. Rindorf, J. B. Jensen, M. Dufva, L. H. Pedersen, P. E. Høiby, and O. Bang, "Photonic crystal fiber long-period gratings for biochemical sensing," *Optics Express*, vol. 14, no. 18, p. 8224, 2006.

[5] H. Shafiee, E. A. Lidstone, M. Jahangir, F. Inci, E. Hanhauser, T. J. Henrich, D. R. Kuritzkes, B. T. Cunningham, U. Demirci, "Nanostructured optical Photonic crystal Biosensor for HIV viral load measurement," *Scientific Reports*, vol. 4, 2014.

[6] S. Padmanabhan, V. K. Shinoj, V. M. Murukeshan, and P. Padmanabhan, "Highly sensitive optical detection of specific protein in breast cancer cells using microstructured fiber in extremely low sample volume," *Journal of Biomedical Optics*, vol. 15, no. 1, p. 017005, 2010.

[7] S. Padidar, V. Ahmadi, and M. Ebnali-Heidari, "Design of high sensitive pressure and temperature sensor using Photonic crystal fiber for Downhole application," *IEEE Photonics Journal*, vol. 4, no. 5, pp. 1590–1599, 2012.

[8] M. A. Eijkelenborg, M. C. J. Large, A. Argyros, J. Zagari, S. Manos, N. A. Issa, I. Bassett, S. Fleming, R. C. McPhedran, C. M. Sterke and N. A.P. Nicorovici , "Microstructured polymer optical fibre," *Optics Express*, vol. 9, no. 7, p. 319, 2001.

[9] Z. Xu, K. Duan, Z. Liu, Y. Wang, and W. Zhao, "Numerical analyses of splice losses of photonic crystal fibers," *Optics Communications*, vol. 282, no. 23, pp. 4527–4531, Dec. 2009.

[10] S. Asaduzzaman, K. Ahmed, M. F. H. Arif, and M. Morshed, "Proposal of a simple structure photonic crystal fiber for lower indexed chemical sensing," *2015 18th International Conference on Computer and Information Technology (ICCIT)*, Dec. 2015.

[11] Y. Hou, F. Fan, Z.-W. Jiang, X.-H. Wang, and S.-J. Chang, "Highly birefringent polymer terahertz fiber with honeycomb cladding," *Optik - International Journal for Light and Electron Optics*, vol. 124, no. 17, pp. 3095–3098, Sep. 2013.

[12] R. Hao, Z. Li, G. Sun, L. Niu, and Y. Sun, "Analysis on photonic crystal fibers with circular air holes in elliptical configuration," *Optical Fiber Technology*, vol. 19, no. 5, pp. 363–368, Oct. 2013.

[13] F. Begum , Y. Namihira , S.M. A. Razzak, S. Kaijage, N. H. Hai, T. Kinjo, K. Miyagi, N. Zou, "Design and analysis of novel highly nonlinear photonic crystal fibers with ultra-flattened chromatic dispersion," *Optics Communications*, vol. 282, no. 7, pp. 1416–1421, Apr. 2009.

[14] H. Ademgil, "Highly sensitive octagonal photonic crystal fiber based sensor," *Optik - International Journal for Light and Electron Optics*, vol. 125, no. 20, pp. 6274–6278, Oct. 2014.

[15] S. M. A. Razzk, Y. Namihira, F. Begum, S. Kaijge, N. H. Hai, and N. Zou, "Design of a Decagonal Photonic crystal fiber for ultra-flattened chromatic dispersion," *IEICE Transactions on Electronics*, vol. E90-C, no. 11, pp. 2141–2145, Nov. 2007.

[16] J. Park, S. Lee, S. Kim, and K. Oh, "Enhancement of chemical sensing capability in a photonic crystal fiber with a hollow high index ring defect at the center," *Optics Express*, vol. 19, no. 3, p. 1921, Jan. 2011.

[17] L. Zhang, G. Ren, and J. Yao, "A new photonic crystal fiber gas sensor based on evanescent wave in terahertz wave band: Design and simulation," *Optoelectronics Letters*, vol. 9, no. 6, pp. 438–440, Oct. 2013.

[18] Z. Tan, X. Li, Y. Chen, and P. Fan, "Improving the sensitivity of fiber surface Plasmon resonance sensor by filling liquid in a hollow core Photonic crystal fiber," *Plasmonics*, vol. 9, no. 1, pp. 167–173, Sep. 2013

[19] D. M. Chow, S. R. Sandoghchi, and F. R. M. Adikan, "Fabrication of photonic crystal fibers," *2012 IEEE 3rd International Conference on Photonics*, Oct. 2012.

Study on Doping Profile and Scaling Characteristics of Gate and Channel Engineered Symmetric Double Gate MOSFET

Md. Arafat Mahmud,[1,2] Md. Tashfiq Bin Kashem,[1] and Samia Subrina[1*]

[1]Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology
Dhaka-1205, Bangladesh

[2]School of Photovoltaic and Renewable Energy Engineering, University of New South Wales
Sydney, NSW 2052, Australia

*samiasubrina@eee.buet.ac.bd

Abstract—**The present work reports the effect of device doping profile and dimension scaling on electronic properties of a gate and channel engineered symmetric double gate MOSFET. Based on a two dimensional mathematical model incorporating Poisson's and basic drift-diffusion equations, the work demonstrates how the basic electronic parameters like surface potential, threshold voltage and drain to source current of a triple material double gate double halo gate stacked MOSFET vary owing to different doping profile (uniform, linear or Gaussian doping profile) and the scaling of device dimensions like channel thickness and effective gate oxide thickness. The variations in electronic properties due to the doping profile and device dimension scaling have also been explained in terms of fundamental device physics. The study illustrates that the selection of a particular device dimension is actually a trade-off among several basic electronic properties of the device. All the analytical results have been validated with a commercially available two dimensional device simulator.**

Index Terms—**halo doping, doping profile, surface potential, threshold voltage, channel thickness, effective oxide thickness**

I. Introduction

Practical applications demand ever decreasing chip sizes, which leads to reduced length of the channel in MOS devices. However, as the channel length is reduced to increase both the operation speed and the number of components per chip, the short-channel effects (SCE) arise [1-3]. The threshold voltage of the device is modified due to the shortening of channel length [4, 5].

To deal with short channel effects associated with downsizing of device, gate and channel engineering have been reported to be effective in earlier literatures. Gate engineering techniques include the insertion of multiple materials [1, 2, 6-8] in the gate electrodes to reduce Drain Induced Barrier Lowering (DIBL) effect and also the inclusion of high-k dielectric material [9, 10] along with the conventional SiO_2 layer to mitigate leakage current [11]. Whereas, the channel engineering refers to the use of higher doping concentration (halo doping) around drain and source region to restrict the depletion regions for the minimization of severe charge sharing effect in a short-channel MOSFET [12]. Thus a double gate MOSFET structure [5, 13, 14] with gate and channel engineering applied on it, is a promising candidate for microelectronic industry because of its excellent capability of short channel effect suppression.

Of late, we have reported an analytical model of gate and channel engineered triple material double gate double halo gate Stacked (TM-DG-DH-GS) MOSFET [11]. However, no study has been reported on the impact of device dimension scaling and doping profile on electronic properties of TM-DG--DH-GS MOSFET in previous works.

In this paper, we have reported and explicated the variation in surface potential, threshold voltage and drain to source current of TM-DG-DH-GS MOSFET due to the scaling in device dimensions like channel thickness and effective oxide thickness and due to various doping profiles like uniform, linear and Gaussian doping profile. We have also validated our analytical results with commercially available ATLAS^{TM} device simulator.

II. Device Structure

A schematic structure of a TM-DG-DH-GS MOSFET device is shown in Fig. 1. The channel length, L can be classified into 5 different regions, where L_1 and L_5 (L_1 and L_5 have been assumed equal for halo doping symmetry) refer to the lengths corresponding to the channel regions having uniform halo doping (doping concentration: N_{Am}) near the source and drain terminal, respectively. A constant substrate doping concentrations of N_{As} is maintained in L_2, L_3 and L_4 regions. Unless otherwise stated, uniform halo doping profile has been considered to show the effect of device dimension scaling on electronic properties. Analytical results using linear [1] and Gaussian [15] doping profile have also been presented to show the effect of doping profiles.

The symmetric top and bottom gate electrodes are comprised of three metals with work functions of Φ_{m1}, Φ_{m2} and Φ_{m3} ($\Phi_{m3} < \Phi_{m2} < \Phi_{m1}$), lying over the channel regions having length of (L_1+L_2), L_3 and (L_4+L_5), respectively. The top and the bottom gate terminals are shorted. The gate terminals are kept at a fixed potential of V_{GS}, while the drain bias, V_{DS} is maintained with respect to ground potential. Both the top and the bottom gate oxides consist of a stack of HfO_2 (dielectric constant, k_{HfO2}=25) and SiO_2. So, the effective oxide thickness can be mathematically expressed as: $t_{eff} = t_{sio2} + \frac{t_{Hfo2} \times \in_{sio2}}{\in_{Hfo2}}$ [16], where, t_{sio2}, t_{Hfo2}, $\in_{sio2}$ and $\in_{Hfo2}$ stand for physical thickness of SiO_2 and HfO_2 layer and permittivity of SiO_2 and HfO_2 respectively.

III. Results And Discussion

In this section, we have lifted up the variations in surface potential, threshold voltage and drain to source current profile of TM-DG-DH-GS MOSFET due to the dimension scaling of

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

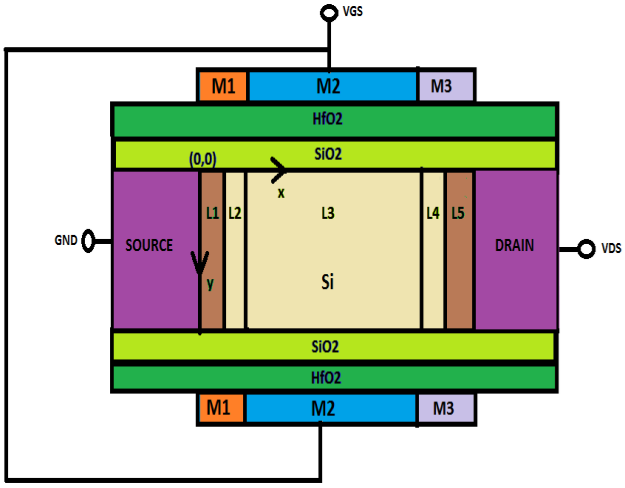


Fig. 1 Schematic view of triple material double gate double halo gate stacked (TM-DG-DH-GS)MOSFET (not drawn to scale)

effective oxide and channel thickness using the two dimensional mathematical model developed in our recently reported work [11]. The model has been established by solving Poisson's and basic drift and diffusion equations. For uniform halo doping, the lengths of the halo pockets near source (L_1) and drain regions (L_5) have been considered as 20 nm. Unless stated otherwise, total channel length, L is assumed to be120 nm. A fixed length ratio for the five channel regions (L_1:L_2:L_3:L_4:L_5=1:1:2:1:1) has been considered throughout the work. The halo doping concentration (N_{Am}) and the substrate doping concentration (N_{As}) are assumed to be 1.2×10^{18} cm^{-3} and 10^{15} cm^{-3} respectively. We have also studied linear [1] and Gaussian [15] halo doping concentration profile to compare the results obtained with uniform halo doping concentration profile. In our study, the work functions considered for both the top and the bottom gate electrode materials are φ_{m1} = 4.8eV, φ_{m2} = 4.6eV and φ_{m3} = 4.4eV.

Fig. 2 shows the variation in the surface potential along the lateral position of the channel with changes in effective oxide thickness of the device. As observed from the figure thicker oxide layer causes smaller values of surface potential. Thicker oxide layer contributes to weaker control of gate over the channel which leads to an increase in energy barrier height. This attributes to the fact that the surface potential is at some lower value, being inadequate to keep the energy barrier lower. Fig. 3 illustrates that the surface potential is higher with thicker silicon channel. The semiconductor channel with larger thickness benefits from more readily available charge carriers for inversion layer formation. Thus, the energy barrier height is smaller and the surface potential remains at a higher value. Fig. 4 lifts up the variation in surface potential along the channel owing to the variation in doping concentration profile. With lower effective doping concentration value, inversion layer formation is easier because of higher electron mobility and less impurity scattering. This indicates the surface potential value is high enough to keep the energy barrier height more lowered compared to the case with higher effective doping concentration. The effective doping concentration is the highest for the uniform doping concentration profile [11] and the lowest for the linear doping concentration profile [1]. As a result, highest value of surface potential is observed with linear doping profile concentration and lowest with uniform doping profile.

Fig. 5 shows the changing pattern of threshold voltage along with the changes in effective oxide thickness for different channel lengths. The influence of gate voltage, being less on the MOS channel with a thick oxide layer, the amount of gate bias required to turn on such a device is higher. So, the threshold voltage is higher with a thicker gate oxide layer.

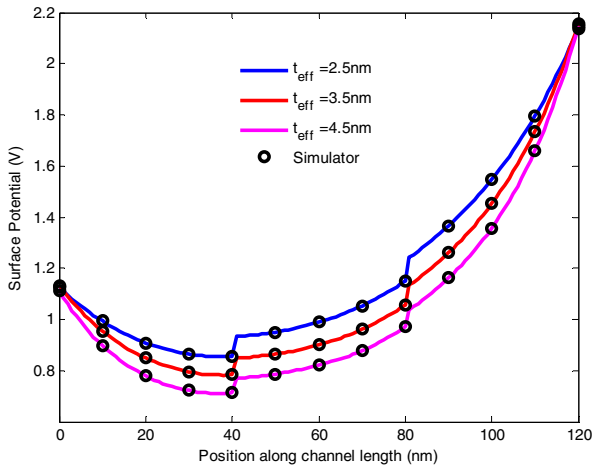


Fig. 2 Surface potential variation with lateral position along channel from source to drain region for effective oxide thickness, t_{eff} of 2.5nm, 3.5nm and 4.5nm at V_{GS}=0.5V and V_{DS}=1.5V.

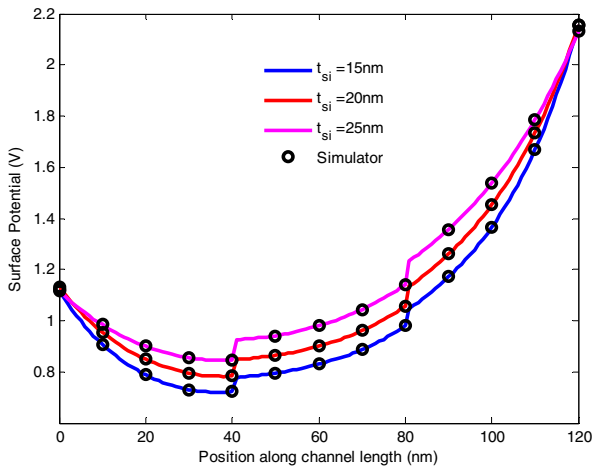


Fig. 3 Surface potential variation with lateral position along channel from source to drain region for channel thickness, t_{si} of 15nm, 20nm and 25nm at V_{GS}=0.5V and V_{DS}=1.5V

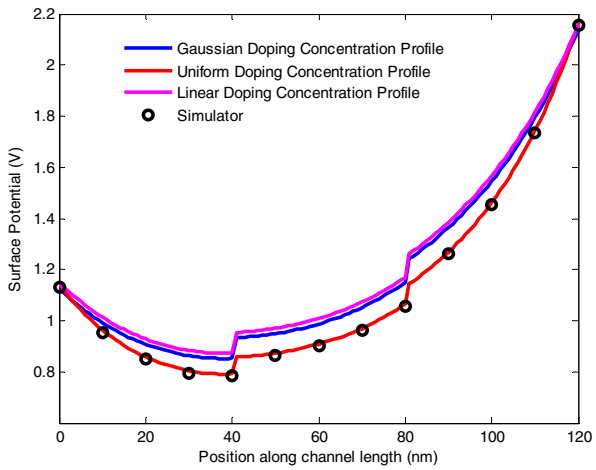


Fig. 4 Surface potential variation with lateral position along channel from source to drain for Gaussian, Uniform and Linear doping profile concentration with V_{GS} =0.5V and V_{DS}=1.5V (t_{eff}=3.5nm and t_{si} =20nm)

Consequently, the highest value of threshold value has been found with thickest oxide layer and vice versa. The higher threshold voltage with a thicker effective gate oxide thickness thus decreases the leakage current in it, however, the current drive in such MOSFET device is expected to be lower than the devices having lower effective oxide thickness. Fig. 6 holds a pictorial representation of the variation in threshold voltage as a function of channel length owing to different channel thickness. It is obvious from the figure that the threshold voltage is lower for channel with larger thickness. This phenomenon is attributed to the fact that the charge carriers are more abundant in a thicker channel device and the energy barrier height is lower likewise. Accordingly, the threshold voltage is lower in a device with higher channel thickness. Thus, a TM-DG-DH-GS MOSFET having thicker silicon channel can have benefit from higher current drive; however, the leakage current is anticipated to be higher in such device structures. Fig. 7 demonstrates the changing pattern of threshold voltage with the variation of doping concentration profile for a set of channel lengths. High doping concentration causes abundance of impurity atoms which leads to a rise in energy barrier height for conduction and increased threshold voltage thereby. Since, uniform doping profile concentration renders the highest value of effective doping concentration among the profiles used [11]; threshold voltage is highest with this concentration profile. Moreover, it is observed that reverse short channel effect is not found in the device with thinnest oxide and thickest channel thickness.

Fig. 8 show the plots of drain to source current with changing gate and drain voltages for various effective oxide thickness. Channel length modulation (CLM) [12] has also been considered in this study. Since, a thick oxide layer passivates the control of gate over channel, inversion layer formation becomes harder. Consequently, the magnitude of drain to source current is lower with a thicker oxide. Fig. 9 illustrates the effect of channel thickness on the drain to source current with various gate and drain biases. As observed from Fig. 9 the highest values of drain to source current are found with thickest silicon channel. Thick channel contains plentiful of charge carriers which make it easier to reduce the energy barrier height. The reduced energy barrier height lowers the threshold voltage of MOS device and raises the magnitude of current. Thus the current rise in Fig. 9 is congruent with the lowering of threshold voltage in thicker silicon channel devices as shown in Fig. 6. Fig. 10 shows the changing trend in drain

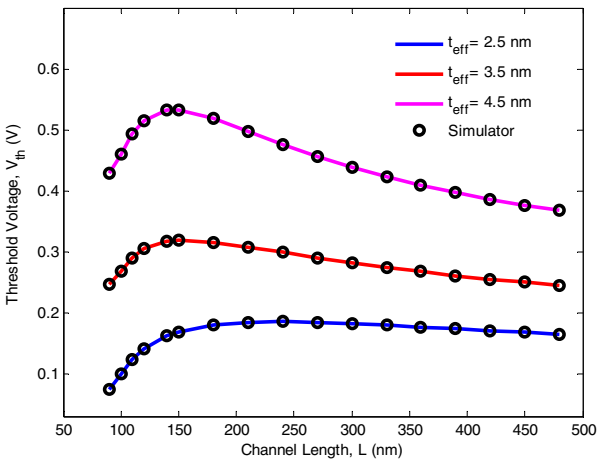


Fig. 5 Threshold voltage of the device with various channel lengths for effective oxide thickness, t_{eff} of 4.5nm, 3.5nm and 2.5nm at V_{DS}=1.5V.

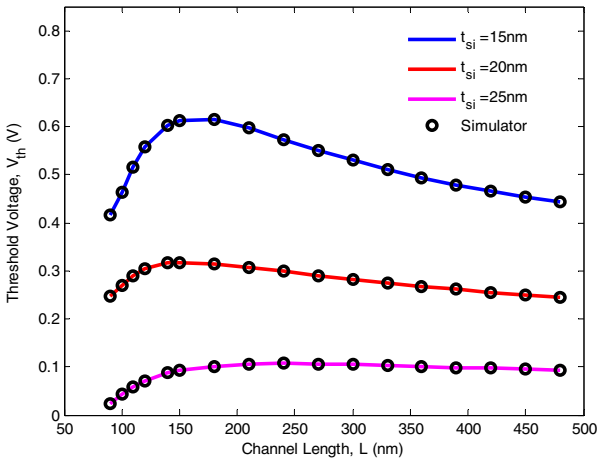


Fig. 6 Threshold voltage of the device with various channel lengths for channel thickness, t_{si} of 15nm, 20nm and 25nm at V_{DS}=1.5V

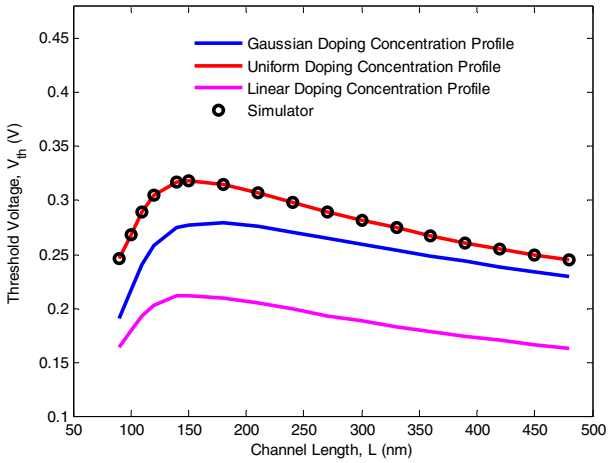


Fig. 7 Threshold voltage of the device along the channel length with Gaussian, Uniform and Linear doping profile concentration at V_{DS}=1.5V

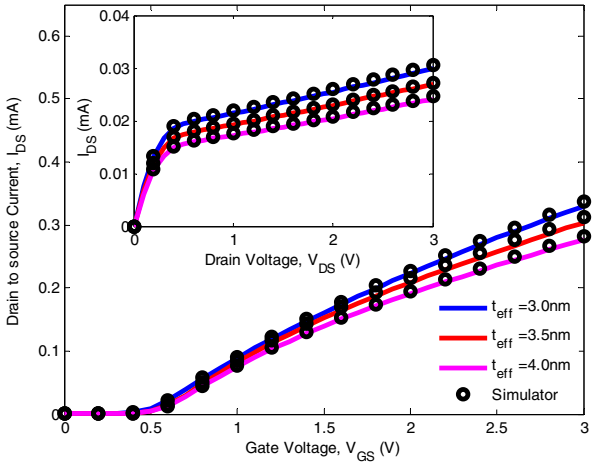


Fig. 8 Drain to source current for effective channel thickness, t_{eff} of 3.0nm, 3.5nm and 4.0nm as a function of gate voltage at V_{DS}=0.5V) and of drain voltage at V_{GS}=0.5V [inset]

to source current with various doping concentration profile. The highest value of current is observed with linear doping concentration profile and the lowest with the uniform doping profile. Because of highest effective doping concentration, uniform doping profile leads to more impurity scattering and less available inversion charge at the surface causing the

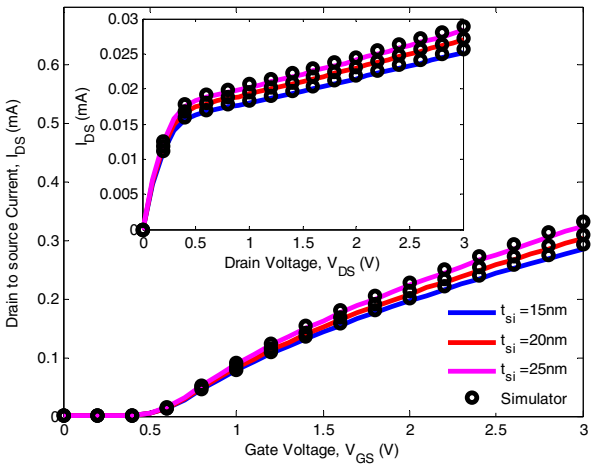


Fig. 9 Drain current for channel thickness, t_{si} of 15nm, 20nm and 25nm as a function of gate voltage (V_{DS}=0.5V) and drain voltage (V_{GS}=0.5V) [inset]

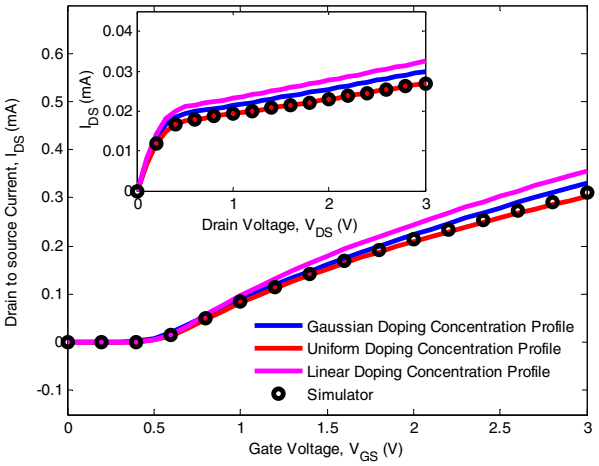


Fig. 10 Drain to source current for Gaussian, Uniform and Linear doping concentration profile as a function of gate voltage (V_{DS}=0.5V) and of drain voltage (V_{GS}=0.5V) [inset]

lowest value of drain to source current. Likewise, the highest value of current is attained with linear doping profile concentration with the lowest effective doping concentration along the channel. All the results obtained from the analytical model show good agreement with the simulation results from device simulator. It is to be mentioned here that validation of our results for Gaussian and linear doping profile were not possible due to our software version.

IV. Conclusions

In summary, we have presented the effects of device dimension scaling and doping profile on the basic electronic properties of a gate and channel engineered triple material double gate double halo gate stacked (TM-DG-DH-GS) MOSFET in this work. The variations in surface potential, threshold voltage and drain to source current profile of TM-DG-DH-GS MOS device have been lifted up due to the scaling in channel and effective oxide layer thickness and various doping profiles. Analysing the outcomes of the work, it can be concluded that the selection of a particular device dimension requires an optimization among different basic electronic properties of the device. Thicker effective oxide layer will significantly decrease leakage current, but the cost of conceding higher energy barrier height leading to lower current drive of the device. Thicker silicon channel may contribute to the easier formation of inversion layer and thus benefit from higher current drive and trans-conductance; but it is let alone that the leakage current will be a concern simultaneously. Uniform doping concentration ensures that the charge sharing effect is less severe which improves threshold voltage roll-off effect but it cannot deny its adverse role in minimizing charge carrier mobility which leads to higher energy barrier and lower transconductance. Hence, the selection of the device physical parameters is to be made considering a number of co-related device phenomena and it is largely application-dependent.

References

[1] A. Pal and A. Sarkar, "Analytical study of Dual Material Surrounding Gate MOSFET to suppress short-channel effects (SCEs)," *Engineering Science and Technology, an International Journal,* vol. 17, pp. 205-212, 12// 2014.

[2] P. Razavi and A. A. Orouji, "Nanoscale Triple Material Double Gate (TM-DG) MOSFET for Improving Short Channel Effects," in *Advances in Electronics and Micro-electronics, 2008. ENICS '08. International Conference on*, 2008, pp. 11-14.

[3] P. K. Tiwari, S. Dubey, M. Singh, and S. Jit, "A two-dimensional analytical model for threshold voltage of short-channel triple-material double-gate metal-oxide-semiconductor field-effect transistors," *Journal of Applied Physics,* vol. 108, p. 074508, 2010.

[4] A. Tsormpatzoglou, C. A. Dimitriadis, R. Clerc, G. Pananakakis, and G. Ghibaudo, "Threshold Voltage Model for Short-Channel Undoped Symmetrical Double-Gate MOSFETs," *IEEE Transactions on Electron Devices,* vol. 55, pp. 2512-2516, 2008.

[5] A. Cerdeira, O. Moldovan, B. Iñiguez, and M. Estrada, "Modeling of potentials and threshold voltage for symmetric doped double-gate MOSFETs," *Solid-State Electronics,* vol. 52, pp. 830-837, 5// 2008.

[6] G. V. Reddy and M. J. Kumar, "A new dual-material double-gate (DMDG) nanoscale SOI MOSFET-two-dimensional analytical modeling and simulation," *IEEE Transactions on Nanotechnology,* vol. 4, pp. 260-268, 2005.

[7] A. Sarkar, S. De, and C. K. Sarkar, "Asymmetric halo and symmetric Single-Halo Dual-Material Gate and Double-Halo Dual-Material Gate n-MOSFETs characteristic parameter modeling," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields,* vol. 26, pp. 41-55, 2013.

[8] C. Li, Y. Zhuang, and R. Han, "Cylindrical surrounding-gate MOSFETs with electrically induced source/drain extension," *Microelectronics Journal,* vol. 42, pp. 341-346, 2// 2011.

[9] B. Cheng, M. Cao, R. Rao, A. Inani, P. V. Voorde, W. M. Greene, *et al.*, "The impact of high-κ gate dielectrics and metal gate electrodes on sub-100 nm MOSFETs," *IEEE Transactions on Electron Devices,* vol. 46, pp. 1537-1544, 1999.

[10] I. Anand, R. Ramgopal Valipe, C. Baohong, and W. Jason, "Gate Stack Architecture Analysis and Channel Engineering in Deep Sub-Micron MOSFETs," *Japanese Journal of Applied Physics,* vol. 38, p. 2266, 1999.

[11] M. A. Mahmud and S. Subrina, "Two-dimensional analytical model of threshold voltage and drain current of a double-halo gate-stacked triple-material double-gate MOSFET," *Journal of Computational Electronics,* pp. 1-12, 2016.

[12] Y. Tsividis, "Operation and modelling of the MOS transistor," *McGraw-Hill : New York,* 1989.

[13] C. Qiang, E. M. Harrell, and J. D. Meindl, "A physical short-channel threshold voltage model for undoped symmetric double-gate MOSFETs," *IEEE Transactions on Electron Devices,* vol. 50, pp. 1631-1637, 2003.

[14] K. Suzuki, T. Tanaka, Y. Tosaka, H. Horie, and Y. Arimoto, "Scaling theory for double-gate SOI MOSFET's," *IEEE Transactions on Electron Devices,* vol. 40, pp. 2326-2329, 1993.

[15] A. S. Swapnadip De, Chandan Kumar Sarkar "Modelling of characteristics parameters for asymmetric DHDMG MOSFET," *WSEAS TRANSACTIONS on CIRCUITS and SYSTEMS,* vol. 11, pp. 371-380, 2012.

[16] B. Toufik, D. Faycal, and B. Abdel Hamid, "Continuous analytic I — V model for GS DG MOSFETs including hot-carrier degradation effects," *Journal of Semiconductors,* vol. 33, p. 014001, 2012.

Characterization of Voltage Sag Based on Point-on-Wave

M.S. Hasan, [1,*] K. M. Muttaqi, [2] and A. Bouzerdoum[3]
[123] School of Electrical, Computer and Telecommunications Engineering,
University of Wollongong, NSW 2522, Australia
[*]msh671@uowmail.edu.au

***Abstract*- Point-on-Wave (POW) is a new characteristic of voltage sag which has a significant influence on the sensitivity of some equipment. This paper investigates the Point of Inception (POI) and the Point of Recovery (POR) of the voltage sag signal due to symmetrical and unsymmetrical faults in power systems. The POI of voltage sag could be any point on the wave where the voltage sag starts. On the other hand, the POR depends on the circuit breaker fault clearing practice. Based on the actual fault clearing angles, three groups of voltage sag signals separated by 180° are identified and characterized. For single-phase faults there are three possible recovery points on the voltage wave. For double phase faults, six PORs are found at three instants of time. Nine possible PORs for three phase faults are obtained at nine different instants in time. This paper demonstrates that depending on the circuit breaker fault clearing time, the voltage sag duration increases significantly, and this can be harmful to some sensitive equipment. This paper also analyzes the current signals at different POI of corresponding voltage sag signals and detects some critical POI of voltage sags for which the highest transients in current signals occur. The analysis has been carried out on two different radial distribution systems for 25 voltage sag signals over a cycle with 15^0 intervals, for which simulation results are reported.**

Index Terms - Voltage sag (or voltage dip), Point-on-wave, Point of sag inception, Point of sag recovery.

I. Introduction

Voltage sag is an important power quality problem because it incurs a massive financial loss for customers. To date, hundreds of research works have been published based on the voltage sag phenomenon. The standard definition of voltage sag has been established based on two major characteristics: the root-mean-square (rms) magnitude of the voltage sag and the length of time it lasts [1], [2]. It has been reported that, apart from sag magnitude and duration, the POW characteristics of sag inception point and sag recovery point also have a significant influence on the behavior of some equipment [3]-[10]. The Point-on-Wave characteristics refer to the exact point on a sinusoidal wave where the voltage signal starts to deviate and starts to return to the original signal [5]. This means that the voltage sag signal has two point-on-wave characteristics: one is the point of sag inception (POI) and the other is the point of sag recovery (POR).

POI characteristics drew research attention when AC contacts were shown to have a link with the unintentional tripping of the industrial process during the voltage sag events. This was confirmed by testing the electromagnetic contacts for different voltage sag magnitudes, durations and different points of initiation [3]. In [3], the experimental results showed that for the same magnitude of voltage sag signals, disengagement of electric and magnetic contact occurred due to the variation in sag inception angles. It was found that for a voltage depth of 90V rms and 0 degree of POI, AC contacts could engage up to 30 ms, while for the same voltage with 90 degree of POI, they disconnected after 12 ms [3]. Relays and motor starters were also found to be vulnerable to different POIs of voltage sag signal. The impacts of the POI on electric contact, personal computers, induction machines and transformers have been reported in [4]-[6].

In an experiment conducted by EPRI [7], AC contacts were found to be sensitive to 90° of POI for sag durations shorter than 4-5 cycles. On the other hand, they disconnected at 0° sag inception angle for sag duration greater than 5 cycles but it was reported that all disconnection of AC contact occurred well before the end of the sag [3]. This means that the sag duration has no influence on the behaviour of sensitive equipment in response to voltage sag sensitivity at different POI.

Similar to the POI of voltage sag, previously published works have shown that the voltage sag recovery points also have a significant impact on the behaviour of some equipment. The voltage sag recovery points for three phase balance and unbalance faults have been studied in [8]. The influence of the voltage sag recovery process on transformers [4], induction machines [9] and DFIG (Doubly-fed induction generator) based wind turbines [10] have been studied in detail.

Most of the existing research works have focused on the equipment immunity testing due to different POIs of voltage sag but the reason why the equipment becomes sensitive at some points of inception of voltage sag, while for the other points they work properly even though the voltage sag magnitudes and durations are the same for all sagged signals has not been identified. In order to determine the sensitivity indices accurately, extensive analysis is required to understand the characteristics properly.

This research aims to characterize the voltage sag signals based on the POW of sag starting and ending. Both of these refer to the angles of the sinusoidal wave at the beginning and ending point of the voltage sag event. For this reason, this characteristic cannot be detected accurately by rms methods [6]. In this paper, we analyse the instantaneous voltage to investigate the POW characteristics. Three types of fault-induced voltage sags are considered in order to analyse the characteristics. Extensive simulations have been completed using MATLAB Simulink software to characterize the voltage sag signals in terms of POW.

II. Point-on-wave Analysis for different types of sag

Point-on-wave of initiation and point-on-wave recovery of voltage sag caused by short circuits directly correspond to the fault inception and fault clearing angles. In the report of the investigation of recorded voltage sag signal, it has been suggested by EPRI that the sag initiation point, POI, is not

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

biased by the voltage peak rather it may occur at any point in a fairly random way [5]. Conversely, the point of recovery occurs when the voltage approaches the positive or negative peak.

A. Calculation of the Angle at POI of Voltage Sag

Let us consider a three-phase voltage signal $v(\theta)$ which is subjected to a voltage sag event due to a short circuit fault. If θ_i is the sag inception angle, then the expression of the sagged voltage signal in terms of the sag inception angle can be written as:

$$v(\theta_i) = V\sin(\theta_i + \delta) \quad (1)$$

Here,

$v \rightarrow v\{v_a, v_b, v_c\}$; $\delta\{\delta_a, \delta_b, \delta_c\}, \theta_i\{0^\circ \text{ to } 360^\circ\}$

At steady state condition:

$\delta\{\delta_a, \delta_b, \delta_c\} \rightarrow \delta\{0^\circ, -120^\circ, 120^\circ\}$

B. Calculation of the Angle at POR of Voltage Sag

The voltage sag recovery point depends on the circuit breaker switching instant. The circuit breaker clears the fault when the fault current has zero crossing [10] in order to reduce the high current transient during the recovery process. Therefore, even though the fault is cleared from the system, the circuit breaker waits for the current zero crossing point. This phenomenon results in about a half cycle of delay for some voltage sag signals increasing the sag duration. If the fault is actually cleared at any point between the negative peak and before the positive peak of the voltage signal, the possible voltage recovery angle will remain on only one particular point in the positive first quarter cycle depending. On the other hand, if the actual fault clearing occurs at any point between the positive peak and before the negative peak of the voltage signal, the recovery points will remain in the negative first half cycle.

Let θ is the voltage angle when the fault is assumed to be cleared from the system. In other words, the circuit breaker is sent a fault clearing signal at the instant when voltage angle is θ. If θ_r is the sag recovery angle, then the expression of the sagged signal in terms of the sag recovery angle can be written as:

$$v(\theta_r) = V\sin(\theta_r + \delta) \quad (2)$$

Here,

$$\theta_r = +\theta_r \quad if -90^\circ \geq \theta < 90^\circ \quad (3)$$

$$\theta_r = -\theta_r \quad if\ 90^\circ \geq \theta < -90^\circ \quad (4)$$

The recovery angle θ_r is a single point on the wave for a particular system and for a particular type of fault which shifts by 180^0 depending on the instant when the fault is actually being cleared from the system and the angle between voltage and current (i.e. the power factor angle). If we consider that the actual fault recovery happens at any instant from 0° to 360^oover the cycle, only three recovery points separated by 180° should be expected. Therefore, for single phase to ground faults three possible recovery points are expected. For phase to phase faults, both phases recover at the same instant of time, so six possible points on two phases at three different times should occur, as two phases are cleared simultaneously. On the other hand, three phase faults and double phase to ground faults are cleared at different times [8] in order to avoid current transients. This means that the voltage sag recovery points for a three phase fault should be found in nine points on the wave at nine different times and for the latter type of fault, six POR should be obtained at six different times. In time domain analysis, the POW angle can be calculated using the following formula:

$$POI, \theta_i(t_i) = (t_i - T * n) * 360/T \quad (5)$$

$$POR, \theta_r(t_r) = (t_R - T * n) * \frac{360}{T} \quad (6)$$

Here n = number of cycles, $T = \frac{1}{f}$

The theoretical analysis is validated in the following sections in two different distribution networks which are extracted from practical distribution systems.

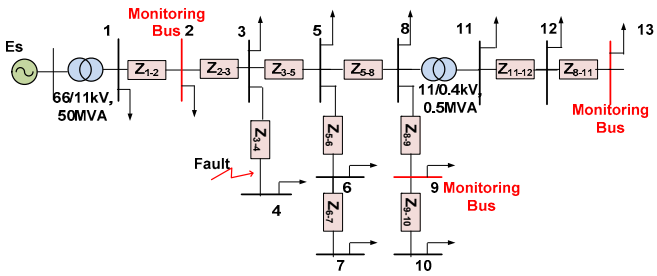


Fig. 1 13- bus radial distribution Network (NSW)

III. Analysis of Voltage Sag at Different POW

A. Case Study I

The voltage sag events are simulated in SimPower system software in a 13 bus radial distribution network which is derived from New South Wales (NSW) power distribution systems, Australia, as shown in Fig.1. To characterize the POW, the voltage sag events have been generated at Bus 4 by applying a three-phase balance fault, a single line to ground fault in phase A and a line to line fault in phase A-B at different POI angles from 0° to 360° degree. With 15^0 intervals, 25 sag signals have been generated for each type of fault in every case study. The duration of voltage sag has been chosen for two cycles (0.04sec for 50Hz power system), because in standard practice, the fault should be cleared from the system within two or three cycles in order to protect the power equipment from damage and to maintain system reliability. The sag duration has been kept constant for each sag signal. Both rms and instantaneous values of voltage and current signals are measured and analysed. Measurements are taken at upstream bus 2. It should be noted that the faults are modelled to be started from 0^0 and ended after two cycles at 0°. Afterwards, the next signals are generated with 15^0 intervals. Therefore, actual fault clearing angles are equal to the fault starting angles.

a) Analysis of Voltage Signal for Sag Type B

Fig. 2 shows the voltage sag signals of type B measured at monitoring bus 4 at different POI angles due to single- line to ground fault on phase-A at Bus 4. It is clearly shown that, for

all the 25 voltage sag signals , even though there are 25 sag inception points , all the sags recover nearly at three particular points: $\theta_{ra1}, \theta_{ra2}$ *and* θ_{ra3} with 180° phase shift between them. The characteristics of the three groups of signals can be explained as follows.

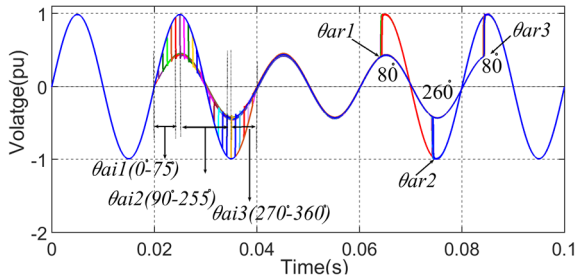


Fig. 2 Type B voltage sag signals for different POW

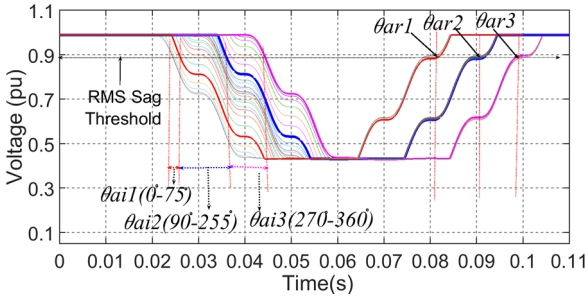


Fig. 3 RMS magnitude of Sag Type B at different POI angles

Group 1→ It contains the voltage sag signals which have the same recovery angle at 80°. The actual fault clearing angle, θ for this group is seen between 0^0 and 75^0.

Group 2→ In this group 12 voltage sag signals with actual fault recovery angles from 90° to 255^0 are included because of their same recovery point at nearly 260^0 .

Group 3 → Group 3 has the same recovery angle as Group 1, except it is found to be shifted to the next cycle. In this group, the voltage sag signals with 270^0 to 360^0 actual fault clearing angle can be seen to be recovered very closely at 80^0 of POR angle.

Fig. 3 illustrates the rms plot of the voltage sag signals at different POIs. For all sag signals, the voltage magnitudes are found to be constant but as the points of recovery are the same for each group of signals, but the sag durations no longer remain constant. Instead, they increase significantly for certain sag signals. This phenomenon has been illustrated in Table I with detail explanation in Case study II.

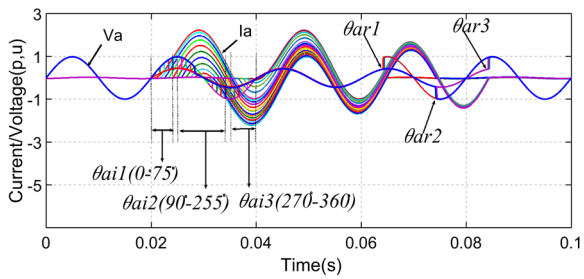


Fig. 4 Instantaneous current and voltage signals at different POI angles

b) Analysis of Current Signals

In Fig. 4 and Fig. 5, the instantaneous current signals and rms magnitude of current signals are presented for the 25 voltage sag signals at 25 POI/POR angles. In the signatures of current, similar recovery characteristics are observed like voltage signals with three groups of signals having three recovery points : $\theta_{ra1}, \theta_{ra2}$ *and* θ_{ra3}. Fig. 4 also implies that the voltage sag signals are recovered only at the points when current signals cross the zero points after the fault is cleared from the system.

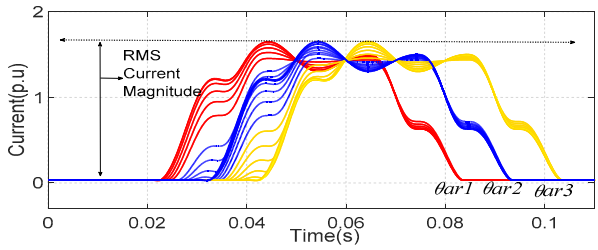


Fig. 5 RMS magnitude of current signals at different POI angles

From the rms magnitude of the current signals, as shown in Fig 5, it can be seen that unlike voltage sag signals, the magnitudes of current signals are found to be different at different POI of voltage sag signals. Highest current magnitude with 2/3 overshoots is identified for the voltage sag signals with zero crossing POI angles (0°, 180° and 360°). On the other hand, for the voltage sag signals which start near the positive and negative peak (75°, 90°, 270° and 255°), the current magnitudes remain at least 20% lower than those of the zero crossing initiations of voltage sag signals.

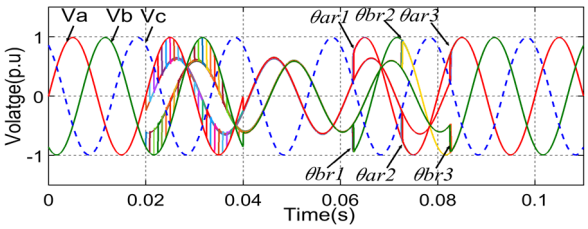


Fig. 6 Sag Type C for line to line fault at different POI angles

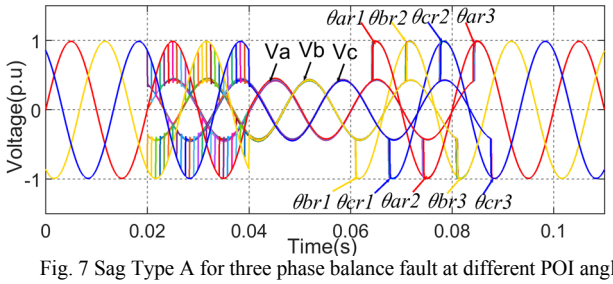


Fig. 7 Sag Type A for three phase balance fault at different POI angles

c) Analysis of Voltage Signal for Sag Type C

The type C voltage sag signals at different POI/POR angles due to the line to line faults between phase-a and phase-b are shown in Fig.6. As line to line faults are cleared at the same instant of time in both phases, six POR angles have been detected for the three groups of signals which are also separated by 180° in each phase. The angles are marked as θ_{ra1} , θ_{ra2} and θ_{ra3} at phase a, θ_{rb1}, θ_{rb2} and θ_{rb3} at phase b and θ_{rc1}, θ_{rc2} and θ_{rc3} at phase c .

d) Analysis of Voltage Signal for Sag Type A

Fig.7 shows the 25 voltage sag signals of type A which are generated from a three phase fault at different POI/POR. Three phase faults are cleared at three different times with a few millisecond gaps between each phase. In Fig.7 it can be seen that, due to the different fault clearing instant in each phase, the sag recovery times are different. Therefore, Group 1, Group 2 and Group 3 signals recover with nine POR angles: $\theta_{ra1}, \theta_{rb1}, \theta_{rc1}, \theta_{ra2}, \theta_{rb2}, \theta_{rc2}, \theta_{ra3}, \theta_{rb3},$ and θ_{rc3} at nine different times. It can also be seen that the 180° shift among the three groups is maintained in each phase.

TABLE I

ANALYSIS OF SAG DURATION AT DIFFERENT POI FOR IEEE-15 BUS TEST NETWORK

POI Deg.	Breaker Switching duration (Sec)	Actual sag Duration	Rise in sag Duration,%
Group-1			
0°- 60°	0.04	0.043150 -0.040117	7.88%-0.29%
Group-2			
75°- 240°	0.04	0.048983-0.040517	22.46%-1.29%
Group-3			
255°-360°	0.04	0.048983 - 0.0438	22.46%-9.50%

B. Case Study II

A similar study has been performed to analyze the POW characteristics in IEEE-15 bus test distribution network. The test results for type B voltage sag starting at 15° intervals over a cycle are presented in Table I. This analysis shows that, although the sag duration is kept constant at 2 cycles for each sag signal, the actual sag duration changes for each sag signal. The maximum increase of 22.46% in sag duration is calculated for the voltage sag signals starting at 75° and 255° which are the starting points of Group 2 and Group 3. It can also be seen that sag duration is comparatively higher when it starts near the negative and positive peak. The lowest sag duration has been identified for the voltage sag signals which are included in the transition points of Group 1 and Group 2. The lowest increase of 0.29% sag duration is calculated at the end of Group 1 for 60° initial angle of the sag signal.

TABLE II

COMPARISON OF POW CHARACTERISTICS BETWEEN TWO DIFFERENT NETWORKS

POI Group	IEEE-15 Bus network	NSW Network
Group 1	0°-60°(5 signals)	0°-75°(6 signals)
Group 2	75°-240°(12 signals)	90°-255°(12 signals)
Group 3	255° - 360°(8 signals)	270°-360°(7 signals)

Table II shows a comparison between the three groups of voltage sag signals which are obtained from the simulation results in two individual networks for type B sag . Between the IEEE -15 bus network and the NSW distribution network, there is a variation of one signal, i.e. Group 1 ends at 60° in the first network, while in the second network it ends at 75°. This difference continues at the end of Group 2 which is 240° for IEEE test network and 255° for the NSW test system. In both of the networks, the starting voltage sag angle of Group 1 and the ending angle of Group 3 are the same but there is a small variation in the starting and ending angle of each of the three groups in the two individual networks. The significant changes in the sag duration are unavoidable. It can also be seen that as there is a gap of 15° between each voltage sag signal and the bandwith of each group will change if more precise simulation with 1° gap is taken.

IV. CONCLUSIONS

This paper analyses the characteristics of voltage sag signal for different POWs of voltage sag initiation and voltage sag ending. A detailed theoretical analysis has been developed which is tested and verified in two individual networks. From the test results, it can be concluded that even though voltage sag starts randomly at any point on the signal, the recovery points are not random because all voltage sag signals wait for the zero crossing of the current signal. Therefore, voltage sag signals are categorized into three groups each of which are separated by 180°. It has been observed that the voltage sag recovery process increases the sag duration significantly for some sag signals and this might be one of the key causes of sensitive equipment disruption. This analysis will be helpful to predict the sag recovery points and the actual sag duration. This research also detects some points of sag initiations for which current transients are found 20% higher than any other points of sag initiations. This variation of current transients due to the starting points of voltage sag is harmful for some sensitive equipment. The trend of voltage sag signals to be recovered closely at three particular points should be considered carefully during voltage sag immunity testing of sensitive equipment. This analysis need to be further extended to analyse the variation in system load, X/R ratio and power factor angle in terms of POW characteristics.

REFERENCES

[1]. Testing and measurement techniques – Power quality measurement methods, *IEC, 61000-4-30: Electromagnetic compatibility (EMC) –Part 4-30:* 2015.

[2]. IEEE Guide for Voltage Sag Indices, IEEE Std 1564-2014,2014: p. 1-59.

[3]. Collins, E.R. and M.A. Bridgwood, "The impact of power system disturbances on AC-coil contactors" , *Textile, Fiber, and Film Industry Technical Conference, IEEE 1997*

[4]. Pedra, J., et al., "Symmetrical and unsymmetrical voltage sag effects on three-phase transformers.", *IEEE Transactions on Power Delivery, 2005.* ***20****(2): p. 1683-1691.*

[5]. *Waveform Characteristics of Voltage Sags: Statistical Analysis, EPRI, Palo Alto, CA. 1999.TR-112692.* 1999

[6]. Djokic, S.Z., J.V. Milanovi!c, and S.M. Rowland, "Advanced voltage sag characterisation ii: point on wave", *IET Generation, Transmission & Distribution, 2007.* ***1****(1): p. 146-154.*

[7]. "The Effects of Point-on-Wave on low-voltage tolerance of industrial process devices", *EPRI Power Quality Testing Network Brief No. 44, 2007 #286*

[8]. Bollen, M.H.J., "Voltage recovery after unbalanced and balanced voltage dips in three-phase systems.", *IEEE Transactions on Industrial Electronics, 2003. 18(4): p. 1376-1381.*

[9]. Guasch, L., F. Corcoles, and J. Pedra, "Effects of symmetrical and unsymmetrical voltage sags on induction machines." *IEEE Transactions on Power Delivery, 2004. 19(2): p. 774-782.*

[10]. Rolán, A., J. Pedra, and F. Córcoles, "Detailed study of DFIG-based wind turbines to overcome the most severe grid faults", *nternational Journal of Electrical Power & Energy Systems, 2014. 62: p. 868-878.*

Prediction of HSC Examination Performance using Socioeconomic, Psychological and Academic Factors

A T M Shakil Ahamed[1], Navid Tanzeem Mahmood[2], Rashedur M Rahman[3]
[1]ahnafshakil@gmail.com, [2]navid.mahmood@hotmail.com, [3]rashedur@northsouth.edu

[1,2,3] Department of Electrical and Computer Engineering, North South University,
Plot-15, Block-B, Bashundhara, Dhaka, Bangladesh.

Abstract—**Understanding how to make education effective is a critical step in educational data mining. We have considered various socioeconomic, psychological and academic factors to fully understand what a person's life is during adolescence and how those factors impact their academic performance. Using pre-processing techniques such as feature selection, data balancing, discretization and normalization, and classification techniques such as Random Forest, Artificial Neural Net (ANN), and Naïve Bayes we have developed different models to predict academic performance and also investigated prominent patterns in this study.**

Index Terms—**Data Mining, Educational data mining, academic performance, machine learning**.

I. Introduction

Educational data mining is a relatively new field where key factors relating to academic performance are being studied. Most of the data has been collected from students who have just given HSC examination; most of whom are fresher at university or are preparing to join in an undergraduate degree. The findings from this research will help the students who plan to sit for their HSC exams.

The main objective of our research is to predict the result of HSC based on socioeconomic, psychological and academic factors and to find out the key connections of the attributes to visualize the relationship of the features, which can be used as prescriptive towards the performance of the HSC. Parents and responsible authorities will get the necessary insight to achieve better performance in HSC exam and avoid failures.

II. Related Work

K-means, model based clustering and other clustering techniques have been used in [1] in order to perceive the skill level of the students and to find out groups according to their skill level. In [2] quiz, midterm, Lab mark, CGPA and final score have been used to predict academic result of a semester of students in the university level. Authors have collected data form 181 students throughout 5 semesters of the students of North South University and used ANN, Naïve Bayes and C4.5 for the classification of the data set. They managed to achieve 82% of accuracy.

III. Designing survey questionnaire and data set

The crucial part of designing predictive model greatly depends on the quality of the data set. We have carefully designed the survey questionnaire by conducting interviews with experts and researching different factors of academic performance to get the relevant attributes pertaining to the academic attainments of students. We have made sure that every attribute included in the survey questionnaire has been proven to be relevant to academic performance through published research. We have managed to survey 423 students and to maintain the versatility in the data set we have surveyed students with different economic status and from different locations. Our data set consists of 33 attributes and mainly divided into 3 categories as shown in table I.

IV. Overview of designing predictive model

Overall design of our predictive model is shown in fig 1. We have used feature selection, data balancing, discretization and normalization as pre-processing before implementing learning algorithms and finally derive the best model.

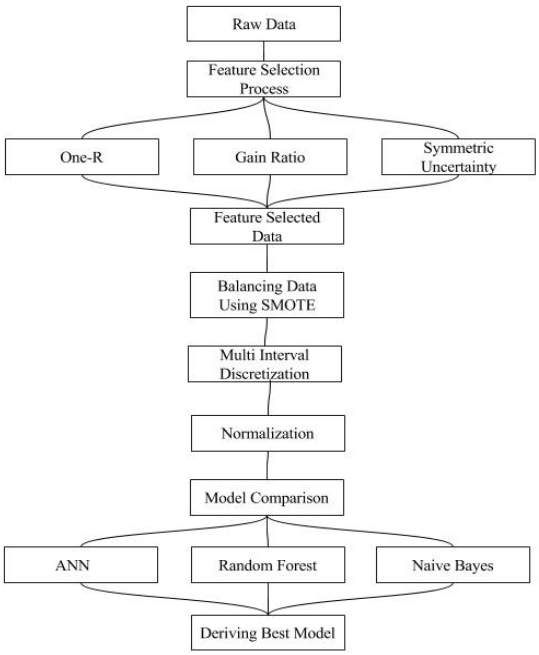


Fig. 1 Overview of our system showing different methodologies

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

TABLE I
ATTRIBUTES USED IN PREDICTIVE MODEL

Socioeconomic and Demographic factors	Psychological factors	Academic Factors
• Sex • Age • Mother's Education • Father's Education • Mother's Job • Father's Job • Average Monthly Family income • Place of Education • Private Tuition	• Family Size • Age difference with most immediate Sibling • Parents' involvement in education • Parent's Marital Status • Absence/Death of Parent(s) • Involvement in Physical Activity • Motivating Teacher • Motivating Parents • Academic Ambitions • Time Spent On Social Media • Romantic Relation • Hanging Out With Friends • Alcohol/Drugs consumption • Smoking • Health Status Of Participant • Health Status of Family members	• Early Childhood Education • Weekly Study Time • Previous Failures • Number of Absence(s) • Extracurricular Activities • Accessibility of the Internet • SSC (Secondary School Certificate) result • HSC result

V. DATA PRE-PROCESSING

A. Feature Selection

Our data set consists of 33 attributes amongst which 32 of them are predictors and one (HSC result) is target class. Our first step is to identify irrelevant attribute from the data set. Having irrelevant features in the data set can cause lower predictive accuracy by provoking the problem of overfitting [3] due to the noise in the data set. To reduce the curse of dimensionality we have used 3 feature selection techniques:

1) *One-R Attribute Evaluation*: It [3] measures the worth of an attribute by evaluating the correlation with the target class using One-R classifier.

2) *Gain Ration Attribute Evaluation:* The method [4] measures the worth of an attribute by evaluation the gain ratio between the target class and attribute.

3) *Symmetrical Uncertainty Attribute Evaluation*: It [5] measures the worth of an attribute by evaluating symmetric uncertainty of the features towards the target class.

All the methods provide the rank of the attribute along with individual score based on the predictive accuracy of the target class. We have converted the scores provided by individual attribute evaluation into 10 and then taken the average of the scores. After that we have arranged the attributes in descending order according to the average score. Then we have used Naïve Bayes to find out optimal classification error by eliminating attribute one at time from the lowest scored attribute.

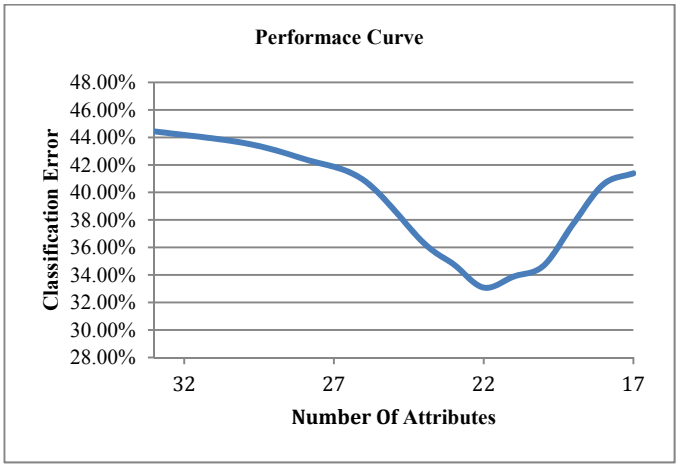


Fig. 2 Performance curve using Naïve Bayes Classifier

The fig.2 shows that while we have used 1st 22 attributes the classification error is optimal. Thus we have used only these 22 attributes to design our predictive model.

B. Data Balancing

Fig.3 shows that our set is imbalanced which can seriously hamper the quality of the predictive accuracy. The problem can be easily visualized in mammographic images, which consist of 98% normal pixels and mere 2% of abnormal pixels, are used as data. By simply guessing pixel as normal one can be correct for 98% of the time. So the minority and majority class needs to be balanced by getting equal entropies to ensure higher quality of the accuracy. We have used Synthetic Minority Over-Sampling technique (SMOTE) to balance the data set. SMOTE [6] uses K nearest neighbour to measure the distance between the instances in the minority class, then generates artificial examples of minority class. Fig. 4 shows the balanced dataset we have got after applying SMOTE.

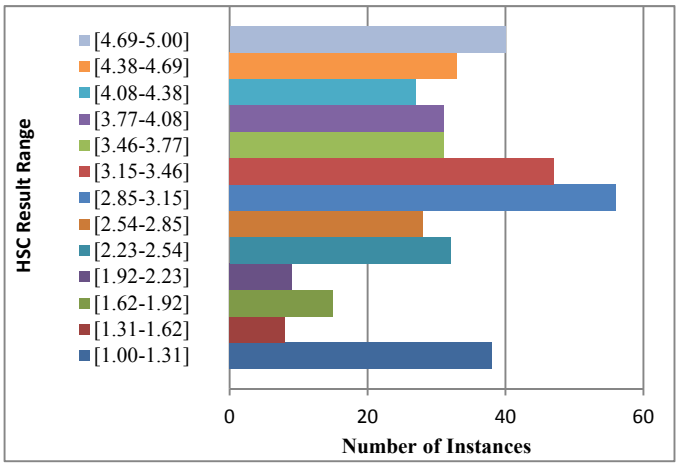


Fig. 3. Instances of target class (HSC Result) before applying SMOTE

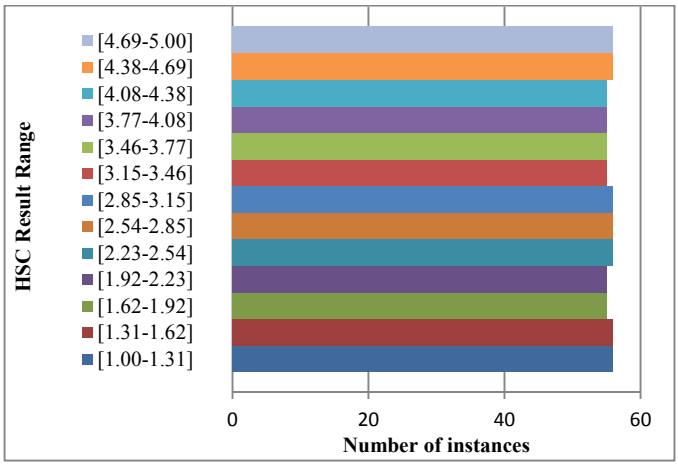


Fig. 4 Instances of target class (HSC Result) after applying SMOTE

C. Discretization

In our data set HSC result and SSC result are continuous valued attribute. In order to improve the predictive accuracy of the system we have used Multi interval discretization (MLD) to discretise the data set. MLD [7] discretises the data set by dynamically search for the cut point to have optimal width of the bin to achieve higher accuracy.

D. Normalization

Our data set consists of attributes of different units. For example, "Weekly study Time" is 'Hours' whereas the unit of "Average family income" is in the unit of 'Taka' and some of the attribute has different scale, for instance, "Family involvement" measures in a scale of 5 conversely "Heath status" is measured in a scale of 10. Normalization ensures the equal contribute of the attribute to predict target class. We have used min-max [8] normalization technique to rescale the attribute. It converts the range of the attributes to a new range by rescaling the feature from 0 to 1 or -1 to 1.

VI. Learning Algorithms

A. Random Forest

Random Forest [9] is a tool that can be used for both classification and regression. It selects random subsets from data and produces decision trees from it. These decision trees are then 'bagged' and a 'forest' is produced. Using a majority vote system on the decision trees, classification is improved and the 'out of bag' error gets reduced.

Random Forest also averages the weight of 'm' trees resulting in the equation (1).

$$\vartheta = \frac{1}{m}\sum_{j=1}^{n}\sum_{i=1}^{n} W_j(x_i, x')y_i = \sum_{i=0}^{n}(\frac{1}{m}\sum_{j=1}^{m} W_j(x_i, x')y_i \quad (1)$$

B. Artificial Neural Network

Artificial neurons work much like their biological counterparts: they transmit information between each other and causing each other to 'fire' based on the inputs. To 'fire' the following activation function (2) must be satisfied.

$$a_i = \sum_{j=1}^{N} W_{ji}x_j + \theta_i \qquad (2)$$

Where, x_j is the output from a neuron or an external input, θ_i is the threshold, W_{ji} is the weight.

Our system uses a 'feed forward' architecture; outputs of a layer of neurons are fed into the inputs of the neurons from the next layer. Between the input and output layer, there are a 'hidden' layers [10]. Our system has the following hidden layer calculated using equation (3).

$$(Number\ of\ Attributes + Number\ of\ Classes)/2 + 1 \qquad (3)$$

= (22+1)/2+1 = 12 (approx.) hidden layers.

C. Naïve Bayes

Naïve Bayes works [11] by applying probabilistic functions for classification as shown in (4).

$$p(C_i|v_1, v_2, \dots\dots\dots, v_n) = \frac{p(C_i)\Pi_{j=1}^{n} p(v_j|C_i)}{p(C_i|v_1, v_2, \dots\dots\dots\dots, v_n)} \qquad (4)$$

This function works under the assumption that all inputs are independent of each other statistically.

VII. Result Analysis

A. Random Forest

While only feature selection pre-processing is applied on data set, Random forest has reached the accuracy of 47.46% as shown in table II. As soon as we have balanced the data set using SMOTE, accuracy increases by 26.27%. Having equal entropies of both minority and majority class bolsters Random Forest to provide better accuracy. After applying MLD for discretization, accuracy has significantly fallen to 57.73% which can be inferred that random forest generates a lot of trees by using random subset of the data and discretization limits its extent for generating trees result in lower accuracy. After introducing normalization accuracy immediately augmented to 82.57%. Rescale the attribute solves the problem caused by discretization.

B. Artificial Neural Network (ANN)

After applying feature selection ANN provides accuracy of 70.34%. Then applying SMOTE to balance the data has increased the accuracy by almost 1%; thus SMOTE does not have a significant impact on the accuracy of ANN. Discretization of continuous valued attribute using MLD bolsters ANN to produce higher accuracy which is as high as 89.34%. Normalization of the data set does not have any impact on accuracy.

C. Naïve Bayes

While feature selection is used as pre-processing technique Naïve Bayes has reached mere 35.59% of the accuracy. As we have introduced SMOTE the accuracy is significantly augmented to 64.98%, balancing data set helps Naïve Bayes to soar the accuracy by 29.39%. After using MLD for discretization the accuracy has further increased to 68.81%. Normalization has not played significant role in case of Naïve Bayes.

VIII. Key Findings

The key findings are generated from decision trees and shown as rules below. One observation is "When parents are separated, romantic relationship hampers academic performance". In contrast, another observation demonstrates romantic relationship enhances GPA when parents are together.

```
ParentStatus = Living Apart
|   RomanticRelation = no: (HSC result) 2.750
|   RomanticRelation = yes: (HSC result) 1
ParentStatus = Living Together
|   RomanticRelation = no: (HSC result) 2.250
|   RomanticRelation = yes: (HSC result) 2.750
```

Students studying rural areas have higher GPA if they have a romantic partner whereas it is the exact opposite for students going to urban schools.

```
PlaceOfEducation = Rural
|   RomanticRelation = no: (HSC result) 2.250
|   RomanticRelation = yes: (HSC result) 2.750
PlaceOfEducation = Urban

|   RomanticRelation = no: (HSC result) 2.500
|   RomanticRelation = yes: (HSC result) 1
```

Ambition for higher studies has a notable effect on GPA that is generated from the rule below.

```
higherStudyAmbition = no
|   WeeklyStudyTime = 1 hour: (HSC result) 3
|   WeeklyStudyTime = 2 hours: (HSC result) 3.250
higherStudyAmbition = yes: (HSC result) 3.500
```

IX. Conclusion

The highest accuracy of our predictive model is 89.48% using ANN along with feature selection, SMOTE, MLD and normalization as pre-processing. As the pre-processing is being introduced the accuracy gets increased most of the time. If the students get the prediction of their academic performance beforehand they could be more cautious about their exam. The predictive model and key findings would give the parents and the responsible authority an instrument to help student get better academic result in HSC exam.

References

[1] E. Ayers, R. Nugent, & N. Dean. A Comparison of Student Skill Knowledge Estimates. In International Conference On Educational Data Mining, Cordoba, Spain, pp.1-10, 2009, J. Breckling, Ed., *The Analysis of Directional*

[2] S.T. Jishan, R.I. Rashu, , A. Mahmood., F. Billah, & R. M. Rahman. Application of Optimum Binning Technique in Data Mining Approaches to Predict Students' Final Grade in a Course. *Advances in Intelligent Systems and Computing*, 159-170. doi:10.1007/978-3-319-13153-5_16, 2015

[3] J. Novakovic, P. Strbac, & D. Bulatovic. Toward optimal feature selection using ranking methods and classification algorithms. *Yugoslav Journal of Operations Research*, *21*(1), 119-135, 2011.

[4] M. A. Hall & L.A. Smith.. Practical feature subset selection for machine learning. Australia, 1998.

[5] C. Gayathri. Feature subset selection using filtering with Mutual information and maximal information coefficient. *International Journal of Innovative Research in Computer and Communication Engineering*, *2*(1), 1350-1354, 2014.

[6] N.V. Chawla, K.W.Bowyer, KW, Hall, LO, & W.P. Kegelmeyer, WP. . SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research, 16*, 321–357, 2002.

[7] U.M.Fayyad, & K. B. Irani. Multi-interval discretization of continuous valued attributes for classification learning. Chambery, France: 13th International Joint Conference on Artificial Intelligence, 1993.

[8] T. Jayalakshmi, & A. Santhakumaran. Statistical normalization and back propagation for classification. *IJCTE*, *3*(1), 89-93. doi:10.7763/ijcte.2011.v3.288, 2011.

[9] L. Breiman, Random Forests. Machine Learning,45,5-32, 2001.

[10] A. Kar (n.d.), "Stock Prediction using Artificial Neural Networks", Dept. of Computer Science and Engineering, IIT, Kanpur, http://people.eecs.berkeley.edu/~akar/IITK_website/EE671/report_stock.pdf

[11] Hall, M. A. . *Correlation-based feature selection for machine learning* (Doctoral dissertation, 1999, The University of Waikato, Hamilton, NewZealand). Retrieved from http://www.cs.waikato.ac.nz/~mhall/thesis.pdf

TABLE II
Performance Comparison of Different Models

Models	Accuracy	Precision	Recall	F-measure
Random Forest (Feature Selection)	47.46%	39.84%	41.55%	40.68%
Random Forest (Feature Selection + SMOTE)	73.73%	74.11%	73.45%	73.78%
Random Forest (Feature Selection + SMOTE+ Discretise)	57.37%	84.48%	91.54%	87.87%
Random Forest (Feature Selection + SMOTE + Discretise + Normalize)	82.57%	82.68%	89.55%	85.98%
ANN (Feature Selection)	70.34%	60.35%	52.69%	56.26%
ANN (Feature Selection + SMOTE)	71.43%	64.43%	51.32%	57.13%
ANN (Feature Selection + SMOTE+ Discretise)	89.45%	88.41%	83.74%	86.01%
ANN (Feature Selection + SMOTE + Discretise + Normalize)	89.48%	88.40%	83.75%	86.01%
Naïve Bayes (Feature Selection)	35.59%	35.86%	29.95%	32.64%
Naïve Bayes (Feature Selection + SMOTE)	64.98%	66.60%	64.98%	65.78%
Naïve Bayes (Feature Selection + SMOTE+ Discretise)	68.81%	66.45%	66.95%	66.70%
Naïve Bayes (Feature Selection + SMOTE + Discretise + Normalize)	68.81%	66.44%	66.94%	66.69%

Towards Concurrent Data Structure Development with Relaxed Synchronization

Naw Safrin Sattar,[1*] Tasmia Aqila[1] and Rifat Shahriyar[1]

[1]Department of Computer Science and Engineering, Bangladesh University of Engineering and Technology, Dhaka-1000, Bangladesh.

[*]1005099.nss@ugrad.cse.buet.ac.bd

Abstract— **Concurrent data structures may introduce a performance and scalability holdup and thus prevent the effective use of parallel hardware. There is a trade-off between scalable performance and precision in implementing concurrent data structures. A remedy to the scalability problem is to relax the interpretation of concurrent data structures. The construal is given by some notion equivalent to the sequential behaviour. The equivalence is determined by a consistency condition, most commonly linearity, and the sequential behaviour is inherited from the sequential version of the data structure (e.g. the sequential behaviour of a concurrent stack is a regular stack). Therefore, relaxing the semantics of a concurrent data structure extents to either weakening the consistency condition or redefining or relaxing its sequential specification. In this paper, we present a framework for relaxing the synchronization of a Stack data structure in a quantitative manner. A fundamental challenge in using relaxed synchronization is guaranteeing that the relaxed program always produces results with a specified quality. We propose a methodology that addresses this challenge in programming with relaxed synchronization. Using our methodology, results that are of the same quality as the original program, can be produced with a speedup in time.**

Index Terms— ***concurrency; read-write locks; synchronization; timeout***

I. Introduction

A. Synchronization

Synchronization in Java is an important concept since Java is a multi-threaded language and an essential tool in concurrent programming. It is the capability to control the access of multiple threads to any shared resource. At times when more than one thread try to access a shared resource, we need to ensure that resource will be used by only one thread at a time thus allowing us to protect, for example, variables or data from being corrupted by simultaneous modifications from different threads. The process by which this is achieved is called synchronization.

Java Synchronization is better option where we want to allow only one thread to access the shared resource. The **synchronized** keyword in java creates a block of code referred to as critical section. Every Java object with a critical section of code gets a lock associated with the object. To enter critical section a thread need to obtain the corresponding object's lock. If we do not use synchronization, and let two or more threads access a shared resource at the same time, it will lead to distorted results.

Using Synchronized methods is a way to accomplish synchronization. Java **synchronized** keyword provides following functionality essential for concurrent programming.

- provides locking which ensures mutual exclusive access of shared resource and prevent data race.
- prevents reordering of code statement by compiler which can cause subtle concurrent issue if we don't use **synchronized** or **volatile** keyword.
- involves locking and unlocking. before entering into synchronized method or block. Thread needs to acquire the lock. At this point, it reads data from main memory than cache and when it release the lock, it flushes write operation into main memory which eliminates memory inconsistency errors.

B. Java Lock

A lock is a thread synchronization mechanism like synchronized blocks except locks can be more sophisticated than synchronized blocks. The package **java.util.concurrent.locks** contain several lock implementations. Here are some methods that we have used in our work.

- ReentrantLock(): A **ReentrantLock** is owned by the thread last successfully locking, but not yet unlocking it. A thread invoking lock will return, successfully acquiring the lock, when the lock is not owned by another thread. The method will return immediately if the current thread already owns the lock.
- TryLock(): The tryLock() method attempts to lock the Lock instance immediately. It returns true if the locking succeeds, false if Lock is already acquired. This method never blocks. It acquires the lock only if it is free at the time of invocation.
- TryLock(timeout, timeUnit): The tryLock(long timeout, TimeUnit timeUnit) works like the tryLock() method, except it waits up the given timeout before giving up trying to acquire the Lock.

C. Difference between Lock Interface and synchronized keyword

- Having a timeout trying to get access to a synchronized block is not possible. Using **Lock.tryLock(long timeout, TimeUnit timeUnit)**, it is possible.
- The synchronized block must be fully contained within a single method. A Lock can have its calls to **lock()** and **unlock()** in separate methods.

D. Relaxed Synchronization

Relaxed Synchronization is the method to avoid synchronization for its huge computational time in a way such that our desired outcome originates within the minimum

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

possible time with the accuracy on which we may count on. An approximate data structure that execute without synchronization is designed with relaxed synchronization. It is an approach in which the approximate data structures are accurate enough for its clients to use successfully. This approach enables developers to reuse the basic data structures of stack and queue in a different form which have been designed to execute successfully without following any synchronized techniques ignoring the possibility of the presence of data losing.

II. Problem Formulation

Stack is the most commonly used data type for manipulating arbitrarily large collections of objects. Two basic operations of stack are:

- Insert a new item
- Remove an item

In our approach, we have suggested an implementation for stack using **Array** which enables thousands of threads to insert and remove items from the stack in the time less than that of synchronization. This helps to implement an approximately correct data structure of stack.

While working with threads, using synchronization in stack implementation, helps us to allow a single thread for inserting an item in the stack or removing an item from the stack at any instance. When the no of threads is small, this approach gives the accurate result quickly.

But when there is a large no of threads, it gives an accurate result at the cost of a huge time. The time needed for the large no of threads for pushing and popping element from the stack is high. As a result the program becomes slow and takes a lot of operating system overhead.

A. Motivation

These are a few reasons why we tried to avoid synchronization to speed-up our work sacrificing our accuracy a little bit.

1) Problems of Java synchronized keyword

- It doesn't allow concurrent read, which can potentially limit scalability. By using concept of lock stripping and using different locks for reading and writing, one can overcome this limitation. **java.util.concurrent.locks.ReentrantReadWriteLock** provides ready-made implementation of ReadWriteLock in Java.
- It can only be used to control access of shared object within the same JVM. If you have more than one JVM and need synchronized access to a shared file system or database, it is not at all sufficient. You need to implement a kind of global lock for that.
- It incurs performance cost. Synchronized method in Java is very slow and can degrade performance. So synchronization is done when it is absolutely required and we consider using synchronized block for synchronizing critical section only.
- Java synchronized code could result in deadlock or starvation while accessing by multiple thread if synchronization is not implemented correctly.
- According to the Java language specification one cannot use Java **synchronized** keyword with constructor. It's illegal and results in compilation error. So one cannot have synchronized constructor in Java which seems logical because other threads cannot see the object being created until the thread creating it has finished it.
- It cannot be applied to variables. Using the **volatile** keyword in Java on variables is a possible solution to this. **Volatile** only synchronizes the value of one variable between Thread memory and "main" memory while **synchronized** synchronizes the value of all variable between thread memory and "main" memory and locks and releases a monitor to boot. Due to this reason **synchronized** in Java is likely to have more overhead than **volatile**.
- It also synchronizes memory. In fact, it synchronizes the whole of thread memory with main memory.

2) Efficiency of Reentrant Lock over Synchronization

- Ability to make lock interruptible
- Ability to timeout while waiting for lock
- Power to create fair lock
- API to get list of waiting thread for lock
- Flexibility to try for lock without blocking

B. Related Works

Synchronization is a major overhead in terms of time. Several works have been done to relax synchronization for efficiency. A few of them are discussed here. Further, we also describe work done in the related area of approximate computing, where the performance improvement is obtained by allowing some degradation of quality of the results.

1) Relaxed Synchronization with Ordered Read-Write Locks

Ordered read-write locks, ORWL [1] provides new synchronization methods for resource oriented parallel or distributed algorithms for which it allows an implicit deadlock-free and equitable control of a protected resource and provides means to couple lock objects and data tightly.

Macro-benchmarks show that the library behaves well on multi-core machines and clusters; it realizes almost perfect computation/communication overlap and weak scaling properties.

2) Programming with Relaxed Synchronization

This work is based on using the original and relaxed versions, together with the quality criteria, in a structured way that guarantees the same acceptable quality of results for all executions of the relaxed version [2].

- It proposes a methodology to systematically relax synchronization and achieve significant performance speedups without sacrificing quality of results
- It formulates a simple program restructuring scheme, Relax and Check (RaC) that casts the problem of relaxing synchronization as a performance optimization, thereby providing a systematic way to produce quality results by sacrificing performance in the expected infrequent cases
- It provides the results of an extensive evaluation of the proposed methodology for relaxing synchronization on a variety of benchmarks and on two major parallel platforms, showing significant speedups

3) Parallel Synchronization-Free Approximate Data Structure Construction

Approximate data structures with construction algorithms that execute without synchronization is

presented here. The data races present in these algorithms may cause them to drop inserted or appended elements. Nevertheless, the algorithms

- do not crash
- may produce a data structure that is accurate enough for its clients to use successfully

It is an approach in which the approximate data structures are composed of basic tree and array building blocks with associated synchronization-free construction algorithms [3]. This approach enables developers to reuse the construction algorithms, which have been engineered to execute successfully in parallel contexts despite the presence of data races, without having to understand the details of why they execute successfully. Moreover, unlike much early research in the field [4], the algorithms do not use reads and writes to synthesize higher-level synchronization constructs.

4) Data Privatization

Privatization [5] [6] is another important technique used to eliminate storage-based dependencies in parallel loops and thereby reduce the number of synchronization events needed. Variable privatization [5] is used by parallelizing compilers to improve exploitation of parallelism in the program. For example, scalar privatization removes access conflicts for a scalar variable modified in different iterations of a parallel loop by allocating a private storage for that variable in each thread. Array privatization [6], an extension of scalar privatization, creates a private copy of the array in each thread when it can be determined that such privatization is safe. Privatization eliminates the synchronization that would be overhead incurred if atomic updates were used instead to share variables. However it comes at the cost of memory for data duplication and possibly communication overhead between threads [7].

5) Unsynchronized Techniques for Approximate Parallel Computing

Techniques for obtaining acceptable unsynchronized parallel computations is presented here. Even though these techniques may generate interactions (such as data races) that the current rigid value system condemns as incorrect, they are engineered to

- preserve key data structure consistency constraints
- produce a result that is accurate enough, often enough

As they contain no synchronization, they eliminate the synchronization overhead, parallelism reduction, and failure propagation typically associated with synchronization mechanisms [8].

Besides these, Relaxed Synchronization Message Passing, Loose Synchronization for Large- Scale Networked Systems, Relaxing Synchronization in Distributed Simulated Annealing, Relaxing Synchronization Constraints in Behavioral Programs [9] etc. are some of the notable works on relaxed synchronization to mention.

III. Design and Implementation

Synchronization inherently limits parallelism. As a result, there is a recent trend towards implementing semantically weaker data structures which reduce the need for synchronization and thus achieve better performance and scalability. We have considered possible relaxations of stacks, and examined their relative computational power.

A. Simple Lock

At first, we used simple write lock to ensure that push operation of stack is done correctly. We used sleep to ensure that if it works correctly for multi-thread.

- Without sleep it gives desired result within a very small time but this is not what we wanted.
- With increasing sleep time, the processing time also increases significantly which is similar to using synchronization along with there exists some error.
- With increasing number of threads error increases.

B. Reentrant Lock

Next, we used reentrant lock with **tryLock()** method.

- We got errors but could not figure out what was causing this. So we tried to identify the threads which were actually doing the push operation by keeping a count of them.
- We tried to identify the timeout at first using a minimum and maximum bound and then guessing the optimal timeout in a trial and error manner.
- Whenever error increased, we increased timeout and vice-versa.
- It was not the right approach to identify the correct timeout. So we adopted an adaptive approach to measure timeout.

C. Use of Adaptive trylock(timeout)Reentrant Lock

We determined timeout depending on error we got. We calculated the error from how many times the push function is actually called and how many times an object is actually pushed in the stack, shown in (1). We kept a threshold error which was the maximum error we would allow for the program. It's time-consuming to check error at every iteration and it's costly too. So we took a sample number and calculated error after that no of iteration. We used a fixed amount of time to add to or deduct from timeout to increase or decrease timeout.

$$currentErrorRate = 100 \times \frac{|functionCalled - pushCount|}{functionCalled} \quad (1)$$

- We got significant amount of errors but were unable to identify what was causing this.
- We had used a fixed amount of time to add or deduct which was not correct. It should be a factor of current error and timeout.
- Java executor service does not allow lock and sleep inside try-lock to work together. This was the main problem leading to error.
- So now instead of sleep we had to use something else.

D. Use of Adaptive trylock(timeout)Reentrant Lock with Simple I/O Operation

Now we have eliminated sleep which was our problem and performed an I/O operation. We did a simple I/O operation to elapse time. A huge file was read at the interval. The time to read the file is set as initial ***timeout***, shown in (2). Again, we have adjusted timeout as a factor of elapsed time in I/O operation, current error and current timeout, shown in (3). This works fine as we wanted in approximate computing.

$$timeout = fileReadTime \quad (2)$$

$$timeChange = |functionCalled - pushCount| \times factor \times currentTimeout \quad (3)$$

- We get a minimum error which is cost efficient.
- The processing time is faster than synchronization.

IV. Experimental Setup

We used a script to run our codes. We had used Linux as the environment. We used Java as the coding platform. We ran our codes in a 32 bit processor with 4 GB RAM. The source code can be found in Github Repository [10].

V. Results

A. Comparison between Non-Synchronized, Synchronized and Lock for Different Sleep Time

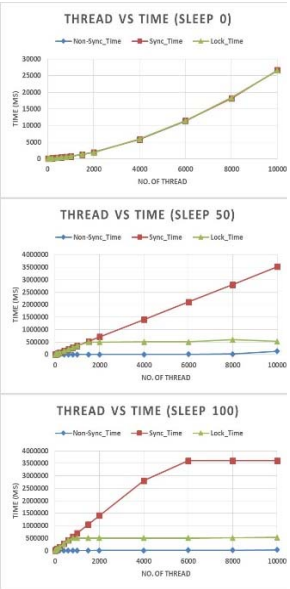


Fig. 1 Execution Time

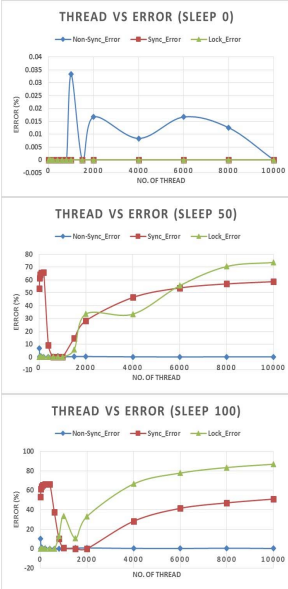


Fig. 2 Percentage Error

For both Fig. 1 and Fig. 2, we have varied sleep time to see what the execution time and percentage error is for several threads. In Fig. 1 synchronization takes the highest time with increasing sleep time. In Fig. 2 we get no error for both lock and synchronization with sleep time 0. As sleep time increases, both lock and synchronization error increases linearly for higher no of threads but for smaller no. of threads there is a deviation. The reason behind it is not clear enough. Lock gives such higher error because of Java Executor Service that doesn't allow lock and sleep in the same try-catch block.

B. Comparison between Lock, Synchronized and Adaptive I/O Method

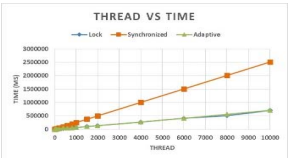


Fig. 3 Execution Time

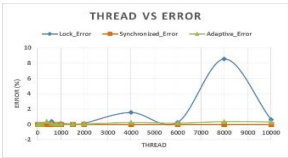


Fig. 4 Percentage Error

In Fig. 3 synchronization takes a huge amount of time but both lock and Adaptive method takes similar time. Why we have chosen adaptive over lock can be clearly seen in Fig. 4. Fig. 4 tells that synchronization gives no error but takes a lot of time that we want to eliminate. As both lock and adaptive take same time, we have to decide depending on error. Adaptive gives the minimum error which is allowable in efficient time that is shown in Table I. It is **3x** faster than synchronization.

TABLE I
Performance of Adaptive Method with I/O

No. of Threads	Pushed	Actually Pushed	Time	Error (%)
100	500	500	7153	0
200	1000	1000	14448	0
1000	4000	3999	54917	0.025
4000	10000	9993	135733	0.07
8000	30000	29949	411520	0.17
10000	50000	39856	558935	0.36

We tried different approach but all of those didn't give us the expected output. At last our last approach Adaptive method with I/O gives us our desired result with a better performance that we sought after.

VI. Conclusions

Synchronization overhead is a major performance-limiting factor in parallel applications. In this work, we have shown that significant performance speedups can be obtained by relaxing the present assumption of synchronizations which are essential to produce good quality results. We proposed an adaptive timeout scheme, which allows to systematically relax synchronization to achieve performance improvement and still produce results that are of the same acceptable quality as the original. The positive results presented are a strong indication of the potential of the general area of approximate computing, where we sacrifice the determinism and preciseness of the general computing exemplar as adept today for improved latency, reduced energy consumption, and lower system cost, providing acceptable results. It inspires us to use relaxed synchronization in other similar data structures such as queue.

References

[1] M. Alexander et al. (Eds.), *Euro-Par 2011: Parallel Processing Workshops*, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, 2012, vol. 7155, pp. 387-397.

[2] L. Renganarayana, V. Srinivasan, R. Nair and D. Prener, "Programming with Relaxed Synchronization", in *Proc.* RACES'12, ACM workshop on Relaxing synchronization for multicore and manycore scalability, 2012, pp. 41-50.

[3] M. Rinard, "Parallel Synchronization-Free Approximate Data Structure Construction", in HotPar '13, the 5th USENIX Workshop on Hot Topics in Parallelism, 2013.

[4] R. L. Ashenhurst Ed., *Communications of the ACM*, NY, USA, ACM New York, Aug 1974, vol. 17, Issue 8, pp. 453-455.

[5] J. Lenfant Ed., *ICS '88 Proceedings of the 2nd international conference on Supercomputing*, NY, USA, ACM New York, 1988, pp. 207-215.

[6] U. Banerjee, Ed., *ACM International Conference on Supercomputing 25th Anniversary Volume*, NY, USA, ACM New York, 2014, pp. 173-182.

[7] M. Gupta, "On privatization of variables for data-parallel execution" in 11th International *Proc*. Parallel Processing Symposium, 1997, pp. 533-541.

[8] M. Rinard, "Unsynchronized Techniques for Approximate Parallel Computing", in *Proc*. RACES'12, ACM workshop on Relaxing synchronization for multicore and manycore scalability, 2012.

[9] K. McMillan et al. (Eds.), Logic for Programming, Artificial Intelligence and Reasoning, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, 2013, vol. 8312, pp. 355-372.

[10] Github Repository [Online] Available: https://github.com/nawsafrin/Towards-Concurrent-Data-Structure-Development-with-Relaxed-Synchronization

Estimation of Optimum Tilt Angle for PV Cell: a Study in Perspective of Bangladesh

Mohammad Nadim,[1,*] Muhammad Rashedul Haq Rashed,[1] Anik Muhury[1] and Sharif Mohammad Mominuzzaman [1]
[1]Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology
Dhaka-1205, Bangladesh

[*]mohammad.nadim.eee09.buet@gmail.com

Abstract— **Bangladesh, a subtropical country located not very far from the equator, has certain geographical advantages in solar energy generation. But, as the solar energy generation is not as robust as other forms of energy, the solar systems must be optimally installed to reap every physical benefits possible. One very important factor in this matter is the tilt angle of the solar panel. This paper focuses on finding the optimum tilt angle for a solar panel in a particular region. The paper finds the optimum tilt angles for various divisions of Bangladesh using two methods: one optimum tilt angle for a whole year, and two tilt angles for two halves of a year. The paper demonstrates the benefit of using two seasonal tilt angles instead of using just one tilt angle throughout the year. The paper additionally checks the extent of benefit that might occur from using more complex systems involving more tilt angles.**

Index Terms— **Optimum tilt angle, PV cell, Bangladesh, solar energy, latitude.**

I. Introduction

Being the eco-friendly substitute of fossil fuel and practically inexhaustible in nature, the renewable energy has seen a revolution in past few decades. Renewable energy sources like hydraulic systems, wind power systems, photovoltaic systems, solar thermal systems and biomass plants are now actively contributing to a significant portion of total power grid in many countries [1]. Solar energy is one of the most popular systems of renewable energy. And with the continuous development in solar technology and the resultant drastic decrease in installation cost, solar energy generation has never seemed more attractive.

Bangladesh, a developing country with a vast population, is fighting its ever increasing energy demand. Following the footstep of many developed countries, Bangladesh is embracing various methods of energy generation to check this demand. Solar energy, being one of the most adopted form of renewable energy system, has been gaining popularity throughout the country. Geography is also favoring Bangladesh in this matter.

To achieve the maximum energy output from a solar system, the proper installation of the panel is very important. Selecting an ideal tilt angle is a key factor in this matter. Bangladesh, like many other countries, follows a convention of using one tilt angle of solar panel throughout the year not regarding the variation of seasons. This default tilt angle is nearly equal to the average latitude of Bangladesh [2]. This method, if done by selecting an optimum tilt angle, can produce an adequate amount of solar energy. However, there is space for much improvement in this system by using the sun tracking method [3]. Sun tracking method, however, is quite complex to install. The technology is also very expensive [4]. Hence, a cheaper and simpler alternative is of demand. One very easy improvement can be done by adopting two tilt angles for two halves of a year.

This paper deals with finding the optimum tilt angle in various latitudes of Bangladesh. It focuses on both the conventional one tilt angle for the whole year system and the proposed two tilt angles for two halves of a year system. Using the standard formulation, a simulation was performed. From the output of the simulation a comparison was done between the conventional and the proposed system. Then the paper takes the study a bit further; it checks how beneficial a more complex system might be. The paper compares its proposed two tilt angles system with a four tilt angles system. Finally, the one best system was proposed considering the minimum increase in complexity of installation and resultant increase in output.

II. Formulation

Standard equations were used for determining the relation between tilt angles and received solar radiation on a tilted surface. The equations are briefly described below.

First, the solar intensity was calculated using Air Mass, which accounts for the direct optical path length through earth's atmosphere. The simplified formula for Air Mass calculation is: [5]

$$\mathrm{AM} = \frac{1}{\mathrm{Cos}(\theta)} \tag{1}$$

where, θ is the angular height of the sun in the sky calculated from the vertical, known as zenith angle. θ is calculated as follows:

$$\theta = 90° - \text{elevation angle} \tag{2}$$

Now, the direct component of solar intensity is calculated using the following equation: [6]

$$\mathrm{I} = 1.353 * 0.7^{\mathrm{AM}^{0.678}} \tag{3}$$

where, I is the intensity of the sunray on a perpendicular plane and calculated in units of kW/m^2. To account for the diffuse radiation component, which is nearly 10% of the direct radiation component, the formula for solar intensity is modified as follows: [7]

$$\mathrm{I} = 1.1 * 1.353 * 0.7^{\mathrm{AM}^{0.678}} \tag{4}$$

Finally, the relation between tilt angles and received solar radiation is established using the following formula [8]

$$\text{Solar Radiation} = \mathrm{I} * \mathrm{Sin}(\alpha + \beta) \tag{5}$$

where, β is the tilt angle of the module calculated from the horizontal, and α is the elevation angle which accounts for the angular height of the sun in the sky calculated from the horizontal. Elevation angle is hemisphere dependent. At countries like Bangladesh, which is in North hemisphere, the elevation angle is calculated as follows:

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

$$\alpha = 90° - \text{latitude angle} + \text{declination angle} \quad (6)$$

The declination angle, δ is calculated using the following formula: [9]

$$\delta = 23.45° * \mathrm{Sin}(\frac{360}{365} * (d - 81)) \quad (7)$$

where, d is the number of a day in a year that starts with January 1 as d = 1.

One last thing taken into account while calculating solar radiation is the amount of daytime in a 24 hours period. The sun hour is calculated using the following formulas:

$$\mathrm{Sunrise} = 12 - \frac{1}{15°}\,\mathrm{Cos}^{-1}\left(\frac{-\mathrm{Sin}\varphi\mathrm{Sin}\delta}{\mathrm{Cos}\varphi\mathrm{Cos}\delta}\right) \quad (8)$$

$$\mathrm{Sunset} = 12 + \frac{1}{15°}\,\mathrm{Cos}^{-1}\left(\frac{-\mathrm{Sin}\varphi\mathrm{Sin}\delta}{\mathrm{Cos}\varphi\mathrm{Cos}\delta}\right) \quad (9)$$

$$\mathrm{Sun\ hour} = \mathrm{Sunset} - \mathrm{Sunrise} \quad (10)$$

where, φ is the latitude of the place and δ is the declination angle.

III. SIMULATION

The simulation was done using scripting language *Jquery* and the output was plotted using *Matlab*. The sample simulation results for Dhaka, the capital of Bangladesh (Latitude- 23.77°), is presented here.

A. Trend of change in tilt angles throughout the year

A simulation was done to observe the trend of change in array's optimum tilt angle, day by day, throughout the year. The result of the simulation is shown in Fig 1. In the figure the optimum tilt angle for an array has been drawn as a function of the days of a year ranging from day 1 to day 365.

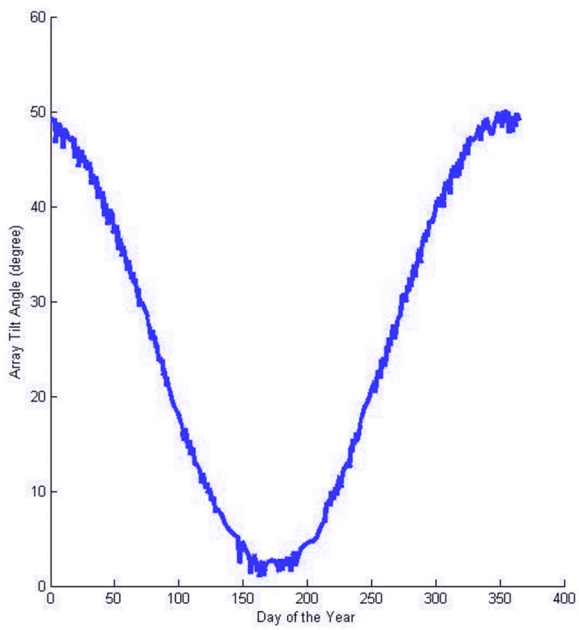


Fig. 1 Array's optimum tilt angle in degree is drawn as a function of different days of the year (ranging from 1 to 365)

B. One optimum tilt angle for the whole year

This part of simulation was done to find one optimum tilt for a whole year. The result of the simulation is presented graphically in Fig 2. In the figure, total radiation for a year was plotted as a function of tilt angles. For various tilt angles ranging from 0 degree to 90 degree, the resultant radiations was calculated and summed up for a year. Then the tilt angle for which the total radiation is maximum was determined. The light blue dot in the figure shows the maximum radiation point. The result of this simulation can be summarized as follows:

Optimum tilt angle= 21.4°

Maximum total radiation = 3275.5 Kwh/m^2/year

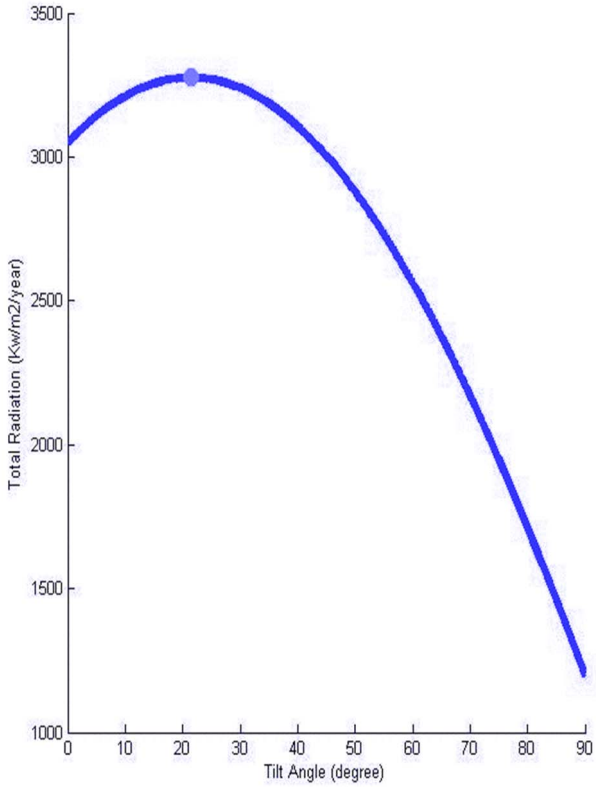


Fig. 2 Total Radiation (kwh/m^2/year) is drawn as a function of Tilt Angles in degree (For a twelve month period from January to December)

The validity of this simulation can be checked comparing with the default tilt angle used in Bangladesh. Bangladesh ideally uses a tilt angle nearly equal to its latitude [2]. The default tilt angle and the simulated tilt angle are very close. Hence, the simulation can be considered valid.

C. Two seasonal optimum tilt angle

This part of simulation was done for finding a system with two tilt angles, each of which will be used for a six months period. For this purpose, the 12 months of a year was divided into 2 periods; first being the period from April to September and the second being the period from October to March.

1) *Period- April to September:* The result of this part of simulation is presented graphically in Fig 3. In the figure, total radiation for six months starting from April was plotted as a function of various tilt angles. For tilt angles ranging from 0 degree to 90 degree, the resultant radiations was calculated and summed up over six months. Then, the tilt angle for which the total radiation is maximum in this period was determined. The light blue dot in the figure shows the maximum radiation point. The result of this simulation can be summarized as follows:

Optimum tilt angle= 8.7°

Maximum total radiation= 1895.1 Kwh/m^2/year

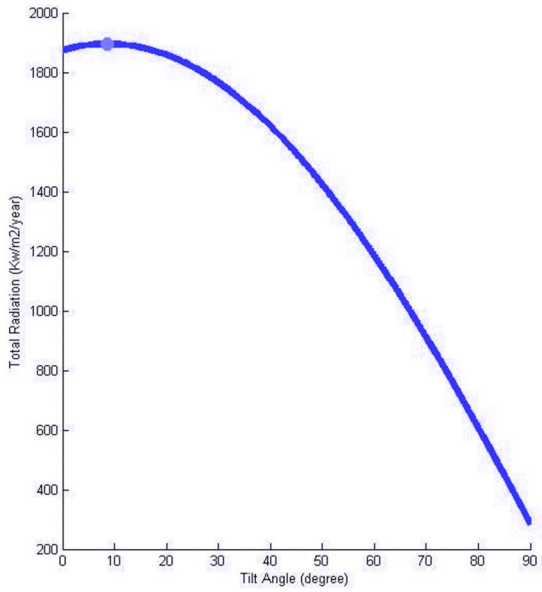


Fig. 3 Total Radiation (kwh/m^2/year) is drawn as a function of Tilt Angles in degree (For a six month period from April to September)

2) *Period- October to March:* The result of this part of simulation is presented graphically in Fig 4. In the figure, total radiation for six months starting from October was plotted as a function of various tilt angles. For tilt angles ranging from 0 degree to 90 degree, the resultant radiations was calculated and summed up over six months. Then, the tilt angle for which the total radiation is maximum in this period was determined. The light blue dot in the figure shows the maximum radiation point. The result of this simulation can be summarized as follows:

Optimum tilt angle= 37.9°
Maximum total radiation= 1487.9 Kwh/m^2/year

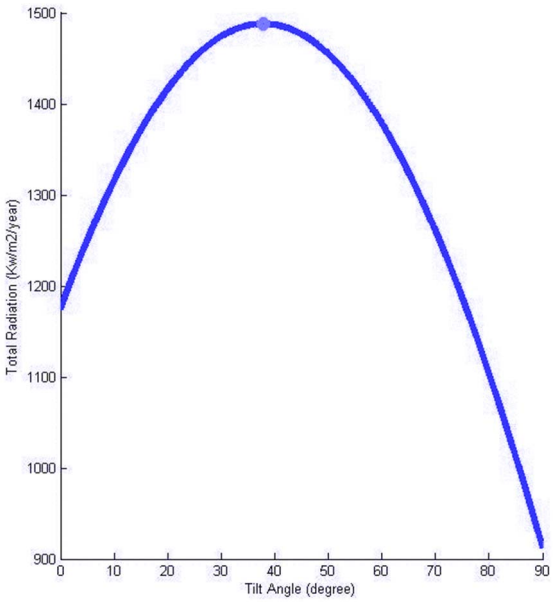


Fig. 4 Total Radiation (kwh/m^2/year) is drawn as a function of Tilt Angles in degree (For a six month period from October to March)

Apart from Dhaka, the same simulation was also performed for the rest of the divisions of Bangladesh. The summery of the simulation is shown in the Table I:

TABLE I
OPTIMUM TILT ANGLE FOR EIGHT DIVISIONS OF BANGLADESH FOR ONE TILT ANGLE SYSTEM AND TWO TILT ANGLES SYSTEM

	One Tilt Angle for the Whole Year System	**Two Tilt Angles for Two Halves of a Year System**	
Division	**Optimum Tilt Angle (January-December)**	**Optimum Tilt Angle (April-September)**	**Optimum Tilt Angle (October-March)**
Dhaka	21.40°	8.70°	37.90°
Chittagong	20.20°	7.20°	36.40°
Rajshahi	22.10°	9.30°	38.50°
Khulna	20.70°	7.80°	37.00°
Barisal	20.60°	7.70°	36.90°
Sylhet	22.50°	9.80°	39.00°
Mymensingh	22.40°	9.70°	38.90°
Rangpur	23.30°	10.70°	39.80°

IV. COMPARISON

The simulation result is compared on the basis of maximum solar radiation received by any system. The total radiation data for the two halves of two tilt angles system was summed up and these data were compared with the ones found using one tilt angle systems. The detailed comparison of the two systems for various divisions of Bangladesh is presented in Table II. The difference in total radiation for two systems are shown in the fourth column of the table. From that column, it can be seen that using two tilt angles system results in almost 107 Kwh/m^2/year increase in total radiation.

TABLE II
TOTAL RADIATION COMPARISON BETWEEN ONE TILT ANGLE SYSTEM AND TWO TILT ANGLES SYSTEM

Division	**Total Radiation for One Tilt Angle System in Kwh/m^2/year**	**Total Radiation for Two Tilt Angles System in Kwh/m^2/year**	**Difference of Total Radiation Between Two Systems in Kwh/m^2/year**
Dhaka	3274.72	3382.21	107.49
Chittagong	3291.75	3400.23	108.48
Rajshahi	3267.82	3374.91	107.09
Khulna	3285.34	3393.45	108.11
Barisal	3286.53	3394.71	108.18
Sylhet	3261.80	3368.52	106.72
Mymensingh	3263.47	3370.29	106.82
Rangpur	3251.16	3357.25	106.09

V. A PEEK INTO MORE COMPLEX SYSTEMS

Now that it is evident that, systems involving more than one tilt angle receives better solar radiation, some more inspection needed to be done to find the extent of this phenomena. The concern here is, whether the received radiation keeps increasing at this rate if the number of tilt angles are further increased. To find this out, one more simulations were performed with respect to Dhaka (Latitude- 23.77°). This simulation involves four tilt angles for four quarters of a year.

A. Four optimum tilt angle for four quarters

For the simulation of four tilt angle system, the months of a year was divided into following quarters: January to March, April to June, July to September, and October to December. The optimum tilt angles for these quarters and the subsequent maximum total radiation data is presented in Table III.

TABLE III
OPTIMUM TILT ANGLE AND RESPECTIVE MAXIMUM TOTAL RADIATION FOR A FOUR TILT ANGLES SYSTEM IN DHAKA (LATITUDE- 23.77°).

Month	Optimum Tilt Angle (degree)	Maximum Total Radiation (Kwh/m^2/year)
January to March	35.30°	760.37
April to June	6.40°	959.04
July to September	11.00°	937.66
October to December	40.60°	729.17
	Total	3386.24

Now, one final comparison can be done between the systems with two tilt angles and four tilt angles for the latitude of Dhaka. Table IV shows such a comparison. The table compares these two systems with the conventional one tilt angle system. It shows how much more solar energy can be received if any of the two systems is used instead of a one tilt angle system.

TABLE IV
COMPARISON BETWEEN THE TWO TILT ANGLES SYSTEM AND THE FOUR TILT ANGLES SYSTEM WITH RESPECT TO ONE TILT ANGLE SYSTEM

	Two Tilt Angle System	Four Tilt Angle System
Total Solar Energy Increase from One Tilt Angle System (Kwh/m^2/year)	107.5	110.7
Percentage Increase in Solar Energy from One Tilt Angle System	3.3%	3.4%

VI. CONCLUSION

From the study shown in this paper, it is evident that- with the increase of number of tilt angles, total radiation of a system also increases. But, from the Table IV, it is also realised that- this increase of solar radiation has no linear relation with number of tilt angles. While using a two tilt angle system results in considerable increase in solar radiation from one tilt angle system, a four tilt angles system shows comparatively less increase in solar radiation. Moreover, the installation and maintenance of latter system will be far more complex than a two tilt angles system. Hence, it can be concluded that, if a solar system adopts a two tilt angles system instead of the default one tilt angle system, it will witness a significant increase in radiation with only a minimum increase in installation complexity. As mentioned at the beginning of this paper, the solar radiation lacks the vigorous nature that many other energy sources have. To compensate for this constraint, this kind of modification in tilt angle can contribute a lot. To help finding the optimum two seasonal tilt angle in any location of Bangladesh, a site [10] has been created. The site has been made free for all.

REFERENCES

[1] Sayigh A. *Renewable energy—the way forward*. Applied Energy 1999;64:15–30. 1999

[2] Alimul Haque Khan ; Maidul Islam ; Asif Islam ; Mohammad Saifur Rahman.*"A systematic approach to find the optimum tilt angle for meeting the maximum energy demand of an isolated area"*. Electrical Engineering and Information Communication Technology (ICEEICT), International Conference, Dhaka, Bangladesh, 2015.

[3] Hossein Mousazadeha; Alireza Keyhania; Arzhang Javadib; Hossein Moblia; Karen Abriniac; Ahmad Sharifib. *A review of principle and sun-tracking methods for maximizing solar systems output.* Department of Agricultural Machinery Engineering, University of Tehran, Renewable and Sustainable Energy Reviews 13, 1800–1818 (2009)

[4] Ryan Wiser; Galen Barbose Carla Peterman; Tracking the Sun: *The Installed Cost of Photovoltaics in the U.S. from 1998-2007*, Lawrence Barkeley Natioanl Laboratory, 2009

[5] Gueymard, C.; Myers, D.; Emery, K. *"Proposed reference irradiance spectra for solar energy systems testing"*. Solar Energy 73 (6): 443-467, (2002).

[6] PVCDROM retrieved, Stuart Bowden and Christiana Honsberg, Solar Power Labs, Arizona State University, 1 May 2011

[7] Meinel, A. B. and Meinel, M. P. Applied Solar Energy Addison Wesley Publishing Co., (1976)

[8] Photovoltaic Education Network website: http://www.pveducation.org/ accessed in 2015.

[9] Cooper PI. *The absorption of radiation in solar stills. Solar Energy* [Internet]; 12:333 - 346. 1969

[10] Author's website: www.nadim.netai.net/Thesis ; created in: 2015

Delay Analysis of SWCNT Bundle Interconnect for Different Technology Nodes

Shouvik Musavvir,* Prattay Chowdhury, and Sharif Mohammad Mominuzzaman

Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology, Dhaka-1000, Bangladesh

* shouvikmusavvir@gmail.com

Abstract— **In this paper the delay occurring in SWCNT bundle interconnect is analysed based on process related parameters and it is also compared to copper wire. The circuit parameters are calculated to develop an equivalent RLC model for both SWCNT and copper considering the practical constraints. This model is used to calculate the delay for local, intermediate and global interconnects by varying the bundle density and contact resistance. The relative effect of circuit parameters is analysed for different technology nodes. This paper also points out the most important parameter affecting the performance. Similarly, the equivalent circuit model for copper interconnect is developed and the delay is determined. The comparison of copper and CNT interconnect shows that the delay for CNT can be up to 2 times more than copper in local length. The delay of CNT interconnects in 16nm, 22nm and 45nm is less than copper in intermediate range. For global interconnects, the SWCNT delay is 0.3-0.5 times less than the delay of copper wire.**

Index Terms—**SWCNT bundle interconnect, RLC equivalent model, DIL method, Delay analysis.**

I. Introduction

Since the birth of semiconductor industry, transistors were the centre of analysis and interconnect effects were considered to be of minor importance. Due to aggressive scaling of technology nodes, the gate length is getting smaller. As a result, the transistor delay is decreasing too. Whereas, interconnect delay due to reduced spacing is on the rise. So decreasing the interconnect delay is a major concern. Copper's limited current carrying capability and electromigration issues are the main impediment in this aspect. Moreover, lattice impurities and defects, grain boundary and surface scattering contribute to high resistivity in nanometer range [1]. For other interconnect metals like aluminium these effects are much worse. In recent years, the carbon nanotube has emerged as the most desirable candidate to replace metal interconnects in the semiconductor industry. It has higher current carrying capability ($>10^9$A/cm^2) [2] and high thermal conductivity up to 3500W/m-K [3], which makes it an ideal candidate for next generation chip fabrication requirements. Furthermore, it is ten times more strong than copper. So thermal stress induced deformation and diffusion is a minor problem. But the imperfect metal-CNT contact resistance is a major problem in this aspect. It can be as high as 116kΩ [4] for a single SWCNT, which severely degrades the performance. Instead, hundreds of SWCNT are used together as a bundle, where they act like a single conductor so the equivalent resistance becomes much smaller. The production of metallic nanotube bundle with high density is also a challenge for implementation. In general, one-third of the total tubes in a bundle are metallic. However, in sophisticated production process up to 90% metallic CNTs is reported [5]. So high contact resistance of metal-CNT bonding and the percentage of metallic CNTs in a bundle still remains a fabrication process related issue.

In comparison with earlier reports [6-7], this work presents a detailed analysis considering practical limitations. Rather than using equivalent resistance and capacitance of just one technology node, five transistor models for present and future implementation are used in the calculation. In this paper, a comparison of SWCNT and copper interconnect delay is done for local, intermediate and global lengths as well as the effect of metallic CNT probability which will be reported for the first time.

The paper is organized as follows– section II is on the development of RLC equivalent model of SWCNT interconnect and copper, bundle density estimation is discussed in section III followed by delay determination and comparison in section IV.

II. Circuit Model For SWCNT Bundle and Copper

The prototype equivalent circuit is developed in accordance with the fabrication and physical constraints. This model takes the effect of limited density of states, one-dimensional flow, scattering effect, capacitive interaction, mean free path, bundle density and contact with metal into account. The basic RLC components are derived by considering a metallic SWCNT on a ground plane [8].

There are four parallel conductive channels in an SWCNT. Each band is spin degenerate which means it can transport a spin-up and spin-down electron. The one-dimensional flow and quantum point contact contribute to quantum resistance. It is determined by two-terminal Landauer-Buttiker formula $R_Q=h/4e^2=6.45k\Omega$, [9] considering perfect contact at both ends. This ballistic conduction occurs for length less than mean free path (λ_{CNT}=1μm) [10]. If the interconnect length is greater than mean free path, the nanotube resistance is $(h/4e^2)L/\lambda_{CNT}$, which consists of the quantum resistance and scattering resistance appearing for every mean-free-path length. Also, the contact resistance (R_c) is added at both ends of the terminal which is dependent on the fabrication process. Now the total resistance is divided by bundle density (N_{CNT}) as the nanotubes are connected in parallel. The equivalent resistance of the bundle for lengths greater than mean free path [11] is calculated by the following equation.

$$R_{bundle} = \frac{\frac{h}{4e^2}\left(\frac{L}{\lambda_{CNT}}\right)+R_C}{N_{CNT}} \quad (1)$$

Two types of inductance are present in CNT. The magnetic inductance is the self-inductance due to current flow variation

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

in the channels of the tube. The kinetic inductance is related to the energy system of the carrier. The value of kinetic inductance is $h/2e^2v_F=16nH/\mu m$, here fermi velocity for CNT is $8x10^5$ m/sec [12]. Another important fact is- kinetic inductance occurs only for zero voltage drop in the interconnect and for length less than mean free than path [13].

The magnetic inductance is expressed as

$$l_M = \frac{\mu}{2\pi}\ln(\frac{H}{d}) \tag{2}$$

where H is the distance from the ground plane and d is the diameter.

The total inductance for a bundle is expressed as

$$L_{Bundle} = \frac{L_{CNT}}{N} = \frac{l_M + l_K}{N} \tag{3}$$

A single nanotube also has quantum capacitance and electrostatic capacitance. Electrostatic capacitance is caused by charge accumulation with respect to a ground plane and quantum capacitance is caused by the low dimensionality of the system and limited charge carriers and states. This is related to the voltage rate change with extra electron flow in the existent system. The value of quantum capacitance is $2e^2/hv_F \approx 100aF/\mu m$ [14] and for a bundle, it is expressed as the following equation for equivalent capacitance in parallel.

$$C_{q(bundle)} = C_q.n_{CNT} \tag{4}$$

The electrostatic capacitance for CNT is not as straightforward as other parameters. The nanotubes in the bundle have the same potential. Due to this, there is no coupling capacitance between tubes in the same bundle. The electrostatic capacitance of the bundle arises mainly from the CNTs at the edge of the bundle rather than those of inside. Researchers used FastCap simulator for capacitance extraction [15] conforming to the geometrical configuration. This led to an equation for the estimation of bundle electrostatic capacitance [16].

$$C_{e(bundle)} = 2C_{en} + \frac{n_w - 2}{2}C_{ef} + \frac{3(n_H - 2)}{5}C_{en} \tag{5}$$

Where the number of metallic CNT along width is n_W and height is n_H. C_{en} and C_{ef} are the capacitances of isolated CNT with respect to near and far neighbouring interconnects respectively.

C_{en} and C_{ef} are estimated as

$$C_{en} = \frac{2\pi\varepsilon}{\ln(s/d)} \tag{6}$$

$$C_{ef} = \frac{2\pi\varepsilon}{\ln(2s/d)} \tag{7}$$

where s is the distance from ground plane equal to the distance between interconnects.

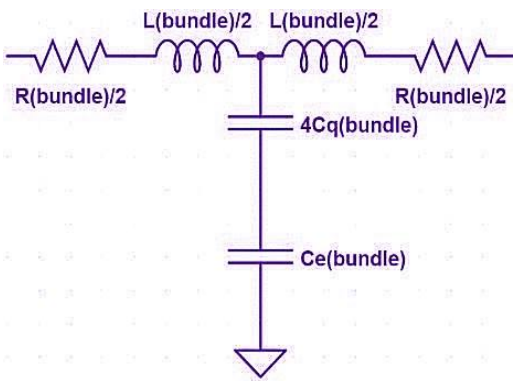


Fig. 1 Equivalent circuit model of SWCNT bundle.

The equivalent circuit is defined as a 'T-Model'. Total bundle resistance and inductance are equally divided at the incoming and outgoing nodes. The central node has quantum and electrostatic capacitance in series. The quantum capacitance is fourfold because of four conduction channel per SWCNT.

For copper interconnect, three segment π model is used as RLC equivalent circuit. The resistivity of copper increases due to surface scattering and grain boundary scattering. Now these two effects are represented by the following equations, where ρ_o is the traditional bulk resistivity of copper, ρ_s and ρ_g is the resistivity related to surface and grain boundary scattering effect respectively, λ is the mean free path of copper, $p=0.6$ is the scattering parameter, w is the width of wire, d is the mean distance between grain boundaries and $R=0.5$ is the reflectivity coefficient [17].

$$\frac{\rho_s}{\rho_o} = 1 + \frac{3}{4}(1-p)\frac{\lambda}{w} \tag{8}$$

$$\frac{\rho_o}{\rho_g} = 1 - \frac{3\alpha}{2} + 3a^2 - 3\alpha^3\ln(1+\frac{1}{\alpha}) \tag{9}$$

$$\alpha = \frac{\lambda}{d}\frac{R}{1-R} \tag{10}$$

The resistance, capacitance and self-inductance are determined by the following formula for copper wire with length l, width w and height t [18],

$$R = \frac{(\rho_s + \rho_g)l}{wt} \tag{11}$$

$$C_g = \varepsilon(0.74(\frac{w}{w+1.59s})^{1.14} + 1.14\frac{t}{s}(\frac{h}{h+2.06s})^{0.09} + 1.17(\frac{s}{s+1.51h})^{0.76}(\frac{t}{t+4.53h})^{0.12}) \tag{12}$$

$$L = \frac{\mu_0}{2\pi}(\ln(\frac{2l}{w+t}) + \frac{1}{2} + \frac{0.22(w+t)}{l}) \tag{13}$$

III. Bundle Density Calculation

The bundle density is one of the most important factors in delay analysis. The densely packed model used by various researchers is incorrect as there is always some distance between two metallic SWCNTs in fabrication process [17]. So the hexagonal sparsely packed model with uniform separation s is used here for maximum density calculation. The distance between two metallic CNTs depends on the fraction of metallic CNT in the bundle. This can be modelled as- $d/\sqrt{p}$, where p is the probability of metallic CNT in the bundle [19].

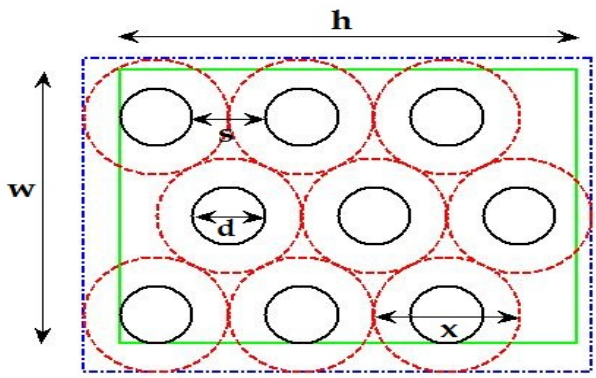


Fig. 2 Sparsely packed bundle.

To determine the number of tubes (N_{CNT}) fit in an area with a width w and height h, $x=d+s$ is defined for the following equations where d is the diameter and s is the separation-

$$N_W = \frac{(w+s)}{x} \tag{14}$$

$$N_H = \frac{h-d}{\frac{\sqrt{3}}{2}x} + 1 \tag{15}$$

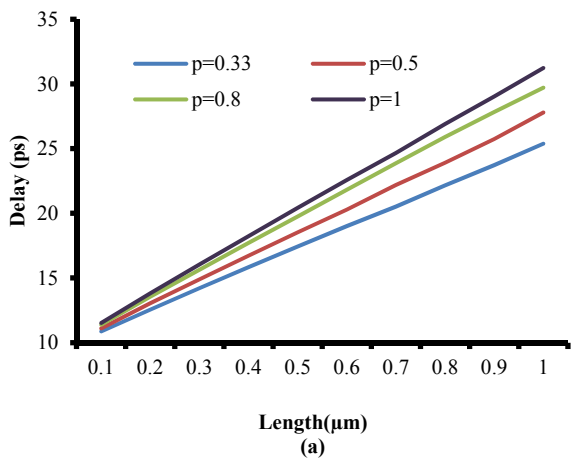
$$\begin{aligned} N_{CNT} &= N_w N_H \ if\ (w+s-N_W x) \geq x/2 \\ &= N_W N_H - N_H/2\ for\ (w+s-N_W x) < x/2 \end{aligned} \tag{16}$$

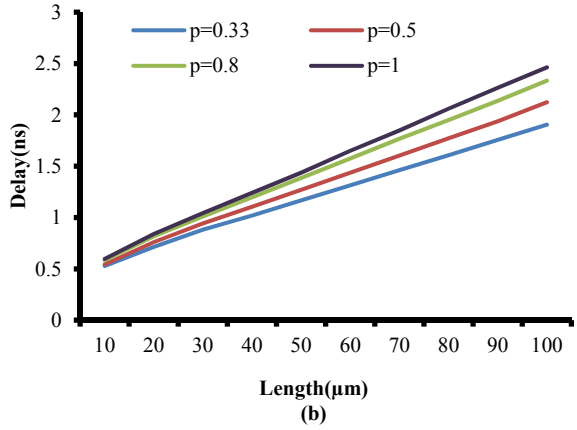
Here N_w is the CNT number along the width, N_H is the CNT number along the height and N_{CNT} is the bundle density of metallic CNT.

IV. Delay Effect

The analysis is done for DIL (Driver interconnect load) method. Here the input signal is fed into a driving inverter. Then the signal is passed through interconnect to another load inverter. At first, the bundle density is calculated using MATLAB program. Then the circuit parameters are calculated. The dimensions of the interconnect is taken as ITRS 2013 predictions [20]. CMOS inverters used for different technology nodes are taken from open source Predictive Technology Model [21]. Finally, the SPICE simulation is done to measure the propagation delay for metallic CNT probability of 0.33, 0.5, 0.8 and 1 and for 16, 22, 32, 45 and 65nm technology. The simulation is performed for local, intermediate and global interconnects.

In this analysis, adding contact resistance 100kΩ shows less than 1% delay increment in local interconnect compared to the perfect contact scenario. In intermediate and global interconnect the resistance becomes much smaller as they are connected in parallel, so no delay increment for adding contact resistance. Moreover, it is much less than the existing scattering resistance. Similarly, varying inductance showed almost no variation in delay. The delay is dominated by electrostatic capacitance as it is much less than quantum capacitance in series.





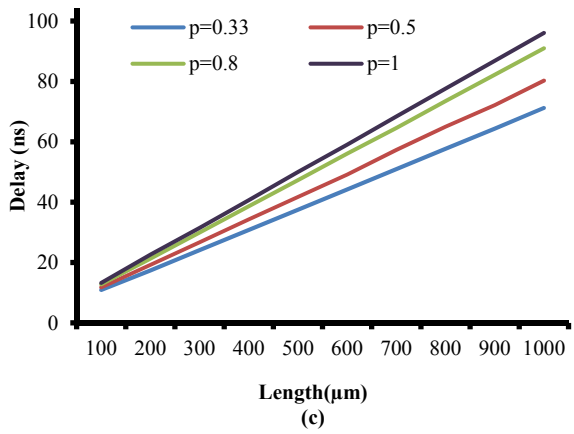
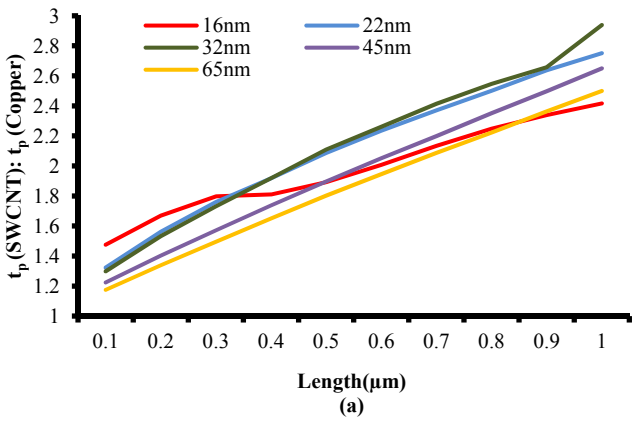
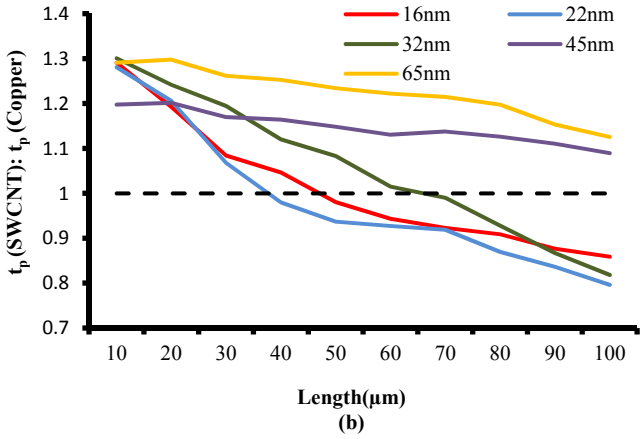


Fig. 3 Propagation delay for a) 45nm node Local interconnect b) 22nm node intermediate interconnect c) 65nm node Global interconnect (p= metallic SWCNT probability).

Increasing the bundle density decreases overall resistance of the bundle which is ideal for low power consumption process. But as the density increases, numbers of CNT around the perimeter increases too. So electrostatic capacitance increases and delay increases with metallic SWCNT density increment. This result can be observed in figure 3. This issue is a challenging problem for achieving desirable performance. One way to solve this problem is to accommodate less dense metallic carbon nanotubes at the perimeter and highly dense carbon nanotubes at the centre. So the number of nanotubes at the perimeter decreases, which eventually reduces electrostatic capacitance.





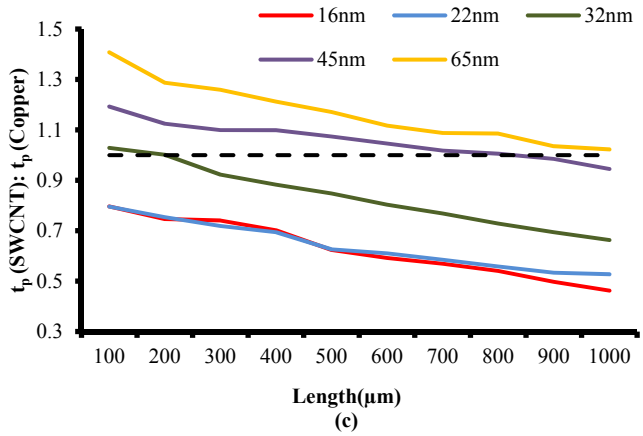


Fig. 4 Propagation delay ratio of SWCNT and copper for different technology nodes for a) Local b) Intermediate and c) Global Interconnect.

For local interconnect, the delay in CNT is worse than copper due to high resistance in CNT. However, the delay of 16nm, 22 nm and 32nm technology nodes is less than copper as we increase the length in the intermediate range, but the delay performance is still worse for 45nm and 65nm transistors. The resistance of copper increases rapidly compared to CNT which accounts for the rise in delay in this range. Global interconnect delay reduces significantly for CNT in comparison with copper but for 65nm technology node the performance is still inferior to copper. This can be attributed to the higher electrostatic capacitance as the interconnect cross section area and CNT density in the perimeter is much higher for 65nm process compared to other transistors in the analysis.

V. Conclusion

In this paper, the delay of SWCNT bundle and copper is determined and compared in accordance with ITRS predictions and up to date transistor models. The equivalent circuit model and necessary equations for SWCNT bundle are derived taking all physical and practical phenomena into account. Then the process dependent parameters are varied for interconnect analysis. Driver Interconnect load method is used for simulating the whole process. It is shown that electrostatic capacitance is the most influential parameter in propagation delay. The effect of contact resistance is negligible. In comparison with copper, CNT gives better performance in global and intermediate range but the performance is still worse in local length for 16nm, 22nm and 32nm technology nodes. So using copper interconnects for local interconnects and SWCNT for intermediate and global lengths will lead to a faster response and delay optimization in future.

References

[1] J. M. Roberts, A. P. Kaushik, and J. S. Clarke, "Resistivity of sub-30 nm copper lines," *2015 IEEE International Interconnect Technology Conference and 2015 IEEE Materials for Advanced Metallization Conference (IITC/MAM)*, pp. 341-344.

[2] X. Lu and Z. Chen, "Curved Pi-Conjugation, Aromaticity and the Related Chemistry of Small Fullerenes (C60) and Single-Walled Carbon Nanotubes," *Chemical Reviews 105 (10),* pp. 3643–3696, 2005.

[3] S. Y. Yue, T. Ouyang, and M. Hua, "Diameter Dependence of Lattice Thermal Conductivity of Single-Walled Carbon Nanotubes: Study from Ab Initio," *Scientific Reports 5:15440*, Oct. 2015.

[4] N. Chiodarelli, S. Masahito, Y. Kashiwagi, and P. Vereecken, "Measuring the electrical resistivity and contact resistance of vertical carbon nanotube bundles for application as interconnects," *Nanotechnology2011;22(8):085302.,* Jan. 2011.

[5] T. Ohashi, H. Iwama, and T. Shima, "Growth of vertically aligned single-walled carbon nanotubes with metallic chirality through faceted FePt-Au catalysts," *Journal of Applied Physics*, Volume 119, Issue 8, 2016.

[6] P. Srivastav and A. Kumar, "Comparative Analysis of Carbon Nanotubes as VLSI Interconnects," *International Journal of Science, Engineering and Technology Research (IJSETR),* Volume 4, Issue 8, August 2015.

[7] M. G. Kumar, R. Chandel, and Y. Agrawal, "Timing and Stability Analysis of Carbon Nanotube Interconnects," in *2015 IEEE International Symposium on Nanoelectronic and Information Systems*, pp. 308-313.

[8] M. K. Rai, R. Khanna, and S. Sarkar,"Crosstalk analysis in CNT bundle interconnects for VLSI application," *IEEJ Transactions on Electrical and Electronic Engineering,* Volume 9, Issue 4, pp. 391–397, July 2014 .

[9] P. L. McEuen, M. S. Fuhrer, and H. Park, "Single-Walled Carbon Nanotube Electronics," *IEEE Transactions on Nanotechnology*, Vol. 1, Issue 1, pp.78-85, 2002.

[10] J. Y. Park, S. Rosenblatt, Y. Yaish, V. Sazonova, H. Üstünel, S. Braig, T. A. Arias, P. W. Brouwer, and P. L. McEuen, "Electron-Phonon Scattering in Metallic Single-Walled Carbon Nanotubes," *Nano Letters*, Vol. 4,Issue 3, pp. 517-520, 2004.

[11] J. D. Chudow, D. F. Santavicca, C. B. McKitterick, and D. E. Prober, "Terahertz detection mechanism and contact capacitance of individual metallic single-walled carbon nanotubes," *Applied Physics Letters 100, 163503*, 2012.

[12] G. Dhillon and K. Singh, "Evaluation and Comparison of Single-Wall Carbon Nanotubes and Copper as VLSI Interconnect," *The International Journal of Innovative Research in Computer and Communication Engineering,* Volume 2, Issue 5, May 2014.

[13] P. L. McEuen and J. Y. Park, "Electron Transport in Single-Walled Carbon Nanotubes," *MRS Bulletin, 29,* pp. 272-275, 2004.

[14] P. J. Burke, "An RF circuit model for carbon nanotubes," *IEEE Transactions on Nanotechnology,*vol. 2, no. 1, pp. 55-58, March 2003.

[15] K. Nabors and J. White, "FastCap: A Multipole Accelerated 3-D Capacitance Extraction Program," *IEEE Trans. On Computer-Aided Design of Integrated Circuits and Systems*, Vol. 10, Issue 11, pp. 1447-1459, 1991.

[16] K. Zhang, X. Zhu, B. Tian, F. Wang, J. Wei, "Electrostatic capacitance extraction for carbon nanotube bundle interconnects," in *Nanoelectronics Conference (INEC), 2011 IEEE 4th International*, pp. 1-2.

[17] N. Srivastava and K. Banerjee,"Performance analysis of carbon nanotube interconnects for VLSI applications," in *IEEE/ACM International Conference on Computer-Aided Design, 2005,* pp. 383-390.

[18] M. Rai, B. Kaushik and S. Sarkar,"Thermally aware performance analysis of single-walled carbon nanotube bundle as VLSI interconnects," *Journal of Computational Electronics*, Vol. 15, pp. 407-418, Feb 2016.

[19] S. Pasricha, N. Dutt, and F. J. Kurdahi, "Exploring Carbon Nanotube Bundle Global Interconnects for Chip Multiprocessor Applications," *22nd International Conference on VLSI Design*, pp. 499-504, 2009.

[20] Predictive Technology Models, 2016. [Online]. Available: http://ptm.asu.edu/

[21] ITRS 2013. [Online]. Available: http://www.itrs2.net/

Numerical Analysis of CdS:O/CdTe Thin Film Solar Cell using Cu_2Te BSF Layer

Mohammad Wahidur Rahman*[1,2,3], Sheikh Ifatur Rahman[1], Shafayat Naznoor Ahmed[1] and Md. Ashraful Hoque[1]
[1]Department of Electrical and Electronic Engineering
Islamic University of Technology
Board Bazar, Gazipur-1704, Bangladesh
[2]Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka-1205, Bangladesh
[3]Department of Electrical and Computer Engineering
Ohio State University, Columbus, OH 43210, USA
*wahidur@iut-dhaka.edu

***Abstract*—In this paper, numerical analysis was done to evaluate the performance of newly modified structure of CdS/CdTe based thin film solar cell. In this proposed cell conventional absorber CdTe layer, back contact Ni was used along with a 500nm thick Cu_2Te BSF layer. The viability of ultra-thin CdTe absorber layer was examined together with newly added CdS:O layer instead of conventional CdS layer. It was found that the proposed ultra-thin cell have a conversion efficiency of 24.03% (*Voc* =0.8033V, Jsc = 34.499mA/cm^2, FF = 0.8260) which is more efficient than the conventional CdTe thin film solar cell.**

Keywords— CdTe solar cell, CdS:O, Ni, SCAPS, Cu_2Te, Thin film.

I. INTRODUCTION

As the Sun's energy reaches the Earth Surface, almost two third of the energy is absorbed by the land and ocean and the rest gets reflected. Due to the scarcity and the negative impact of the non-renewable energy resources, the world is adapting to more eco-friendly energy sources to meet their demand. Solar Cell is the most widely used renewable energy source till date. It is an electrical device which consists of photovoltaic cells which converts the solar radiation from the sun to useful Electrical energy using semi-conductors. This method of energy conversion is called Photovoltaic [1]. Thin-Film CdS/CdTe based Photovoltaic cells are one of the most promising candidates for low cost PV energy conversion as it has high efficiency, low manufacturing cost and better stability than conventional solar cell technology [2]. Compared to the non-thin film solar cell, the cost of the thin-film solar cell has rapidly decreased over the last decade and on a large scale production thin-film solar cell are cheaper, thus making it more favorable to adapt [5]. Parameters that affects the working principal of the solar cell are Direct Bandgap of the semiconducting material, optical absorption coefficiecnt, lattice and electron matching with other materials, production of n-type and p-type material [15]. Problems associated with the conventional cell is that it uses a lot of material, making it expensive. It has low efficiency and does not have a stable back contact. CdTe solar cells can overcome these issue up to a limit. The CdTe solar cells can be deposited using various low cost techniques such as close-space sublimation (CSS), vapor transport deposition (VTD), physical vapor deposition (PVD), metal organic chemical vapor deposition (MOCVD), electro-deposition and screen printing [2, 8, 11, 12, 16]. The layer has the direct optimum band gap with high absorption coefficient. The highest practical efficiency of CdTe Solar cells is 17.3%, which are quite low compared to theoretical efficiency 29% [12, 16]. In 1972 Bonnet et. al. reported an efficiency of 6%. Tyan et.al. achieved an efficiency of 15.8%. In 2001, Xuanzhi Wu recorded an efficiency of 16.5% for CdS/CdTe solar cell [10, 11, 16]. The best cell efficiency achieved was 16.5% whereas the commercial module is 11% efficient [14].

New window materials for CdTe has been researched and all these initiatives resulted in improved short circuit current density (Jsc) but reduced open circuit voltage (Voc) and fill factor (FF). Higher band gap window layer have large lattice mismatch with CdTe absorber than the poly-CdS film. To overcome the problems associated with this window layer, oxygenated cadmium sulfide (CdS:O) layer was introduced [7, 16].

II. MODELLING AND SIMULATION

Numerical analysis has been done to provide the cell performance and to verify its efficiency and feasibility of the proposed structure. In this work a traditional baseline structure has been chosen which is showed in Figure 1(a). Problems associated with the CdS layer was that one, it has a low band gap which will block lights of certain wavelengths to pass through and lowers quantum efficiency. Two, the layer chemically diffuses with CdTe layer to produce a waste layer that has even lower band gap. Third, there is also a 10% lattice mismatch between the two layers which cause high defect density. Finally it is overall impractical due to the excessive pinhole formation and electrical shunting which decreases the

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

output parameter [9]. Therefore in search for a better window material that has higher optical band gap and better lattice match with CdTe absorber, a n-CdS:O layer is used instead of CdS as it provide variable band gap and allow us to use reduce thickness. This layer is formed during the sputtering growth process, where the poly-CdS can be converted into nanocrystalline CdS:O by introducing to oxygen. Due to its large bandgap, it absorbs less light and most of it passes through it. Due to this reduction of absorption loss in window layer there is an improvement of short circuit current density, Jsc [7]. However reduction will allow forward leakage current to pass through. To prevent the undesired forward leakage current a high resistive buffer layer was needed and that has higher optical band gap, better lattice match with CdTe absorber. There go, we use ZnO in between SnO_2 and CdS layer [6, 16]. The presence of the smoother high-resistive buffer layer also improves the CdS film morphology by providing large grains during CBD process [4].Theoretically minimum thickness required for CdTe thin films to absorb 99% of the incident photons with energy greater than the band gap is approximately 1~2µm and hence the thickness was varied from 1-6µm to monitor the effect it has on the overall efficiency [4].

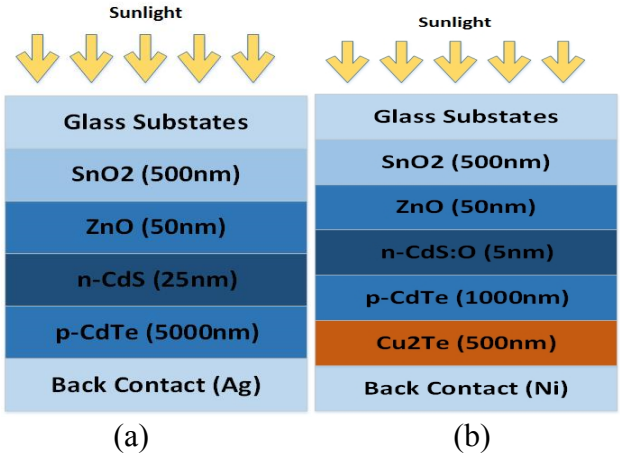


Fig. 1. Structures of CdTe solar cells: (a) Baseline Structure (b) Proposed Modified Structure

A 500nm of SnO_2 is used as the front contact layer for the modified baseline structure. The doping concentration of CdTe layer was changed to monitor the effect of the changes so that highest efficiency can be achieved [10].

Next modification is the addition of a BSF layer between the absorber layer and the back contact. This is done in order to reduce recombination loss at the back contact of the final modified structure [4]. Here, Cu_2Te has been used which has a low band gap that would be able to bounce back the carriers from the CdTe/Cu_2Te junction and thus contribute in the enhancement of carriers [6, 11].

Silver (Ag) was the back contact in the baseline structure. Other materials were needed that forms a low resistance and low barrier back-contact which is crucial for high performance CdTe based solar cells. Theoretically the work function metal of around 5.9 eV is required to make good ohmic contact to p-CdTe [9, 10]. However, most metals do not have sufficiently high work functions and form Schottky-barrier contacts to CdTe absorber layer [6, 9, 14] Generally, Cu is used but they are not stable and leads to efficiency degradation with time due to Cu diffusion to the front contact which causes shunting effect [9, 10]. Upon all consideration Nickel is used as the back contact which has a work function of 5.15 eV and it is relatively cheap.

TABLE I. PERFORMANCE PARAMETER USED IN NUMERICAL ANALYSIS

Parameter	SnO_2	ZnO	n-CdS:O	p-CdTe	Cu2Te
Thickness(µm)	0.5	0.1	0.05	5	0.1
Dielectric ratio, ε/εo	9	9	9	9.4	10
Electron mobility,µn (cm^2/Vs)	100	100	350	320	500
Hole mobility, µp (cm^2/Vs)	25	25	50	60	100
Carrier concentration, n/p (cm^{-3})	$1x10^{20}$	$1.5\ x10^{18}$	$1x10^{17}$	$2x10^{14}$	$1x10^{21}$
Bandgap, Eg (eV)	3.6	3.3	2.8	1.5	1.18
Density of states, NC (cm^{-3})	$2.2x10^{18}$	$2.2\ x10^{18}$	$1.8\ x10^{19}$	$8.0\ x10^{17}$	$7.8x10^{17}$
Density of states, NV (cm^{-3})	$1.8\ x10^{19}$	$1.8\ x10^{19}$	$2.4\ x10^{18}$	$1.8\ x10^{19}$	$1.6x10^{19}$
Electron affinity, χ (eV)	4.5	4.5	4.5	4.28	4.20

Considering all the proposal mentioned above, the final modified structure is showed in Figure 1(b) and the software used to fine the output parameter and overall efficiency for this structure was the Solar Cell Capacitance Simulator (SCAPS). The SCAPS program can vary numerous parameters and it has been developed to realistically simulate the electrical characteristics of the thin film hetero-junction solar cells.

III. RESULTS AND DISCUSSION

Numerical Analysis has been done aiming to realize different properties of the new proposed structure in comparison with the baseline structure considered. In the new structure, the first modification was done in the absorber layer. The thickness of the p-CdTe absorber layer was varied from 1000nm to 5000nm for both our base line CdTe/CdS structure and our modified CdS/CdTe/Cu2Te structure. From figure 2 it is evident that the optimum CdTe layer thickness will be smaller than 1000nm, and below this value the reduced absorption will cause a reduced efficiency and a reduced photo-current. The carrier concentration of the absorber layer was considered as 2×10^{14} cm^{-3}.

With the introduction of defect in our work, we hoped to achieve a more practical result and so the defect density of the absorber layer was considered as $2x10^{12}cm^{-3}$. High defect density in the CdTe layer will increase the re-combination rate in the space-charge region (SCR) and limit the Voc [10] where increase in defect density decreases all the output parameter of the overall efficiency, Current Density, and Open Circuit Voltage of the cell.

Next the CdS:O layer has been optimized as it provides variable band gap which is dependent on the O2/Ar ratio. The bandgap of the layer may vary from 2.42eV to 3.17eV depending on the ratio of O_2 and Ar. In this work, a band gap of 2.75eV has been taken for the CdS:O layer [7]. The thickness of the CdS:O layer has been varied from 5nm to

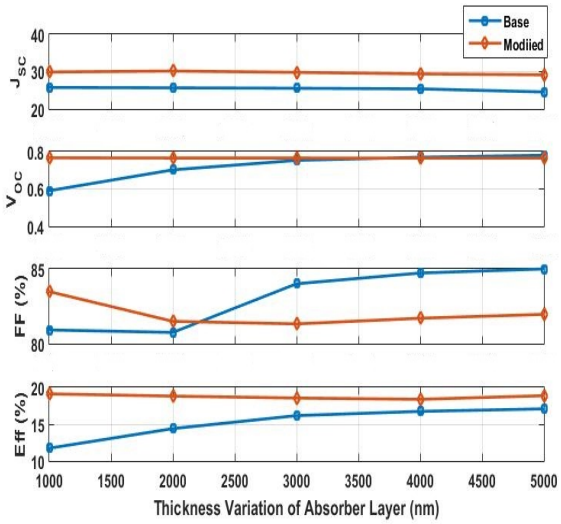


Fig. 2. Changes of performance parameters with the variation of thickness of the absorber layer.

50nm to observe the variation of different output parameters of the solar cell structure. The results for the change of the thickness of the Window layer has been shown in Figure 3. From the graphs we can come to a conclusion that a small thickness of 5nm can give a more acceptable efficiency although the variation for other thicknesses don't vary the efficiency to a great extent and short circuit current density and open circuit voltage also doesn't change much alongside. The corresponding Jsc, Voc and the overall efficiency of the cell is tested to find out the best combination circuit voltage decrease and the overall efficiency increases with increase in the parameters. Therefore to get the best optimized result thickness of 5nm n-CdS:O layer shows the most promising result.

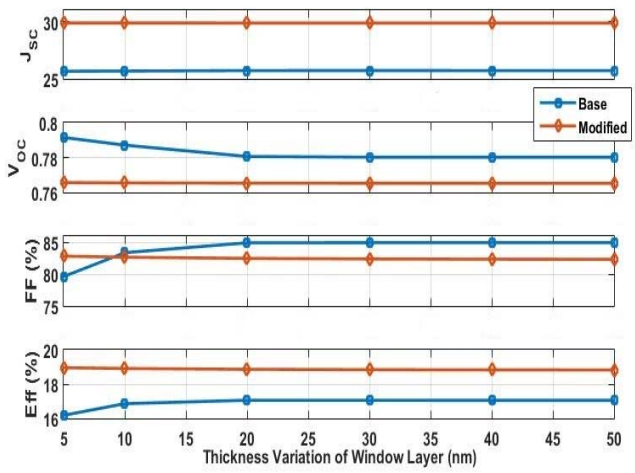


Fig. 3. Changes of performance parameters with the variation of thickness of window layer.

Another parameter that plays an important role is the BSF layer Cu_2Te. The thickness is varied from 100nm to 500nm and the corresponding efficiency, Jsc and Voc of the cell is found out which is shown in Figure 4. The result of these graph states that with increase in thickness of the BSF layer, both Current density and open circuit voltage increases and hence the efficiency increases. The thickness 500nm is chosen for best outcome.

The effect of temperature on the efficiency has been investigated from the range of 290K to 350K and it is found that the efficiency almost linearly decreases with the increase of temperature (Figure 5). The temperature coefficient of the base and optimized structure are -0.073985%/K and -0.065086%/K respectively. The graph also indicates the degree of stability of the cells at higher operating temperature or in stressed conditions [7].

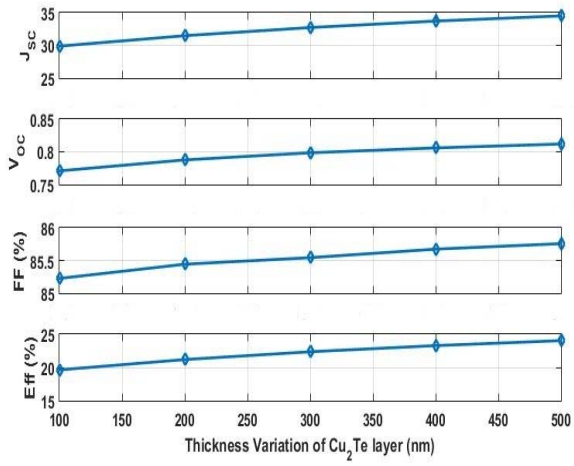


Fig. 4. Effect of thickness variation of Cu_2Te layer on cell efficiency

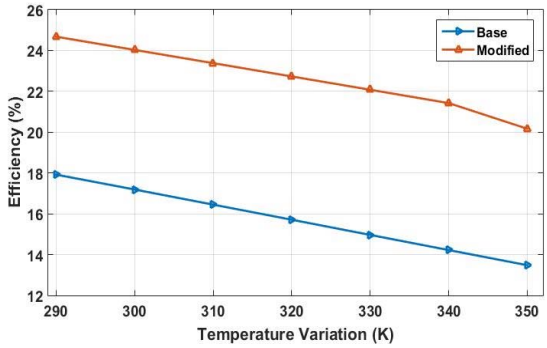


Fig. 5. Effect of Temperature variation on cell efficiency

In the analysis of the spectral response (SR), the frequency specific performance is found. From Figure 6, it is evident that the new proposed structure has a wider range of spectral response than the baseline structure and also a better quantum efficiency ranging from 300nm wavelength to 850nm wavelength of light. In extension our proposed structure can also absorb light of wavelengths of wider range than the base line structure. It also signifies that the thicker CdTe layer has higher SR and higher conversion efficiency but it also means more CdTe material is required and more cost is involved. A generous light absorption is observed from wavelengths of 300 nm to 900 nm [6].

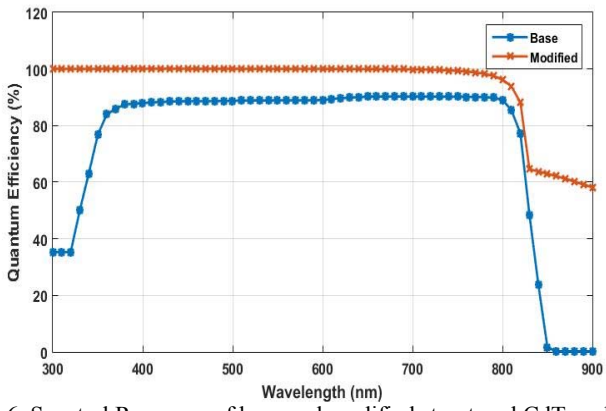


Fig. 6. Spectral Response of base and modified structured CdTe cell.

IV. Conclusions

It was found from the CdS/CdTe baseline solar cell that, the CdTe layer, within an acceptable range of thickness without any BSF layer, shows cell efficiency of 17.19%. The replacement of 25nm CdS layer with only 5nm thick CdS:O shows improved result. Additionally, 500 nm thick BSF layer Cu_2Te was introduced which enhanced the cell absorbance efficiency up to 24.03% which is 6.84% higher than the baseline structure and shows the significant improvement in the cell structure. This modified structure has Jsc of 34.51mA/cm^2, Voc of 0.8119 V and fillfactor of 85.75% whereas baseline structure has JSc of 25.9551mA/cm^2, Voc of 0.7804 V and fillfactor of 84.88%. Besides our proposed structure is more cost effective as the thickness has been reduced by 3520 nm. This proposed thin film solar cell can be fabricated using any standard CdTe cell fabrication techniques.

References

[1] K. L. Chopra, P. D. Paulson,and V. Dutta, "Thin-Film Solar Cells: An Overview," in Progress in Photovoltaics Research and Applications, New Delhi, 2004.

[2] N. A. Khan, K. S. Rahman, F. Haque, N. Dhar, M. A. Islam, M. Akhtaruzzaman, K. Sopian and N. Amin, "Design Optimization of CdTe Thin Film Solar Cells from Numerical Analysis," in *8th International Conference on Electrical and Computer Engineering*, Dhaka, 2014.

[3] M. A. Matin, "Ultra Thin High Efficiency CdS/CdTe Thin Film Solar Cells from Numerical Analysis," in *Proceedings of the 8th WSEAS Int. Conf. on Non-Linear Analysis, Non-Linear Systems and Chaos*, Selangor.

[4] Mrinmoy Dey. M. A. Matin, "High Performance Ultra-Thin CdTe Solar Cell with Lead Telluride BSF," in *3rd International Conference on Informatics, Electronics & Vision*, Chittagong, 2014.

[5] Mitchell RL, Witt CE, King R, Ruby D. PVMat, "Advances in the photovoltaic industry and the focus of future PV manufacturing R&D," in *Proceedings of the 29th IEEE Photovoltaic Specialists Conference*, 2002.

[6] Md. Sharafat Hossain, Nowshad Amin, M. A. Matin, M. Mannir Aliyu, Takhir Razykov and Kamaruzzaman Sopian, "A Numerical Study On The Prospect of High Efficiency Ultra Thin ZnxCd1-xS/CdTe Solar Cell," in *Chalcogenide Letters*, Selangor, 2011.

[7] Aminul Islam, M. A. Matin, Mohammed Mannir Aliyu, Yusuf Sulaiman and Nowshad Amin, "A Numerical Analysis on CdS:O Window Layer for Higher Efficiency CdTe Solar Cells," in *A Numerical Analysis on CdS:O Window Layer for Higher Efficiency CdTe Solar Cells*, Selangor, 2009.

[8] M. N. Imamzai, M. J. Rashid, N. A. Khan, Q. Huda, M. Akhtaruzzaman, K. Sopian and N. Amin, "Prospects of Novel CdZnTe Thin Film Solar Cells from Numerical Analysis," in *2nd International Conference on Green Energy and Technology*, Selangor, 2014.

[9] Md. Sharafat Hossain, Nowshad Amin, M. M. Aliyu, M.A. Matin, M. K. Siddiki, T. Razykov and K. Sopian, "ZnxCd1-xS as Prospective Window Layer in CdTe Thin Film Solar Cells from Numerial Analysis," Selangor.

[10] Maitry Dey, Mrinmoy Dey, M. A. Matin, Nowshad Amin, "Design of High Efficient and Stable Ultra-Thin CdTe Solar Cells with ZnTe as a Potential BSF".

[11] Nipu Kumar Das, Mrinmoy Dey, M. A. Matin and Maitry Dey on "Germanium Telluride as a BSF material for high efficiency ultra-thin CdTe solar cell," in *The 9th International Forum on Strategic Technology (IFOST)*, Chittagong, 2014.

[12] Mohammad Reza Aghaei, Mohammad Noor Imamzai and Nowshad Amin, "Numerical Modeling and Analysis of CdS/Cd1-XZnX Te Solar Cells as a Function of CdZnTe Doping, Lifetime and Thickness".

[13] Saeed Khosroabadi and Seyyed Hossein Keshmiri on "Design of a high efficiency ultratrin CdS/CdTe solar cell using back surface field and backside distributed Bragg reflector"

[14] Yusuf Sulaiman, M A Islam, Shrarafat Hossain, Nowshad Amin, M. Mannir Aliyu and Kamaruzzaman Sopian, "Investigation for Optimum Structure of CdS:O/CdTe Solar Cell from Numerical Analysis," in *Advanced Materials Research Vols 622-623 (2013) pp 1194-1198© (2013) Trans Tech Publications, Switzerland*, 2012.

[15] Hodges, Deidra Ranel, "Development of CdTe thin film solar cells on flexible foil substrates," 2009.

[16] Mohammad Wahidur Rahman, Shafayat Naznoor Ahmed, Sheikh Ifatur Rahman and Md. Ashraful Hoque, "Numerical Analysis of CdTe Thin Film Solar Cells with CdS:O Window Layer and ZnO Buffer layer," Submitted in International Conference on Advances in Electrical, Electronic and Systems Engineering, Malaysia, 2016.

‘Δ’ Slotted Compact Printable Orientation Insensitive Chipless RFID Tag for Long Range Applications

Md. Aminul Islam,[1,*] Yixian Yap,[2] and Nemai Karmakar[2]

[1]Electrical, Electronic and Communication Engineering (EECE),
Military Institute of Science and Technology (MIST), Mirpur, Dhaka 1216, Bangladesh
[2]Electrical and Computer Systems Engineering (ECSE), Monash University,
Clayton, VIC 3800, Australia
[*]aminul9274@gmail.com

Abstract—**A novel design concept of a compact printable chipless RFID tag is presented. The tag consists of a triangular patch loaded with multiple triangular ‘Δ’ shaped slot resonators. Six copies of the identical patches are placed together to achieve a very compact long range chipless RFID tag. This symmetric frequency domain based tag has the advantage to be read from any orientation with the reader antennas. The tag can be read by chipless RFID tag reader in both near field and far-field measurements. This tag does not have a ground plane and has higher data density compared to most of the existing printable chipless tags.**

Index Terms—**Frequency selective surfaces (FSS), radar cross section (RCS), radio frequency identification (RFID), RFID tags, slot antennas.**

I. Introduction

Radio Frequency Identification (RFID) system is a wireless data capturing technology that uses radio frequency (RF) waves for extracting the encoded data from remotely placed tags. This system consists of two main elements, the RFID tag, where data are encoded, and the RFID reader, which is used for extracting the encoded data from the tags. RFID tags have the potential to replace barcodes, but they are still not used in low-cost applications because of their higher prices compared to that of barcodes. The cost of the widely-used passive RFID tags depends mainly on the silicon chips used in them [1]. Therefore, research has been focused on developing chipless printable RFID tags.

A number of printable chipless RFID tags have been reported using simple or hybrid of time, frequency, phase domain and image-based encoding techniques. Among the time domain based tags, only SAW-based tags have the required data capacity [2], but the main drawback of this SAW tag is that it uses costly piezoelectric substrate. Moreover, using other printable techniques in time domain, it is not feasible to encode sufficient bits using microstrip delay lines [3] within a credit card sized area. Therefore, time domain-based tags cannot be used for high data density tag manufacturing. In phase domain-based tags only the tag using stub-loaded multiple patch antennas (SLMPA) [4] is printable and data is encoded in the cross-polarized backscattered phase. But if the backscattered phase is varied due to multipath, then it may cause bit encoding errors [5] and sufficient reliable phase resolution has not been achieved yet. Image-based tags are still in the experimental phase and need costly submicron-level printing.

If the frequency domain-based tag is considered, it is seen that these tags have higher data density than time-domain based tags. A 35-bit tag using spiral resonators and cross-polarized monopole antennas in the frequency domain has been presented in [6]. Low cost chipless RFID system for secure near-field data transfer is presented in [7] with a 9-bit tag having an area of a conventional ID card. It uses frequency shifting principle by length variation of the dipoles and complex higher order mode detection techniques. Another printable chipless tag was proposed in [8], where Multi-Resonant Dipole Antenna (MRDA) was attached to a monopole antenna to encode bits. But it requires both side printing alignment and number of dipoles increases linearly with the number of bits. Chipless tag using capacitively-tuned dipoles in [9] have problems with size and parasitic mutual coupling. Another printable tag using space filling curve [10] has difficulties in encoding data. Electromagnetic code is implemented in [11], where two Split Resonators (SRRs) are used to encode each bit. A compact tag is proposed in [12] where frequency shifting technique is used by varying the length of coplanar strips. Another compact tag is proposed in [13] using dual polarization technique to double the encoding capacity of the existing frequency domain based tags.

However, all the frequency domain based tags described above have a common drawback, that they are needed to be aligned with the linearly polarized reader antennas. Circularly polarized UWB reader antennas are required to overcome the misalignment problem. But circularly polarized antennas usually have lower bandwidth which decreases the encoding capacity of the frequency domain based tags. If a compact fully-printable orientation independent linearly polarized tag with sufficient data capacity can be developed, it will gain wide acceptance in many passive chipless RFID tag applications.

In this paper, a novel compact easily-printable triangular ‘Δ’ shaped slot resonator loaded chipless RFID tag is proposed to overcome the limitations of the other chipless tags suggested in the previous studies such as printing complexities using conductive ink due to the alignment mismatch between tag’s two sides, total tag area, manufacturing cost, orientation dependency and reading

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

range. Like most other designs, the size of our tag will not increase linearly with the number of bits as the bit representing slots will be cut inside the patches. More importantly, multiple copies of similar patches can be placed in a very compact hexagonal array, which will increase the reading range of the chipless tag. Furthermore, this tag is independent of rotation around orthogonal axis with the reader antenna and it does not require a ground plane like most other proposed chipless tags. Hence, this orientation independent single sided tag can be used in ID or credit cards, tickets, banknotes and also can be directly printed on paper or plastic packets for item-level tagging.

Rest of the paper is organized as follows: Section II presents the related theory and operating principle of the proposed chipless RFID tag. Section III presents the design of the tag. Section IV presents the simulation results followed by conclusion in section V.

II. Theory of Operation

A. Triangular 'Δ' Shaped Slot Resonators

Triangular 'Δ' shaped slot resonators of different lengths are used to create different frequency signatures for this tag, where each signature is used to represent one bit of data. It has two main advantages over other resonant structures. Firstly, resonant notches at the 2nd harmonics are negligible and at the 3rd harmonics are not present in the backscattered signal from a triangular 'Δ' shaped slot resonator loaded patch. Due to the 'Δ' resonator's continuous structure, even and odd harmonics' current paths are not prominent (

Fig. 1). So we can encode data using the whole UWB band (3.1 to 10.6 GHz), as the generated 2nd harmonic for the 3.1 GHz resonant slot at 6.2 GHz will be negligible and the 3rd harmonic will not appear at 9.3 GHz for the 'Δ' slot resonators and so on.

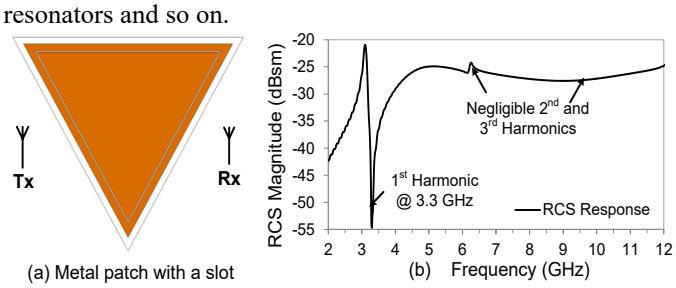


Fig. 1. Backscattered response of a 'Δ' slot resonator inside a patch

Secondly, due to its continuous structure we will get very close frequency signature at any orientation of the resonator with the reader antenna. Hence we can have an orientation independent tag using this 'Δ' slot resonators.

B. Working Principle of the Proposed Tag

The working principle is shown in Fig. 2. The tag is excited by a linearly-polarized transmitter antenna (Tx) and the frequency encoded backscattered signal from the tag is received by another linearly-polarized receiver antenna (Rx).

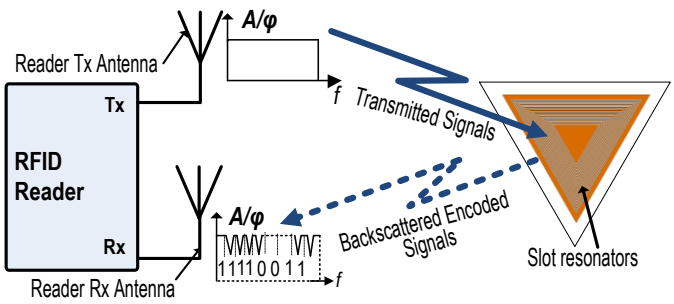


Fig. 2. Working principle of the proposed chipless RFID tag

III. Chipless RFID Tag Design

The proposed tag was designed and simulated in CST Microwave Studio 2012 and the simulation setup is shown in Fig. 3. Linearly-polarized plane wave is used for the excitation of the tag and E-field (far-field) probes and RCS probes are placed 200 mm away from the tag to receive the far-field backscattered signal from the tag.

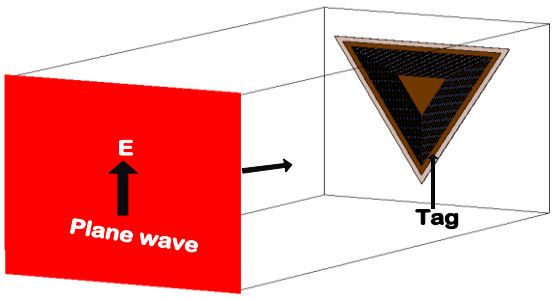


Fig. 3. Simulation setup for the slot resonator loaded chipless tag

The layout of an N-bit orientation independent 'Δ' slot resonator loaded tag is shown in Fig. 4 with dimensions as parameters. Where L_{patch} is the length of one side of the equilateral 'Δ' patch, L_1 is the length of one side of the longest 'Δ' resonator and L_n is the length of the shortest 'Δ' resonator, S_s is the separation between the resonators, W_s is the width of the slot resonators. In order to keep the design simple and easily-printable, we kept W_s and S_s the same for every 'Δ' slot resonator. The slots are numbered according to their lengths and slot number 1 is the longest one with the lowest frequency, which is denoted as MSB. Slot number 'n' with the highest frequency is denoted as LSB

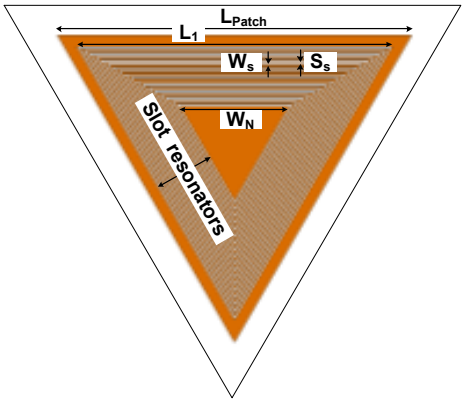


Fig. 4. N-bit 'Δ' slot resonator loaded tag

When a plane wave excites this 'Δ' slot resonator loaded patch, it shows a frequency selective behavior with peaks

followed by deep notches at the resonant frequencies of the resonators. We can increase the number of bits by increasing the number of resonators within the patch. We can change the logic state of a bit simply by removing the slot resonator which will remove the resonant frequency of the slot. The removal of notch frequency is denoted as logic '0' where the appearance of a notch is denoted as logic '1'

IV. Results

A. Simulation Results of the Proposed Tag

For proving the design concept, we started with a 'Δ' metallic patch loaded with eighteen 'Δ' slot resonators as shown in Fig. 5(a). The dimensions of the tag are chosen as W_s = 0.15 mm, S_s = 0.15 mm, L_{patch} = 31.35 mm, R_1 = 27.88 mm and R_{18} = 10.22 mm.

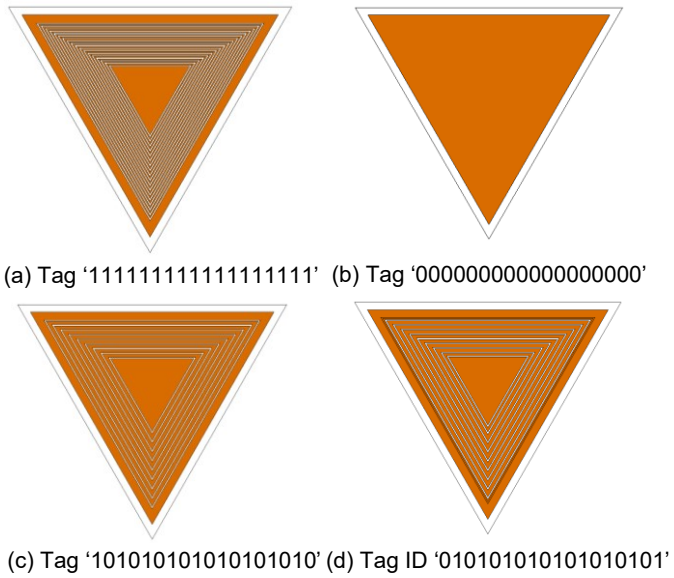


Fig. 5. Simulated tags with different bit combinations

Tags with different bit combinations as shown in Fig. 5 were simulated on Taconic TLX-8 (permittivity, ε_r = 2.55, thickness, h = 0.13 mm, loss tangent, tan δ = 0.0019).

The simulation results are shown in Fig. 6. In Fig. 5(a) the 18-bit tag has all 1's. Eighteen frequency signatures are present in the backscattered signal from the tag, which is used to encode eighteen bits. The simulation result is shown in Fig. 6(a). All 18 slots are removed as shown in Fig. 5(b) to get the tag with all 0's and the simulation result is shown in Fig. 6(a) where all the frequency signatures are removed compared to all 1's. Later tags with IDs '101010101010101010' and '010101010101010101' as shown in Fig. 5(c) and (d) are simulated and the results are shown in Fig. 6(b). It is observed that the corresponding frequencies of 0's are removed in the backscattered signal from the tags. Due to the mutual coupling between the slot resonators, there is a very little shift in the adjacent frequencies after the removal of the slots, which can be overcome by simple signal processing.

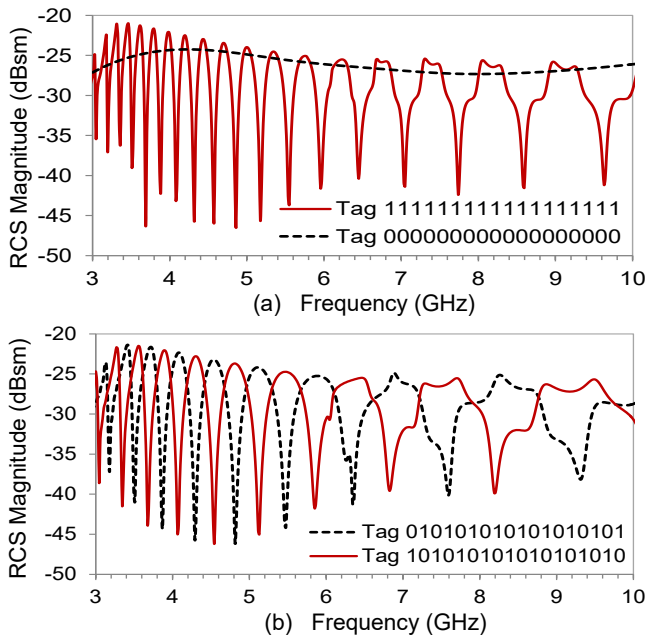


Fig. 6. Simulation results for the tags of Fig. 5

B. Measurement Techniques

The tag can be read in both near-field and far-field of the reader antennas by placing two linearly-polarized antennas on the same side of the tag; one as a transmitter and the other one as a receiver and measuring the S21. Based on its different applications, the tag can also be measured by placing the transmitter antenna in front of the tag and the receiver antenna behind the tag as it has no ground plane, which is shown in Fig. 7.

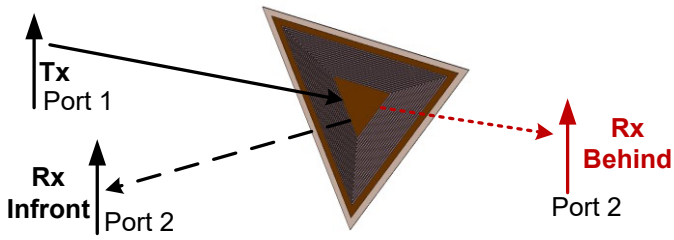


Fig. 7. Measurement setup with receiver antennas in-front of and behind of the tag

Different applications like the proximity cards (access control cards, public transport fare cards etc.) and other secure paper documents can be detected by placing it just in front of the waveguide-based near-filed reader and measuring the return loss (S11) at port 1**(**Fig. 8a**).**

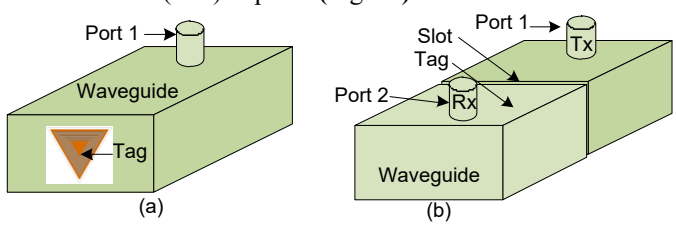


Fig. 8. (a) Proximity and (b) Slot card measurement setups

Credit cards and banknotes can be read by inserting them in a slot cut inside a waveguide and measuring insertion loss (S21) **(**Fig. 8b**)** of the two ports as described in [14].

C. Long Range Tag Design

Multiple copies of this newly designed tag can be placed side-by-side in a very compact way to get a long range chipless RFID tag. Six identical tags are placed to form the hexagonal symmetric structure as shown in Fig. 9.

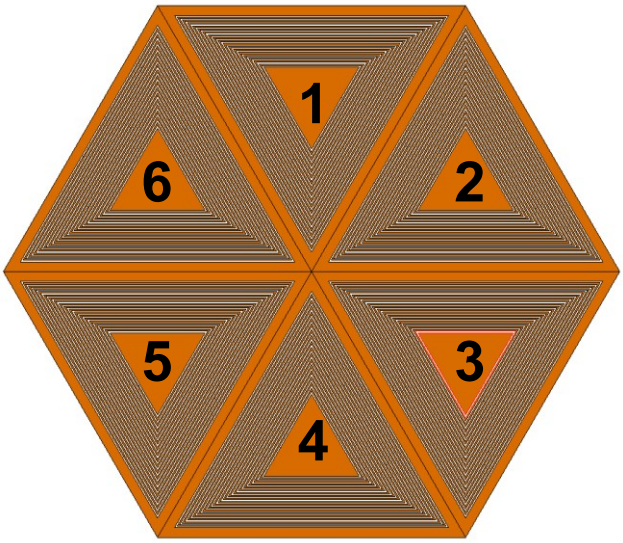


Fig. 9. Long range tag with six identical patches

The respose from this 6 patches is compared with 1 patch in Fig. 10. It is seen that the received signal level increases about 3-13 dB for 6 patches compared with single patch. The signal strength increase more at higher frequencies which will help the detection at higher frequenciess more reliably which is unlikely with the other frequecy domain based tags. Moreover, the effect of second harmonic distortion at higher frequencies will also reduces, when similar copies are placed together.

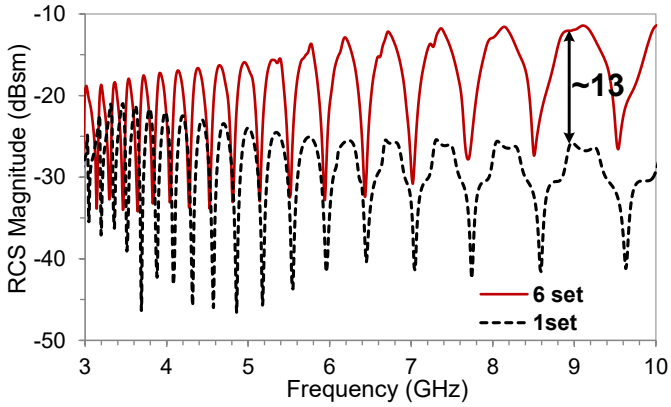


Fig. 10. Backscattered response comparison between 6-patche loaded tag with single patch loaded tag.

V. Conclusion

A novel compact chipless RFID tag concept is presented here, which uses 'Δ' slot resonators to create frequency signatures in the backscattered signal. Later, six patches are put together in a very compact way to increase the reading range of the tag. This symmetric tag will give similar frequency signature for all orientation with the reader antennas and hence can be used as an orientation independent tag. In future, the concept will be proved by measurement after the fabrication of the tag. This single sided compact tag has a great potential to be printed directly on many items like barcode using conductive ink and can also be used in many products like ID or credit cards, banknotes and paper documents for proximity measurement with higher security.

Acknowledgment

This work was supported in part by the Australian Research Council Linkage Project grant (LP0991435: Back-scatter based RFID system capable of reading multiple chipless tags for regional and suburban libraries). Support from ECSE departmental scholarship of Monash University is also acknowledged.

References

[1] U. Kaiser and W. Steinhagen, "A low-power transponder IC for high-performance identification systems," *IEEE Proc. of Solid-State Ckt,* vol.30, no.3, pp:306-10, March 1995.

[2] C. S. Hartmann, "A global SAW ID tag with large data capacity," in *IEEE Ultrasonics Symp.*, 2002, pp. 65-69.

[3] S. Botao, C. Qiang, Y. Amin, D. S. Mendoza, L. Ran, and Z. Li-Rong, "An ultra-low-cost RFID tag with 1.67 Gbps data rate by ink-jet printing on paper substrate," in *IEEE Asian Solid State Ckt. Conf.*, 2010.

[4] I. Balbin and N. C. Karmakar, "Phase-Encoded Chipless RFID Transponder for Large-Scale Low-Cost Applications," *IEEE Microwave and Wireless Components Letters,* vol. 19, pp. 509-511, 2009.

[5] S. Mukherjee, "Chipless Radio Frequency Identification by Remote Measurement of Complex Impedance," in *European Conference on Wireless Technologies*, 8-10 Oct. 2007, pp. 249-252.

[6] S. Preradovic and N. C. Karmakar, "Design of fully printable planar chipless RFID transponder with 35-bit data capacity," in *European Microwave Conference (EuMC)*, Sept.-Oct., 2009, pp. 013-016.

[7] V. Deepu, A. Vena, E. Perret, and S. Tedjini, "New RF identification technology for secure applications," in *IEEE Int. Conf. on RFID-Technology and Applications(RFID-TA)*, 17-19 Jun.2010, pp. 159-63.

[8] I. Balbin and N. Karmakar, "Novel chipless RFID tag for conveyor belt tracking using multi-resonant dipole antenna," in *European Microwave Conference(EuMC)*, Sept. 29 2009-Oct. 1 2009, pp. 1109-1112.

[9] I. Jalaly and I. D. Robertson, "Capacitively-tuned split microstrip resonators for RFID barcodes," in *European Microwave Conference*, 4-6 Oct. 2005, pp. 4, Vol. 2.

[10] J. McVay, A. Hoorfar, and N. Engheta, "Space-filling curve RFID tags," in *IEEE Radio and Wireless Symposium*, 17-19 Jan. 2006, pp. 199-202.

[11] J. Hyeong-Seok, L. Won-Gyu, O. Kyoung-Sub, M. Seong-Mo, and Y. Jong-Won, "Design of low-cost chipless system using printable chipless tag with electromagnetic code," *IEEE Microwave and Wireless Components Letters,* vol. 20, pp. 640-642, 2010.

[12] A. Vena, E. Perret, and S. Tedjini, "Novel Compact RFID Chipless Tag," in *Progress In Electromagnetics Research Symposium Proceedings, Morocco*, March 20-23, 2011, pp. 1062-1066.

[13] M. Aminul Islam and N. Karmakar, "Design of a 16-bit ultra-low cost fully printable slot-loaded dual-polarized chipless RFID tag," in *Asia-Pacific Microwave Conference Proceedings (APMC)*, 5-8 Dec. 2011, pp. 1482-1485.

[14] M. Aminul Islam, M. Shakil Bhuiyan, and N. Karmakar, "A novel compact chipless RFID tag and near-field reader," in *Asia-Pacific Microwave Conference Proceedings (APMC)*, 5-8 Dec. 2011, pp. 1518-1521.

Ćuk Topology Based High Performance Three Phase to Single Phase Switch Mode Cycloconverter with V/f motor control

Ifte Khairul Amin,[1,*] M. Nasir Uddin,[1] and Mohammad Ali Choudhury[2]

[1]Department of Electrical Engineering, Lakehead University
Thunder Bay, Ontario, Canada

[2]Department of EEE, Bangladesh University of Engineering and Technology
Dhaka, Bangladesh

[*] iamin@lakeheadu.ca

***Abstract* - In this paper, a high performance Switch Mode Power Supply (SMPS) using a three-phase to single-phase AC cycloconverter is proposed. The proposed SMPS based cycloconverter is implemented using the Ćuk topology, which has higher reliability, low switching losses and thereby increased efficiency as the number of switches is reduced. It also provides means to have high input power factor and low input current Total Harmonic Distortion (THD) with use of simple input and output filter. Incorporation of SMPS topology enables the proposed circuit to operate in either buck or boost voltage mode. The proposed topology ensures high performance by maintaining reduced input current THD (below 10%), improved input power factor (above 0.85), and better efficiency (greater than 85%) as the duty cycle of the PWM control signal is varied for resistive load while the voltage gain is varied between 50% to 150%. The improvement in converter performance is compared with the conventional 12-SCR three to single phase AC cycloconverter. Furthermore, a closed loop control circuit is designed to provide V/f control while the cycloconverter is employed to drive machines with constant flux at low speed. The integration of control circuit provides enhanced reliability and better grip over the output voltage and frequency of the cycloconverter.**

***Index Term*- Switch Mode Power Supply (SMPS), Cycloconverter, Three phase to single phase conversion, V/f control and High efficiency.**

I. INTRODUCTION

A cycloconverter is a static frequency changer (SFC) designed to convert AC voltage with constant magnitude and frequency to an AC supply of variable voltage and variable frequency without any DC link. Cycloconverters are vastly used for controlling ac motors at low speed, especially in high power application [1]. Cycloconverters are also used in solid state industries for ac-ac conversion to get output frequency lower than the supply- frequency. They are now being used in heavy industries like cement industries [2], rolling mills, electric traction and ship propellers [3] . Traditional three-phase to single-phase cycloconverters require at least 12 switches. Additional switches are required to obtain better input current shaping. In conventional cycloconverters, input current is distorted and contains higher order harmonics and Substantial amount of development and researches have been carried out in the analysis and the application aspects of cycloconverter drives. The control strategies have been developed so far includes cosine-wave control [4] and current feedback method [5] etc. A few publications have appeared in recent years documenting the configuration of SMPS topology based cycloconverters. In [6], the authors proposed a three to single phase Cycloconverter regulated and controlled by Buck-Boost converters. However, in this scheme, input current THD is above 20% and efficiency is less than 80% for buck mode operation with resistive load that demonstrates the limitation of the circuit to maintain satisfactory performance in terms of power quality. The output for resistive, R-L and motor load is demonstrated for a three to single phase Ĉuk topology based cycloconverter by the authors in [8]. Though the power quality parameters are selected to be in satisfactory range but comparison between the conventional and proposed converter is not shown in that work. Additionally, no control circuit is designed for controlling output parameters. A buck–boost topology based three phase to single phase conversion is implemented in [7] that can operate in four quadrant mode at voltage–frequency plane. Fuzzy logic control is incorporated in the design to provide better stable output voltage and frequency and microcontroller is used to drive the switches. However, closed loop control of the converter is overlooked in the design.

Therefore, in this paper a three-phase to single-phase AC cycloconverter is proposed using Ćuk converter based SMPS topology. A conventional three-phase AC to single-phase AC SCR cycloconverter requires at least 12 SCR switches, whereas only four switches are needed in the proposed topology. The reduction of the number of switches minimizes the switching losses and thereby increases the converter efficiency. The number of gate/base drive is reduced and the proposed topology has higher reliability as the number of switch is reduced. It also provides means to have high input power factor and low input current THD incorporating small input and output filters. The performance of the proposed topology is evaluated in terms of THD, input power factor, output voltage, output frequency and efficiency as the duty cycle of the control signal is varied. For controlling purpose, the output to input voltage ratio is selected for 5 different values (50%, 75%, 100%, 125% and 150%). The control approach is selected as open loop sliding mode control. A V/f control circuit is also designed for the proposed cycloconverter based single phase induction motor drive. The constant V/f ratio at the cycloconverter output guarantees constant rotating magnetic field vector for single phase induction motor operated even in low speed.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

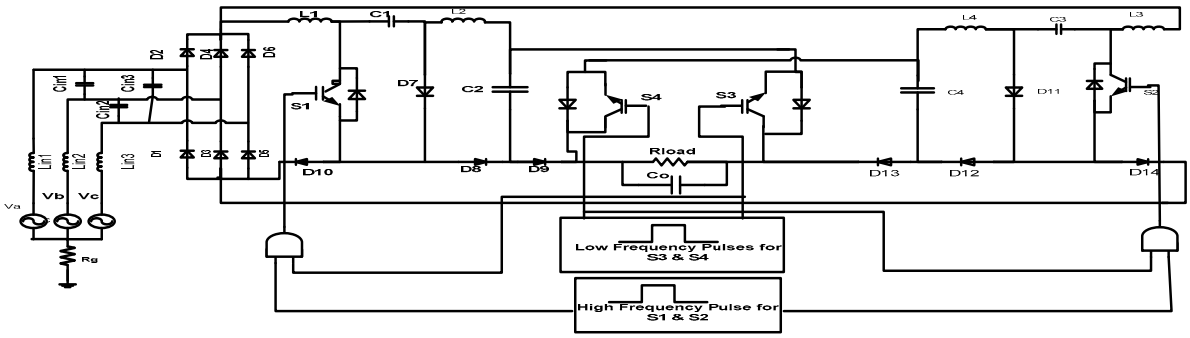


Fig. 1: Proposed Ćuk topology based three phase to single phase AC cycloconverter.

II. Proposed Circuit Configuration

Fig. 1 shows the proposed scheme of the Ćuk topology based 3-phase to 1-phase AC cycloconverter. It changes the voltage and the frequency of input AC signal. Because of high frequency switching, small input filter is sufficient to reduce the low frequency harmonic components with improved input current shaping.

A full-bridge three-phase uncontrolled rectifier circuit is used for three-phase AC to DC conversion. For three phase input voltage supply, 300V amplitude and 50 Hz frequency supply is chosen. For input filter, inductor and capacitor values are chosen as 40mH and 12µF. IGBT modules are employed as switches S1, S2, S3 and S4 to provide paths for the corresponding P and N-converters having 0.001 Ω on state resistance and 1V forward voltage drop.

III. Circuit Operation

A Ćuk converter basically operates by charging and discharging of capacitor during switch ON and OFF state while inductor maintains unidirectional flow of current in both states. Combining these two topologies Ćuk topology based cycloconverter operates in four different states:

1. P converter-S1 ON, S3 ON state
2. P converter- S1 OFF, S3 ON state
3. N converter- S2 ON, S4 ON state
4. N converter- S2 OFF, S4 ON state

Only operation of state 1 and 4 is explained with circuit configuration in following section.

A. P-Converter Switches S1 and S3 ON State

Fig. 2 shows the direction of current flow during ON state of Ćuk P-converter switches S1 and S3. Current from inductor L1 and L2 flow through switches as shown in this figure. The capacitor C1 discharges (which was charged during earlier state when S1 was OFF and S3 was ON) and capacitor C2 is charges. At the same current flows through load, R_{load}, which also charges the capacitor C_o.

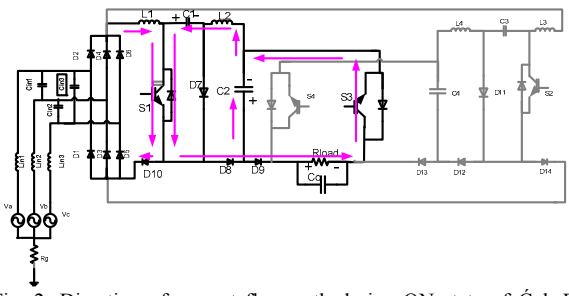


Fig. 2: Direction of current flow path during ON state of Ćuk P-converter switches S1 and S3.

B. N-Converter Switches S2 OFF, S4 ON State

When the switch S2 is turned OFF with S4 ON, the switch off state of the N converter occurs. The changed current directions are shown in Fig. 3. For this, L3 and L4 will continue to carry current in the previous directions. The polarity of C3 changes as it starts charging up. C_o keeps the output voltage constant.

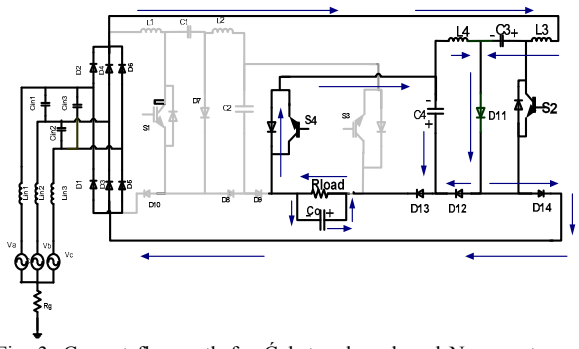


Fig. 3: Current flow path for Ćuk topology based N-converter switches S2 OFF, and S4 ON state.

Both positive and negative converters can generate voltages at either polarity, but the positive and negative converter can only supply positive and negative current individually. Thus, the cycloconverter can operate in four quadrants: (+v, +i) and (-v, -i) rectification modes and (+v, -i), and (-v, +i) inversion modes.

IV. Simulation Results

The simulation for open loop sliding mode control of the cycloconverter has been performed using PSIM and Orcad Pspice software. The closed loop control circuit is simulated using MATLAB-Simulink software.

Switch mode cycloconverters show good performance at variable frequency operation by providing non-integer frequency division at output in terms of input frequency. In the proposed Ćuk topology based cycloconverter, 33.33 Hz, 25 Hz, 20 Hz, 16.66 Hz ,14.28 Hz, 12.5 Hz, 11.11 Hz, 10 Hz or any fraction to 50 Hz conversion of waveform from input side to output side is achievable. Non-integer multiple of supply cycle at the output are obtained without any significant change in supply input current.

As the switching frequency increases the duty cycle of the switching pulses should be kept lower to regulate the performance of the converter within the acceptable limit. Considering this analysis as the basis, open loop variable switching frequency control has been proposed in the paper. For a given load, the performance of the proposed cycloconverter is examined for a varied range of duty cycles and switching frequencies. As Ćuk cycloconverter can

produce variable output voltages with amplitude higher or lower than the input rms value. Five different points have been selected for output voltage with values 64%, 75%, 100%, 125% and 150% of the input voltage. To produce these output voltages, the frequency and the duty cycle have been varied. As the objective is to design a high performance cycloconverter, the operating points are selected in such a way that the input power factor remains higher than 0.85, input current THD is confined below 10% and the efficiency of the cycloconverter is maintained higher than 85%.Table-I depicts the performance of the converter for both buck and boost mode operations at different voltage gains.

Figs. 4 (a)-(c) depicts the phase-a input voltage and input current, input current spectrum and output voltage respectively for 150% voltage gain (Boost mode). It is found from Fig. 4 that for boost mode the input current THD is 9 %, and input power factor is 0.91, which are within the acceptable limits. Figs. 5 (a)-(c) depicts the phase-a input voltage and input current, input current spectrum and output voltage respectively for 75% voltage gain (buck mode). It is found from Fig. 5 that for buck mode operation the input current THD (8%,) and input power factor (0.85) are also within the acceptable limits.

Table II shows the performance of the proposed Ćuk cycloconverter with variable resistive load at different switching frequency. The corresponding optimum duty cycle at each frequency is obtained by trial and error to achieve high performance (i.e., low input current THD, high input PF and high efficiency) of the proposed converter.

TABLE I
SWITCHING CHARACTERISTICS OF THE CONVERTER FOR HIGH PERFORMANCE

Load		R_L=50 Ω			R_L=100 Ω			R_L=200 Ω		
D	Switching frequency f(Hz)	Input PF	%THD	%η	Input PF	%THD	% η	Input PF	%THD	% η
0.41	6000	0.75	0.05	93.74	0.91	0.09	90.36	0.91	0.09	83.52
0.48	4500	0.94	0.07	94.42	0.95	0.08	88.39	0.95	0.08	86.40
0.60	4000	0.93	0.09	88.01	0.92	0.10	87.46	0.91	0.10	84.85
0.70	3510	0.86	0.08	85.43	0.86	0.08	85.12	0.84	0.09	84.49
0.80	3100	0.65	0.07	79.16	0.65	0.07	73.51	0.64	0.07	70.39

TABLE-II
PERFORMANCE OF THE PROPOSED CYCLOCONVERTER AT VARIABLE RESISTIVE LOAD

Switching Freq, f(Hz)	**Duty cycle**	**Input pf**	**Input current THD**	**Output rms Voltage**	**Vout/ Vin**	**Efficiency (%)**
6000	0.4	0.91	0.09	772.56	150%	90.35
4500	0.47	0.94	0.07	730.82	125%	88.39
4000	0.6	0.92	0.09	629.69	100%	87.46
3510	0.7	0.85	0.08	507.62	75%	85.12
670	0.87	0.87	0.09	487.47	64%	86.60

V. COMPARISON WITH CONVENTIONAL CYCLOCONVERTER

The conventional 12-SCR cycloconverter is simulated with gating pulses having different phase angles. The results are obtained for a resistive load of 100 Ohm. The maximum input voltage in all three-phase is 300 volt with a supply frequency of 50 Hz. The proposed Ćuk topology based cycloconverter is selected to be operated at definite duty cycle with corresponding determined switching frequency to provide high performance. Figs. 6 (a) and (b) show the performance comparison between the conventional SCR and proposed Ćuk topology based 3-phase to 1-phase cycloconverter on the basis of same output voltage for input power factor and input current THD.

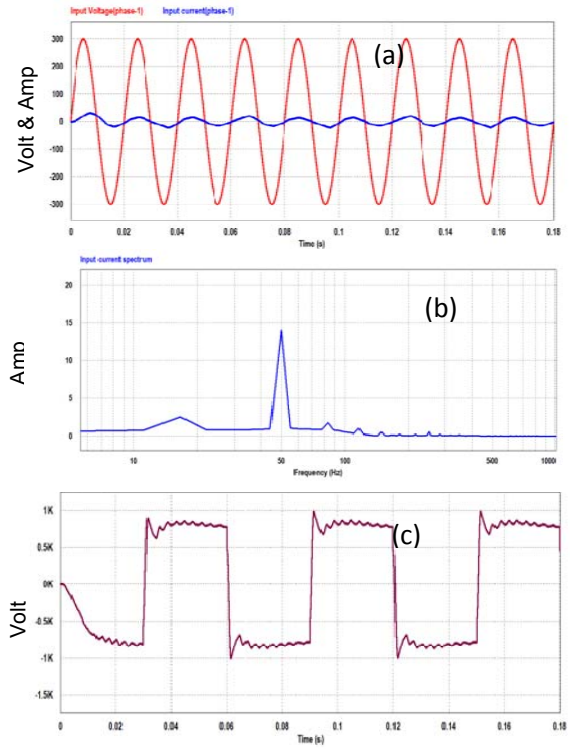


Fig. 4: (a) Input current and input voltage, (b) Input current frequency spectrum and, (c) Output voltage waveform for 150% converter voltage gain

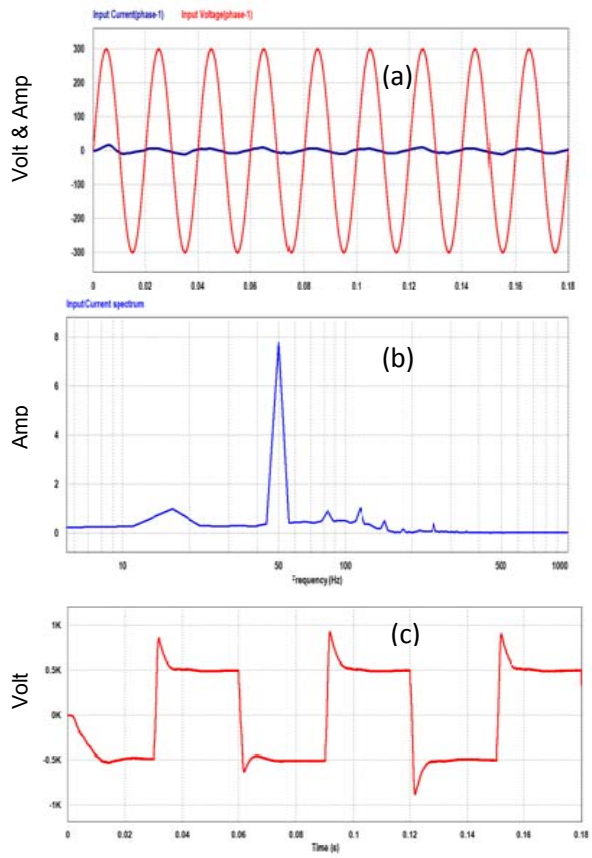


Fig. 5: (a) Input current and input voltage, (b) Input current frequency spectrum and, (c) Output voltage waveform for 75% converter voltage gain

The comparison reveals that the input power factor for the proposed Ćuk topology based 3-phase to 1-phase than that of conventional 3-phase to 1-phase SCR based cycloconverter.

The SMPS topology based converter provides the facility of controlling output voltage through the duty ratio of the high frequency switching pulse and the number of switches of these converters are less as compared to the conventional twelve SCR based 3-phase to 1-phase AC cycloconverters. From all the figures and charts it can be stated that the proposed Ćuk topology based cycloconverter shows superior performance in terms of number of switches, efficiency, input PF and input current THD to the conventional cycloconverter.

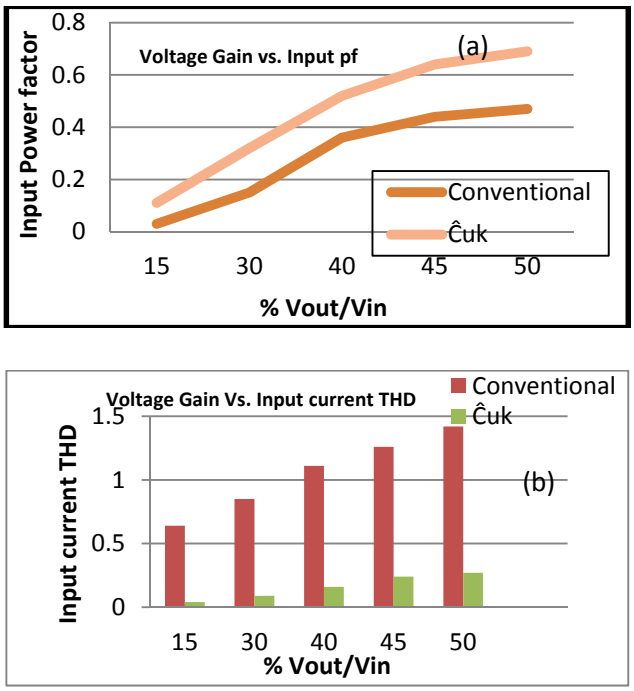


Fig. 6: Comparison between the proposed and conventional cycloconverters in terms of: (a) input Pf (b) input current THD.

Operation of a cycloconverter is best at high input frequency to low output frequency ratios, when a high quality signal can be obtained. In this paper, a controller circuit is designed to serve this goal. The constant V/f principle is achieved through regulation of slip speed to obtain accuracy in control [9]. Fig. 7 depicts the control circuit for adjustable speed operation of the motor. PI control is utilized to regulate the slip speed of the motor to keep the motor speed at its set value determined by output voltage. Since square waves with variable frequency and controlled duty cycle are required to drive the IGBTs, two Voltage Controlled Oscillator (VCO) blocks are designed where PI control is utilized to regulate the slip speed of the motor to keep the motor speed at its set value determined by output voltage. The VCOs generate variable frequency fine-tuned sawtooth waves. In final stage, comparator blocks are employed to yield pulses for converter switches by comparing between sawtooth wave and variable dc signal obtained from modified controller input signals. Table III shows the control block output for different motor torque.

TABLE III
PERFORMANCE DATA OF V/F CONTROL CIRCUIT

Applied Torque(N-m)	Converter Output Freq(Hz)	Converter RMS Output Voltage(Volt)	V/f ratio
2	31	160	5.16
4	29	154	5.31
8	27	141	5.22
10	24	128	5.33

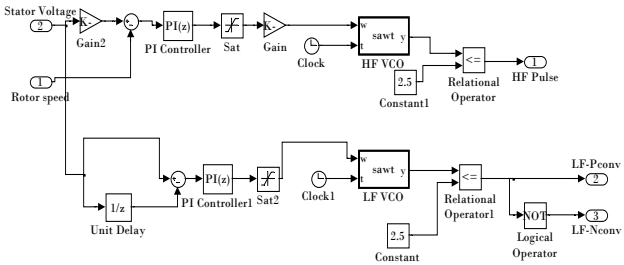


Fig. 7: Detail of control circuit for pulse generation of constant V-f control

VI. CONCLUSION

A high performance switch mode three-phase to single-phase AC cycloconverter using Ćuk topology has been proposed in this paper. The performance of the proposed cycloconverter circuit has been evaluated in terms of power quality parameters, namely, power factor, total harmonic distortion and efficiency. The proposed circuit exhibited satisfactory performance having near sinusoidal input current with low THD (less than 10%), high input power factor (higher than 0.85), better control of voltage and frequency and high efficiency (greater than 85%) for both buck and boost mode operations with voltage gain of 50%-150%. The performance of the proposed Ćuk topology based cycloconverter has also been found superior to the conventional 12-SCR 3-phase to 1-phase cycloconverter in terms of power quality and also used less number of switches. A closed loop control circuit has also been designed to serve the purpose of controlling output voltage and output frequency of the proposed cycloconverter to maintain V/f operation of a single-phase IM load. It has been found that the proposed control of cycloconverter can maintain satisfactory performance to maintain the constant flux operation of 1-phase IM drive.

REFERENCES

[1] V. Goyal and S. Wadhwani, "Simulation Of Six Pulse Cycloconverter Excited Induction Machine", International Journal of Electrical Engineering and Technology (IJEET),Volume 3, Issue 2, pp. 76-83, July – September (2012),.

[2] C. P. LeMone, M. Ehara and L. Nehl, "AC adjustable speed application for the cement industry", Proceedings of IEEE Cement Industry Technical Conference, Salt lake city,UT, pp. 335-362, 1986.

[3] Liu, Y., Heydt, G.T. and Chu, R.F. "The power quality impact of cycloconverter control strategies'', IEEE Trans. Power Delivery, Volume 20(2), pp. 1711–1718 , 2005.

[4] A. Azam, "Three to single-phase high-power quality switch-mode cycloconverter" IET Power Electronics Journal,vol. 7,issue 6,pp. 1603-17, June 2014.

[5] L. Gyugyi and B. R. Pelly, "Static Power Frequency Changers, Theory, Performance and Applications", New York, Wiley, 1976.

[6] G. Seguier and R. Bausiere, "Power Electronic Converters", Vol. 2: AC-AC Conversion, pp. 211-276, North Oxford Academic Publisher, 1987.

[7] M.A. Azam, A.N. Azad, H.T. Imam, M.A. Kabir, M.N. Uddin and M.A. Choudhury, "Three to Single Phase Buck-Boost Regulated High Power Quality Cycloconverter", Conference on Innovative Smart Grid Technologies - Asia (ISGT Asia), IEEE,pp. 1-6,June 2012.

[8] A. N. Azad, M. H. T. Imam, M.A. Azam, M.A. Kabir, M.N. Uddin and M. A. Choudhury, "Ĉuk Topology Based New Three Phase to Single Phase Cycloconverter With Input Current Shaping and Input Power Factor Improvement"- IEEE Conference on Innovative Smart Grid Technologies - Asia (ISGT Asia),June 2012.

[9] Zhenyu Yu and David Figoli, "AC Induction Motor Control Using Constant V/Hz Principle and Space Vector PWM Technique with TMS320C240"- Application Report: SPRA284A,Texas Instrument, April 1998.

GRAM-POSITIVE BACTERIAL PROTEIN SUBCELLULAR LOCALIZATION PREDICTION USING FEATURES FUSION STRATEGY

Julia Rahman*[1], Md. Nazrul Islam Mondal[1], Md. Khaled Ben Islam[1,2], Md. Al Mehedi Hasan[1], S. M. Sabbir Amin[1]
[1]Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh
[2] Department of Computer Science & Engineering
Pabna University of Science & Technology, Pabna, Bangladesh
*juliacse06@gmail.com

Abstract— **Prediction of protein subcellular localization is the most challenging field for the researchers because of its importance in different branch of molecular biology and drug discovery. Last two decades, a large number of machine learning approaches have been tested into sequence based features for the prediction of subcellular localization. Single features like amino acid composition (AAC), pseudo amino acid composition (PseAAC) and physiochemical property model (PPM)) contain insufficient information due to their single perspectives. To overcome this problem, the main contribution of our work is to propose two feature fusion representations AACPPM and PAACPPM which can be fused PPM with AAC and PseAAC respectively. Support Vector Machine (SVM) is applied as a classifier on to both single and fused feature representations of Gram-positive bacterial dataset. The actual accuracy of AACPPM is 72.4% which is 2% higher than single feature representations and 6% higher than X. Qu et al [1]. The locative accuracy of both AACPPM and PAACPPM is 73.2% which is also 2% higher than single feature representations.**

Index Terms— **subcellular localization prediction, AAC, PseAAC, PPM, AACPPM, PAACPPM, SVM.**

I. INTRODUCTION

As the second largest element of living body protein acts as a worker of living activities and responsible for whole life process. In different field of Molecular Biology like protein function prediction, role of biological process, drug discovery, genome annotation and so on, protein subcellular localization plays a vital role. Due to this importance of protein with correct location researchers feel attraction to predict its localization. A traditional biochemical experiment for protein subcellular localization predictions are expensive and faces uncertainty in time boundary to fulfil the research demands. More likely, the explosion of protein sequences makes it more challenging. This in fact shows that, computational methods are required as an alternative choice for the prediction of protein subcellular localization automatically and more accurately.

A best subcellular localization predictor using machine learning approaches solely depends on the knowledge of protein biological attributes better known as features. Features are sampled from protein sequences in number as much as possible. The selection for enriched feature is vital for any classification system. Informative feature representations provide abundant discriminant information to improve prediction accuracy [2]. Different features dealt with single location and multi-location proteins have been identified by building developed numberless predictors. Cai et al. have proposed a model according to amino acid composition (AAC) in which feature vectors have been constructed by calculating occurrence frequency of 20 native amino acid residues [3]. Park and Kanehisa have proposed to apply SVM classifier onto five features based on amino acid compositions, amino acid pair compositions and gapped amino acid compositions [4]. Dipeptide composition (DC) with 400 dimensions has been used as a feature for protein subcellular localization prediction where each value of vector represents the occurrence of 2 consecutive residues from protein sequence string [5]. Pseudo-Amino Acid Composition (PseAAC) feature has included not only protein sequence composition but also its order information to predict subcellular location [6]-[8]. Li et al. have improved the prediction accuracy by combining Gene Ontology (GO) annotations and Amphiphilic pseudo amino acid composition (APseAAC) [9]. A number of classifiers were developed based on Gene Ontology Annotation [9]-[12]. X. Qu et al. have used fusion feature extraction methods based on pseudo amino acid composition (PseAAC), physiochemical property model (PPM), amino acid index distribution (AAID) and N-terminal signal to improve the accuracy [1]. S. Wang and S. Liu have proposed two fusion feature representations by integrating PSSM with DipC and PseAAC, respectively as features and KNN as classifier for the prediction of sub-nuclear location of protein [2].

With informative feature representation researchers also have used different machine learning approaches which more likely includes classifiers like K nearest neighbour (KNN) [2], multi-label K nearest neighbour (ML-KNN) [1], artificial neural network [3], fuzzy-K nearest neighbour (Fuzzy-KNN)[5], support vector machine (SVM) [4], [10], [13].

For accomplishing better accuracy of a classifier, enriched feature construction is the most important. In our paper we have proposed two feature fusion representations for Gram-positive by fusing two single feature representations respectively and then applying SVM classifier as a predictor. Feature fusion representations have been used as an alternative due to the lack of insufficiency in information regarding single feature representation that has proven to be confusing in the context of a classifier and ends up producing worse performance issues while predicting protein subcellular localization. Very first, from protein sequence, Amino Acid Composition (AAC), Pseudo-Amino Acid Composition (PseAAC) and Physicochemical Properties Model (PPM)

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

feature representations has been extracted. Then both AAC and PPM are fused to form a new representation AACPPM taking account on both PseAAC and PPM to construct another representation PAACPPM. After wards SVM classifier is used in all of the single and fused features. This has been observed that the actual accuracy and locative accuracy of feature fusion representations are at least 2% higher than single feature representations alone.

II. Dataset and Feature Extraction Methods

A. Dataset

We have chosen Gram-positive bacterial dataset [1], [12], [13] because of its importance in drug discovery. Proteins of this dataset have ⩽ 25% pairwise sequence similarity to any other proteins in a same subcellular location. This dataset contains 519 unique proteins and 523 total proteins with 4 different subcellular locations. We have only considered protein sequences with 20 standard amino acids and excluded sequences containing B, Z, X, J symbols because of their ambiguity [14]. After pre-processing this dataset we have got the following statistics:

TABLE I.
Proteins in Gram-Positive Dataset After Preprocessing

No. of Unique protein	No. of Locations	No. of 1-locative proteins	No. of 2-locative Proteins	Total proteins
518	4	514	4	522

B. Feature Extraction Methods

Sequences from proteins have been extracted. Protein sequences are expressed by the 20 native amino acid residues: $P = (p_1, p_2, \ldots\ldots\ldots, p_N)$, where $p_i \in A$ and A = 20 amino acids. Initially for feature extraction, we have used AAC, PseAAC, PPM schemes.

1) Amino Acid Composition (AAC): Amino Acid Composition, the most popular method of feature extraction techniques based on calculating the proportion of each amino acid in protein sequences [15] as expressed in the following equation:

$$X_i = \frac{\text{count(i)}}{N}; \qquad i \in [1, \ldots\ldots\ldots, 20] \qquad (1)$$

where, count(i) computes the number of occurrences of ith amino acid within N length protein sequence.

2) Pseudo-Amino Acid Composition (PseAAC): Pseudo-Amino Acid Composition [1, 16] feature extraction method not only based on sequence composition but also its sequence order information. The feature vector is expressed as:

$$F = [v_1, v_2, \ldots, v_{20}, v_{20+1}, \ldots\ldots, v_{20+\eta}] \quad (\eta < N) \qquad (2)$$

Here, first 20 components represent the occurrence frequency of one of the 20 native amino acids and rest of the components represents the protein position information.

The sequence order k-correlation factor of each residue in protein P is defined as:

$$\delta_k = \frac{1}{N-k}\sum_{i=1}^{N-k}\Omega_{i,\,i+k} \;; \qquad (k < N) \qquad (3)$$

Here, $\Omega_{i,\,i+k}$ is coupling factor for k most adjacent residues expressed as:

$$\Omega_{i,\,i+k} = \frac{1}{3}\left\{\left[S_1(P_{i+k}) - S_1(P_i)\right]^2 + \left[S_2(P_{i+k})S_2(P_i)\right]^2 + \left[T(P_{i+k}) - T(P_i)\right]^2\right\} \qquad (4)$$

Here, S_1, S_2 and T are the value of normalized hydrophobicity, hydrophilicity and the side chain mass for amino acid residues. P_i and P_{i+k} are i-th and (i+k)-th amino acid residue respectively.

If the k correlation factor is calculated then PseAAC feature vector can be represented as:

$$P_u = \begin{cases} \dfrac{f_u}{\sum_{i=1}^{20} f_i + \omega\sum_{k=1}^{\eta}\delta_k} & (1 \le u \le 20) \\[2ex] \dfrac{\omega\delta_{u-20}}{\sum_{i=1}^{20} f_i + \omega\sum_{k=1}^{\eta}\delta_k} & (20+1 \le u \le 20+\eta) \end{cases} \qquad (5)$$

Where, weight factor $\omega = 0.05$, f_i is the occurrence frequency of the residues in this protein sequence. The value of η is L_{min}- 1, where L_{min} is the length of the shortest protein sequence in the dataset.

3) Physicochemical Properties Model (PPM): In Physicochemical Properties Model [1], amino acid residues in all proteins are divided into neutral, hydrophobic and polar groups according to their seven physicochemical properties. The seven physicochemical properties include hydrophobicity, normalized vander Waals volume, polarity, polarizibility, charge, secondary structures and solvent accessibility.

According to the seven properties of amino acids residues, compute occurrence frequency of the residues has been manifested as polar, neutral and hydrophobic in a protein sequence. The calculation formula is [1]:

$$f_{i,\,polar} = \frac{n_{polar}}{N} \qquad (6)$$

$$f_{i,neutral} = \frac{n_{neutral}}{N} \qquad (7)$$

$$f_{i,\,hydrophobic} = \frac{n_{hydrophobic}}{N} \qquad (8)$$

Where, i =1, 2, … , 7 (seven properties), $f_{i,\,j}$ is the frequency of amino acid characterized by polar / neutral / hydrophobic. n_j represents the total number of polar / neutral / hydrophobic characters present in protein sequence. N is the length of the protein sequences. However, 21 dimension feature vector for each protein sequence is created according to this feature extraction model. The distribution situation of amino acids properties are showed in TABLE II.

TABLE II.
Distribution Situation of Amino Acids Properties [1]

Property	polar	neutral	hydrophobic
hydrophobicity	RKEDQN	GASTPHY	CLVIMFW
normalized vander Waals	GASCTPD	NVEQIL	MHKFRYW
Polarity	LIFWCMVY	PATGS	HQRKNED
Polarizibility	GASDT	CPNVEQIL	KMHFRYW
charge	KR	ANCQGHILMFPS TWYN	DE
secondary structures	EALMQKRH	VIYCWFT	GNPSD
solvent accessibility	ALFCGINW	RKQEND	MPSTHY

III. METHODOLOGY

At first we have introduced two feature fusion representation and then the classifier which is used to predict subcellular localization. The procedure of protein subcellular localization prediction is shown by the following Fig. 1:

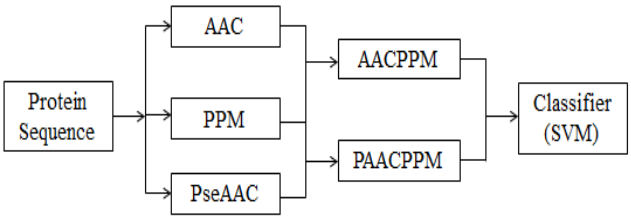


Fig. 1: Flowchart of the prediction procedure

A. Feature Fusion

AAC and PseAAC both represent the information about each native amino acid occurrence frequency whereas PseAAC additionally reflects upon the sequence order information from protein sequences. On the other hand, physicochemical property attributes appear in PPM feature representation where these attributes are absent in AAC and PPM. As a result our proposal is to combine these different properties of amino acid residues by applying feature fusion techniques. Two feature fusion representations namely AACPPM and PAACPPM combined with PPM with AAC and PseAAC respectively are produced as a result.

Parallel feature fusion strategy is used for the production of fused feature data sets. This strategy forms complex feature vector. Let, consider two different feature vectors A and B of same sample ξ with m and n dimension accordingly. Complex vector λ= A+iB (i is an imaginary unit) is used to represent the combination of A and B and is named parallel combined feature of ξ [17]. Zero padding is occurred with lower dimension feature space until m = n when dimension is mismatched. Following this procedure we develop two fused feature vectors AACPPM and PAACPPM.

B. Support Vector Machine (SVM)

In this work we use one-against-rest based Support Vector Machine (SVM) in which decision about each subcellular location is taken separately. In this case, each location is transformed to two class labels, either +1 (support the location), or -1 (does not support the location). SVM tries to find the optimal hyper plane in the feature space with maximum margin between the two classes.

For each target protein P_t, decision about j^{th} location will be made by SVM as:

$$S_j(x^t) = \sum_{i=1}^{n} \alpha_i y_i K(x^i, x^t) + b \qquad (9)$$

where, $K(x^i, x^t) = exp(-\gamma\, ||x^i - x^t||^2)$

$y_i \in (+1, -1)$,

$\gamma = \frac{1}{2\sigma^2}$

σ is the width of the function,

α_i is the lagrange multipliers.

The subcellular location of target protein P_t will be predicted as:

$$Loc(x^t) = \bigcup_{j=1}^{d} \{j : S_j(x^t) > 0\} \qquad (10)$$

If $Loc(x^t) = \emptyset$, then the number of subcellular locations is set to one and the location is given by-

$$Loc(x^t) = arg_{j=1}^{d} min f_j \qquad (11)$$

where, f_j is the functional value returned by each SVM classifier. This idea of avoiding zero prediction is an adaptation from wan's work [10].

IV. RESULT AND DISCUSSION

Features as well as classifier with or without proper parametric evaluation bring performance issues with the classifier as predictor. This in fact is true in the case of using proper classifier. Though SVM classifier with RBF kernel is used, yet the best combination for γ and regularization coefficient c for obtaining highest accuracy is needed. For achieving better accuracy we tune the parameters $\gamma = \{2^{-8}, 2^{-7}, \ldots, 2^{0}, \ldots, 2^{8}\}$ and $c = \{2^{-2}, 2^{-1}, 2^{0}, 2^{1}, 2^{2}\}$ to select best combination.

The two parameters that are actual accuracy and locative accuracy can be penetrated for the performance matrices so as to facilitate performance of multi-label problem. Actual accuracy is defined as the exact match of all locations without any over prediction or under prediction with the actual labels of a target protein [10].

$$AA = \frac{\sum_{i=1}^{N} \Delta[M(P_i),\ L(P_i)]}{N} \qquad (12)$$

Where,

$$\Delta[M(P_i), L(P_i)] = \begin{cases} 1, & if\ M(P_i) \equiv L(P_i) \\ 0, & otherwise \end{cases}$$

$L(P_i)$ represents true label set for i^{th} protein

$M(P_i)$ represents predicted label set for i^{th} protein

If the predicted label $M(P_i)$ of target protein P_i matches with any label of the true label set $L(P_i)$ then the accuracy is considered as Locative Accuracy. It can be defines as [10]:

$$LA = \frac{\sum_{i=1}^{N} \Delta|M(P_i) \cap L(P_i)|}{N} \qquad (13)$$

8-fold cross validation has been performed in Gram-positive bacterial dataset ensuring each fold containing one protein sequence of each class at least. Here, the experimental result of single features and fused feature sets is given bellow:

TABLE III.
PERFORMANCE FOR DIFFERENT FEATURES USING SVM

Features	AA	LA
AAC	366	370
PseAAC	339	355
PPM	350	372
AACPPM	375	382
PAACPPM	365	382

It can be observed from the following graph Fig. 2 that the best actual accuracy 72.4% is achieved by AACPPM which is 2%, 7% and 5% greater than AAC, PseAAC and PPM respectively. AACPPM as a fused form of AAC and PPM feature representation has achieved the best performance while $\gamma = 2^{-4}$, $c = 2^{-1}$.

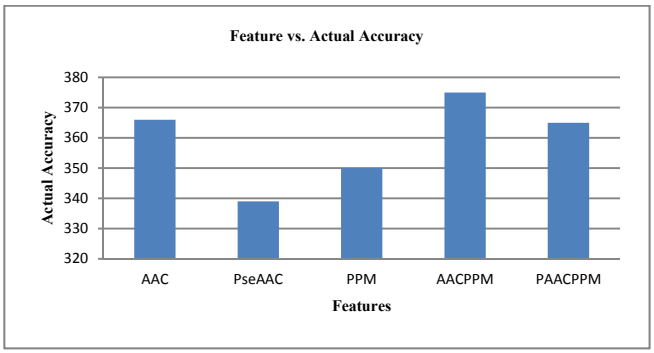


Fig. 2: Actual Accuracy for different Features

In the following Fig. 3, we observe that best locative accuracy 73.2% is achieved by both AACPPM and PAACPPM which are also feature fusion representations. The locative accuracy of both AACPPM and PAACPPM is 3%, 5% and 2% heigher than AAC, PseAAC and PPM respectively.

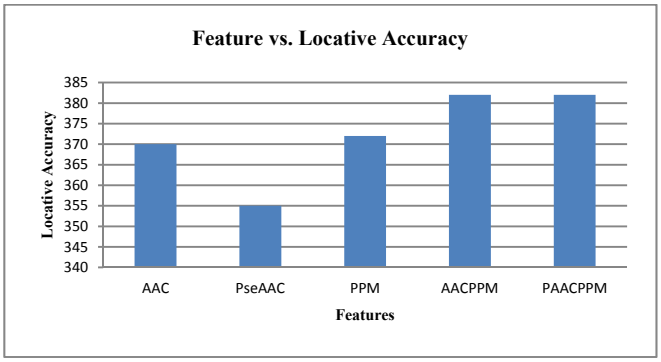


Fig. 3: Locative Accuracy for different Features

The feature fusion representations perform better actual and locative accuracy than single features as illustrated in Fig. 2 and Fig. 3. In paper [1] we have seen that their actual accuracy of combinational representation of N terminal, PseAAC and PPM for ML-KNN classifier is 66.1568% in Gram-positive dataset. In our work we observed that combination of AAC and PPM performed 6% better than paper [1] because of their informative, less redundant feature and SVM classifier with proper combination of RBF parameters.

V. Conclusions

Informative feature extraction methods from protein sequence mostly affect the performances of protein subcellular localization prediction. The accuracy is decreased instead, if inappropriate, noisy or less informative feature extraction methods are selected. SVM classifier with RBF kernel has been applied on the five features and amongst them three features are from single representation and the rest are from feature fusion representation. It can be observed from the result of the work that fused features to be proven remarkably enough for performances. Evaluation of the performance issues can be triggered as emphasis that is given on actual accuracy as well as actual accuracy of feature fusion representation. It is also shown that AACPPM has acquired actual accuracy noted 6% higher than it was before comparatively in X. Qu et al [1] and has obtained accuracy 2% higher than single feature representations only. While locative accuracy for both AACPPM and PAACPM has been staged higher than 2%.

References

[1] X. Qu, Y. Chen, S. Qiao, D. Wang and Q. Zhao, "Predicting the subcellular localization of proteins with multiple sites based on multiple features fusion" ICIC, 2014, Springer International Publishing Switzerland, pp. 456–465.

[2] Shunfang Wang and Shuhui Liu, "Protein sub-Nuclear localization based on effective fusion representations and dimension reduction algorithm LDA," *International Journal of Molecular Sciences*, vol. 16, pp. 30343–30361, December 2015.

[3] YD Cai, Xiao-Jun Liu and Kuo-Chen Chou,"Artificial neural network model for predicting protein subcellular location," *Computers and Chemistry, Elsevier Science Ltd*, vol. 26, pp. 179–182, January 2002.

[4] KJ Park and M. Kanehisa, "Prediction of protein subcellular locations by support vector machines using compositions of amino acids and amino acid pairs," *Bioinformatics*, Oxford University Press, vol. 19, Issue 13, pp. 1656–1663, 2003.

[5] Ying Huang and Yanda Li, "Prediction of protein subcellular locations using fuzzy k-NN method," *Bioinformatics*, Oxford University Press, vol. 20, Issue 1, pp. 21–28, January 2004.

[6] H.B. Shen and K.C. Chou, 'Nuc-PLoc: A new web-server for predicting protein subnuclear localization by fusing PseAA composition and PsePSSM," *Protein Engineering, Design & Selection*, vol. 20, no. 11, pp. 561–567, 2007.

[7] L.Q. Li, S.J. Yu, W.D. Xiao, Y.S. Li, M.L. Li, L. Huang, X.Q. Zheng, S.W. Zhou and H. Yang, "Prediction of bacterial protein subcellular localization by incorporating various features into Chou's PseAAC and a backward feature selection approach," *Biochimie*, 104, pp.100–107, June 2014.

[8] Chao Huang and Jingqi Yuan, "Using radial basis function on the general form of Chou's pseudo amino acid composition and PSSM to predict subcellular locations of proteins with both single and multiple sites," *BioSystems*, pp. 50– 57, April 2013.

[9] Liqi Li, Yuan Zhang, Lingyun Zou, Changqing Li, Bo Yu, Xiaoqi Zheng, and Yue Zhou, "An ensemble classifier for eukaryotic protein subcellular location prediction using Gene Ontology categories and Amino Acid Hydrophobicity," *PLoS ONE*, vol. 7, Issue 1, January 2012.

[10] Shibiao Wan, Man-Wai Mak and Sun-Yuan Kung, "mGOASVM: Multi-label protein subcellular localization based on gene ontology and support vector machines," *BMC Bioinformatics* 2012.

[11] Chin-Sheng Yu, Chih-Wen Cheng, Wen-Chi Su, Kuei-Chung Chang, Shao-Wei Huang, Jenn-Kang Hwang and Chih-Hao Lu, "CELLO2GO: A web server for protein subcellular localization prediction with functional Gene Ontology Annotation," *PLOS ONE*, vol 9, Issue 6, June 2014.

[12] Xiao Wang, Jun Zhang and Guo-Zheng, "Multi-location gram-positive and gram-negative bacterial protein subcellular localization using gene ontology and multi-label classifier ensemble," *BMC Bioinformatics,* 2015.

[13] Abdollah Dehzangi, Rhys Heffernan, Alok Sharma, James Lyons, Kuldip Paliwal, Abdul Sattar, "Gram-Positive and Gram-Negative Protein Subcellular Localization by Incorporating Evolutionary-based Descriptors into Chou's General PseAAC," Journal of Theoretical Biology, pp: 284–294, 2015.

[14] IUPAC — IUB Commission on Biochemical Nomenclature, "A One-Letter Notation for Amino Acid Sequences (Definitive Rules)," *Pure Appl. Chem.,* vol. 31, Issue 4, pp. 639-646, 1972.

[15] Loris Nanni, Alessandra Lumini and Sheryl Brahnam, "An Empirical Study of Different Approaches for Protein Classification", *Hindawi Publishing Corporation, Scientific World Journal*, Volume 2014.

[16] Kuo-Chen Chou, "Prediction of Protein Cellular Attributes Using Pseudo-Amino Acid Composition," *PROTEINS: Structure, Function, and Genetics* 43, pp. 246–255, 2001

[17] Jian Yanga, Jing-yu Yanga, David Zhangb and Jian-feng Lua, "Feature fusion: parallel strategy vs. serial strategy," *Pattern Recognition*, vol. 36, pp. 1369 – 1381, 2003.

Fruit Fly Algorithm based Clustering Protocol in Wireless Sensor Networks

Anamika Dey[1*], Tamal Sarkar[2], and Sharafat Ali[3]

[1]Department of CSE, University of Asia Pacific
Dhaka, Bangladesh

[2]Department of EEE, Bangladesh University of Engineering & Technology
Dhaka, Bangladesh

[2]School of Computer & Communication Engineering, University of Malaysia Perlis
Perlis, Malaysia

[*]anamika.dey.bd@ieee.org

***Abstract*—In recent days, wireless sensor networks (WSN) are catching the spotlights in networking and other emerging fields like large data communication, artificial intelligence, automation systems etc. The major constraint wireless sensor networks are struggling with limited energy supply issues for their sensor nodes. This paper is focused to put emphasis on implementing a clustering protocol inspired from Fruit Fly (FF) movement to reduce the power usage of the sensor nodes by using shortest paths possible for data transmission and in turn enhancing the life time of the network. The whole network is partitioned into several areas by the FF algorithm. At first place, energy efficiency was the vital motto to thrive for implementing FF algorithm. Network lifetime analysis using CH selection method and data transmission efficiency were tested using FF algorithm based clustered in WSN. A comparative analysis was then took place to compare existing and commonly practiced algorithm such as Cuckoo Search, LEACH, LEACH-C algorithm based protocols against the proposed FF algorithm.**

***Index Terms*- Fruit Fly (FF) algorithm, Wireless Sensor Network (WSN), Cuckoo Search (CS), Cluster Head (CH), Base Station (BS), Smell Function.**

I. Introduction

Wireless sensor network is an emerging technology that facilitates the solution of many shortcomings of automation system, security problems, smart grids and cyber security issues. This is a really handy solution for efficient and reliable data communication. WSN consists of numerous sensor nodes which consume little power and miniature in terms of size. Every node is capable enough to perform as autonomous devices and can be deployed in different environments. All nodes have the capabilities of sensing, computing, reporting the acquired data and monitoring through the wireless transceiver to the base station. Sensor nodes are deployed in the monitoring area. The primary issue is lifetime extension of the deployed sensor nodes. The sensor nodes are provided with some measure of detecting, control, information preparing and communicating with base station. WSNs are critically resource constrained by memory, limited power supply, processing performance and communication bandwidth. Accordingly, moving on each situation, sensor nodes depend on restricted supply of energy (utilizing batteries). Considering these energy sources in the field is normally not achieved, and a WSN must be kept alive for a given time or as far as might be feasible. The current works in WSNs area are handling the energy effectiveness.

For highly energy efficient operation of WSNs, cluster based routing protocols are considered as well as–established techniques. They also have special advantages related to effective communication and scalability. In recent days, numerous cluster-oriented protocols have been presented with a view to extending the network lifetime. Collecting data from each sensing node and transmitting those data to the base station properly is a major concern. But collecting the sensed data from all the sensor nodes is difficult when the number of sensor nodes get increased. It takes a long time to collect all the sensed data from all sensor nodes individually. It will increase the data receiving time and also reduce the efficiency of sensor node's battery. This is the main reason of reducing the network lifetime. LEACH is a cluster based protocol that uses a distributed clustering development algorithm. Here, sensing data from every node consumes energy in a fast manner. [1] The choice of CHs is constructed on a predefined possibility. Other node pick the cluster to join by approximating the adjacent distance to the CHs selection. LEACH-C uses the centralized clustering protocols. The clustering protocol of LEACH-C improved the clustering procedure and scattering the nodes of CH all through the network. Both LEACH and LEACH-C do not discover the optimal choice of cluster based on any objective function. CS is a population based optimization technique that produced better response in terms of energy consumption and data transmission compared to LEACH and LEACH-C.

Transmitting the data to the base station through a network in shortest possible time is desirable. CHs collect the data from every cluster nodes and send those information to the BS by using TDMA process. Updating the information and tracking the nodes' data is vital for data transmission. As far as data transmission system's concern, lossless data transmission is the most vital feature.

Various applications of WSN in different environments alike management systems during disasters, smart transport system using artificial intelligence, healthcare or even at battle fields reduces physical efforts to a larger extent.

A new prominent method for realizing universal optimization is Fruit fly optimization algorithm which is based on the foraging comportment of the fruit fly. Compared to the sense of smell and vision, fruit fly is way better in terms of sensory perception. The olfactory organ of a fruit fly can receive

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

various smells from the atmosphere, and even from a victuals source located 40 km away. [2,3] Henceforth, the fruit fly hovers to the source, uses its sharp vision to search for the victuals and where its fellows gather, and then it flies to the direction, which is depicted in fig 1. The protocol divides the entire network into some clusters by Fruit Fly algorithm. The implementation of FF algorithm in wireless sensor networks is the novel feature of the research work. The algorithm deliberates energy consumption of the nodes in the optimization procedure. The established protocol later was verified against the earlier established protocol like LEACH, LEACH-C and CS which are the recently used bio inspired optimization algorithms.

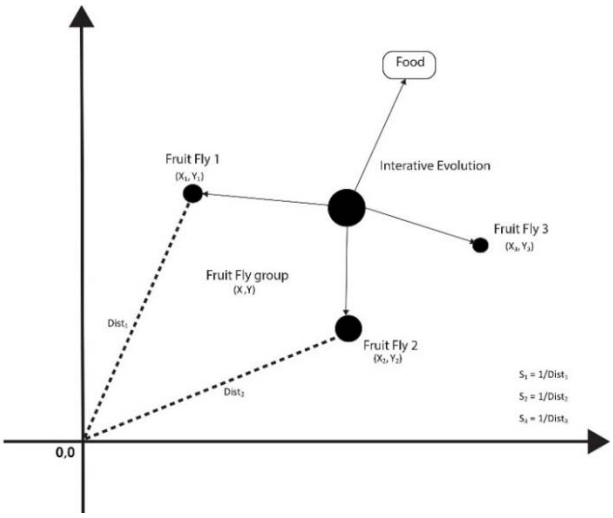


Fig 1 : Iterative search for food of fruit fly swarm [3]

II. The System Model

A. Network Structure

We considered a network arrangement comparable to those utilized in [4], with the associated features: the base station is unchanged and located in the sensor network area. Every sensor node is fixed and perform sensing. The sensor nodes complete their task and send the data information to the BS. Every node is suitable for working in cluster head (CH) mode and detecting mode relying upon the level of residual energy. [11]

B. Wireless Energy Model

The wireless energy model for the sensors utilized as a part of our protocol depends on the first order wireless model as utilized as a part of [5,6].In this model, to accomplish an agreeable Ratio of Signal-to-Noise (SNR) in transmitting an l-bit data over a separation d, the transmitting energy is given by:

$$\begin{aligned} E_{TX}(l,d) &= l \cdot E_{elec} + l \cdot \varepsilon_{FS}\, d^2, \quad \text{if } d < d_0 \\ &= l \cdot E_{elec} + l \cdot \varepsilon_{TR}\, d^4, \quad \text{if } d \geq d_0 \end{aligned} \tag{1}$$

Where, E_{elec} is the degenerated energy for each bit of information which is utilized to run the transmitter or the receiver. ε_{FS} and ε_{MP} shifts as for transmitter amplifier model that is utilized. d_0 indicates the threshold estimation of the transmission distance. Thus, to get l bit data packet, the energy consumed by the radio is:

$$E_{RX}(l) = l \cdot E_{elec} \tag{2}$$

III. Detailing of Fruit Fly Algorithm

FF is one of the swarm optimization algorithm based on the foraging behaviour of fruit fly using the vision and smell function. [7] FF is mainly used to select the minimum distance between CHs and cluster nodes. Initially, it measure the distance between the neighbours CHs and then takes the reciprocal of that value. After that it compare with each measured value and take the minimum one.

FF algorithm sums the value along the particular distance and then it applies the two main functions of fruit flies. From the entire possible distance depends on the vision function, it only selects the minimum distance.

Fruit Fly algorithm procedure is segmented below:

- Initialize the algorithm variables.
- Defining the algorithm Variables.
- Estimate the distance and concentration value.
- Calculating fitness function
- Determine maximum concentration.
- Keeping best concentration value.
- Checking and repetition.

These steps are represented in the next seven subdivisions:

Step 1: The random initial position of cluster is in (0,0) coordinates such that,

$$\text{Init } X_{ax} = 0; \quad \text{Init } Y_{ax} = 0$$

Step 2: Controlling the direction of searching the neighbours distance from each CH individuals.

$$Xi = X_ax + \text{Rand Val}$$
$$Yi = Y_ax + \text{Rand Val}$$

Step 3: The location of the neighbour CHs can be known, then the distance (D) to the neighbour CH nodes are estimated before the decision value of smell concentration (S) is calculated, this value is the reciprocal distance.

$$D_i = \sqrt{(Xi^2 + Yi^2)}$$
$$S_i = 1 / D_i$$

Step 4: The smell concentration value substituted in the smell concentration function or fitness function to work out the smell in the position of the fruit fly individual.

$$\text{Smell}_i = \text{Func}\ (S_i)$$

Step 5: Determine the distance having minimum node count and maximum smell concentration among the paths.

[bSmell pNode] = min (node)
[bSmell bNode] = max (smell)

Step 6: Retain the best smell value and node sending the collected data through the network by using fitness value.

Smellb = bNode

Step 7: Enter into iterative optimization, for each node position the sensor nodes are in controlling manner.

IV. Cluster Setup Using FF Algorithm

The FF algorithm is an optimization clustering algorithm. . Let us assume a framework of sensor nodes of N numbers and that nodes are arranged in P clusters: C_1, C_2, ..., C_p, where the BS is fixed and the nodes divided into many clusters based on positional data and energy level of the nodes. The process of data sending is divided into rounds. In each round, the CHs collect information from every single group part and transmit

to the BS. The entire job completes in 2 stages: first one is clustering model and another is data transmission. [11]

A. Process of Clustering Model

In this clustering framework portion, all sensor nodes transmit data to the BS with the data of their physical position and the energy level. Taking into account this information to the BS will process the normal energy level of every single dynamic node. Just the nodes which have remaining energy upper than the normal level are qualified for being an immediate CH for the running round or circle. To improve the determination of CHs, FF Algorithm is used to discover optimal p cluster heads, the target function in is utilized to calculate the minimum, as appeared underneath:

$$f_1 = max_p \ \{\sum_{\forall i \in Cp} \ d(n_i, CH_p)/|C_p|\} \quad (3)$$

$$f_2 = \sum_{i=1}^{N} E\,(n_i) \,/\, \sum_{p=1}^{K} E\,(CH_p) \quad (4)$$

$$f(x) = \alpha \times f_1 + (1-\alpha) \times f_2 \quad (5)$$

In this cost function, *f1* is the most extreme of the Euclidean space between nodes, node i∀i ∈ cluster C_p, to their cluster heads CH_p and $|C_p|$ is the number of nodes that fit in with cluster C_p. Meanwhile f2 is the ratio of initial energy of every active node, E(nodei) in this network frame with the aggregate existing drive of the running cluster head, $E(CH_p)$, in the running round. The steady α demonstrates the commitment of f1 and f2 in the target capacity f(x). This target function has a tendency to deteriorate the intra-cluster average distance among sensor nodes and their cluster head and in addition to upgrade the energy efficiency of the structure of the network. Once the cluster is structured, the non-cluster head nodes transmit information towards the BS through the cluster heads. The procedure of selecting clusters is rehashed each cycle of exchanging information among sensor nodes. Amid the transmission from the nodes to CH, remaining energy of every node are joined to the information packet, this data helps the BS to choose the most appropriate cluster head and cluster at the following round. Taking into account the number of the dynamic nodes inside of the cluster, the CH makes a timetable in view of Time Division Multiple Access (TDMA). [11]

B. Data Transmission

Initially, the CHs are configured. The routine of the transmission process is made, the nodes of each sensor start to transmit data to the CHs. The non-cluster head nodes transmits their data to the CH. The CH accumulates the total data of a cluster and sends the total data to the BS. The total data sending process from a CH works as a cycle after completing one cycle in new CH is selected based on highest node energy. Furthermore, TDMA method is used; non-cluster head nodes changes their transmitting node amid the process and kill subsequent to complete transmission. Information gathering and accumulation is prepared at the CHs, along these lines the measure of data is decreased, and the CHs just lead the concentrated information to the BS. [11]

C. Flow Chart of CH Selection Using FF Algorithm

In the fig 2, the flow chart of cluster head selection in Wireless Sensor Network is illustrated below:

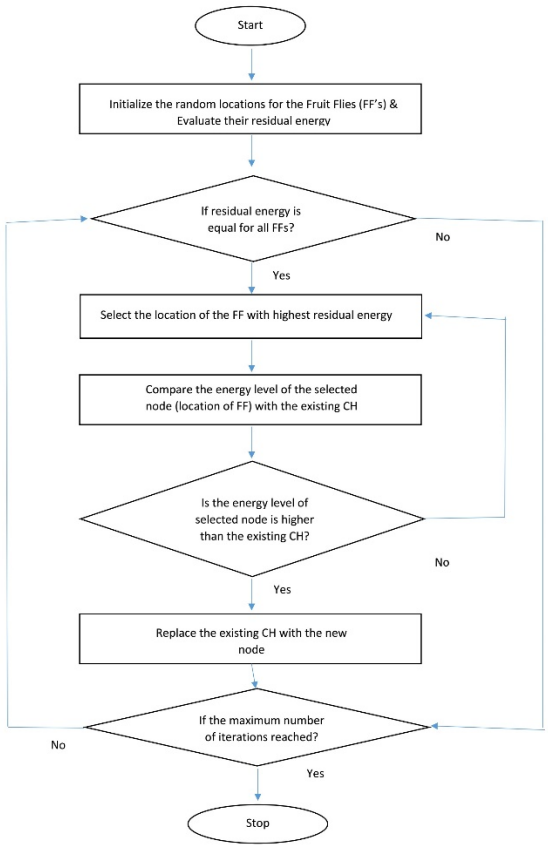


Fig 2: Flow Chart of CH selection using FF algorithm

V. *Simulation Result and Discussion*

Fig 3 and 4 describes the network lifetime of total nodes and the rounds completed by number of nodes. CH collect data from sensor nodes and send the data to BS is called one round. The lifetime measured the characteristics of nodes over cycle (1000 rounds, 100 nodes). It depicts that the FF algorithm is better than the other protocols network life as LEACH, LEACH-C, and CS, by about 54 %, 38 %, and 8 % individually. This change depends on two reasons. Firstly, with least intra cluster distance and optimal CH dispersion over the network, the FF convention indicates enhanced system apportioning. Thus, the energy scattered by all nodes for correspondence is lessened. Besides, the FF system received in the convention creates an arrangement of good exchange offs where the estimations of the cost capacity are middle of the road to the network prerequisites.

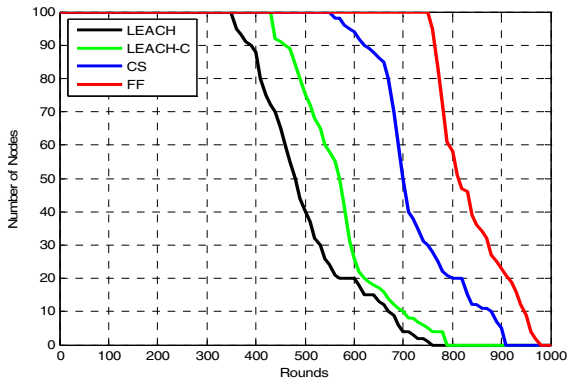


Fig 3: Network lifetime for the systems: LEACH, LEACH-C, CS, FF

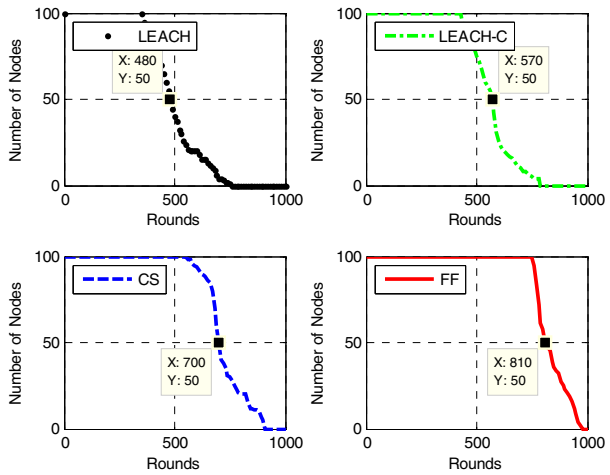


Fig 4: Network lifetime for individual protocols

Fig 5 demonstrates the aggregate sum of information messages got at BS by every one of the conventions and algorithms. Fruit Fly (FF) enhances the information conveyance by elements of 11% over CS, 54% over LEACH-C and 161% over LEACH. The thought behind this is, our protocol can take the advantage of selecting the high energy hub as a CH by mulling over the leftover energy of the CH hopefuls furthermore the base distance between the nodes and their CHs by actualizing advanced cost capacities. Subsequently, more information messages are conveyed to the BS.

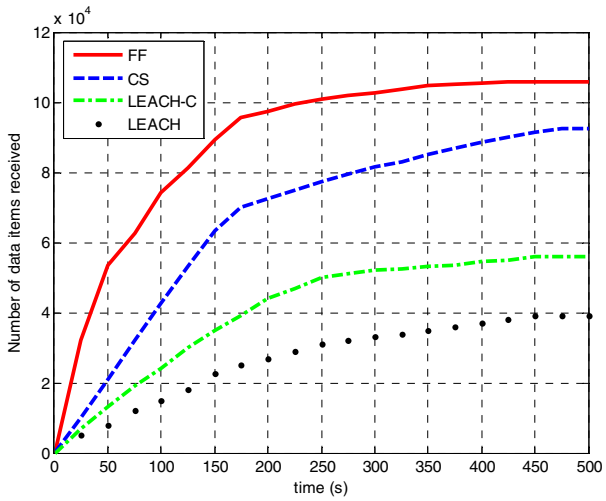


Fig 5: Entire data received at BS with respect to time

VI. Conclusion & Future Scope

This write-up delivers a hierarchical clustering protocol for a low energy-consuming WSN which is inspired by the FF algorithm. The FF algorithm used for finding the minimum intra-cluster distance and paves an efficient way for cluster head selection that ultimately sums up as an energy-efficient setup. Different functions such as cost function were used in a manner with a view to selecting minimum distance among non-CH nodes and the relevant CH node; residual energy of the nodes plays the role of selection parameter for choosing cluster heads. The simulation results show that WSN using FF algorithm has better active lifetime and larger capacity of transmitting data compared with Cuckoo Search, LEACH-C or LEACH based WSNs. In addition to these, the proposed algorithm reshapes the clustered network in an energy-efficient way with better performance in terms of data efficiency. The scope for further investigation includes cross-layer optimization, multi-hop communication and new directing techniques in the emerging area of WSN.

References

[1] *Q. Liao and H. Zhu (2013), 'An Energy Balanced Clustering Algorithm based on LEACH Protocol', In Proceedings of the 2nd International Conference on System Engineering and Modeling,* (ICSEM-13),pp.1272-1280.

[2] *S.Anandamurugan, "An Enegy – Efficient Min-Max Optimization with RSA Security in Wireless Sensor Networks,"* IJMTST, vol. 2, Issue 4, April 2016.

[3] *Pan, W.T. (2011), "A new Evolutionary Computation Approach: Fruit Fly Optimization Algorithm", Conference on Digital Technology and Innovation Management.*

[4] *W.R., Heinzelman, A., Chandrakasan, H., Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Microsensor Networks"* 33rd Hawaii International Conference on System Sciences. pp. 1–10. IEEE, Hawaii (2000).

[5] *W.B., Heinzelman, A.P., Chandrakasan, H., Balakrishnan,: "An application-specific protocol architecture for wireless microsensor networks"* IEEE Transactions on Wireless Communications. 1, 660–670 (2002).

[6] *E. Amiri, H. Keshavarz, M. Alizadeh, M. Zamani, T. Khodadadi, "Energy efficient routing in wireless sensor networks based on fuzzy ant colony optimization, International Journal of Distributed Sensor Networks",* 2014 (2014)1–17.

[7] *N.S.Choubey, "Fruit Fly Optimization algorithm for Travelling Salesperson Problem"* IJCA, Vol. 107, No 18, December 2014.

[8] *P.Kuila, P. Jana, "Energy Efficient Clustering and routing algorithms for Wireless Sensor Network: Particle Swarm Optimization Algorithm approach",* Journal on Communication System, 2014 Vol. 18, No. 6, PP.1016-1025.

[9] *M.Baskaran and C.Sadagopan, "Synchronous Firefly Algorithm for Cluster Head Selection in WSN"*, Journal on Sensor Networks, 2015 Vol. 79, No. 6, pp. 780-789.

[10] *R.Kumar and D.Kumar, "Hybrid Swarm Intelligence Energy Efficient Clustered Routing Algorithm for Wireless sensor Networks",* 2015, Journal on Sensors, Vol.10, No.5, pp.1155-1174.

[11] *A. Dey, T. Sarkar, A. Ullah, N. Nahar, "Implementation of Improved Harmony Search Based Clustering Algorithm in Wireless Sensor networks"*, In proceedings of ICA-ICT 2016, May 2016, Chittagong, Bangladesh.

Analysis of Path Loss Characteristics in Body Area Network for Different Physical Structures

Hasanur Rahman Chowdhury,[1*] Md. Sahil Hassan,[1] and Anis Ahmed[1]

[1]Department of Electrical & Electronic Engineering (EEE), University of Dhaka
Dhaka-1000, Bangladesh

[*]hasan114@ymail.com

Abstract— **Wireless Body Area Network (WBAN) is one of the latest technologies in wireless communication, and is being widely studied for its applications in military, healthcare, sports, and other sectors. In this paper transmission coefficients S_{21} are measured for several on-body radio channels and the effects of two different body structures on path loss coefficients are determined. Considering static position of body in the laboratory, measurements are carried out for Line-of-Sight (LOS) and Non-Line-of-Sight (NLOS) radio channels by using two identical tri-slot microstrip patch antennas of 2.48 GHz connected with a vector network analyzer (VNA). Using necessary statistical method path loss coefficients are calculated from the measured data and compared to explore the effect of different types of bodies on WBAN performance. Finally, we have tried to justify the reasons for considering LOS and NLOS measurements separately in establishing an accurate channel model for BAN.**

Index Terms—**Body Area Network (BAN); Path Loss; Path loss coefficient; ISM band; Line of Sight (LOS).**

I. Introduction

BAN (Body Area Network) or WBAN (Wireless Body Area Network) is the network consisting of wearable/ implanted sensors and devices on or in human body. IEEE has standardized BAN for low power devices and operation on, in or around the human body to serve a variety of applications including medical, consumer electronics, personal entertainment and other [1]. Among three types of BAN scenarios, this research focuses on 'on-body channel' scenario. Among the numerous research works on BAN, the contemporary progress and challenges in this field have been described in works [2], [4]. The path loss measurements for different links on body under different circumstances are described in works [3], [5]-[11]. Among these works, different statistical or analytical models for BAN are proposed in [3], [5], [8] and [9]. The channel models provided in these works do not take into account the body structures where antennas are mounted. So, deviation from expected performance may arise when applying these models on different types of bodies. It is also necessary to separate LOS and NLOS link measurements so that shadow fading effect from NLOS link measurements does not get merged with LOS link measurements. But the measurements in LOS and NLOS links were not evaluated separately in the above mentioned works either. To the authors' knowledge, there has not been any work that demonstrates the effect of different body structures on BAN performance, or proposes path loss models for BAN considering LOS and NLOS channels separately. The comparison is very important as it's the key to understand the performance of BAN on different types of bodies.

In this paper, the variation in BAN due to the structural difference of body is focused. Two test bodies with significantly different physical structure are taken and same sets of measurements are performed on them with the same instruments in the same environment to understand the effect of body structure on BAN. For this experiment, two tri-slot patch antennas were used with resonance frequency at 2.48 GHz which is nonhazardous to human body. A fair number of on-body radio channels were examined by a Vector Network Analyzer (VNA) to measure the forward transmittance S_{21}. The measured data were post-processed by MATLAB to obtain the path loss coefficient on different bodies and compared to conclude the effects of different human body structures on BAN.

In section II, the measurement environment is described along with the important parameters. In Section III the measurement setup for different on-body links are shown. Section IV and V contain the measurement results and conclusion respectively.

II. Measurement Scenario

A. Measurement Room and Test Person

The experiment was conducted in a laboratory of size 8m x 6m x 4m, with floor and walls made of concrete. Two adult male persons of significantly diverse physical structures were taken as test bodies. The specifications of their body structures are given in Table I. Table I reveals that Male-1 represents a short and skinny body while Male-2 represents a comparatively large and fat body. Both the persons were wearing cotton T-shirt, jeans and no shoes. All metallic objects, e.g. ring, belt clasp and watch were removed from the bodies during the measurements. Room temperature was 24 degree Celsius during the experiment.

TABLE I
Body Specifications of test bodies

Parameter	Male-1	Male-2
Height	162 cm	167.5 cm
Weight	49 kg	90 kg
Chest	80 cm	114 cm
Waist	73 cm	111 cm
Thigh	42 cm	55 cm

B. Measuring Equipments and Antenna under Test:

The experiment was conducted using a handheld Vector Network Analyzer (VNA) (R&S ZVH8) which covers the ISM frequency band of our interest. Two miniaturized tri-slot patch antennas, fabricated on FR4 substrate of thickness 1.6 mm and relative permittivity 4.3 and having identical characteristics were chosen as reference antennas. One of

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

these was used as Tx and the other as Rx antenna. Fig. 1 shows the return loss of Tx antenna placed in air at the laboratory.

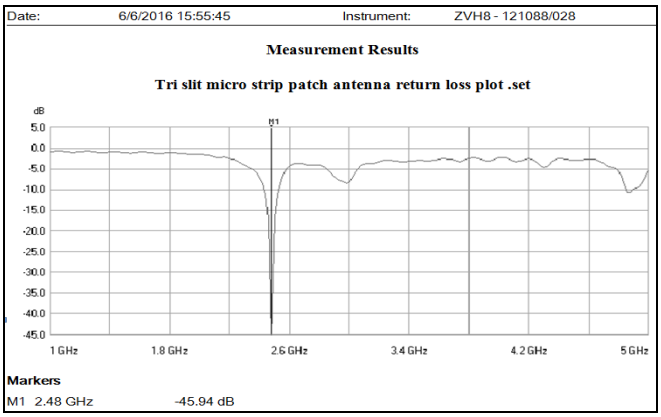


Fig. 1 Return loss (S11) of Tx antenna.

Both the antennas were connected to the VNA via two identical co-axial cables of model LMR-200 having 3m of length each. Full 2 port calibration was performed on VNA by R&S FSH-Z28 calibration kit before using the VNA for measurements. All the equipments specifications are gathered in Table II.

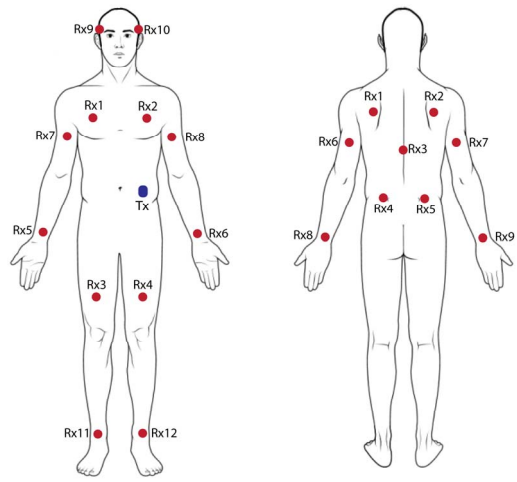
TABLE II

MEASUREMENT EQUIPMENT SPECIFICATION [12][13]

Equipment	Parameter	Value
VNA	Model	R&S ZVH8
	Frequency Range	100 kHz to 8 GHz
Antenna	Type	Tri-slot patch
	Resonant Frequency	2.48GHz
	Return loss	-45.94 dB
	Impedence	(51.3-j1.38) Ω
	VSWR	1.01
	Bandwidth	95 MHz
Cable	Model	LMR-200
	Length	3m
	Frequency Range	Dc to 5 GHz
	Connector Type	N to SMA (Male)
Calibration kits	Name	R&S FSH-Z28
	Calibration Type	Full 2 port Cal

III. MEASUREMENT SETUP

While taking the measurements, Rx antenna was placed on different parts of the body whereas Tx was kept fixed on left abdomen all the time. Fig. 2 shows the different Rx positions for both LOS and NLOS measurements. In order to keep a constant gap between the antenna and the body, a 5 mm thick piece of hard paper was kept between the antenna and body. In practice, the antennas were fixed on the body using elastic belts having plastic made clamps together with masking tape. While recording the measurements, the body was in a static standing position and arms were hanging along the sides.



(a) (b)

Fig. 2 Different Tx and Rx positions for (a) LOS and (b) NLOS links.

Antenna orientation was kept head to head. Fig. 3 shows an exemplary view of antenna placement on body during measurement. The links were mounted one-by-one on both test bodies and 10 snapshots of S_{21} were captured for each Tx-Rx link by the VNA at an interval of 1 second. The measurements were repeated by re-mounting the antennas in order to reduce the effect of the variation of antenna positioning on the results.

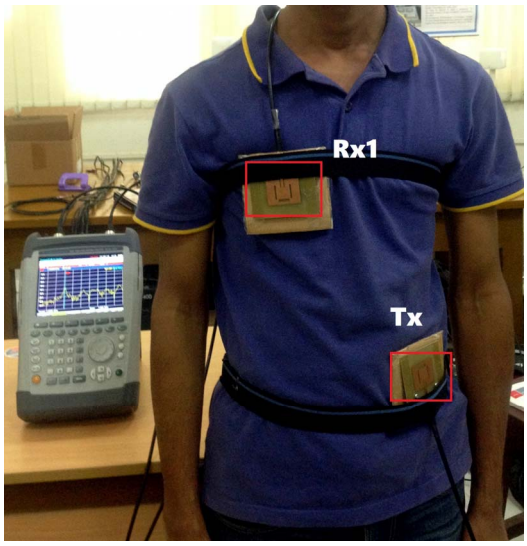


Fig. 3 Measurement Setup.

IV. MEASUREMENT RESULTS

A. Path Loss Coefficient Calculation

The power received at Rx antenna for distance d separation between Tx and Rx follows the inverse square law in free space, according to Frii's equation:

$$P_r \propto 1/d^2$$

But in reality the power received is inversely proportional to distance by exponent of a variable γ and proportional to transmitted power. So,

$$P_r \propto P_t / d^{\gamma}$$

$$P_r / P_t \propto 1/d^{\gamma}$$

Considering antenna gains as unity and taking dB on both sides, we find [14]

$$10\log_{10}(P_r / P_t) = const. + (-\gamma)10\log_{10} d \qquad (1)$$

where, γ is the path loss coefficient and $const. = 10\gamma\log_{10}(4\pi/\lambda)$

The left hand side of Eq. (1) represents the path loss PL, we can define

$$PL(dB) = 10\log_{10}(P_r / P_t) \qquad (2)$$

Considering a reference distance d_0 (usually taken 1cm), Eq. (1) can be rearranged with the help of Eq. (2) as [15]

$$PL(dB) = PL_0(dB) - 10\gamma\log_{10}(d/d_0) \qquad (3)$$

Where, PL_0 is the path loss at reference distance d_0. Eq. (3) is used to calculate path loss coefficient.

Mean Path loss (dB) data for various distances were first plotted in a scatter diagram in Matlab. A linear least square (LS) polynomial fitting was then applied to the plotted data to draw the best fitting line. The slope of this line indicates the value of path loss coefficient γ and the intercept of the line is a measure of PL_0.

B. Analysis of measured and calculated data

The path loss coefficients on different bodies are determined from measured data through calculation. The bar charts of Fig. 4 and Fig. 5 show the comparison of path loss ($-S_{21}$) at different links for Male-1 and Male-2 for LOS and NLOS channels, respectively. It can be seen that most of the links on Male-2 body suffers more attenuation compared to Male-1 body, as expected.

From Fig. 4 it can be seen that, among LOS links, for Rx antenna position at right chest, right arm, right and left head, attenuation of greater than 10 dB was caused by bigger body compared to smaller one. But for Rx positions on right arm, and left arm, path loss on smaller body was greater than the bigger one by around 5 dB. Rest of the LOS links show slight increase in path loss for larger body, but they are less than 5dB.

For NLOS link data shown in Fig. 5, left and right shoulders, right lower back and left arm are the receiver antenna positions where path losses are more than 10 dB greater for larger body. At right arm position path loss is more than 5 dB greater on larger body. In case of Rx antenna on left wrist, path loss for Male-1 body is greater by an amount less than 5 dB. Rest of the Rx positions exhibit slight increase in path loss for Male-2 body.

In order to determine the distance dependency of path loss on both bodies, the obtained S_{21} are plotted in Fig. 6 and Fig. 7 using Matlab for LOS and NLOS, respectively. Using least square polynomial fitting method, two straight lines are drawn in both figures to find out the slope (which is the path loss coefficient). For Male-1, the path loss coefficient increases by 1.0481 in NLOS channel compared to LOS, while for Male-2 the increase is 1.87. So it is revealed that the attenuation in NLOS is more severe if the user body is bigger or fatter. This implies that calculation of path loss model while considering both LOS and NLOS link data under same category should lead to biased results depending on test body structure. So, it is necessary to consider both channels separately.

If we compare the LOS data for Male-1 and Male-2, the difference in path loss coefficient between both bodies is 0.1809 as show in Fig. 6. But the difference increases remarkably for NLOS channel and becomes 1.0069 as shown in Fig. 7. In both cases, data on Male-2 show greater path loss coefficient than Male-1. So if there is LOS between both on-body antennas, difference in body structure does not cause much variation in channel performance. But the on-body channel performance is more sensitive to different body structures for NLOS communication. Thus it is clear that path loss measurements on different types of bodies are needed to be considered during the development of an accurate on-body path loss model.

In Table III the measured path loss coefficient γ and PL_0 for all the 4 cases are shown. It is observed that our path loss coefficients for LOS links in Table III are close to the one calculated in [10]. By taking the average of path loss coefficient γ for LOS and NLOS (Male-1) case we found 2.80 which is close to the values found in [7] in office environment and in [9] (for loop antenna and with error correction). So, our calculated data can be used for channel characterization of BAN.

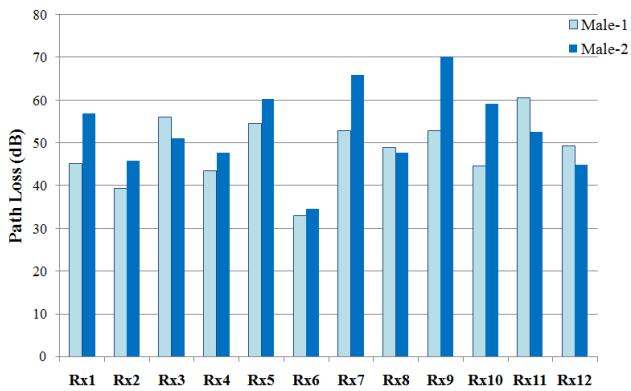


Fig. 4 Comparison between path losses in LOS links for Male-1 & Male-2.

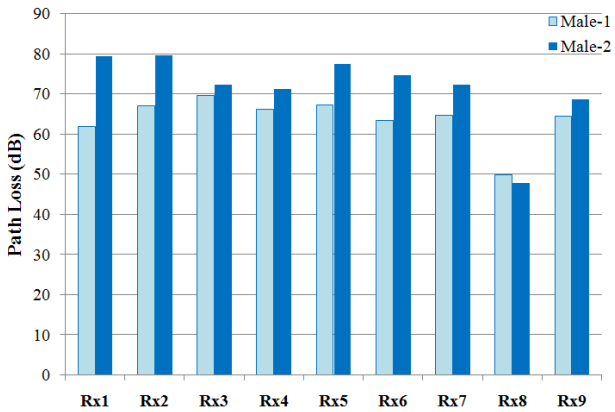


Fig. 5 Comparison between path losses in NLOS links for Male-1 & Male-2.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

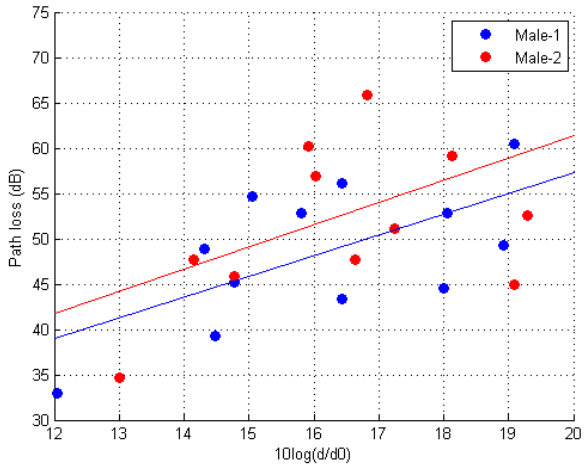


Fig. 6 Comparison of distance dependence of mean path loss for LOS channel.

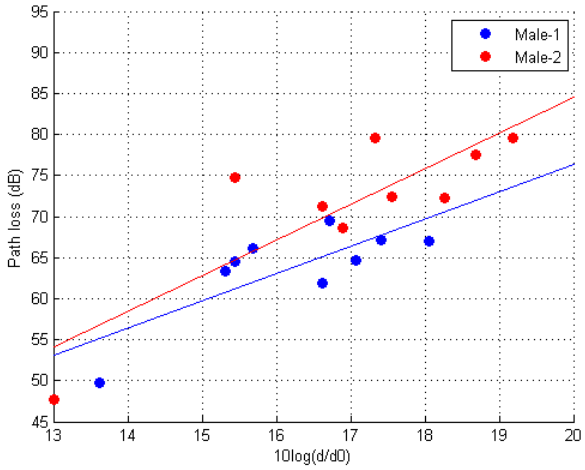


Fig. 7 Comparison of distance dependence of mean path loss for NLOS channel.

TABLE III

MEASURED PATH LOSS COEFFICIENT AND PL_0 FOR ALL CASES

Link Type	Test body	Path Loss coefficient, γ	PL_0 (dB)
LOS	Male-1	2.2817	11.6034
	Male-2	2.4626	12.1758
NLOS	Male-1	3.3298	9.7704
	Male-2	4.3367	-2.2063

V. CONCLUSION

The main objective of this study was to investigate the impact of the body structure on path loss characteristics. It is clear from Table III that the path loss coefficients obtained for the obstructed channel (NLOS) are observed to be higher than that obtained for unobstructed (LOS) channel for both bodies. This implies that presence of body organ in propagation path causes more attenuation or weakening of the signal power compared to LOS links. The path loss coefficient for Male-1 increases by 1.0481 in NLOS channel compared to LOS, while for Male-2 the increase is 1.87. So we may say that the attenuation in NLOS is more severe if the user body is bigger or fatter. This implies that calculation of path loss model while considering both LOS and NLOS link data under same category should lead to biased results depending on test body structure. So, it is necessary to consider both channels separately. Table III shows that our data agrees well with some of the published works on BAN. Thus for accurate channel characterization of BAN, it is essential to consider LOS and NLOS data separately on different types of bodies.

ACKNOWLEDGEMENT

This work was conducted in the Microwave & Optical Fiber Communication Laboratory funded by HEQEP, in the Department of Electrical and Electronic Engineering (EEE), University of Dhaka.

The authors would like to thank Dr. Khaleda Ali, Assistant Professor, ULAB, for her fruitful guidance and Mr. Shibli Sadique, M.S. thesis student of this laboratory for fabrication of the tri-slot microstrip patch antennas used in this work.

REFERENCES

[1] Erik Karulf. (2008). A survey paper on Body Area Network (BAN). [Online]. Available: http://www.cse.wustl.edu/~jain/cse574-08/ftp/ban/index.html

[2] A. Pellegrini, A. Brizzi, L. Zhang, K. Ali, Y. Hao, X. Wu, C. C. Constantinou, Y. Nechayev, P. S. Hall, N. Chahat, M. Zhadobov and R. Sauleau, "Antennas and propagation for body-centric wireless communications at millimeter-wave frequencies: A Review", *IEEE Antennas and Propagation Magazine,* vol. 55, pp. 262-287, Aug. 2013.

[3] A. Vorobyov, and A. Yarovoy, "Human body impact on UWB antenna radiation", *Progress In Electromagnetics Research M,* vol. 22, pp. 259-269, 2012.

[4] R. Kohno, K. Hamaguchi, Huan-Bang Li and K. Takizawa, "R&D and standardization of Body Area Network (BAN) for medical healthcare", *IEEE International Conference on Ultra-Wide Band,* vol. 3, pp. 5-8, Sept. 2008.

[5] L. Liu, R. D'Errico, L. Ouvry, P. De Doncker and C. Oestges, "Dynamic channel modeling at 2.4 GHz for On-Body Area Networks", *Advances in Electronics and Telecommunication,* vol. 2, pp. 18-27, Dec. 2011.

[6] T. B. Welch, R. L. Musselman, B. A. Emessiene, P. D. Gift, D. K. Choudhury, D. N. Cassadine and S. M. Yano, "The effects of the human body on UWB signal propagation in an indoor environment", *IEEE Journals on Selected Areas in Communication,* vol. 20, pp. 1778-1782, Dec. 2002.

[7] T. Zasowski, F. Althaus, M. Stager, A. Wittneben abd G. Troster, "UWB for nonivasive Wireless Body Area Networks: Channel measurements and results", in *IEEE Conference on Ultra Wideband Systems and Technologies,* 2003, p. 285.

[8] David B. Smith and Leif W. Hanlen, *Ultra-Low-Power Short Range Radios,* Switzerland, Springer International Publishing, 2015.

[9] T. Kumpuniemi, T. Tuovinen, M. Hämäläinen, K. Y. Yazdandoost, R. Vuohtoniemi and J. Iinatti, "Measurement-based on-body path loss modelling for UWB WBAN communications", in *7th International Symposium on Medical Information and Communication Technology,* 2013, p. 233.

[10] A.-M. Gao, Q. H. Xu, H.-L. Peng, W. Jiang and Y. Jiang, "Performance evaluation of UWB on-body communication under WIMAX off-body EMI existence", *Progress in Electromagnetics Research,* vol. 132, pp. 479-498, 2012.

[11] Hall P S and Hao, Y, "Antennas and propagation for body centric communications", in *First European Conference on Antennas and Propagation,* 2006, pp. 1.

[12] "R&S ZVH8 Vector Network Analyzer datasheet", Rohde & Schwarz GmbH & Co KG, Munich, Germany.

[13] "LMR200 N-Type Male to RP-SMA Female Pigtail Cable datasheet: TL-ANT 200PT", TP-LINK Technologies Co. Ltd., Shenzhen, Guangdong, China.

[14] Lorne C. Liechty, "Path Loss measurement and Model analysis of a 2.4 GHz Wireless Network in an outdoor environment", M.S. thesis, Dept. Elect. Eng., Georgia Institute of Technology, Georgia, USA, Aug. 2007.

[15] Theodore S. Rappaport, *Wireless Communications: Principles and Practice,* 2nd ed., India, Pearson, 2010.

Designing a Transcranial Magnetic Stimulator Coil for Deep Brain Stimulation

Md Mahmudul Hasan,[1,*] Shahed Md. Abu Sufian[2], Hasan Mehdi[3], Khondkar Siddique-e-Rabbani[4]

[1234] Department of Biomedical Physics and Technology, University of Dhaka, Dhaka - 1000, Bangladesh

[*]mahmudulrussel@bmpt.du.ac.bd

Abstract—**Human brain is a very complex organ and may suffer from various diseases and disorders. Therefore, techniques are required to diagnose the disorders and provide necessary treatment. Electrical or magnetic deep brain stimulation has been successfully used for both these purposes. This research develops a type of therapy that uses stimulation to treat neurological disorders, diseases of the central and peripheral nervous system which comprises of the brain, spinal cord, nerves, neuromuscular junction and muscles. When medications have become less effective and/or when side effects of the medications interfere with daily activities, doctors may prescribe Deep Brain Stimulation (DBS). Transcranial Magnetic Stimulation (TMS) is a recent non-invasive clinical treatment usually reserved for those suffering from severe treatment-resistant depression. Deep TMS can be effective in common diseases associated with neurological disorder: Parkinson's disease (PD), multiple sclerosis, Alzheimer and other traumatic disorders of the nervous system. This paper focuses on designing a suitable transcranial magnetic coil for DBS. Deep TMS, in the field of diagnosis and therapy, is a new line of research in Bangladesh as well as a recent one in the world perspective.**

***Index Terms*— transcranial magnetic coil, deep brain stimulation, neurological disorder, neuromuscular disease, Parkinson's disease, multiple sclerosis.**

I. Introduction

The electromagnetic stimulation of brain is a type of therapy that uses stimulation to treat neurological disorders, diseases of the central and peripheral nervous system which encompasses the brain, spinal cord, cranial and peripheral nerves, nerve roots, autonomic nervous system, neuromuscular junction, and muscles. The earliest technique used is Electro Convulsive Therapy (ECT) which uses large surface electrodes on the skull to stimulate a relatively large area of the brain, sometimes producing deliberate seizures, and was used as a therapy for severe mental disorders. A recent technique is Deep brain stimulation (DBS) which uses surgically implanted electrodes which are connected to battery powered pulse generators placed surgically inside the upper chest. This was first developed as a treatment for Parkinson's disease to reduce tremor, stiffness, walking problems and uncontrollable movements. However, it has also been found to improve depressions and Obsessive Compulsion Disorders (OCD). DBS can be effective in common diseases associated with neurological disorder: Parkinson's disease (PD), multiple sclerosis, Alzheimer disease, epilepsy, essential tremor, and other cerebrovascular diseases including stroke, migraine and other headache disorders, neuro-infections, brain tumours, traumatic disorders of the nervous system due to head trauma, and certain other neurological conditions [1]. Movement-related symptoms of PD and other neurological conditions are caused by faulty electrical signals in the areas of the brain that control movement. When medications have become less effective and/or when side effects of the medications interfere with daily activities, doctors may prescribe DBS. Unlike earlier surgical procedures that destroyed brain tissue in order to eliminate the faulty signals, DBS does not damage brain tissue. Instead, it blocks the faulty signals that cause tremors and other movement related symptoms.

Transcranial Magnetic Stimulation (TMS) is a recent non-invasive clinical treatment conserved for those suffering from severe treatment-resistant depression. TMS can focus a small region of the brain which is an advantage and can stimulate to a depth of about 2 inches or 5cm from the skull surface. A repetitive version of TMS is rTMS where pulsating stimulations are applied repetitively for a certain periods, the typical time period of this stimulation is 30 to 60 minutes. In 2008, for major depression of patients who do not respond to at least one antidepressant medication in the current episode, rTMS was approved for use by the FDA as a treatment. It is also used in various countries as a treatment for depression in patients who have not responded to medications and/or who might otherwise be considered for ECT [2]. Recently a version of TMS called Deep TMS have been claimed to be effective in Parkinson's disease as well [3, 4]. Deep TMS can stimulate brain regions about 6cm deep, about a cm more than normal TMS. Again, TMS can also be used in diagnosis, disease classification and in the monitoring of symptomatic and disease-modifying therapies [5]. It can be used to stimulate deep peripheral nerves that lie far beneath the muscles and specially, the brain.

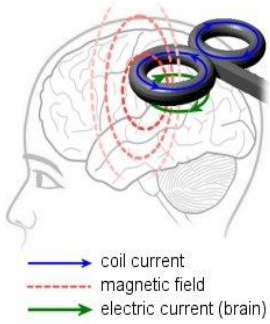


Fig. 1 Trancranial magnetic coil performing deep brain stimulation

Since TMS is recent, its full potentials have not yet been explored and there is much scope for its use in different diagnostic and therapeutic uses in neurological problems.

One of the applications of these developments is the measurement of nerve conduction using artificial electrical stimulation of peripheral nerves. Therefore, designing a transcranial magnetic stimulator coil would be a useful supplement to the development of the full system of transcranial magnetic stimulator for deep brain stimulation,

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

which can be used for conduction velocity measurements of the nerves in the head and neck regions – giving a means to detect Multiple Sclerosis – a special neurological problem.

II. Objectives Of This Work

The aim of this research is to design a transcranial magnetic coil for deep brain stimulation. The objectives of the work are briefly given below,

i. To study neurological disorders where TMS may have potential applications in diagnosis and therapy.
ii. To study non-invasive deep brain stimulation technique.
iii. To design a suitable coil for transcranial magnetic stimulator (TMS) which could have use for conduction velocity measurement of the nerves in the head and neck regions, giving a means to detect Multiple Sclerosis – a special neurological problem, Parkinson's disease through deep brain stimulation.

III. Literature Review

There are numerous cranial stimulation technologies currently being studied within the scientific community.

A. *Direct Deep Brain Stimulation(DDBS)*

DDBS is a FDA-approved neurosurgical procedure. In this process electrodes are penetrated in the brain to send mild electrical signals to the respective area that controls movement. These electrodes are connected to a stimulator which is an implantable pulse generator that is embedded under the collar bone, similar to pace maker [6] for lucid understanding.

B. *Electroconvulsive Therapy(ECT)*

This therapy is previously known as electroshock therapy. ECT is a non-invasive brain stimulation technique, used in controlled clinical and research purposes. It is reserved to treat only severe psychiatric disorder which have proven resistant to other forms of treatment and is usually considered as a 'last resort' treatment option. Unlike Transcranial Direct Current Stimulation. (tDCS) which implements a maximum current of 2 milliamps, ECT uses extremely high current level, around 800 milliamps, which induces an intentional, controlled and short-term seizure. Due to the severity of the effects produced by Electroconvulsive Therapy, patient is almost always sedated during stimulation, which lasts for an average of 1 to 6 seconds. The required use of sedation greatly increases the potential and severity of any negative side effects [6].

C. *Transcranial Magnetic Stimulation (TMS)*

TMS is a non-invasive clinical stimulation technique, usually reserved for those suffering from severe treatment-resistant depression. Recently TMS is the most commonly practiced brain stimulation technique in the medical field [6, 7]. Repetitive Transcranial Magnetic Stimulation (rTMS) is a special type of electromagnetic TMS therapy. When reviewing the differences in TMS with rTMS, the fundamental difference lies in the name itself "repetitive". As described previously, during TMS treatment the device's magnetic coils switch its polarity creating a steady electromagnetic pulse (figure a).

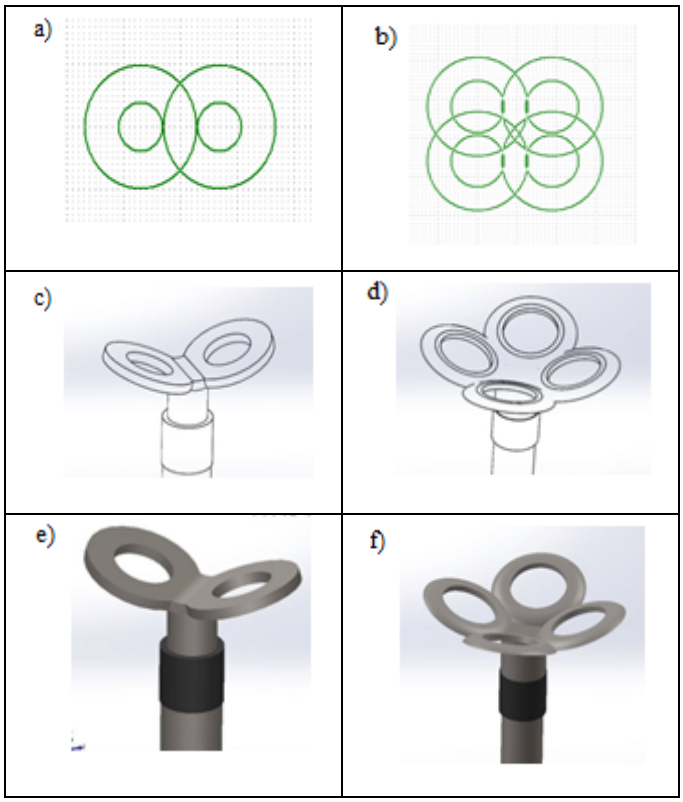


Fig. 2 Perspective view of TMS Coil

On the other hand rTMS, the speed at which the magnetic coils change polarity is rapidly increased, generally switching between positive and negative polarities in just fractions of seconds and it is usually in microseconds order. This produces "repetitive" electromagnetic pulses, which in turn creates stronger electromagnetic induction. In order to the increase in strength, rTMS has the potential to long lasting changes in the brain [6].

To generate an electric current across the scalp and skull, TMS uses electromagnetic induction without any physical contact as depicted in Fig. 1. Insulated coil of wire is positioned next to the skull and when activated, produces a magnetic field oriented orthogonally to the plane of the magnetic coil. The magnetic field passes unimpeded through the skin and skull, inducing an oppositely directed current in the brain that activates nearby nerve cells in much the same way as currents applied directly to the cortical surface. It is difficult to model the path of this current because the shape of brain is irregular; electricity and magnetism are not conducted uniformly throughout its tissues. The strength of magnetic field and pulse generally reaches into the brain based on TMS requirements [14]. Deep TMS can reach up to 6 cm into the brain to stimulate deeper layers of the motor cortex, such as that which controls leg motion [8].

IV. Coil Design

For TMS coil designing, Faraday's Law is considered as a principal theory. From the Biot–Savart law and using Maxwell-Faraday equation we can calculate the strength and direction of magnetic flux. The time varying magnetic field **B** and electric field **E** causes a resultant electromagnetic flux. This field causes a change in the level of trans-membrane current of the neuron that leads to the depolarization or hyperpolarization of the neuron and the firing of an action potential [9].

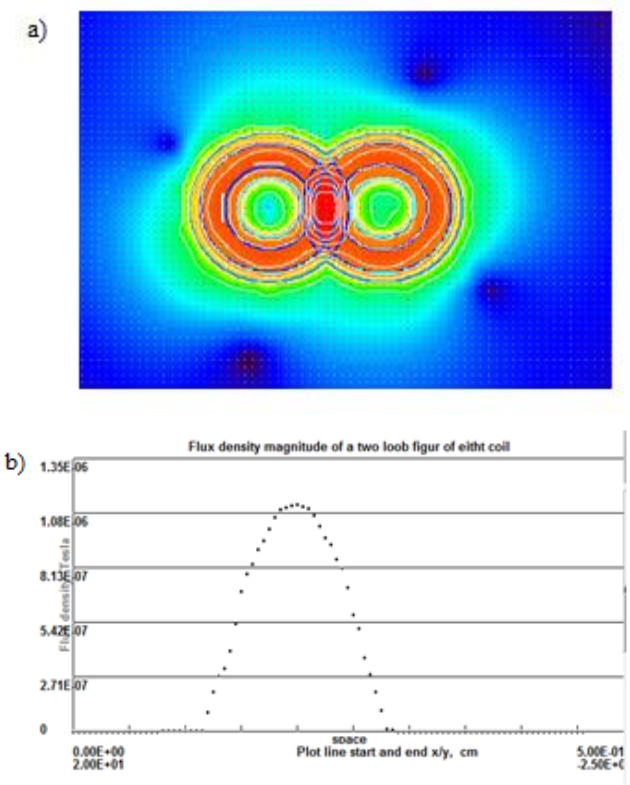


Fig. 3 Figure of eight coil and its flux density magnitude

The most important considerations include the type of material used to construct the core of the coil then the orientation geometry of the coil configuration and the biophysical characteristics of the pulse produced by the coil. For our design we are considering solenoid based insulated coil. This coil posse a ferromagnetic active material, solid core design.

Based on different types of magnetic field pattern, different types of coils can be considered. Some examples:

- Round coil: The simple type of TMS coil
- Figure-eight coil: Commonly known as butterfly coil. Results in a more focal pattern of activation.
- Double-cone coil: Conforms to shape of head, useful for deeper stimulation.
- Four-leaf coil: For focal stimulation of peripheral nerves system[10]
- H-coil: Overhead mount type TMS coil.

Coil is needed to create an electromagnetic pulse to induce current in brain tissue. For this purpose we need to prepare a coil with multi turns of heavy duty wire which should be insulated for ensuring necessary safety. Getting knowledge from a major qualifying project report submitted at Worcester Polytechnic Institute on TMS coil design [11], the O-shaped coil with 35mm radius has close proximity with double magnetic field magnitude as compared to the O-shaped coil with 70mm radius.

The ‘8’ shaped coil with opposite current directions has magnetic field magnitude 15 times stronger than the ‘8' shaped coil with same current directions. As a result, the ‘8’ shaped coil with opposite current directions has the strongest magnetic field magnitude among 4 different coil designs.

Keeping in mind those pros and cons this paper is proposing a special type of hybrid coil that can fulfil our basic requirements with necessary feature. The coil is designed with ferromagnetic core, figure of eight coil considering oblique loop position which can cope up with the irregular shape of human scalp.

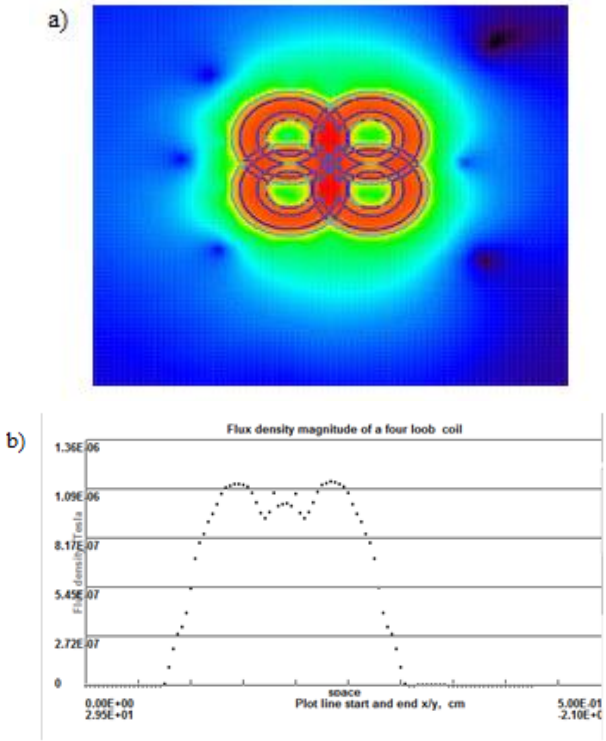


Fig. 4 Four leaf coil and its flux density magnitude

The distance between two loops can be optimized by minimum distance that points the electromagnetic focus area and increase the stimulation depth as shown in Fig. 3. For more specific purpose to enhance the strength of magnetic flux a four leaf coil provides the best solution as presented in Fig 4.

Considering the minimum flexibility, ferromagnetic toroid coil is chosen having inner and outer radius about 4 cm and 10 cm respectively, where number of turns used is100.

Using oblique toroid alignment shown in figure 2, we can get improved focus region with maximum flus density. An upgraded design pattern, special oblique four leaf coil provides a wide equal flux density perpendicular to the toroid. This design pattern is also improving the total flux magnitude. The magnetic flux density is distributed by VIZIMEG electric and magnetic fields simulation tool. Moreover, toroid coil with helping aid is visualized by SolidWorks design.

V. Conclusions

As mentioned before, neurological disorders are prevalent widely in Bangladesh. Cases of depression have increased over time. There are cases of Parkinson’s and Alzheimer diseases as well. These affect the social and economic life of the patient, besides the trauma that the whole family faces.

Therefore, particularly this designed deep TMS coil can help many patients and bring back their normal life activity; thus contributing to an improvement of the overall socio economic condition of the country. As per our domestic

perspective, locally made equipment would cost a fraction of the price of similar imported equipment. Some of the medical equipment made by the group members of the team at the Department of Biomedical Physics & Technology is working for over 20 years. Any repair is done locally as expertise and spares – both are available locally. Therefore, the socio economic importance of this research work would be huge in the long run.

VI. ACKNOWLEDGMENT

The authors would like to acknowledge Department of Biomedical Physics and Technology, University of Dhaka for technical support during the experimental procedures.

REFERENCES

[1] European Commission DG Health and Food SafetyPublic health Major and chronic diseases Diseases Brain and neurological diseases: http://ec.europa.eu/health/major_chronic_diseases/diseases/brain_neurological/index_en.htm#fragment0

[2] Repetitive Transcranial Magnetic Stimulation, Transforming the understanding and treatment of mental illnesses, National Institute of Mental Health: http://www.nimh.nih.gov/health/topics/brain-stimulation-therapies/brain-stimulation-therapies.shtml#part_152879

[3] Hovington CL, McGirr A, Lepage M, Berlim MT. Repetitive transcranial magnetic stimulation (rTMS) for treating major depression and schizophrenia: a systematic review of recent meta-analyses. Ann Med. 2013; 45(4): 308-21; PMID 23687987 :https://en.wikipedia.org/wiki/Deep_transcranial_magnetic_stimulation

[4] NationalPerkinsonfoundation :http://www.parkinson.org/understanding-parkinsons/treatment/surgery-treatment-options/Deep-Brain-Stimulation

[5] Mult Scler Relat Disord. 2015 Sep;4(5):430-6. doi: 10.1016/j.msard.2015.06.014

[6] The Brain Stimulator affordable tDCS Device menual :https://thebrainstimulator.net/brain-stimulation-comparison/

[7] NiCE. January 2014 Transcranial magnetic stimulation for treating and preventing migraine

[8] Zangen, A.; Roth, Y.; Voller, B.; Hallett, M. (2005). "Transcranial magnetic stimulation of deep brain regions: Evidence for efficacy of the H-Coil". Clinical Neurophysiology 116 (4): 775–779. doi:10.1016/j.clinph.2004.11.008. PMID 15792886.
(2) Huang, YZ; Sommer, M; Thickbroom, G; Hamada, M; Pascual-Leonne, A; Paulus, W; Classen, J; Peterchev, AV; Zangen, A; Ugawa, Y (2009). "Consensus: New methodologies for brain stimulation". Brain Stimulation 2 (1): 2–13. doi:10.1016/j.brs.2008.09.007. PMID 20633398

[9] V. Walsh and A. Pascual-Leone, "Transcranial Magnetic Stimulation: A Neurochronometrics of Mind." Cambridge, Massachusetts: MIT Press, 2003

[10] Roth, BJ; MacCabee, PJ; Eberle, LP; Amassian, VE; Hallett, M; Cadwell, J; Anselmi, GD; Tatarian, GT (1994). "In vitro evaluation of a 4-leaf coil design for magnetic stimulation of peripheral nerve". Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section 93 (1): 68–74. Doi: 10.1016/0168-5597(94)90093-0. PMID 7511524

[11] A Major Qualifying Project Report:Submitted to the Faculty of the WORCESTER POLYTECHNIC INSTITUTE in partial fulfillment of the requirements for the Degree of Bachelor of Science By: Yi Yu Tsai Date: May 10, 2011 :https://www.wpi.edu/Pubs/E-project/Available/E-project-051011-115628/unrestricted/MQP.pdf

Crystallographic Orientation-Dependent Performance of 445nm InGaN Blue Laser

Sourav Roy, [1,*] Md. Rafiqul Islam, [1] Md. Mahbub Hasan [1] and Sheikh Abid Hossain[1]

[1] Department of Electrical and Electronic Engineering, Khulna University of Engineering and Technology
Khulna-9203, Bangladesh

[*]souravroy2k9@gmail.com

Abstract— **A numerical approach is presented to study the crystallographic orientation-dependent performance of 445nm InGaN/GaN quantum well (QW) Blue Laser by solving a six-band k.p Hamiltonian for wurtzite crystal using finite difference method. Euler's rotation technique is used to modify the wave vector and Hamiltonian matrix in conventional (0001) crystal orientation. It is found that there is a substantial correlation of the energy band dispersion profile, peak gain and peak emission wavelength with change of crystal orientation. The regular optical gains are inspected as 3845, 4460, 4178, 4880 and 4750 cm^{-1} corresponding to peak emission wavelength of 440, 445, 450, 454 and 458nm at (0001), $(10\bar{1}0)$, $(11\bar{2}0)$, $(11\bar{2}2)$ and $(10\bar{1}\bar{1})$ orientations which shows that highest optical gain and lowest effective mass are attained for $(11\bar{2}2)$) orientation.**

Index Terms— **Crystal orientation, Blue laser, InGaN, Valence band, Optical gain**

I. Introduction

The wurtzite (WZ) III-N semiconductors (GaN, InN and AlN) have received considerable attention in the last decades due to its unique characteristics such as high electron mobility and superior carrier saturation velocity. Consequently, its application scope broadens up to the operation as the active region of high-power optoelectronic devices in the blue and ultraviolet regions. Blue-green lasers based on strained $In_xGa_{1-x}N$/GaN quantum-well (QW) structures have been realized recently with a lifetime of many thousand hours [1]. For true blue lasing emission, a wavelength of 445~450 nm is required. Recent research demonstrates that 445 nm blue laser imaging provides excellent endoscopic images of upper gastrointestinal lesions. Commercial applications of Blue lasers are in high-definition Blu-ray players, DLP and 3LCD projectors, telecommunications, environmental monitoring and magnetic levitation devices etc. which led us to concentrate on 445nm blue laser [2]-[4].

In comparison with GaAs-based QW's, WZ GaN-based QW's of c-plane (0001) crystal orientation need high carrier densities to produce optical gain [5]. This is due to the heavy effective masses in the valence and conduction bands and strong interaction between heavy-hole (HH) and light-hole (LH) bands at the Γ point. Also, the pseudomorphic biaxial compressive strain and the quantum size effect in (0001) WZ QW's shifts both the HH and LH bands by almost the same amount, and the in-plane effective masses remain almost the same as those in the unstrained case. Moreover, the QW's grown on c-plane faces polarization induced strong quantum confined stark effect (QCSE) that reduces the overlap between the electron and hole wave functions and results in the reduction of the internal quantum efficiency so, less optical gain [6]-[7].

Recently, it has become apparent that energy band dispersion profile of strained WZ III-N materials is strongly co-related with crystal orientations. High performance optoelectronic devices can be realized in nonpolar and semipolar orientations which may solve the QCSE problem from the conventional c-plane QW growth. In favorable band structure for higher optical gain, the splitting between HH and LH is increased and effective mass of holes is reduced [8]. The growth of non-(0001) crystal orientated QW is possible by modern crystal growth techniques such as molecular beam epitaxy and metal-organic chemical vapor deposition. The growth of $(1\bar{1}00)$ GaN films on $(1\bar{1}00)$ SiC substrates has been reported [9]. The valence sub-band structures of WZ GaN QW's were studied at $(10\bar{1}2)$, $(20\bar{2}1)$ and $(10\bar{1}0)$ orientations [10]. Experimental observation for the absence of QCSE in nonpolar AlGaN/GaN $(10\bar{1}0)$ and $(11\bar{2}0)$ QWs has been reported. A reduced QCSE has also been observed in $(11\bar{2}2)$ and $(10\bar{1}\bar{3})$ semipolar InGaN/GaN QWs [11]-[12]. So far our knowledge, orientation-dependent optical performance of true-blue laser was reported for cubic InGaN [13]. In this study, the orientation-dependent optical performance of 445nm WZ InGaN/GaN QW blue laser is presented in details. Here the optoelectronic performance of (0001)/c-plane, nonpolar $(10\bar{1}0)$/m-plane and $(11\bar{2}0)$/a-plane and semipolar $(11\bar{2}2)$ and $(10\bar{1}\bar{1})$ oriented $In_{0.17}Ga_{0.83}N$ laser is investigated by using the 6*6 k.p matrix. To investigate the optoelectronic properties in non-(0001) orientation, unitary transformations are used to modify the wave vectors and Hamiltonian matrix in conventional crystal orientation. The conduction and valence band wave functions and Eigen energies of the proposed WZ InGaN/GaN QW structure are evaluated using the finite difference method. The momentum matrix element and optical emission profile are studied using the calculated energy band dispersion profiles and wave functions in different crystal orientations.

The paper is outlined as follows: the theoretical details to calculate optical properties are illustrated in section 2. Results and discussions on the energy band structure and optical gain performances are delineated in section 3. Conclusively, a summary is depicted in section 4.

II. Theory

A. Band gap energy calculation

The peak emission wavelength of laser mainly depends on the band gap energy of the material used in active region. So, the selection of composition of the semiconductor material used in the active region and barrier layer is very important. In this structure GaInN is used in the active region and GaN is used in the barrier of QW [14].

$$E_g(In_xGa_{1-x}N)\ (eV) = 0.7x+3.44(1-x)-1.43x(1-x) \quad (1)$$

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

B. Device structure

Fig. 1 shows the proposed laser structure where 8nm $In_{0.17}Ga_{0.83}N$ is used in the active layer and 10nm GaN is used in the guiding/barrier layer. $Al_{0.3}Ga_{0.7}N$ and $Al_{0.1}Ga_{0.9}N$ is used in the capping and cladding layers. The proposed laser structure is grown on GaN substrate. Above the upper AlGaN cladding layer, an oxide aperture is used for current and index confinement. Highly doped contact layers are grown at the top and the bottom of the surface for the formation of ohmic contact.

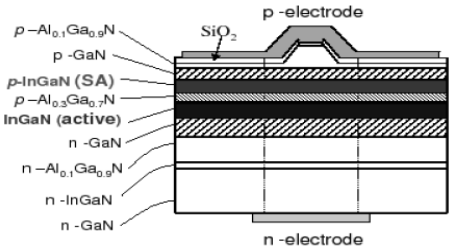


Fig. 1 Structure of InGaN/GaN Blue LASER

C. The k.p Method of Band-Structure near Γ-point

The electronic band structure is solved in the envelope approximation using a six-band **k.p** Hamiltonian. The 6×6 Hamiltonian matrix H for (0001) WZ crystals can be asserted by [15]: $H\Psi_h(z) = E\Psi_h$ (2)

The components of H are:

$$H = \begin{bmatrix} F & -K* & -H* & 0 & 0 & 0 \\ -K & G & H & 0 & 0 & \Delta \\ -H & -H* & \lambda & 0 & \Delta & 0 \\ 0 & 0 & 0 & F & -K & H \\ 0 & 0 & \Delta & -K* & G & -H* \\ 0 & \Delta & 0 & H* & -H & \lambda \end{bmatrix} \begin{matrix} |u_1> \\ |u_2> \\ |u_3> \\ |u_4> \\ |u_5> \\ |u_6> \end{matrix}$$

$$\Psi_h(z) = \begin{bmatrix} \Phi_{1/2,1/2}(z) \\ \Phi_{1/2,-1/2}(z) \\ \Phi_{3/2,1/2}(z) \\ \Phi_{3/2,-1/2}(z) \\ \Phi_{3/2,3/2}(z) \\ \Phi_{3/2,-3/2}(z) \end{bmatrix} = \begin{bmatrix} g^{(1)}(z) \\ g^{(2)}(z) \\ g^{(3)}(z) \\ g^{(4)}(z) \\ g^{(5)}(z) \\ g^{(6)}(z) \end{bmatrix}$$

Where, ϕ_{j,m_j} or $g^{(j)}$ is the valence band wave function component and E is the energy of three valence bands.

The matrix elements are:

$F = \Delta_1 + \Delta_2 + \lambda + \theta$

$G = \Delta_1 - \Delta_2 + \lambda + \theta$

$\Delta = \sqrt{2}\,\Delta_3$

$\lambda = \frac{\hbar^2}{2m_0}[A_1k_z^2 + A_2(k_x^2+k_y^2)] + D_1\mathcal{E}_{zz} + D_2(\mathcal{E}_{xx} + \mathcal{E}_{yy})$

$\theta = \frac{\hbar^2}{2m_0}[A_3k_z^2 + A_4(k_x^2+k_y^2)] + D_3\mathcal{E}_{zz} + D_4(\mathcal{E}_{xx} + \mathcal{E}_{yy})$

$K = \frac{\hbar^2}{2m_0}A_5(k_x + ik_y)^2 + D_5\mathcal{E}_+$

$H = \frac{\hbar^2}{2m_0}A_6(k_x + ik_y)k_z + D_6\mathcal{E}_{z+}$

$\mathcal{E}_+ = \mathcal{E}_{xx} - \mathcal{E}_{yy} + 2i\,\mathcal{E}_{xy}$

$\mathcal{E}_{z+} = \mathcal{E}_{zx} + i\,\mathcal{E}_{yz}$ (3)

Here the Ai's are the valence-band effective-mass parameters, which are similar to the Luttinger parameters in a zinc blende crystal; the Di's are the deformation potentials for wurtzite crystals; k_i is the wave vector; $\mathcal{E}_{ij}$ is the strain tensors; Δ_1 is the crystal-field-split energy; Δ_2 and Δ_3 account for the spin-orbit interaction. For $\Delta_1 > \Delta_2 > 0$, the three bands from top to bottom are labeled as HH, LH, and CH, respectively. The basis functions are shown in Fig. 2. The material properties of GaN and InN are listed in Table 1. The parameters of $In_{0.17}Ga_{0.83}N$ have been calculated by using Vagard's law.

$$\left.\begin{aligned} |u_1\rangle &= -\tfrac{1}{\sqrt{2}}\langle (X+iY)\uparrow\rangle \\ |u_2\rangle &= \tfrac{1}{\sqrt{2}}\langle (X-iY)\uparrow\rangle \\ |u_3\rangle &= \langle Z\uparrow\rangle \\ |u_4\rangle &= \tfrac{1}{\sqrt{2}}\langle (X-iY)\downarrow\rangle \\ |u_5\rangle &= -\tfrac{1}{\sqrt{2}}\langle (X+iY)\downarrow\rangle \\ |u_6\rangle &= \langle Z\downarrow\rangle \end{aligned}\right\}$$

Here, |X>, |Y> and |Z> are the p_x, p_y and p_z wavefunctions at the Γ point with their dipoles along the $(10\bar{1}0)$, $(11\bar{2}0)$ and (0001) directions.

Fig. 2 Basis Function for the Hamiltonian

For a biaxially strained InGaN/GaN QW of the (0001) orientation, the strain tensor in the well region contains:

$\mathcal{E}_{xx} = \mathcal{E}_{yy} = (a_{well} - a_{barrier})/a_{barrier}$

$\mathcal{E}_{zz} = (-2\,C_{13}/C_{33})\,\mathcal{E}_{xx}$

$\mathcal{E}_{xy} = \mathcal{E}_{yz} = \mathcal{E}_{zx} = 0$

Here, C_{13} and C_{33} are the stiffness constant of GaN well layer.

The energy dispersion profile along (0001) orientation is calculated as a function of k for the 6-band k.p theory by diagonalizing the Hamiltonian using finite difference scheme.

D. Crystal orientation dependent wave vector and strain

The following expression of the rotation matrix is used to obtain the Hamiltonian for arbitrary crystal orientation [16]:

$$U = \begin{pmatrix} \cos\theta\cos\emptyset & -\sin\emptyset & \sin\theta\cos\emptyset \\ \cos\theta\sin\emptyset & \cos\emptyset & \sin\theta\sin\emptyset \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \quad (3)$$

Where,

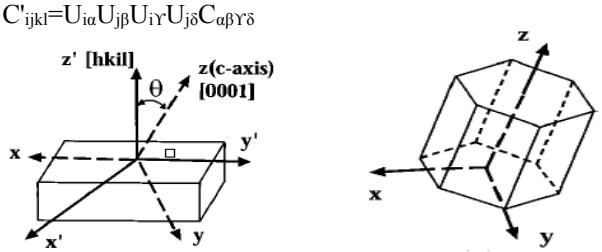
$$\theta = \tan^{-1}\frac{\sqrt{(h^2+k^2+i^2)}}{l} \text{ and } \emptyset = \tan^{-1}\frac{ik}{h}$$

The rotation of the Euler angles θ and Ø transforms physical quantities from (x,y,z) coordinates to (x',y',z') coordinates. The z axis corresponds to the c axis [0001], and the growth axis z' is normal to QW plane [hkil] as shown in Figure 3. The relations between the coordinate systems for vectors and tensors are expressed as [8]:

$k'_i = U_{i\alpha}k_\alpha$

$\mathcal{E}'_{ij} = U_{i\alpha}U_{j\beta}\,\mathcal{E}_{\alpha\beta}$

$C'_{ijkl} = U_{i\alpha}U_{j\beta}U_{i\Upsilon}U_{j\delta}C_{\alpha\beta\Upsilon\delta}$ (4)



(a) (b)

Fig. 3 (a) Configuration of the coordinate systems (x',y',z') in (hkil)-oriented crystals. (b) The coordinate system in a wurtzite primitive cell. The angle θ is defined as the angle between the growth direction z' and the c axis

Here, k_i' is the wavevector, $\mathcal{E}'_{ij}$ is the strain tensor and C'_{ijkl} is the stiffness coefficient tensor in the primed coordinate system. When the crystal is translated, the hexagonal primitive translational vectors are:

$\alpha_i'' = a_i\hat{x}''$

$\beta_i'' = (a_i/2)\hat{x}'' + (\sqrt{3}a_i/2)\,\hat{y}''$

$\Upsilon_i'' = c_i\hat{z}''$

Where *a* and *c* are the lattice constants and *i* labels the well and barrier layer.

$$\left.\begin{aligned} \hat{x}'' &= (1+\mathcal{E}'_{xx})\,\hat{x} \\ \hat{y}'' &= (1+\mathcal{E}'_{yy})\,\hat{y} \\ \hat{z}'' &= (1+\mathcal{E}'_{zz})\,\hat{z} \end{aligned}\right\} \quad \begin{aligned} \mathcal{E}'_{xx} &= (a_{well} - a_{barrier})/a_{barrier} \\ \mathcal{E}'_{yy} &= (c_{well} - c_{barrier})/c_{barrier} \\ \mathcal{E}'_{zz} &= -(C_{12}/C_{11})\,\mathcal{E}'_{yy} - (C_{13}/C_{11})\,\mathcal{E}'_{xx} \end{aligned}$$

Hence we obtain a 6×6 Hamiltonian in the (x',y',z') coordinates by substituting the transformation relation for the wave vector k in (4) and the strain coefficients for a general crystal orientation into the Hamiltonian (2) for the (0001) orientation.

TABLE I
MATERIAL PARAMETERS FOR BLUE LASER [17]

Parameters	GaN	InN
Lattice Constant(A^0)		
a	3.189	3.544
c	5.185	5.718
Energy parameters		
E_g at 300K(eV)	3.44	0.7
Δ_1(meV)	21	17
Δ_{so}(meV)	11	3
$\Delta_2 = \Delta_3 = \Delta_{so}/3$(meV)	3.67	1
Valence band effective mass parameters		
A_1	-7.24	-15.803
A_2	-0.51	-0.497
A_3	6.73	15.251
A_4	-3.36	-7.151
A_5	-3.35	-7.060
A_6	-4.72	-10.078
Deformation Potentials(eV)		
D_1	-4.08	-3.62
D_2	0.7	-4.60
D_3	2.1	2.68
D_4	1.4	1.74
D_5	-0.7	-2.07
D_6	-0.7	-2.07
Elastic Stiffness Constant (10^{11} dyne/cm^2)		
C_{11}	29.6	19
C_{12}	13	10.4
C_{13}	15.8	12.1
C_{33}	26.7	18.2

E. Optical Gain Profile

The optical gain as a function of energy for quantum well can be expressed by [16]:

$$g(E) = \frac{2q^2\hbar}{n\varepsilon_0 c m_0{}^2 LE} \times \sum_{n,m}\int_0^\infty \frac{\frac{k_t M_{nm}(k_t)\Gamma}{2\pi}}{(E_{cn}(k_t)-E_{kpm}(k_t)-E)^2+(\Gamma/2\pi)^2}(f_c{}^n - f_v{}^m)\,d\,k_t \quad (5)$$

Here, q is denoted as free electron charge, $\hbar$ is the reduced Planck constant, n is the index of refraction, ε_0 is the free space dielectric constant, c is the speed of light, E is the photon energy, E_{cn} is the nth conduction sub-bands, E_{kpm} is the mth valence sub-bands, and M_{nm} is the momentum matrix element in strained quantum well architecture. $\Gamma = \hbar/\tau$ where, τ is the photon relaxation time. F_c and F_v are the Fermi levels of conduction and valence bands. The momentum matrix element for band to band and band to sub-bands transition can be written as:

$$M_{nm}(k_t) = \frac{3}{4}\left[\langle\Psi_{n(k_t=0)}|g_m{}^{(1)}\rangle^2 + \frac{1}{3}\langle\Psi_{n(k_t=0)}|g_m{}^{(2)}\rangle^2\right] M_b{}^2$$

$$\langle\Psi_{n(k_t=0)}|g_m{}^{(1)}\rangle = \int_{-\infty}^{\infty}\Psi_{n(k_t=0)}(z)g_m{}^{(1)}(z)dz$$

M_b is the bulk average crystal momentum matrix that can be noted by (E_g=band gap and Δ=split-off band separation):

$$M_b = \left(\frac{m_0}{m_e} - 1\right)\frac{m_0 E_g(E_g+\Delta)}{6(E_g+\frac{2}{3}\Delta)} \quad (6)$$

III. RESULT AND DISCUSSIONS

Based on the mathematical model presented in section 2, the electronic and optical properties of 445nm compressively strained WZ InGaN/GaN blue laser in different crystal orientations are studied. The simulation is carried out in MATLAB environment at room temperature. The valence band dispersion profiles are estimated in (0001), nonpolar $(10\bar{1}0)$ and $(11\bar{2}0)$ and semipolar $(11\bar{2}2)$ and $(10\bar{1}\bar{1})$ orientations and shown in Fig. 4. It is demonstrated that, for nonzero value of k (wave vector in the plain direction); the increasing coupling from distant bands gives rise to strong non-parabolicities at the top valence band structure. For c-plane/ (0001) orientation, strong intermixing between HH and LH and piezoelectric field induced strong QCSE effectively ceases quick transition of charge carriers causing reduction in quantum efficiency. It is seen that, the $(10\bar{1}0)$ orientation demonstrates its superiority over the $(11\bar{2}0)$ orientation in optical characteristics. On the other hand, semipolar $(11\bar{2}2)$ and $(10\bar{1}\bar{1})$, forming an angle of 47 and 62 degree respectively, shows better dispersion profile in terms of reduced interband coupling and QCSE. So, they have a reduced density-of-states near the valence band edge leading to reduced hole effective mass. Table 2 shows the energy gaps between conduction(C) to HH, LH and CH (crystal-field split-off hole) at the band edges as a function of the crystal orientation. It is observed from the table that, crystal orientation changes the energy splitting. At (0001) and $(11\bar{2}0)$ orientations, the energy difference between the HH and LH bands is very small. For semipolar cases, the increased band gap results in a significant reduction of the effective mass of the top valence band along the transverse direction. As the QW is compressively strained, the maximum transition occurs between C to HH. The momentum matrix element for the transitions between C-HH, C-LH and C-CH for different orientations is calculated from band-edge Eigen values and wavefunctions and shown as function of θ in Fig. 5. It is found that with an increasing angle θ, the optical matrix element for the C-LH transition rapidly decreases, and is nearly zero above θ=30° and with θ, the optical matrix element for the C-HH transition rapidly increases and approach a constant value from θ=30°. The constant value is about two times as large as that of the (0001) orientation. The C-CH transition isn't so

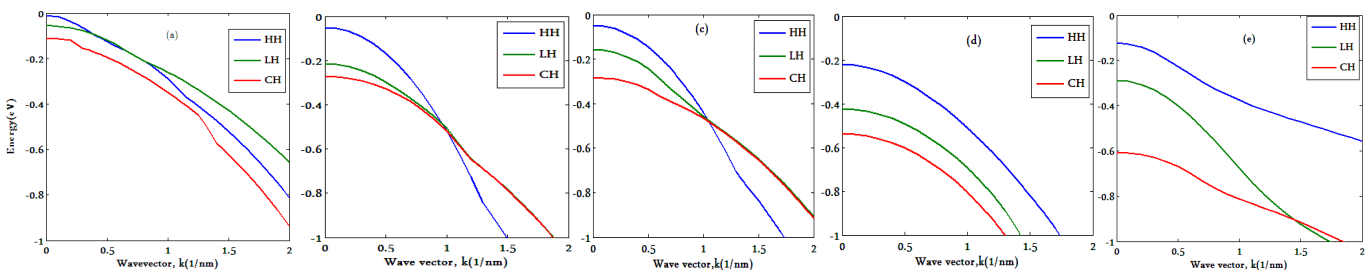


Fig. 4 Valence band dispersion profile of (a) (0001), (b) $(10\bar{1}0)$, (c)$(11\bar{2}0)$, (d) $(11\bar{2}2)$ and (e) $(10\bar{1}\bar{1})$ of InGaN/GaN QW blue laser.

dominant. Using "Equation (6)" the gain of the laser is determined at 300k for different crystal orientations. "Fig. 6" shows the orientation dependent optical gain of the laser. It is seen that the peak gain and peak emission wavelength have crystal orientation dependence. For the injection carrier density of $3x10^{18}$ cm^{-3}, the peak gains are evaluated to be 3845, 4460, 4178, 4880 and 4750 cm^{-1} with wavelength of 440, 445, 450, 454 and 458nm at (0001), $(10\bar{1}0)$, $(11\bar{2}0)$, $(11\bar{2}2)$ and $(10\bar{1}\bar{1})$ orientations. Maximum gain is evaluated in the $(11\bar{2}2)$ semipolar direction due to reduced density of states and large gap between heavy hole and light hole band, so reduced QCSE and reduced effective mass.

TABLE III
ORIENTATION-DEPENDENT ENERGY SEPARATION

Crystal orientation	C-HH(eV)	C-LH(eV)	HH-LH(eV)
(0001)	2.772	2.783	0.04
$(10\bar{1}0)$	2.803	3.023	0.16
$(11\bar{2}0)$	2.725	2.732	0.09
$(11\bar{2}2)$	2.864	3.114	0.21
$(10\bar{1}\bar{1})$	2.812	2.992	0.19

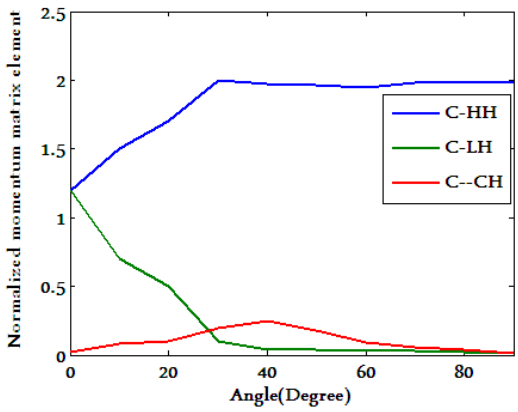


Fig. 5 Orientation dependent momentum matrix element

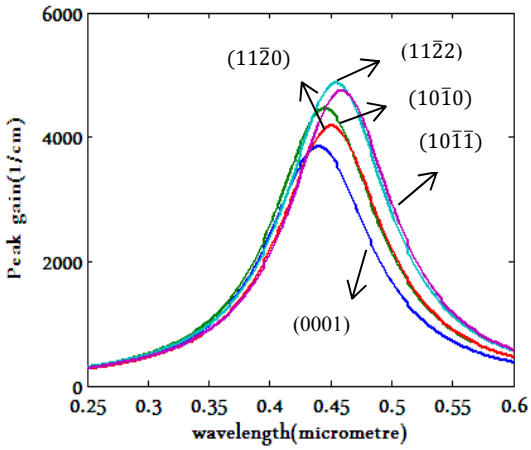


Fig. 6 Optical emission profiles of $In_{0.17}Ga_{0.83}N$/GaN QW blue laser in different orientations.

IV. Conclusions

The change in crystal orientation on the band gaps, optical momentum matrix, and optical gain characteristics of InGaN/GaN QW 445nm blue laser is numerically investigated using k.p model. The peak gain is found to have crystal orientation dependence and maximum gain is evaluated at $(11\bar{2}2)$ orientation due to reduce band mixing effect. This is because of higher energy separation between the valence sub-bands is observed in this orientation. It is also seen that the peak emission wavelength of the proposed laser shows tunable characteristics and it varies from 440nm to 458nm when the orientation changes from (0001) to $(10\bar{1}\bar{1})$. This numerical result demonstrates that semipolar $(11\bar{2}2)$ orientation is of special interest and can be incorporated into the active layer of blue laser in order to have best performance.

References

[1] S. Nakamura and G. Fasol, *The Blue Laser Diode*, Springer, Berlin (1997).

[2] H. Osawa et al., "Blue Laser Imaging provides excellent endoscopic images of upper gastrointestinal lesions", *Video journal and Encyclopedia of GI Endoscopy*, vol. 1, pp. 607-610(2014).

[3] S. Srinivasan, M. Stevens, F. A. Ponce, H. Omiya, and T. Mukai, "Carrier dynamics and electrostatic potential variation in InGaN quantum wells grown on GaN pyramidal planes", *Appl. Phys. Lett. 89*, 231908(2006).

[4] A.E.Romanov et al., "Strain-induced polarization in wurtzite III-nitride semipolar layers", *Journal of Appl. Phys. 100*, 023522(2006).

[5] Yu Hongbo, L.K.Lee, Taeil Jung and P.C.Ku, "Photoluminescence study of semipolar InGaN/GaN multiple quantum wells grown by selective area epitaxy", *Appl. Phys. Lett. 90*, 141906 (2007).

[6] You-Da Lin et al., "Blue-Green InGaN/GaN Laser Diodes on Miscut m-Plane GaN Substrate", *Appl. Phys. Express 2*, 082102(2009).

[7] Y.C.Yeo, T.C.Chong, and M.F.Li, *IEEE journal of quantum electronics*, vol. 34, no. 7, (1998).

[8] S.H.Park and S.L.Chuang, "Crystal-orientation effect on piezoelectric field and electronic properties of strained wurtzite semiconductors", *PHYSICAL REVIEW B*, vol.59, No. 7(1999).

[9] Horino et al., "Growth of (1100) Oriented GaN on (1100) 6H-SiC by Metalorganic Vapor Phase Epitaxy", in Proc. of International Symposium on Blue Laser and Light Emitting Diodes, Chiba Univ., Japan, Mar. 5-7, 1996, pp. 530-533.

[10] S.C.Lee, X.Y.Sun, S.Harsee and S.R.Brueck, "Orientation-dependent nucleation of GaN on a nanoscale faceted Si-surface", *Journal of Crystal Growth*, vol. 279, issues 3-4, pp. 289-292(2005).

[11] T. Ohtoshi, T. Kuroda, A. Niwa, and S. Tsuji, "Dependence of optical gain on crystal orientation in surface-emitting lasers with strained quantum wells", *Appl. Phys. Lett.*, vol. 65, no. 15, pp. 1886-1887, 1994.

[12] A. Niwa, T. Ohtoshi, and T. Kuroda, "Orientation dependence of optical properties in long wavelength strained quantum-well lasers", *IEEE J. Select. Topics Quantum Electron*, vol. 1, pp. 211-217, 1995.

[13] M.S.Sarker, M.M.Islam, M.M.Hasan, and M.R.Islam," Crystal orientation dependent performance of cubic InGaN QW blue-violet laser", in Proc. of ICCIT'18, 2015, pp. 167-171.

[14] Deng et al.,"Theoretical study on InGaN/GaN quantum dots solar cell", *Physica B 406*, pp.73-76, Elsevier (2011).

[15] S.L. Chuang and C.S.Chang, "Effective-mass Hamiltonian for strained wurtzite GaN and analytical solutions", *Appl. Phys. Lett.*, vol. 68, pp. 1657-1659, (1996).

[16] M.M.Hasan and M.R.Islam," Crystal orientation-dependent optical properties of mid-infrared GaInSb/GaInAlSb quantum well laser", in Proc. of ICCCE'10, 2010, pp. 11-13.

[17] I. Vurgaftman, J.R.Meyer and L.R.Ram-Mohan, "Band parameters for III-V compound semiconductors and their alloys", *J. Appl. Phys. 89*, 5815 (2001).

Estimation of Exciton Confinement in III-Nitride Quantum Wires

Md. Rakibul Islam, [1,*] Md. Soyaeb Hasan, [1] and Md. Rafiqul Islam[1]

[1]Dept. of Electrical and Electronic Engineering (EEE), Khulna University of Engineering & Technology (KUET)
Khulna-9203, Bangladesh

[*]irokib069@gmail.com

Abstract— **The optical properties of one dimensional (1D) structures are dominated by the excitonic behaviour which are formed by optical absorption. The excitonic properties in III-nitride based cylindrical quantum wires (QWRs) have been investigated theoretically. The confinement of the electron and the hole is subjected for the peculiarities of the exciton in a QWR where confinement of the particles can be governed through the size and shape of the QWR as well as through the selection of structure and impediment materials to produce various band offsets. In this paper, it has been traced the energy levels and wave functions for confined electrons and holes from numerical solution of Schrodinger-like equations in the effective mass approximation. Finally, the energy of confined exciton have been calculated from the above result. For the different composition in InGaN/GaN, AlGaN/GaN, InAlN/AlN based cylindrical QWRs, the excitonic behaviour have been investigated individually and the combined effects have also been pursued, where considering radius of the wire within the range of 30nm to 70nm.**

Index Terms—**exciton, quantum wire, composition, confinement.**

I. Introduction

III-Nitride semiconductors such as aluminium nitride (AlN) , indium nitride (InN) gallium nitride (GaN) and their alloys are took into account as key materials for practicable optoelectronic devices in green, blue and ultraviolet spectral range. As well, they are fascinating volition for power electronics as wide band gap and break-down electric field. When Shuji Nakamura from Nichia Corporation exhibited the first commercially enduring high-brightness blue light emitting diode (LED) in 1993, from that the attainment began the story of nitrides [1]. AlN has a band gap resembling to profound UV emission, inasmuch as that of InN remain in near infrared, thus possess nitrides to cover a spacious range of wavelengths with a single material system. Excepting III-Nitride semiconductors such as InN, AlN, and GaN have great consideration in LED [2] to the applications for lasers and LEDs, thermoelectricity [3, 4], high power transistors [5] for solid state lighting [2], and solar cells [6,7].

During the 1980s, with the help of improved growth techniques experimentalists qualified to fabricate systems of subordinate dimensionality. In contrast to quantum wells (QWs), where carrier's localization in the point of the compass perpendicular to the layers but movement fairly in the layer plane, in quantum wires (QWRs) carriers localization is in two directions and movement is fairly towards the wire axis. In quantum dots (QDs) they are determinate in all three directions resulting in a totally discrete energy spectrum.

Intensively semiconductor QWRs has been practised worldwide for a wide band spectrum of materials. Such one-dimensional (1D) nanostructures are not only pompous for fundamental research due to their unique structural and physical properties relative to their bulk counterparts, but also captivating potential for future technological applications. Considering QWRs are promising candidates for optoelectronic and microelectronic devices, following the next propagation of classical heterostructures [8]. Than the 2D or 3D density of states (DOS), the 1D DOS becomes more prolonged at the bottom of its band edge, to show improvement in lasing performance a QWR laser is expected [9-11]. In QWRs enhancement of excitonic effects occurs, this spatial confinement leads, among others, due to the subsidence average distance of the constituent electron and hole. Excitonic effects become hence more pronounced in subordinate dimensions, excitons survive and dominate the optical properties near the fundamental band edge at even normal temperature. Due to excitonic interactions reducing the spectral density considerably, again early expectations that QWR devices would profit from the diverging one-dimensional density of states had to be revised [12, 13].

Takahisa Ohno et al. [14] investigated exciton effect for silicon QWRs. D. S. Citrin [15] studied about the exciton effect on GaAs QWRs and their lifetime. E. W. S. Caetano et al. [16] showed ground state exciton confinement in InGaN/GaN QWRs considering the indium composition 20% and wire radius as variable.

In this paper, ground state confined exciton properties have been investigated for different promising nitride ternary alloys (InGaN/GaN, InAlN/GaN, AlGaN/GaN) with different composition of the constituent elements in cylindrical QWRs and showed a comparison which one is better for QWR application.

II. Mathematical Model

To calculate the excitonic states in a QWR we use the single-band effective mass approximation. Through the wire forming direction (here taken to be the z-direction), for the particles in the wires there is no confinement, while in the xy-plane the electron and hole are highly confined. For what reason we separate the z-motion from the xy-motion. The Coulomb interaction term which we also consider as a perturbation, for that we can separate the electron and hole wave functions. Next, we introduce the relative coordinates r

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

$= r_e - r_h$ and we calculate the single electron and hole states in the xy-plane, where m^*_e (m_h) is the effective mass. Hamiltonian of exciton is:

$$\hat{H} = E_g + \hat{H}_e + \hat{H}_h - \frac{e^2}{4\pi\varepsilon\varepsilon_0|r_e - r_h|}$$

Where,

$$\hat{H}_e = -\frac{h}{2}\nabla_e\left(\frac{1}{m_e(r_e)}\right)\nabla_e + V_e(r_e)$$

and $$\hat{H}_h = -\frac{h}{2}\nabla_h\left(\frac{1}{m_h(r_h)}\right)\nabla_h + V_h(r_h)$$

Where the electron and heavy hole effective mass and confinement potential are denoted by m_e (m_h) and V_e (V_h). Considering the effective masses are isotropic which calculated by averaging the parallel and perpendicular direction [0001] growth axis. Confinement potential which depends on the band offset which is taken here Q_e=0.7, and Q_h=1-Q_e=0.3 and assumed the bowing parameter assumed to be *b=02.6* eV. All other parameters are shown in table 1.

TABLE I
BAND PARAMETERS

Parameters	InN	GaN	AlN	AlAs	GaAs
E_g (eV)	0.64	3.4	6.1	3.09	1.59
m_e/m_o	0.12	0.20	0.27	0.15	0.063
m_h/m_o	1.547	1.081	6.33	0.28	0.165
Ɛ	14.60	10.31	11.2	10.3	12.3

For the ground exciton various wave function we have used is:

$$\psi_{eh} = \psi_e(r_e)\psi_h(r_h)e^{-\frac{\sqrt{r_e^2+r_h^2-2\beta r_e r_h}}{\lambda}}e^{\frac{z_e^2+z_h^2}{\lambda}}$$

Where the carrier coordinate is z_e (h) along with the quantum wire axis, where the growth direction is [0001]. $\psi_i(i = e, h)$ is the ground state solution of Schrodinger equation, λ and β are parameters. The exciton ground state energy is obtained though the following expression:

$$E^0_{exc} = \frac{\langle\psi_{eh}|\hat{H}|\psi_{eh}\rangle}{\langle\psi_{eh}|\psi_{eh}\rangle}$$

Where, $E^0_B = E^0_{EXC} - E_g - E_e - E_h$ is the binding energy of exciton.

III. RESULT AND DISCUSSION

We have calculated the ground state exciton binding energy of QWRs as a function of wire radius and indium composition of wurtzite (WZ) InGaN/GaN, InAlN/AlN, and AlGaN/GaN, then have been showed a comparison between all the three ternary alloys. For simplicity we ignored the complicated valence band structure and existing internal electric field due to spontaneous and piezoelectric polarizations which are highly important in WZ nitrides [8, 9].

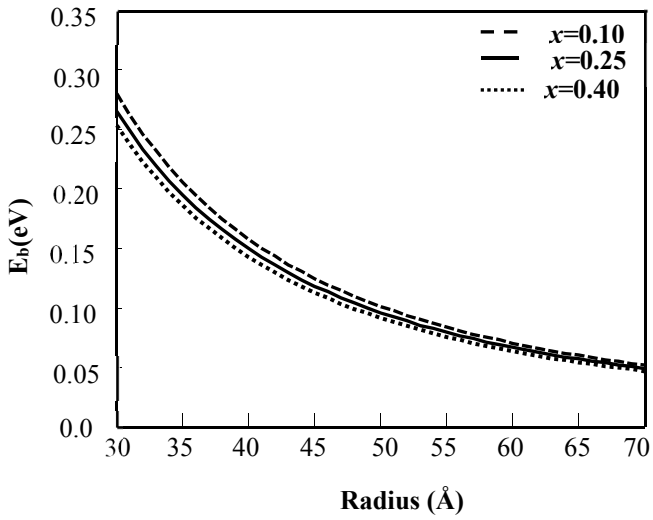


Fig. 1 Ground state exciton binding energy of $In_xGa_{1-x}N$ for various Indium composition.

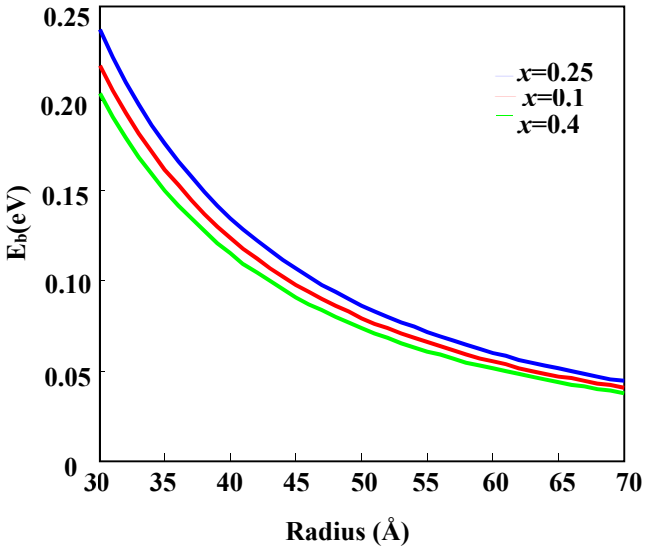


Fig. 2 Ground state exciton binding energy of $Al_xGa_{1-x}N$ for various Aluminium composition.

In Fig. 1 it is investigated the ground-state exciton binding energy of WZ InGaN/GaN QWRs as a function of wire radius and composition. Numerical result shows that the blue shifts on the exciton total energy are more when the radius of the wire is reduced more than 50 Å due to significantly enhancement of exciton confinement, which spanned the visible spectrum. When In composition is varied significantly Binding energy also varied, and Fig. 1 also shows that due to increase Indium composition binding energy also increased.

Fig. 2 shows the same investigation for AlGaN/GaN, that shows when the radius of QWR is reduced more than 50 Å the binding energy of exciton slightly minimized than InGaN/GaN,and binding energy also comparing with InGaN/GaN for each corresponding wire radius and shows revers character when varying the composition, as increasing the composition binding energy is decreased where increasing composition binding energy also increased in InGaN/GaN.

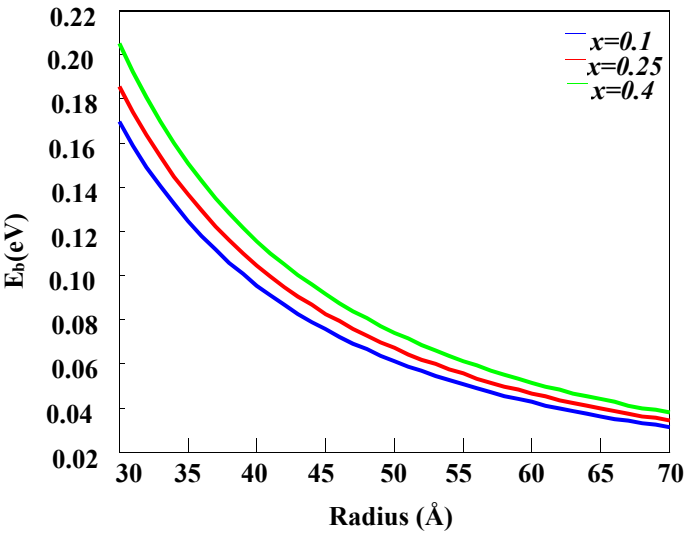


Fig. 3 Ground state exciton binding energy of $In_xAl_{1-x}N$ for various Indium composition.

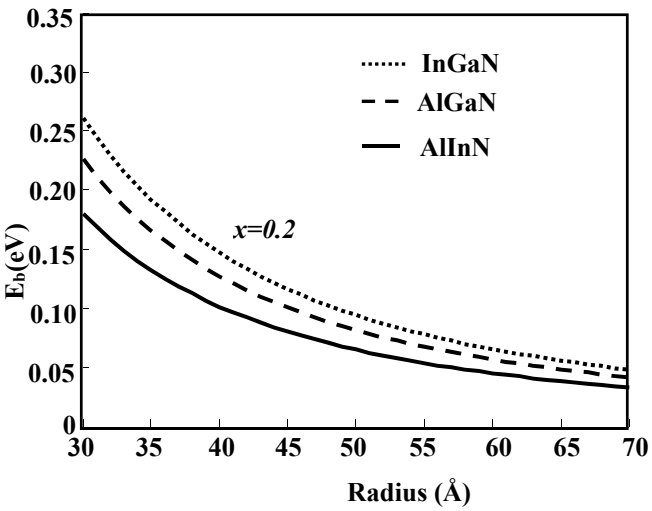


Fig. 4 A simple comparison between all the three optical ternary alloys where considering composition is x=0.2.

In Fig. 3 it shows the same investigation about InAlN/Aln, which shows also nature of binding energy corresponding with the variation of wire radius. Fig. 3 shows that decreasing the wire radius binding energy of exciton is also gradually decreased but up to 50 Å of wire radius decrement, diminution of exciton energy is more rapid as InGaN/GaN and InGaAs/GaAs. With the variation of indium composition binding energy of InAlN/AlN also varie, and this variation is similar with InGaN/GaN as increasing Indium composition binding energy also increased.

Fig. 4 shows comparison between the three InGaN/GaN, AlGaN/GaN, and InAlN/AlN ternary alloys in the view of excitonic effect, in which considered the composition of all the three alloys is *x=0.2*.this comparison shows that for a particular wire radius binding energy of InGaN/GaN is higher than AlGaN/GaN and InAlN/AlN, which means that exciton effect is higher in InGaN/GaN then other two alloys. But in the case InAlN/AlN binding energy of exciton is lower ofr same wire radius and composition. It shows that, we will get more monochromatic light from InAlN/AlN.So InAlN/AlN is more preferable.

IV. Conclusions

In summary, the excitonic properties InGaN/GaN, AlGaN/GaN, InAlN/AlN based cylindrical QWRs have been investigated theoretically for various composition of the constituent element. For InGaN/GaN QWR, with the increase of Indium composition, the energy of exciton increases and the binding energy decreases. It is evident that the change in exciton energy is found to be more significant when the wire radius is less than 50 Å. Same phenomena has been observed for InAln/AlN QWRs with better performance, but for AlGaN/GaN QWRs, with the increase of Aluminium composition an inverse behaviour is observed in compared to the InGaN/GaN QWRs. It is found that the binding energy of exciton in AlGaN/GaN QWRs is greater than InAlN/AlN and less than InGaN/GaN QWRs. From this observation, it can be concluded that among the III-nitride based QWRs, InAlN/AlN has comparatively better performance.

References

[1] S. Nakamura, T. Mukai, and M. Senoh, "Candela-class high-brightness InGaN/AIGaN double-heterostructure blue-light-emitting diodes," Appl. Phys. Lett. 64, 1687 (1994).

[2] S. M. H. Crawford, "LEDs for solid-state lighting: performance challenges and recent advances," IEEE J. Sel. Top.Quantum Electron. 15(4), 1028–1040 (2009).

[3] B. N. Pantha, R. Dahal, J. Li, J. Y. Lin, H. X. Jiang, and G. Pomrenke,"Thermo electric properties of InxGa1-xN alloys," Appl. Phys. Lett. 92(4), 042112 (2008.

[4] J. Zhang, H. Tong, G. Y. Liu, J. A. Herbsommer, G. S. Huang, and N.Tansu,"Characterizations of see beck coefficients and thermoelectric figures of merit for AlInN alloys with various in-contents," J. Appl. Phys.109(5), 053706 (2011).

[5] *U. K. Mishra, P. Parikh, and Y. F. Wu, "AlGaN/GaN HEMTs-an overview of device operation and applications," Proc. IEEE 90(6), 1022–1031 (2002).*

[6] R. Dahal, B. Pantha, J. Li, J. Y. Lin, and H. X. Jiang, "InGaN/GaN multiple quantum well solar cells with long operating wavelengths," Appl. Phys. Lett. 94(6), 063505 (2009).

[7] Puron E, Martinez-Criado G, I. Riech, J. Almeida-gracia and A. Cantariro " Gr- owth and Optical characterization of indirect-gap $Al_xGa_{1-x}As$ alloys," j. Appl. Phys. Lett. 86(1), 091902 (2010).

[8] Zh. I. Alferov," The double heterostructure concept and its applications in physics, electronics, and technology," Rev. Mod. Phys 73, 767 (2001).

[9] Y. Arakawa and H. Sakaki, "Multidimensional quantum well laser and temper- ature dependence of its threshold current,"Appl. Phys. Lett. 40, 939 (1982).

[10] E. Kapon, D. M. Hwang, and R. Bhat," Stimulated emission in semiconductor quantum wire heterostructures," Phys. Rev. Lett. 63, 430 (1989).

[11] W. Wegscheider, L. N. Pfei®er, M. M. Dignam, A. Pinczuk, K. W. West, S. L. McCall, and R. Hull," Lasing from excitons in quantum wires," Phys. Rev. Lett. 71, 4071 (1993).

[12] T. Ogawa and T. Takagahara "optical properties of low dimension materials" Phys. Rev. B, 44,8138, 1991.

[13] S. Benner and H. Haug. "Influence of the external electronic and magnetic field on excitons" Phys. Rev. B, 47,15750, 1993.

[14] Takahisa Ohno, Kenji Shiraishi, and Tetsuo Ogawa "Intrinsic origin of visible light emission from silicon quantum wires: Electronic structure and geometrically restricted exciton" Phys. Rev. Lett. 69, 2400, 1992

[15] D. S. Citrin "Long intrinsic radiative lifetime of excitons in quantum wires" Phys. Rev. Lett. 69, 3393, 1992

[16] E. W. S. Caetano, V. N. Freire, G. A. Farias, E. F. da Silva Jr "Exciton confinement in InGaN/GaN cylindrical quantum wires" Braz. J. Phys. vol.34 no.2b June 2004

Many Body Corrections for Higher Optical Transitions in Semiconducting SWCNTs

G. R. Ahmed Jamal[1,2,*] and S. M. Mominuzzaman[2]

[1]Department of Electrical and Electronic Engineering, University of Asia Pacific
Dhaka, Bangladesh

[2]Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology
Dhaka, Bangladesh

[*]ahmed.eee@uap-bd.edu

Abstract— **In this work, Coulomb effect on determining optical transition energies of semiconducting single wall carbon nanotubes is discussed. Due to their quasi-one-dimensional structure, electron-electron and electron-hole interaction and corresponding self energy and exciton binding energy are important in SWCNTs and the difference between these two energies gives many body effect. Here, at first, a brief review of nature of correction in the conventional single particle electronic picture of SWCNTs to include many body effect in first four optical transitions is presented. Then, many body corrections needed for next three higher optical transitions i.e. 5th, 6th and 7th transitions of semiconducting SWCNTs are investigated. Experimental values of these higher transitions are collected from recent experimental reports so as to explain those data by extending single-particle picture corrected for nanotube curvature and chirality effect along with many body corrections. Our result shows that these three transitions excellently follow the proposed corrected picture with less than 0.5% average absolute error which proves that excitonic behaviour is strong in 5th, 6th and 7th optical transitions of semiconducting SWCNTs unlike 3rd and 4th transitions as reported in some recent works.**

Index Terms—**SWCNT, optical transition energy, electron, exciton, many body effect,**

I. Introduction

The optical properties of semiconducting carbon nanotubes are intrinsically important for potential applications in photonics and also provide insight into their structural and electrical properties [1]. Early optical spectra in SWCNTs had been interpreted in terms of free electron–hole carriers. Indeed, the electronic structure of SWCNTs, predicted by tight-binding Hamiltonian models, provides equally spaced sub-bands of valence and conduction bands with diverging density of states at the edges, known as van Hove singularities, which result from one-dimensional (1D) confinement conditions [2]. Most of the experimental results have been discussed and analyzed in terms of these interband transitions and associates the observed optical transitions with van Hove resonances.

However, a number of theoretical calculations of the optical spectra [3-17] and recent experiments [18-29] suggest that the observed transitions correspond to exciton energies, not interband transitions. An exciton is a bound pair of a photoexcited electron (e) and a hole (h). In conventional semiconductors, an exciton exists only at low temperature, below 10° K. Since the exciton binding energy for a SWCNT is very large (up to 1 eV) due to their quasi-one-dimensional nature, an exciton can exist even at room temperature [3, 4, 9, 12]. Recent transient spectroscopy and nonlinear absorption have unambiguously revealed that the photophysics of SWCNTs is dominated by strongly bound excitons rather than free particles [2].

In particular, the optical transition energies of semiconducting nanotubes, along with their dependence on the nanotube diameter and chiral angle, have been studied in a number of fluorescence and Raman spectroscopy experiments [9]. Several observations from those experiments have attracted attention. First, optical gaps in SWCNTs are greater than those predicted from the tight binding (TB) model. Second, the ratio of the threshold energy corresponding to the second optical transition polarized along the SWCNT axis to that of the first such transition is less than the value 2 predicted within the TB model for wide SWCNTs [5]. Kane and Mele [5-7] pointed out that this ‘‘ratio problem’’ is a signature of electron-electron interactions [13]. More elaborate tight-binding calculation including orbital overlap and interactions beyond nearest neighbours partly account for the observed dependence of optical spectra on chirality. Though some aspects of the experiments can be interpreted within the context of a simple, noninteracting electron model, it is now established [9] from both theoretical calculations [3-17] and experimental measurements [18-29] that a simple band-structure calculation is insufficient for an accurate description of optical transition energies in SWCNTs. This is due to the omission of two highly important effects in low-dimensional semiconductors: (i) quasi-particle corrections to the simple band-structure and (ii) exciton effects. Both quasi-particle and exciton shifts are due to screened Coulomb interactions. The former describes the repulsive energy needed to add an additional electron to the system and, hence, raises the energy of the conduction band or, equivalently, increases the band gap. In contrast, the exciton shift describes the attractive Coulomb interaction between electrons and holes, which lowers the excitation energy. The experimental results in Refs [19,20] demonstrate that, in fact, the overall effect is a blue-shift so that the positive quasi-particle correction actually dominates over the negative exciton binding energy [9].

Starting from the pioneering study by Ando [3] a large amount of theoretical and experimental works has confirmed strong excitonic effects in SWCNTs. Ando [3] originally predicted that the many-body interactions would shift the band gaps of semiconducting SWCNTs to higher energies and create excitons with significant electron-hole binding energy in the excited electronic state. Such strong many-body interactions are supported by further theoretical work [4-17] and several recent experiments [18-29]. Dukovic et al. [24] has shown experimentally that for a (6,5) SWCNT with a

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

diameter of 0.76 nm, the exciton binding energy is roughly 0.42 eV, which is a significant fraction (around 33%) of the band gap. This is about one hundred times larger than that for bulk semiconductors, but comparable to other 1D materials. Thus, excitonic effects cannot be largely neglected or treated as a small perturbation to the bandgap of SWCNTs as it was assumed in earlier studies [2]. Such result also suggests that excitonic effects dominate absorption and emission spectra and hence, all aspects of the optical properties of carbon nanotubes: optical absorption, fluorescence, Raman, and Rayleigh scattering of SWCNTs [24].

Current understanding of the photophysical properties of semiconducting carbon nanotubes are based mostly on experimental results for the first (E_{11}) and second (E_{22}) optical transitions in a small diameter range. Recently, third and fourth optical transitions (E_{33} and E_{44}) of semiconducting SWCNTs also ganied some theoretical and experimental attractions [26-29]. In fact, higher optical transitions are important for the optics of large diameter semiconducting SWCNTs, since for $d_t > 1.3$ nm, E_{22} is already in the infrared range.

The objective of this work is to review the nature of many body effect in first four optical transitions and then to derive an equation that can give complete description of experimentally observed values of next three higher optical transitions i.e. 5th, 6th and 7th transitions of semiconducting SWCNTs under the many body electronic picture. The many body effect on these three transitions will be revealed partially through this attempt.

II. Method and results

The transition energies for parallel polarization to a semiconducting SWCNT are denoted by E_{ii} (i = 1, 2, 3, 4, . . .) for a transition between i-th conduction (an electron) and i-th valence (a hole) energy bands [30]. Originally E_{ii} denoted a van Hove singular energy for the joint density of states for the one-particle energy bands of a SWCNT. Here same notation will be used for the transition energies of the exciton [11].

For a better description of the electronic structure within the tight binding method one has to consider an extended tight-binding model that considers the effects of chirality and curvature, plus the blue-shift correction due to many-body effects. The transition energy of a SWCNT can be considered as the sum of a one-particle energy and a many-body energy which is further decomposed into an exciton binding energy and a self-energy. The one particle energy is calculated by the energy difference between the i-th conduction and the i-th valence energy bands. The exciton binding energy and self energy are the Coulomb interaction for a photo-excited electron with a hole and with other electrons in the valence energy bands, respectively [11]. The latter interaction increases the energy of the electron (decreases the energy of a hole), and this interaction thus contributes to the increase in the energy gap compared with the single-particle energy difference between an electron and a hole. This energy difference is called the self-energy. The exciton excitation energy is thus given by (single-particle energy) + (self-energy) – (exciton binding energy) ≡ (quasi-particle energy) – (exciton binding energy). Here, (self-energy) – (exciton binding energy) is the many-body effect.

Earlier pioneering work by Kane and Mele [7] predicted that both electron–electron and electron–hole long range one-dimensional coulomb interactions (with length scales larger than tube circumference) strongly depend on the inverse tube diameter, competing and nearly cancelling each other [6,7,10,11]. As a result of this cancellation, the observed optical transitions in carbon nanotubes are dominated by short range two-dimensional graphene self-energy effects, leading to a logarithmic correction to the electronic self-energy of the first (E_{11}) and second (E_{22}) related levels [7, 29], commonly known as KM correction in literatures [10,17,27] as given by,

$$\Delta E = .55\frac{2p}{3d_t}\log\left(\frac{3}{2p/3d_t}\right) \qquad (1)$$

Where, p =1, 2, 4, 5, 7, 8 with i =1, 2, 3, 4, 5, 6, respectively, for 1st, 2nd, 3rd, 4th, 5th and 6th optical transitions (E_{ii}) of semi-conducting SWCNTs.

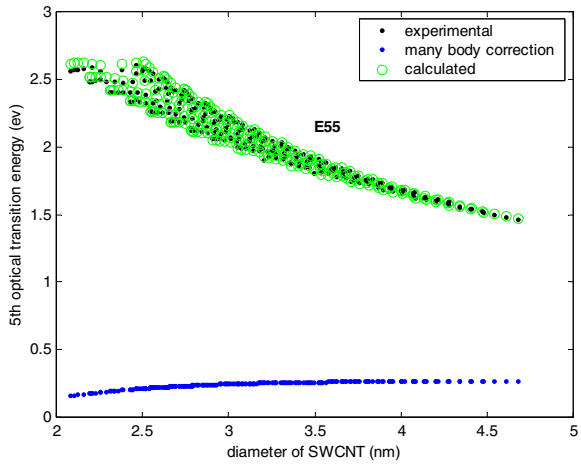
The ability of KM correction, given by Eqn. (1), to fit first and second optical transitions (E_{11} and E_{22}) of semiconducting SWCNTs is already proven [17,27]. The many-body correction to E_{11} and E_{22} energies is reported to be around 0.2 eV, while the correction to E_{33} and E_{44} energies can be 0.6 eV [10,11,17]. Moreover, the many-body correction does not contribute to the family spread in E_{11} and E_{22} in the Kataura plot, while it contributes significantly to the family spread in E_{33} and E_{44} [10,11]. Experimental values of third and fourth optical transitions (E_{33} and E_{44}) of semiconducting SWCNTs are found to deviate from same scaling law and the deviation from the extended tight binding method can be successfully fit by summing a $\Delta E \approx 0.3/d_t$ dependence to the above many-body logarithmic corrections [26-29]. The E_{33} and E_{44} are, therefore, blue-shifted from the excitonic scaling law, and the blueshift goes with inverse diameter, like the exciton binding energy whose value is found to be similar like an earlier experiment [24]. Authors of those works [26-29] interpreted this result as weaker(or even null) exciton binding energy for E_{33} and E_{44} transitions when compared to the lower lying levels.

However, so far there is no significant report that revealed the nature of many body corrections for fifth, sixth or higher optical transitions in semiconducting SWCNTs. Here, KM correction term will be used to fit experimentally reported higher optical transitions, E_{55}, E_{66} and E_{77} of semiconducting SWCNTs to reveal the excitonic picture in these transitions.

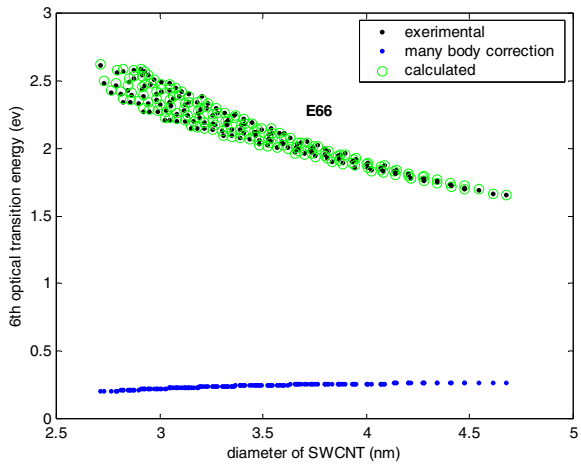
As a first approximation, considering the linear dispersion of the graphene and wave vector quantization, the optical transition energies in carbon nanotubes are given by, $E_{ii} = (4p/3d_t) \times \hbar v_F$, where, $\hbar$ is plank's constant (=6.582×10^{-16} eV.s) and v_F =9.3×10^5 m/s, is the Fermi velocity in semiconducting SWCNT [31] and p =1, 2, 4, 5, 7, 8 for E_{ii} with i=1, 2, 3, 4, 5, 6, respectively [28-33]. The resultant approximation is $E_{ii} \approx p(0.8/d_t)$ [31,32]. The E_{ii} values also exhibit a dependence on the nanotube's curvature and chiralily which can be expressed by a term $\mathrm{Cos}(3\theta)/d^2$ comprising diameter and chiral angle to account for curvature effect and chirality effect, respectively [33-41]. Finally, to complete the picture we need to add the many body correction to optical transitions as derived by Kane and Mele [6,7,10,11,17]. The complete equation is represented by Eqn (2) along with two fitting parameters A and B.

$$E_{ii} = \frac{A(\approx 0.8p)}{d_t} + B\frac{\cos 3\theta}{d^2} + .55\frac{2p}{3d_t}\log\left(\frac{9d_t}{2p}\right) \qquad (2)$$

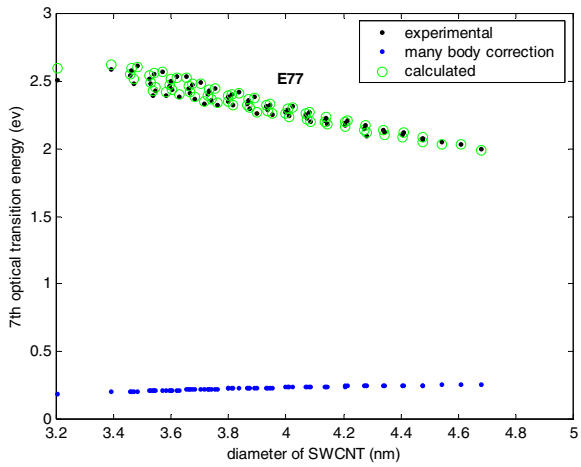
Equation (2) describes three higher optical transitions E_{55}, E_{66} and E_{77}, for p= 7, 8 and 10, respectively, where, first term of



(a)



(b)



(c)

Fig. 1 Experimental value(black dot), calculated value from this work (green triangle) and many body correction (blue dot) of 5th , 6th and 7th optical transitions in (a), (b) and (c), respectively, plotted vs diameter of semiconducting SWCNTs.

the equation carries the linear dependence of E_{ii} on p/d_t expected from quantum confinement of the 2D electronic structure of graphene, second term accounts for curvature and chirality dependence of optical transitions and the third term stands for many-body logarithmic corrections. The number of samples for each mod type and best fit of the fitting parameters A and B to the reported experimental data for E_{55}, E_{66} and E_{77} are given in Table I. In fact, the value of A is expected to be close to 0.8p as shown earlier which is reflected from the values of A for three transitions given in Table I. Hence, mainly B is the only fitting parameter and it can be observed that it takes opposite sign for mod 1 and mod 2 type semiconducting SWCNTs and this alters with odd and even transitions. It was observed that, with these fitting parameters, our calculated values for 5th, 6th and 7th optical transitions in semiconducting SWCNTs excellently match with experimentally reported values over the full diameter range. The average absolute error for all three transitions is below 10 meV and the corresponding %average absolute error was found to be less than 0.5%. The error in each transition rapidly decreases further for higher diameter SWCNTs, as expected from the theory.

TABLE I
FITTING PARAMETERS IN EQN (2) FOR 5TH, 6TH AND 7TH OPTICAL TRANSITIONS OF SEMICONDUCTING SWCNTS

E_{ii}	p	Mod	Total Samples	A ≈0.8p	B	Average Error	% Error
E_{55}	7	1	137	5.63	-1.1	0.01	0.47%
		2	109	5.63	1.1	.008	0.35%
E_{66}	8	1	82	6.5	1.2	.008	0.36%
		2	77	6.45	-1.4	.008	0.39%
E_{77}	10	1	37	8.12	-1.6	.012	0.54%
		2	33	8.2	1.6	.007	0.29%

Figure 1 (a), (b) and (c) show the plot of experimental data as well as values calculated in this work for 5th, 6th and 7th optical transitions, respectively, along with their many body corrections against their diameters. It can be observed from these three plots that, for all three optical transitions, calculated values predicted by Eqn. (2) are almost inseparably too close to the experimental values over the full diameter range under consideration.

As many body correction in the single particle electronic picture of semiconducting SWCNT almost accurately reproduce experimental result of these three optical transitions, it indicates that, unlike 3rd and 4th optical transitions [26-29], Coulomb effect is significant in 5th, 6th and 7th transitions and must be considered for calculating these transitions. This also means, exciton binding energy is not negligible for these three transitions. Future experiments may reveal the validity of this conclusion by measuring the exciton binding energies of these three transitions. Nevertheless, this work can be taken as a guide to understand the optical process involved in these higher transitions of semiconducting SWCNTs. Future experiments are needed to through more light on this matter.

III. CONCLUSIONS

Due to their quasi-one-dimensional structure, electron-electron and electron-hole interaction and corresponding self energy and exciton binding energy are important in SWCNTs. In this work, a brief review of Coulomb effect on determining optical transition energies of semiconducting single wall carbon nanotubes is presented. The difference between electron-electron and electron-hole energies gives many body effect and corresponding correction is called many body correction. Proposed correction by some authors in the conventional single particle electronic picture of SWCNT to include many body effect in first four optical transitions has been discussed. Here, we investigated next three higher optical transitions of semiconducting SWCNTs namely, i.e. 5th,

6th and 7th transitions. Experimental values of these higher transitions are collected from recent experimental reports. An equation is formed to explain those data by extending single-particle picture corrected for nanotube curvature and chirality effect along with many body corrections. Our result shows that these three transitions excellently follow the proposed corrected picture with less than 0.5% average absolute error. This result is an indirect proof that excitonic behavior is strong in 5th, 6th and 7th optical transitions of semiconducting SWCNTs unlike 3rd and 4th transitions as reported by some authors. This also suggests a significant value of exciton binding energies in these higher transitions. Future experiments are needed to through more light on this matter.

REFERENCES

1. S. K. Choi, J. Deslippe, R. B. Capaz and S. G. Louie, "An Explicit Formula for Optical Oscillator Strength of Excitons in Semiconducting Single-Walled Carbon Nanotubes: Family Behavior", *Nano Lett.*, 13 (1), pp 54–58, 2013.
2. S. Kilina and S. Tretiak, "Excitonic and Vibrational Properties of Single-Walled Semiconducting Carbon Nanotubes", *Adv. Funct. Mater.* 17, 3405–3420, 2007.
3. T. Ando, "Excitons in carbon nanotubes", J. Phys. Soc. Jpn. Vol. 66, p.1066, 1997.
4. C. D. Spataru, S. Ismail-Beigi, L. X. Benedict, S. G. Louie, "Excitonic effects and optical spectra of single-walled carbon nanotubes", Phys. Rev. Lett. 92, 077402, 2004.
5. C. L. Kane and E. J. Mele, "Ratio Problem in Single Carbon Nanotube Fluorescence Spectroscopy",Vol. 90, No. 20, 207401, 2003.
6. E. J. Mele and C. L. Kane, "Many body effects in carbon nanotube fluorescence spectroscopy", Sol. Stat. Commn., Vol. 135, pp. 527–531, 2005.
7. C. L. Kane, E. J. Mele, "Electron interactions and scaling relations for optical excitations in carbon nanotubes", Phys. Rev. Lett. 93, 197402, 2003.
8. V. Perebeinos, J. Tersoff, P. Avouris, "Effect of exciton-phonon coupling in the calculated optical absorption of carbon nanotubes", Phys. Rev. Lett. 94, 027402, 2005
9. Thomas G. Pedersen, "Exciton effects in carbon nanotubes", *Carbon* 42, pp.1007–1010, 2004
10. J. Jiang, R. Saito, Ge.G. Samsonidze, A. Jorio, S.G. Chou, G. Dresselhaus, M.S. Dresselhaus, "Chirality dependence of the exciton effects in single-wall carbon nanotubes:Tight-binding model", Phys. Rev. B 75 035407, 2007
11. K. Sato, R. Saito, J. Jiang, G. Dresselhaus and M.S. Dresselhaus, "Chirality dependence of many body effects of single wall carbon nanotubes", *Vibrational Spectroscopy,* Vol. 45, Issue 2, pp.89–94, 2007.
12. H. Zhao and S. Mazumdar, "Excitons in semiconducting single-walled carbon nanotubes", Synth. Met, 155 pp.250–253, 2005
13. H. Zhao and S. Mazumdar, "Electron-Electron Interaction Effects on the Optical Excitations of Semiconducting Single-Walled Carbon Nanotubes", Phys. Rev. Lett., Vol. 93, No. 15, 157402, 2004
14. Z. Wang, D. Psiachos, R. F Badilla and S. Mazumdar, "Electron–electron interaction effects on the photophysics of metallic single-walled carbon nanotubes", J. Phys.: Condens. Matter 21, 095009, 2009
15. T. Ando, "Excitons in carbon nanotubes revisited: Dependence on diameter, Aharonov–Bohm flux, and strain", J. Phys. Soc. Jpn. 73, 335, 2004
16. T. Ando, "Family Effects on Excitons in Semiconducting Carbon Nanotubes", J. Phys. Soc. Jpn. Vol. 78, No. 10, pp.104703, 2009.
17. A. Jorio, C. Fantini, M.A. Pimenta, R.B. Capaz, Ge.G. Samsonidze, G. Dresselhaus, M. S. Dresselhaus, J. Jiang, N. Kobayashi, A. Gruneis, R. Saito, "Resonance Raman spectroscopy (*n, m*) dependent effects in small diameter single-wall carbon nanotubes", Phys. Rev. B 71, 075401, 2005.
18. Z. Wang, H. Pedrosa, T. Krauss and L. Rothberg, "Determination of the Exciton Binding Energy in Single-Walled Carbon Nanotubes", Phys Rev Lett., 96, 047403, 2006
19. M Ichida, S Mizuno, Y Saito, H Kataura, Y Achiba, A Nakamura. "Coulomb effect on the fundamental optical transition in semiconducting single-walled carbon nanotubes: divergent behavior in the small-diameter limit", Phys Rev B 65, 241407, 2002.
20. M Ichida, S Mizuno, Y Tani, Y Saito, A. Nakamura, "Exciton effects of optical transitions in single-wall carbon nanotubes", J Phys Soc Jpn 68, pp.3131–3, 1999.
21. J. Maultzsch, R. Pomraenke, S. Reich, E. Chang, D. Prezzi, A. Ruini, E. Molinari, M.S. Strano, C. Thomsen, C. Lienau, "Exciton binding energies in carbon nanotubes from two-photon photoluminescence", Phys. Rev. B 72, 241402, 2005.
22. J. Maultzsch, R. Pomraenke, S. Reich, E. Chang, D. Prezzi, A. Ruini, E. Molinari, M. S. Strano, C. Thomsen, and C. Lienau, "Excitons in carbon nanotubes", phys. stat. sol. (b) 243, No. 13, pp. 3204–3208, 2006.
23. H. Lin, J. Lagoute, V. Repain, C. Chacon, Y. Girard, J.-S. Lauret, F. Ducastelle, A. Loiseau and S. Rousset, "Many body effects in electronic bandgaps of carbon nanotubes measured by scanning tunnelling spectroscopy", Nature Nanomaterials, Vol 9, pp. 235-238, 2010
24. G. Dukovic, Feng Wangm Daohua Song, Matthew Y. Sfeir, Tony F. Heinz and Louis E. Brus, "Structural dependence of excitonic optical transitions and band-gap energies in carbon nanotubes", Nano Lett. 5, pp. 2314-2318, 2005.
25. F. Wang, G. Dukovic, L. E. Brus, T. F. Heinz, "The optical resonances in carbon nanotubes arise from excitons", *Science* 308, pp.838-841, 2005.
26. T. Michel, M. Paillet, J. C. Meyer, V. N. Popov, L. Henrard and J. L. Sauvajol, "E_{33} and E_{44} optical transitions in semiconducting single-walled carbon nanotubes: Electron diffraction and Raman experiments", Phys. Rev. B 75, 155432, 2007.
27. T. Michel, M. Paillet, J. C. Meyer, V. N. Popov, L. Henrard, P. Poncharal, A. Zahab and J. L. Sauvajol, "Raman spectroscopy of (n,m)-identified individual single-walled carbon nanotubes", phys. stat. sol. (b) 244, No. 11, pp.3986–3991, 2007.
28. P. T. Araujo, S. K. Doorn, S. Kilina, S. Tretiak, E. Einarsson, S. Maruyama, H. Chacham, M.A. Pimenta, A. Jorio, "Third and fourth optical transitions in semiconducting carbon nanotubes", Phys. Rev. Lett. 98, 067401, 2007.
29. A. Jorio, P.T. Araujo, S. K.Doorn, S. Maruyama, H. Chacham, M.A. Pimenta, "The Kataura plot over broad energy and diameter ranges", Phys. Stat. Sol. (b) 243 (2006) 3117.
30. R. Saito, G. Dresselhaus, M. S. Dresselhaus, "Trigonal warping effect of carbon nanotubes", Phys. Rev. B 61, p2981, 2000.
31. Y. Zhao, A. Liao and E. Pop, "Multiband Mobility in Semiconducting Carbon Nanotubes", *IEEE Electron Device Lett.*, Vol. 30, No. 10, pp.1078-1080, 2009.
32. S. Kim and S. Mohammadi, *Single-walled carbon nanotube transistors,* F. Rahman, Ed., Nanostructures in electronics and photonics, Singapore, Pan Stanford Publishing Pte. Ltd., 2008.
33. K. Liu, J. Deslippe, F. Xiao, R. B. Capaz, X. Hong, S. Aloni and F. Wang, "An atlas of carbon nanotube optical transitions", Nature Nanotechnology 7, 325–329, 2012.
34. H. Yorikawa, S. Muramatsu, "Electronic structure characteristic of carbon nanotubules", *Z. Phys. B, Condensed Matt.*, vol. 104, pp. 71–76, 1997.
35. H. Yorikawa and S. Muramatsu, "Energy gaps of semiconducting nanotubles", *Phys. Rev. B*, Vol.52, No.4, pp.2723-2727, 1995.
36. H. Yorikawa and S. Muramatsu , "Chirality-dependence of energy gaps of semiconducting nanotubules" , *Solid State Communications,* Vol. 94, Issue 6, Pages 435-437, 1995.
37. S. M. Bachilo, M. S.Strano, C. Kittrell, R. H. Hauge, R. E. Smalley, R. B. Weisman, "Structure-Assigned Optical Spectra of Single-Walled Carbon Nanotubes", *Science,* Vol 298 No. 5602, pp.2361, 2002.
38. J. Maultzsch, H. Telg, S. Reich, and C. Thomsen, "Radial breathing mode of single-walled carbon nanotubes Optical transition energies and chiral-index assignment", *Phys. Rev. B* 72, 205438, 2005.
39. A. Jorio, C. Fantini, M. A. Pimenta, R. B. Capaz ,Ge. G. Samsonidze, G. Dresselhaus, M. S. Dresselhaus, J. Jiang, N. Kobayashi, A. Grüneis and R. Saito, "Resonance Raman spectroscopy (*n*,*m*)-dependent effects in small-diameter single-wall carbon nanotubes", *Phys. Rev. B* 71, 075401, 2005.
40. G. Lanzani, L. Luer, "Carbon Nanotubes: Electronic Structure and Spectroscopy", *Comprehensive Nanoscience and Technology,* Vol. 1, pp. 23–39, 2011.
41. O. Gülseren, T. Yildirim and S. Ciraci, "A systematic ab-initio study of curvature effects in carbon nanotubes", *Phys. Rev. B* 65, 153405, 2002.

Energy Harvesting Based on Cantilever Beam for Different Piezoelectric Materials from Acoustic Vibration by Simulation Process

Mohammad Rayhan Ahmad,[1] and Md Rifat Al Amin Khan [1], Md Ataur Rahman,[2,*]

[1] Department of Electrical and Electronic Engineering, International Islamic University Chittagong
Chittagong 4203, Bangladesh

[2] School of Electrical and Telecommunication Engineering, RMIT University
Victoria 3001, Australia
*ataur024@gmail.com

Abstract—**This paper includes the output of a cantilever beam based piezoelectric energy harvester for various piezoelectric materials including Lead zirconate titanate (PZT), Potassium niobate ($KNbO_3$) and poly(vinylidenefluoride-co-trifluoroethylene) (PVDF-TrFE) through some simulation process using numerous characteristics of the materials like as dielectric constant, coupling coefficient, Young modulus etc. in different excitation frequencies. The effect of the physical parameter likes piezoelectric layer length, thickness and width on voltage generation from acoustic vibration of the beam is also described in some simulated figure. The maximum output voltage is found over 7 V for $KNbO_3$ among other materials in our study.**

Index Terms—**Energy harvester, piezoelectric materials, cantilever beam.**

I. Introduction

The method of scavenging little energy from ambient sources like as natural lighting, ambient radio frequency, mechanical or thermal resources, body movement etc. is called energy harvesting [1]. This tiny energy is capable of operating small electronic devices without help of any external power source or battery especially in remote area using oscillation energy available in environment [2]. Self-powered microscale device system like micro-electro mechanical system (MEMS) can be built with a piezoelectric microgenerator or nanogenerator where piezoelectric materials can convert available natural ambient energy into electrical energy and the system has no bad impact on environment [3]. There are many researches and works are now ongoing to make this technology useful and available in numerous field for example in industries, unmanned aerial vehicles, drones, handled or wearable devices as well as biomedical application. Piezoelectric materials like as PZT is widely used piezoelectric ceramic which have high piezoelectric voltage, dielectric constants but it has extreme fragile structures [4]. On the other hand, PVDF-TrFE has better flexibility, good chemical resistance yet it need electrical poling for polarization and electrical output is also very low like 1-20 mV output voltage from the device consist of such material [5]. $KNbO_3$ which is known as a ferroelectric material of lead free energy harvesting process has high dielectric constant, high electromechanical coupling, and high curie temperature can be the alternatives of other piezoelectric materials [6]. Here we demonstrate the output voltage from a cantilever beam model energy harvester for this three materials and give comparison among them based on simulation result from MATLAB software.

II. Analytical Modelling

We consider a piezoelectric energy harvester cantilever beam of linear stiffness, k along with residual stress, σ_0 which is fixed in one end and the other end is free to move under an external force due to acoustic vibration. Piezoelectric material is attached to the substructure in upper and lower side which can band with the deflection of the beam. There is a proof mass attached in the free end of the beam as shown in Fig.1.

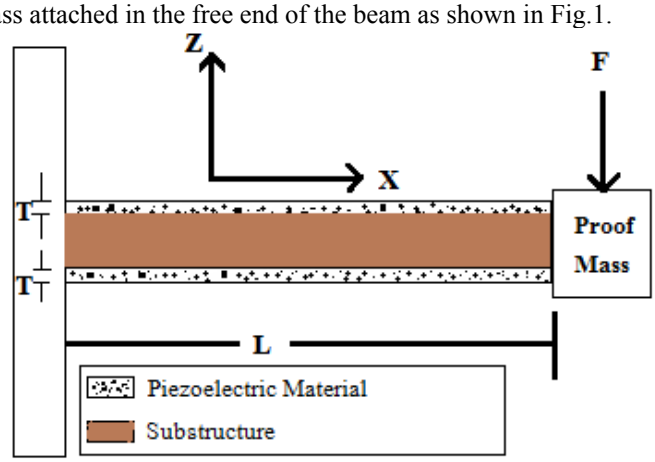


Fig.1 Cantilever beam energy harvester configuration with proof mass.

On the fixed end strain is maximum in bending of the beam by natural oscillation or acoustic vibration which causes potential difference between the two sides of piezoelectric layer thus create piezoelectricity. Due to oscillation of the beam the total strain energy for the beam is given in (1) [7].

$$W_t = \sigma_0 w \int_{-t}^{t} dz \int_{-l/2}^{l/2} \varepsilon_a \, d\varepsilon \qquad (1)$$

Here, ϵ is strain of the beam and l, t, w be the length, thickness and width respectively. Considering beam deflection and assuming some trail solution of the equation the liner stiffness of the beam can be found as given in (2) [8].

$$K = \left(\frac{\pi^2}{2}\right)\left(\frac{\sigma_0 \omega t}{l}\right) + \left(\frac{\pi^4}{6}\right)\left(\frac{Ywt^3}{l^3}\right) \qquad (2)$$

Where Y is Young modulus of piezoelectric material and ω is the excitation frequency. Now interfacing of the piezoelectric energy harvester is shown in Fig.2 with an external load of dielectric loss Rg where I is the total current.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

Now because of potential difference between the two end of the piezoelectric material there is a voltage induced in it and it is given in (3).

$$V^2 = \frac{2t}{\varepsilon A}\left[\frac{k_c\, m^2 Y^2\, \omega}{2\,k}\,\frac{\omega^4}{[1-\lambda^2]^2 + [2\zeta\lambda]^2}\right] \quad (3)$$

Here, A and k_c are the area and electromechanical transform coefficient of the piezoelectric material; m is proof mass, ω_n is natural frequency, ζ is damping factor, ε is dielectric constant and λ is the ratio of ω and ω_n. Now if R_g is dielectric loss and Z_L is the load then from the fundamental law of current and power and by combining them the total power from the energy harvester, P is given in (4).

$$P = \left(\frac{IR_g}{R_g + Z_L}\right)^2 Z_L \quad (4)$$

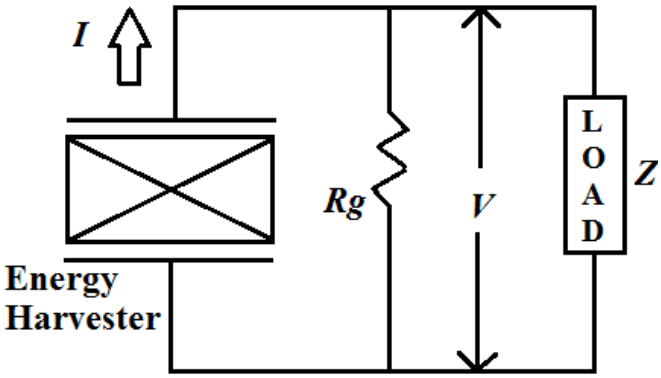


Fig. 2 Parallel configuration of piezoelectric energy harvester with external load.

Using the voltage and the power equation and putting the values of the corresponding characteristics given in Table 1 we measured the voltage generation and output power from the piezoelectric energy harvester by MATLAB simulation.

TABLE I
PROPERTIES OF PIEZOELECTRIC MATERIALS

Piezoelectric Material	Young Modulus Y (GPa)	Coupling Coefficient (K_{33})	Dielectric Constant (ε_r)	Ref.
PZT	60	0.6	1700	[8]
$KNbO_3$	91	0.53	500	[6]
PVDF-TrFE	3.1	0.3	15	[3]

III. RESULT AND DISCUSSION

Under an acoustic vibration or natural oscillation the beam deflected due to downward external force hence piezoelectric layer bends causes voltage generation across it. We assumed a piezoelectric layer having length of 8 mm, width 5 mm and thickness 5.7 µm for each material. When the beam oscillates at about 86 Hz among the three materials for $KNbO_3$ the energy harvester has the highest voltage generation which is around 2.4 V for these physical parameters of the piezoelectric layer. At same frequency PZT gives only 900 mV output and PVDF-TrFE gives about 300 mV output. Fig. 3 shows the output voltage of the energy harvester for three materials.

Voltage generation from energy harvester is mainly depends on the Young modulus of the material. But it is possible to generate more potential difference between the two end of piezoelectric material by changing its layer length, width and thickness to adjust resonance frequency of the beam.

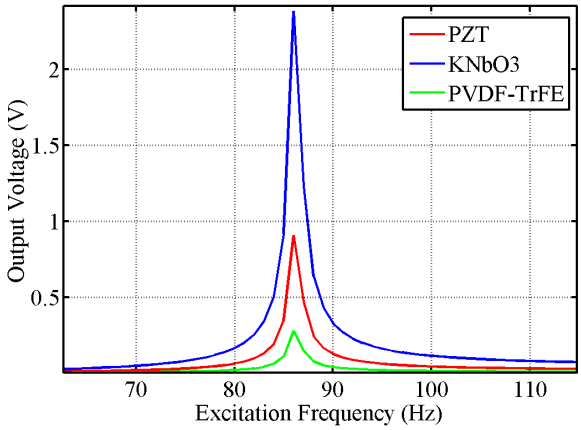


Fig. 3 Voltage generation in different excitation frequencies.

For length variation voltage generation is gradually decreases with the increase in piezoelectric layer length which is given in Fig.4. For example in case of $KNbO_3$ for 7, 9 and 11 mm length output voltage is 3 V, 1.9 V and 1.1 V approximately. On the other hand due to change in layer width and thickness output voltage increases which is described in Fig.5 and Fig.6. When layer width is 4 mm output voltage from the harvester is 0.5 V and it reaches to 1.1 V for PZT.

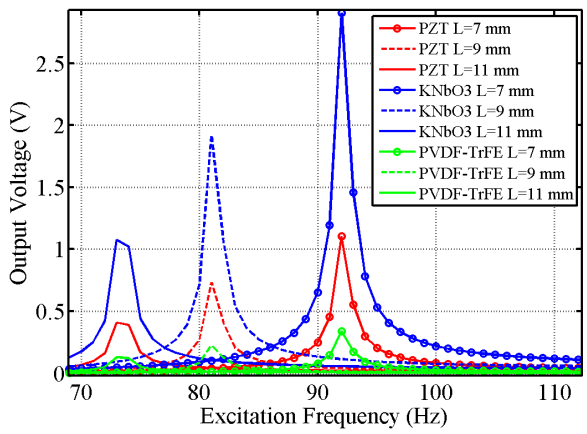


Fig. 4 Effect of changing piezoelectric layer length on output voltage for piezoelectric materials.

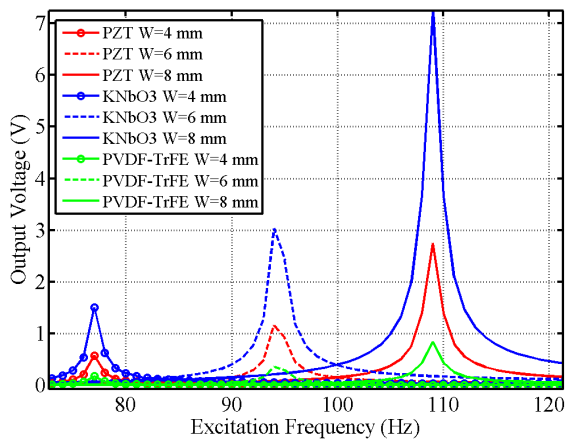


Fig. 5 Effect of changing piezoelectric layer width on output voltage for piezoelectric materials.

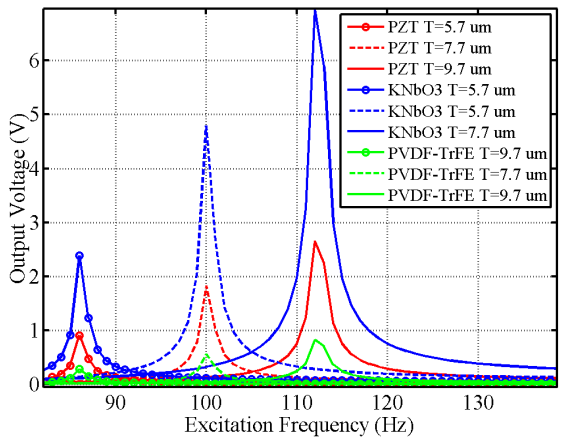


Fig. 6 Effect of changing piezoelectric layer thickness on output voltage for piezoelectric materials.

Simulating the power curve against mega ohm load we also measured the power for piezoelectric materials illustrated in Fig.7. For $KNbO_3$ output power is 700 nW which is the highest value among other two materials. PZT and PVDF-TrFE have the output power of 81.2 nW and 11.2 nW respectively.

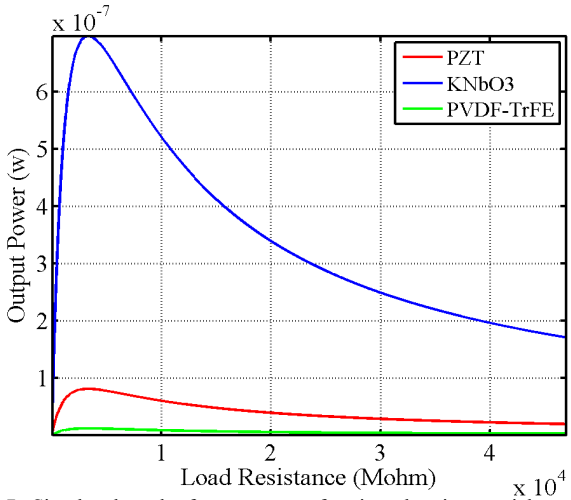


Fig. 7 Simulated result of output power for piezoelectric materials.

IV. Conclusions

By changing the physical parameter of piezoelectric layer we can modify the output voltage of the energy harvester as well as output power. Again this change can allow us to harvest more energy from different types of ambient environment and sources at different excitation frequencies for all the piezoelectric materials. In our future work we will study deeply on other types of the cantilever beam to make an efficient piezoelectric energy harvester.

References

[1] R. Calio, U. B. Rongala, D. Camboni, M. Milazzo, C. Stefanini, G. Petris and C. M. Oddo, "Piezoelectric Energy Harvesting Solutions," Sensors 14, 4755-4790, Mar. 2014.

[2] A. Erturkand D. J. Inman, Smart Mater. Struct, "Issues in mathematical modeling of piezoelectric energy harvesters," 17, 65016, Oct. 2008.

[3] S. Crossley, R. A. Whiter and S. K. Narayan, "Polymer-based nanopiezoelectric generators for energy harvesting applications," Materials Science and Technology 30 (13a), 1613-1624, Jul. 2014.

[4] Xi Chen, Shiyou Xu, Nan Yao, and Yong Shi, "1.6 V nano generator for mechanical energy harvesting using PZT nano fibers," Nano Lett. 10, 2133–2136, May 2010.

[5] L. Persano, C. Dagdeviren, Y. Su, Y Zhang, S. Girardo, Dario Pisignano, Yonggang Huang & John A. Rogers, "High performance piezoelectric devices based on aligned arrays of nanofibers of poly(vinylidenefluoride- co-rifluoroethylene)," Nature Communications 4(1633), 1-6, May 2013.

[6] J. H. Jung, C. Chen, B. K. Yun, N. Lee, Y. Zhou, William Jo, L. Chou and Z. L. Wang, "Lead-free $KNbO_3$ ferroelectric nano rod based flexible nano generators and capacitors," Nanotechnology 23, 375401-406, Aug. 2012.

[7] S. D. Senturia, *Microsystem Design*, Springer, 2nd ed., Berlin, Germany: Springer-Verlag, 2004.

[8] Arman Hajati, Sang-Gook Kim, "Ultra Wide-Bandwidth Micro Energy Harvester," Appl. Phys. Lett. 99, 083105, Aug. 2011.

Active Management of Wind Turbine-Diesel Generator Hybrid System Considering Economic Criteria

Hasnain Ahamad Ziad, Md. Rafiqul Islam
Department of Electrical and Electronic Engineering (EEE)
Khulna University of Engineering & Technology (KUET)
Khulna, Bangladesh
ziad.kuet86@gmail.com, rafiq043@yahoo.com

Abstract— **A simulation on developing the voltage control with active management system of Wind Turbines (WT)-diesel generator (DG) hybrid system considering economic criteria is exhibited in this paper. The hybrid system model is developed from mathematical equations of WTs as well as diesel generators and evaluated the system performances. Minimizing the overall cost and power losses the proposed active management method includes the On Load Tap Changers (OLTC) voltage regulator in addition to reactive power control of WTs and DG. The performance of this method is verified on a 30-bus grid associated 11-kV distribution network comprising WTs and DG, consistent with time-varying characteristics of the load and WTs power generation.**

Keywords-Active management, hybrid system, energy curtailment, economic criteria, wind turbines, diesel generator

I. INTRODUCTION

The ever increasing energy demands, falling sources of fossil fuels and concern about growing pollution levels in the environment has been driving force behind electricity generation by renewable power supplies. The utilization of conventional energy sources causes to increase of electrical energy generation cost besides their pollutant effects to environment [1]. Tapping the energy available from the renewable sources allows the attainment of remarkable reductions in the pollution levels and worrying climate changes. Another is the accessibility of energy from renewable sources unpredictable and is pretended by factors ahead of human control. So these renewable energy schemes should be supported by conventional energy sources or energy storage mechanisms to cater the continuity supply. In order to optimum and economic operation of the renewable energy schemes the model of hybrid systems has been projected. The distributed generations of the smart grid which is the compounding of conventional and renewable energy source considering the energy and voltage management, is called hybrid system (HS). The HS consisting of WTs, and a DG is focused in this paper.

The important factors of HS are voltage control and energy management with optimum and economical operation. Newly, voltage instability has been creditworthy for major network collapses [2]. As a result, the terms voltage management and voltage collapse are coming out more appear in the literature and in discussions of system planning and operation. In order to attain efficient and authentic operation of power system, the control of voltage as well as reactive power should satisfy some specific conditions and that is given in [3].

Many works on voltage control with distributed generators, such that voltage control and discusses a sort of active voltage management systems based on harmonized voltage control. These structures can be used to progress the voltage profile in 11kV radial distribution system and increase their ability to put up distributed generation [4] [5]. Another investigates an active voltage control technique by coordinating the reactive power of circulated generation and the operation of OLTC with sensitivity model of the V-Q combined in distribution network is used in the optimization of the corresponding voltage regulation [6]. The description of selected power control features of shunt power arrangement, STATCOM systems, static reactive power systems and DG generators that combine both these resolutions, which are referred to as Static Var Compensator based on STATCOM where not left out [3][7].

On the above discussion it is seen that all studies of voltage control of HS are related to OLTC active management control and reactive power control with proper device as well as DG without considering total power cost benefit and power losses. In this paper, the voltage control is discussed with coordinate OLTC voltage control, WTs and DG according to cost benefit, power losses as well as Energy Curtailment (EC). The system could be designed to operate in grid adjunctive mode, through OLTC. The simulation model of HS is developed from mathematical models of WT and DG with MATLAB SimPowerSystems. For the purpose of voltage control of this distributed system, Coordinate OLTC transformer over and above WT and DG reactive power control is used. The soaring cost DG is used to backup for active and reactive power of this model. Another main apprehension is to reduce the overall cost including hourly cost of this HS. The coordinated voltage regulation of OLTCs and the reactive power control of WTs and DG are best suitable for cost efficiency as well as less power loss.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

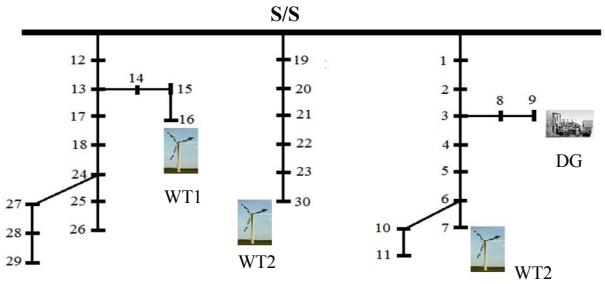


Figure 1. 30 Bus radial distribution system

TABLE I. NETWORK LOADING

Load Band	Active Power [MW]	Reactive power [MVAR]
Minimum	1.51	1.13
Medium	2.16	1.60
Normal	2.59	1.93
Maximum	3.05	2.28

II. METHOD DESCRIPTION

A. Active Management

Active management means a marginal method to facilitate suitable goals for renewable energy and increase the penetration of WTs into the existing distribution networks [8] [9]. It has, indeed, the prospective to maximize DG penetration level while minimizing DG-related network assistances [10]. Active management can be appreciated for instance, through generation dispatch, transformer tap changing and reactive power compensators or device. The EC of WT's generation during minimum load, reactive power supervision using DG and area-based OLTC coordinated voltage control have been used in the active management.

B. Energy curtailment

In order to improve the over-voltage problem and to maintain equality constrain, it may be essential to curtail some amount of inserted wind energy into the network [9]. In this active management, wind energy may be curtailed during certain periods in order to relieve any voltage or thermal constraint violation. For a certain episode, there are different probable combinations of load as well as wind power and any amount of wind energy may be curtailed at the combination of minimum demand and maximum wind power.

C. Discrete Variable Handling

Using both discrete and continuous controls converts the optimal power flow into a miscellaneous discrete-continuous optimization problem. Since, in the painstaking mixed-integer nonlinear problem, the sole integer variable is the OLTC, the solution of the mixed-integer nonlinear programming problem was executed in a two-step strategy [11] [12].

D. Constraints

The equality constraints h(x, u) represent the static load flow equations for instance the Kirchhoff's current law ∀b ∈ B, where B is the set of busses (b is indexed), and the Kirchhoff's law of voltage ∀v ∈ V, where V is the set of lines (v is indexed) [13]. The inequality constraints g(x, u) are scheduled in the subsequent [14]:

TABLE II. WT GENERATED ACTIVE POWER

Wind velocity [m/s]	0	6	8	10	12
Active power [MW]	0	.11	.30	.59	1.05

1) Active and reactive power constraints for the WTs and DG ∀w ∈ NW

$$P_w^- \le P_w \le P_w^+$$
$$Q_w^- \le Q_w \le Q_w^+ \quad (1)$$

Where P_w and Q_w are the active and reactive power outputs of w, respectively, and P^-_w /P^+_w and Q^-_w/Q^+_w are the min/max values that they can assume.

2) Voltage level constraints ∀b ∈ B

$$V_b^- \le V_b \le V_b^+ \quad (2)$$

Where V_b is the voltage at node b and $V^+_b$and V^-_b are the max/min values that it can assume.

3) Voltage stability constraint

$$\lambda_* - \lambda_p \ge \Delta\lambda_{min} \quad (3)$$

$\Delta\lambda_{min}$ correspond to the minimum acceptable edge of stability for the system, and it is defined by the system operator. The constraint λ is a "loading factor," and λp is given a fixed value [14].

III. CASE STUDIES

A. Study system

This suggested mechanism was applied to a 30-bus 11-kV radial distribution system, shown in Fig. 1, obtained by a modified data given in [16].

The three feeders are delivered by a 6-MVA 33/11-kV OLTC transformer; the tap position allows nine different voltages with a step ΔU=0.0235 p.u.

The dynamic models for the WTs, the DG, the OLTC, and the other elements of distribution system (transformers, wires and control systems) have been put into practiced in MATLAB SimPowerSystems, and the distribution structure has been simulated by using the phasor explanation process [15].

The loads of each bus across one year are thought about: maximum, normal, medium, and minimum load [10] can be seen in Table I. The three-phase maximum load levels which are scaled down for every bus are given in [16].

In the experiment network, three wind turbines (WT1, WT2 and WT3) are attached at 16, 7 and 30 bus respectively. Each WT generates about 1.05 MW at 12 m/s wind speed, in service from 0.85 leading power factor to lagging [17]. The power take out from a WT is a function of the obtainable wind power [17] given in Table II. The high cost DG producing an active power within 1.705 MW to 600 kW [14] is connected at bus 9.

B. Power Cost Definition

It has been scrutinized that the active and reactive powers supplied by the grid are cheaper than those supplied by the DG during all load hours in [18], as shown in Table III. The WTs have very low functional costs since they do not rely on fuel; it

TABLE III. ECONOMIC DATA [$/MWH]

Diesel Generator		Imported from the Grid		Power Losses
Active Power	*Reactive Power*	*Active Power*	*Reactive Power*	
173	17.3	55	5.5	75

TABLE IV. NETWORK VOLTAGE [P.U.] AT WIND SPEED 8 M/S FOR MAXIMUM LOAD

Bus No.	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
3	.92	.93	.97	.99	1.00
21	.93	.94	.97	1.00	1.00
29	.92	.93	.96	.99	.99
30	.93	.94	.97	1.00	1.00

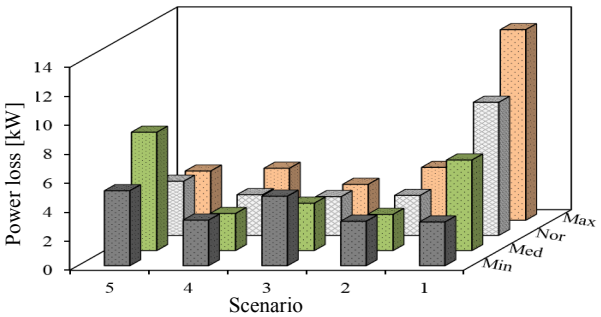


Figure 2. Power losses at wind speed 12m/s

has been supposed that it has no costs for generations. The costs are received from the commodities market electricity price.

C. Scenario 1: Distribution network without WT and DG

In this scenario renewable energy and active management option are not used. The voltage level of this network is 0.93 p.u. of maximum and normal load state shown in Table IV. For insufficient reactive power supported from the grid, the voltage level in this network is minimum. The power loss is high at normal and maximum load, shown in Fig. 2, due to line loss over long distance to transported active and reactive power. This scenario shows that the power loss increases with increasing the load. It is occurred for power loss is the I^2R loss. Figure 3 shows that the total cost of the generated power is maximum as all power are supported by the grid. Total power imported from grid is 2.65 MW at normal load.

D. Scenario 2: Distribution network with WT

The three WTs are used as the renewable energy sources at this time. Table IV shows that the voltage of the system is slightly increased (about 1%) due to WT support few amount of reactive power. Figure 2 shows that the loss of this network during minimum load is 1.66% increased at 12 m/s wind speed, as the injected wind power is much superior to the total load demand of the network. At normal and maximum load, power loss is decreased about 70% with respect to scenario 1, here the injected wind power is no longer higher than load demand of the network and the network absorbed most of active and

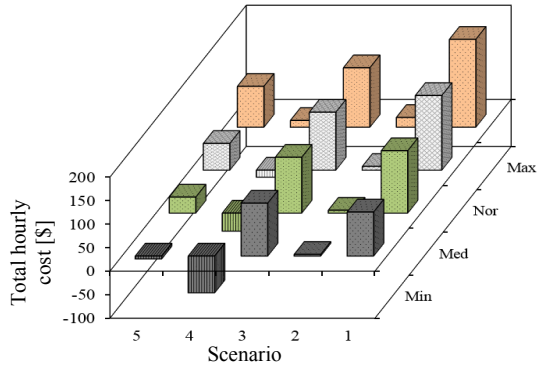


Figure 3. Total cost at wind speed 12m/s

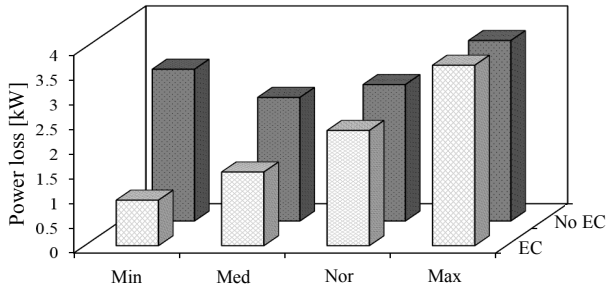


Figure 4. Power losses in Scenario 2

reactive power from WT which are connected in this distributed network as distance losses is minimum. It is seen from Fig. 3 that the total cost is minimum and decreased to 95% from scenario 1 owing to the maximum active and reactive power of this network are supported by three WTs. During normal and minimum load at wind speed 12m/s wind power are larger than total load. In this situation no OLTC are used so bidirectional power flow cannot possible at this condition. But it is necessary to maintain equality constrains as 1.44 MW and .29 MW energy must be curtailed. Fig. 4 shows that the losses are decreased when EC option are implemented.

E. Scenario 3: Distribution network with WT and DG

In this circumstance the DG is used for reactive power sustain to the distribution network. The voltage of the network is 3%-4% increased, shown in Table IV, due to reactive power supported by DG. It can be seen from Fig. 2 that the power losses are slightly reduced with respect to scenario 2 during normal and maximum load at wind speed 12m/s as long distance reactive power loss near to zero. When the high cost DG is operating at minimum output power of 600 kW the operating cost is high of 103.8 $/MWh. The total cost of this network is increased 508% during maximum load and 3000% during minimum load shown in Fig. 3. As high operating cost it is not economical operation.

The bidirectional power flow cannot possible as scenario 2 is same. Here no power exported to the grid. So it is needed EC operation to maintain equality constrain. During minimum and normal load operation 2.07 MW and .97 MW power must be curtailed. Fig. 5 shows that the losses are decreased when EC option are applied.

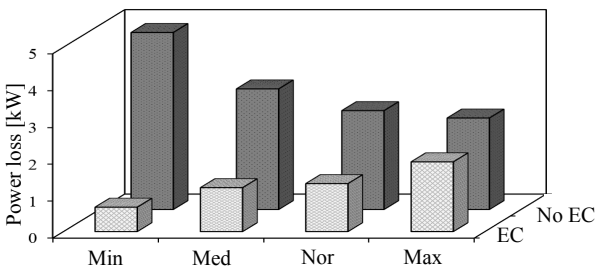


Figure 5. Power losses in Scenario 3

F. Scenario 4: Distribution network with WT and OLTC

Only the OLTC are used here to support the voltage at required level of the network. Voltage level of the distribution network increased about 4% to 6% by changing tap -1 to -3 from scenario 2 shown in Table IV. As the bidirectional power flow can possible no energy curtailment operation are required. The total network loss is increased 2% at during maximum wind speed shows in Fig. 2. As the increment of the losses is higher than the increment of the produced energy. Fig. 3 shows that total cost of the network is minimum during maximum load at maximum wind speed and it is 14.34 $/MWh. During minimum and normal load at maximum wind speed, energy is sold overseas to the grid. During minimum load and maximum wind speed 1.48 MW are exported to the grid. To make batter reliability of the smart grid with renewable energy OLTC transformer must be used.

G. Scenario 5: Hybrid system

In this scenario both DG and OLTC voltage control are used simultaneously. The OLTC used for maintaining the voltage of the network at desired level and the DG to support the active and reactive power during minimum wind speed. The voltage level of the network is faintly increased from scenario 4, shown in Table IV, because the DG supports reactive power. During minimum load and utmost wind speed the total loss shown in Fig. 2 is maximum than other scenario because generated energy is much superior to the total load demand. In Fig. 3 the total cost of scenario is 6 times higher than scenario 4 due to the high operating cost DG. The total power exported to the grid during minimum load and maximum wind speed is 2.08 MW, on the other hand during maximum load and minimum wind speed, 1.06 MW power is imported from the grid although the DG supplied its maximum output of 1.7 MW. The entire power loss and cost is high when the OLTC and DG are used simultaneously. So here the DG are used for support the active and reactive power during peak hour.

IV. CONCLUTION

This paper has proposed to develop the active management of hybrid system for optimum and economic operation. The 30-bus 11 kV radial distribution system includes the OLTC voltage regulator and reactive power control of WT and DG. The result recommended that the combination of WT and OLTC as active management are suitable for economic operation. Hence the voltage of the network is 4-6% increases as well as power loss 2% decreases for minimum load compared to the scheme where active management system are not implemented. It is also found that 1.48 MW active powers are exported to the grid at maximum wind speed and minimum load. To maintain equality constraint the energy curtailment option are applied when the area based OLTC are not used as active management system. It is also seen that the maximum power loss are reduced to 87% for minimum load hour. The result also proofs that the high cost diesel generator used only for supported active power during low wind speed and maximum load hour.

REFERENCES

[1] E. Kabalci, "Design and analysis of a hybrid renewable energy plant with solar," Energy Conversion and Management, vol. 72, pp. 51-59, 2013.

[2] CIGRE Task Force 38-02-10, Modelling of Voltage Collapse Including Dynamic Phenomena, 1993.

[3] Okwe Gerald Ibe, “Concepts of Reactive Power Control and Voltage Stability Methods in Power System Network”, IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p- ISSN: 2278-8727. Volume 11, Issue 2, PP 15-25, May. - Jun. 2013.

[4] M. Fila, D. Reid, G. A. Taylor, P. Lang, M. R. Irving, “Coordinated voltage control for active network management of distributed generation” Power & Energy Society General Meeting, PES '09. IEEE,2009.

[5] Tengku Juhana, Tengku hasim, Azah mohamed, Hussain Shareef, “A review on voltage control methods for active distribution networks”, Przegled Elektrotechniczny (Electrical Review), R. 88 NR , june, 2012.

[6] He Yujun, Marc Petit, Supélec E3S, “Active voltage control using distributed generation on distribution networks”, PowerTech (POWERTECH), 2013 IEEE Grenoble, june 2013.

[7] M. A. Mahmud, M. J. Hossain, H. R. Pota, and A. B. M. Nasiruzzaman, “Voltage control of distribution networks with distributed generation using reactive power compensation”, IECON 2011 - 37th Annual Conference on IEEE Industrial Electronics Society, November, 2011

[8] Pudjianto D., Castro M., Djapic P., Transmission investment and pricing in systems with significant penetration of wind generation’. IEEE Power Engineering Society General Meeting, pp. 1–3, 2007.

[9] Liew S.N., Strbac G., ‘Maximising penetration of wind generation in existing distribution networks’, IEE Proc. Gener. Transm. Distrib., 149, pp. 256–262, 2002.

[10] P. Siano, P. Chen, Z. Chen, and A. Piccolo, “Evaluating maximum wind energy exploitation in active distribution networks,” IET Gener. Transm. Distrib., vol. 4, no. 5, pp. 598–608, 2010.

[11] F. Capitanescu and L. Wehenkel, “A new heuristic approach to deal with discrete variables in optimal power flow computations,” in Proc. IEEE PowerTech Conf., Bucharest, Romania, pp. 1 8, Jun. 28–Jul. 2, 2009.

[12] W. F. Tinney, J. M. Bright, “Some deficiencies in optimal power flow,” IEEE Trans. Power Syst., vol. PWRS-3, no. 2, pp. 676–683, May 1988.

[13] P. Siano, L. F. Ochoa, G. P. Harrison, and A. Piccolo, “Assessing thestrategic benefits of distributed generation ownership for DNOs,” IET Gener. Transm. Distrib., vol. 3, no. 3, pp. 225–236, Mar. 2009.

[14] C. Cecati, C. Citro, A. Piccolo, and P. Siano, “Smart operation of wind turbines and diesel generators according to economic criteria,” IEEE Trans. Ind. Electron., vol. 58, no. 10, pp. 4514–4525, 2011.

[15] SimPowerSystems for Use With Simulink—User’s Guide, The MathWorks, Natick, MA, ver. 4, 2009.

[16] D. Das, “A fuzzy multiobjective approach for network reconfiguration of distribution systems,” IEEE Trans. Power Del., vol. 21, no. 1, pp. 202–209, Jan. 2006.

[17] Carlo Cecati, Costantino Citro, Pierluigi Siano, “Combined operations of renewable energy systems and responsive demand in a smart grid” IEEE Transactions On Sustainable Energy, vol. 2, no. 4, October 2011.

[18] “Bangladesh Energy regulatory commission” Available at: www.sari-energy.org.

Teaching Prim's Algorithm via Mobile App

Debzani Deb, M. Muztaba Fuad

Department of Computer Science, Winston-salem State University
Winston-salem, North Carolina, USA
debd@wssu.edu, fuadmo@wssu.edu

***Abstract*—Computer Science instructors have been exploiting learning technology such as Algorithm Visualization for last few years to explain hard-to-understand algorithms to the learners through simulations and animations. While much progress was made to improves learner's lower-level cognitive skills in the revised Bloom's taxonomy such as "Remembering" and "Understanding", very few research initiatives were undertaken to enhance higher order skills that require learners to apply their knowledge to a different scenario. The research presented in this paper implements an active classroom to teach Prim's Minimum Spanning Tree (MST) Algorithms by exploiting Mobile Response System (MRS) and interactive exercise (IE) Apps. MRS is a software environment that facilitates in-class interactive problem solving exercises and their real-time assessment utilizing mobile devices. By having interactive mobile App exercises rather than traditional paper-based exercises, our goal is to make the problem-solving exercise more visual and appealing to the students and allow them to realize the effect of their interactions at the different stages of the exercise. This paper presents the design and experiences of utilizing a Prim's MST App in classroom. Initial results show the increased learner's satisfaction and interest toward this approach.**

***Index Terms*— Interactive learning environments; learning technology; immediate assessment; mobile technology**

I. Introduction

Active learning is an educational approach where learners become engaged in interactive learning activities in the classroom and has been proven to be an effective mechanism to enhance learner's mastery of skills such as critical thinking and problem solving [1,2]. In a traditional lecture-oriented classroom with minimal learner participation, it is hard to engage learners at levels above "understand" of the revised Bloom's taxonomy [3]. In the discipline of Computer Science (CS), Algorithm is a required course for all CS majors where they acquire skills such as problem solving and critical thinking by being engaged at higher-order cognitive activities in Bloom's taxonomy such as creating, applying, evaluating and analyzing solutions for a problem. Computer Science instructors have been exploiting learning technology such as Algorithm Visualization (AV) [4] for last few years to explain hard-to-understand algorithms to the learners through simulations and animations. While much progress was made to improves learner's lower-level cognitive skills in the revised Bloom's taxonomy such as "Remembering" and "Understanding", very few research initiatives were undertaken to enhance higher order skills that require learners to apply their knowledge to a different scenario. To date, in terms of engaging learners in technology-enhanced interactive problem solving activities, both the number of efforts and our knowledge of its effectiveness are limited.

The research presented in this paper implements an active classroom to teach one of the important Algorithm such as Prim's Minimum Spanning Tree (MST) by exploiting Mobile Response System (MRS) and interactive exercise (IE) Apps. MRS [5-7] is a software environment that facilitates in-class interactive problem solving exercises and their real-time assessment utilizing mobile devices. This client-server software allows the instructor to dynamically prompt the learners with IE Apps, synchronized with the lecture material in their mobile devices. Learners are able to actively interact with the problem while recognizing the effect of their interactions visually at different stages of the problem and send their solutions back to the server computer. MRS then facilitates grading of the exercises automatically, by comparing the learner made sequence of interactions with the correct sequence of interactions. Since the delivery, time keeping and grading of the IEs are automated in this technology-enhanced learning environment, instructor could use more of it in the classroom and learners could actively participate in more exercises where they can receive feedback.

The interactive learning activities and their formative assessments allows the instructor to have real-time evidence of learners' comprehension of covered lecture materials on that particular class and also helps instructor to identify the concepts that need to be repeated or reinforced. By utilizing MRS software, instructor is also able to capture screens from learner submissions and discuss/question on those screens if context-sensitive feedback is needed. On the other hand, these hands-on problem-solving activities allow the learners to obtain faster and frequent feedbacks that reinforce their learning, provide self-assessment and help them to identify misconceptions and problem areas. The visual and active interaction with a problem via multiple steps while going back and forth and seeing the consequence of their choices at each step is expected to enhance learners' analytical and problem solving capabilities and encourage them to perform higher-level activities.

Few researches leverages technology while supporting in-class interactive exercises and integrated assessment. TRAKLA2 [8] is the pioneering learning environment that supports algorithm simulation exercises, automatic assessment and visual feedback. JHAVÉ [9] is a java application that renders algorithm visualizations and allows learner explorations by pop-up questions. VisuAlgo [10] is an online collection of AVs that generates questions for the learners to answer based on the content. OpenDSA [11] and related JSAV [12] supports algorithm visualization and proficiency exercises. However all of them support content delivery via web browser and therefore utilized technologies such as HTML, Java, JavaSript etc. but none to our knowledge do so via native mobile App like MRS.

The detail design, implementation and deployment of the MRS environment is out of scope of this paper. Interested readers are advised to explore related articles and resources [5-7] for downloadable modules and to learn more about MRS

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

technology. This paper focuses on design, development and deployment of Prim's MST App in a classroom and describes the way it is exploited to create a technology-rich interactive learning environment for CS students leaning Algorithms. The rest of the paper is organized as follows: section two briefly describes the MRS system and interactive exercises, section three presents Prim's MST App and section four concludes the paper.

II. MRS System and IE Apps

In MRS, instructor computer runs the server component of the software, which hosts questions, manages users, and maintain communication and synchronization. The client component executes in learner's mobile devices, which allows learners to login to the system and facilitates interactive exercise solving. The server has been developed in Java and the client has been developed in Android. The MRS server is designed as a multi-process, multi-threaded entity to satisfy simultaneous invocation from users and to provide real time responses to in-class activities. Instructor running the MRS server can import learner credentials, which will allow the learners to login to the sever. Instructor can also import the interactive exercise (IE) and broadcast them to learners' client devices. When the MRS client receives a new IE problem from the server, the corresponding app that renders the given problem into interactive exercise is executed and allows learner interaction with the App. The App captures learner interaction at every stage (screen) and sends it back to the server for grading. When the server receives answers back from all the clients, it uses runtime reflection to find the corresponding server-side grading component of the IE and loads and executes the corresponding method in that component dynamically in order to grade learner submissions.

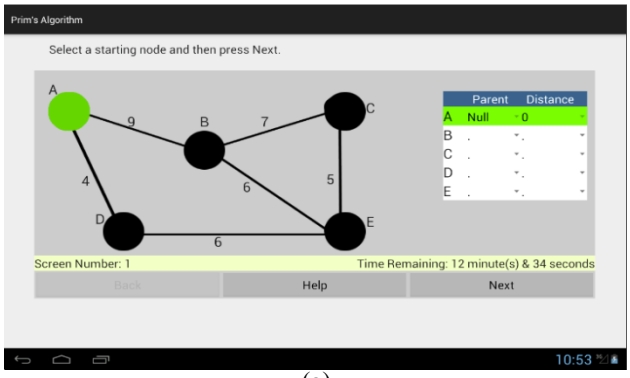
The most important aspect of the MRS software is the facilitation of interactive problem solving. An interactive problem solving activity is defined as an Interactive Exercise with corresponding grading components and associated rubrics. In an interactive exercise, learners require to directly work on a visual representation of a problem and develop the answer following a set of steps guided by a particular algorithm or process. In each step, learners make key choices (for example clicking the table or array indices for selection or swap, selecting from a drop-down menu, selecting a tree node or an edge of a graph etc.) that will impact their next step of interaction. During these interaction steps, learners can go back and forth (by utilizing "Back" and "Next" button) and see the impact of their different choices. Interactive exercises can be offered as solving a whole problem from bottom-up or top-down fashion or solving certain steps of a particular problem in order to give learners different perspective on the problem and to assess their problem solving skills. Each problem has a rubric that grade partially correct answers to gauge learner's problem solving skills and cognitive models.

In this research, the interactive exercises are designed as dynamic entities rather than static entities to support greater diversity and to allow learners to practice with different variations of the same problem. The exercises are therefore parameterized, where parameters can be populated with either randomly generated values or instructor generated values to create different instances of a problem. An IE definition is stored as an XML file and contains exercise parameters such as problem components, time to answer, special instructions etc., which are used to render the corresponding IE App in the client device. The file additionally contains different rubric and grading parameters such as correct answers, grading weights, etc. to support server-side automated grading.

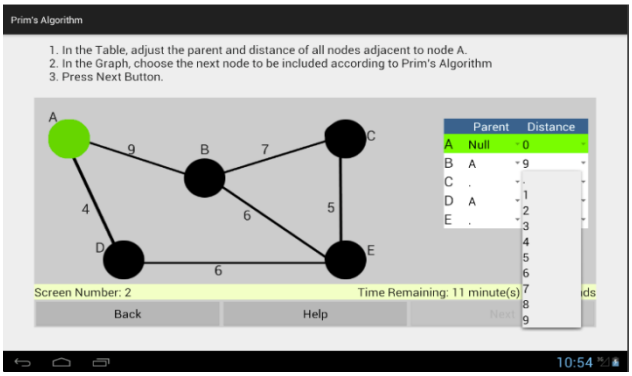
III. Prim's MST Algorithm App

In computer science, Prim's algorithm is a greedy algorithm that finds a minimum spanning tree (MST) for a weighted undirected graph. Prim's algorithm builds the tree up by adding leaves one at a time to the current tree. The concept of Prim's algorithm is difficult for learners to learn and visualize. Usually learners trace psuedocode and draw diagrams with pen to understand how the algorithm works for a particular graph. In this case-study, we describe an IE App that helps learners clarify their concept of Prim's algorithm and aseess their own understnading via visual presentation, active interaction and hands-on nature of the approach. Few researches [8,10,11] developped java-based interactive prim's exercise, where, given a graph, learners require to select (click) the next edge to be included in the MST following prim's algorithm. Our experience while teaching prim's algorithm shows that students often get confused with the part where they need to find the "shortest edge" between any vertex of already existing MST and vertices that are not part of the MST. As "shortest edge" changes with evolving MST, Having a visual graph representaton and allowing the learners to interact by clicking edges after edges is neither adquate for clarifying their concept, nor helpful for applying their skill for different scenario.

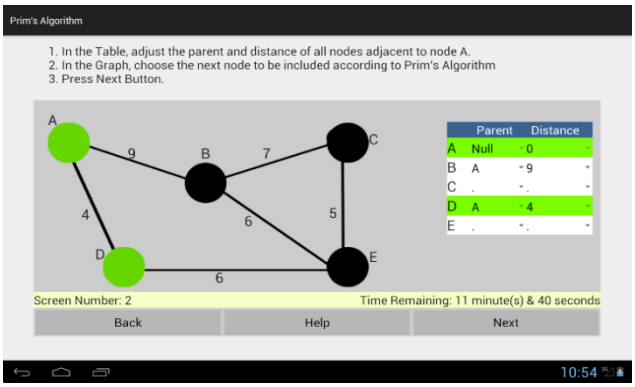
In this proposed Prim's MST App, along with the graph, we also provide a table as part of the visualization where leraners can keep track of the shortest distnace to a non-tree vertex (via updating its parent and distnace) everytime the MST evolves. Therefore, the learner interaction in this App is attained by clicking a graph node, manipulating drop-down menu items in a table, and by clicking navigation buttons (i.e Next, Back, Help, Submit Answer etc.). In step one (screen one), leraner selects the starting node of the MST as instructed by the instructor and click the "Next" button. Fig. 1(a) shows the scenario when a leraner clicks node A (starting node) in the graph and is about to press "Next" to see the next interactive screen. It should be noted that when a node is included to MST, the corresponding node in the graph and entry in the table are marked as color "Green" to signify that and their parents and distnaces can no longer be modified unless learner utilizes "Back" button. In each following step (screen), a learner must work on the table first to adjust parents and distances of non-tree vertices that are adjacent to the vertices that are part of existing MST. Fig. 1(b) shows the intermediate screen where a learner is modifying node entries in the table that are adjacent to node 'A'. In this particular instance, she has completed modifying node 'B' (parent: 'A', distnace:9) and is about to modify distnace of node 'D' by utilizing drop-down menu. Based on the information in the distance column of the table, the learner then applies prim's greedy strategy to find out the next node to be included to the MST and click that node in the graph. Fig. 1(c) shows the situation when the learner decides to include node 'D' next, correspondingly clicks that node and is about to press "Next" to see the next interactive screen. Once "Next" button is pressed, the App automatically connects the appropriate edge based on the table information to include the recently chosen node to the MST and reflects that in the next screen (Fig.



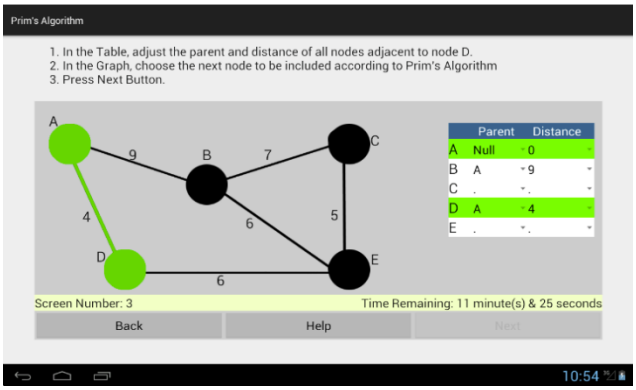
(a)



(b)



(c)



(d)

Fig. 1 Prim's MST App Interactions

1(d)), where the leraner will repeat the same procedure until all graph nodes become part of the MST.

At any step, learner can click "Back" button to undo the previous selection and repeat that particular step. Therefore, it is possible for a learner to identify a mistake at the very last step, yet start from a clean slate again by utilizing "Back" button multiple times. Pressing the "Help" button opens the

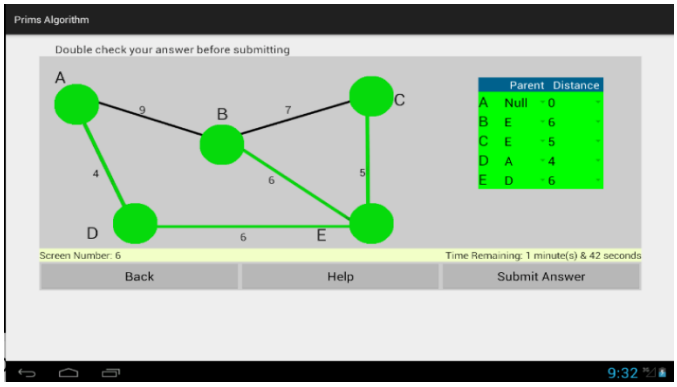


Fig. 2 Button "Submit Answer" is available as the MST is complete.

prim's algorithm in a separate window for learners to utilize during problem solving. The App captures learner activities at each step (screen) by storing the intermediate tables and graphs. The "Submit Answer" button becomes active only when the MST is complete (Fig. 2) and submits learner interactions at each step of the App to the server for automatic grading. The grading component of each App contain rubrics that can be customized to assess partially correct learner answers based on assigned weights. Each learner receives an email immediately containing the correct answer (i.e. the graph and the table) and a capture of the final screen of her submission.

During Spring of 2016, Prim's algorithm App was deployed for the first time in the Algorithm class at Winston-salem State University. During this deployment, in an 75 minutes long class, the instructor spent the first 30 minutes to explain Prim's algorithm and solving example problems. It should be noted that students were familiar with greedy strategy and the use of MRS system and Apps at that point as previous classes were used to explain them. Then in the next 10 minutes, students applied their understnading of Prim's algorithm by interacting with the mobile App and solve an exercise (Attempt #1) which is very similar to the already covered examples in terms of degree of difficulty. As all students received instant email with their answers and correct answers, instructor then spent the next 5 minutes pointing out the mistakes and clarifying the concepts more. The following 10 minutes were spent by the students solving an exercise (Attempt #2) that is of higher order of difficulty than the examples and exercises covered previously. The next 10 minutes were spent by the instructor to enhance student understnading and correct their misconceptions (if any). The last 10 minutes of the class were spent by the students to perform a difficult exercise (Attempt #3) that requires critical thinking and involves unfolding scenarios that were not explicitly taught so far.

During this particular intervention, 18 students were pressent on that class and their avergae grades for Attempt #1 was 94, for Attempt #2 was 73 and for Attempt #3 was 100. It is clearly evident from this sequence of performance data that most students solved the first exercise successfully as it was very similar to the lectured examples, however struggled for the second attempt as that was of higher order of difficulty, and finally did well for the last attempt once instructor was able to spend additional times correcting misconceptions and enhancing understanding. The automated assesment makes it possible for the instructor to offer more hand-on-activities to

the students during the classtime and utilize the extra time on correcting misconceptions and improving understnading of problems with higher order of difficulties. A pre- post- survey was conducted to understand the potential and pitfall of the proposed approach. Table 1 shows the survey data for three specific questions administered for 18 students to understand learner's attitude toward interactive classroom technology. The data reveals that once exposed to the MRS environment, substantially more students agreed that the software enhanced their learning and they are willing to utilize more of it in their future classes.

Table 1: Survey data related to student attitude toward interactive classroom.

Question	**Response**			
Q1	Strongly agree	Agree	Disagree	Strongly disagree
Pre	74%	26%		
Post	89%	11%		
Q2	Strongly agree	Agree	Disagree	Strongly disagree
Pre	47%	47%	5%	
Post	78%	17%	6%	
Q3	Strongly agree	Agree	Disagree	Strongly disagree
Pre			32%	68%
Post		6%	17%	78%

Q1: I would like my professors to integrate more interactive classroom technologies in my courses.
Q2: Using interactive classroom technology enhanced my learning.
Q3: I do NOT find the use of interactive classroom technologies to be helpful or motivating

IV. Conclusions

The research presented in this paper implements an active classroom to teach Algorithms to CS students by exploiting Mobile Response System (MRS) and interactive exercise (IE) Apps. MRS allows instructor to exploit mobile technology to offer students with interactive problem solving exercises with different learning outcomes and by receiving an immediate feedback on their performance and mental models. Initial evaluation shows enhanced learning experiences as students are applying their knowledge to solve problems with higher order of difficulty in this interactive learning environment.

Acknowledgment

This research was supported by National Science Foundation Award # 1332531.

References

[1] J. M. Braxton, W. A. Jones, A. S. Hirschy, & H. V. Hartley, "The role of active learning in college persistence", New Directions for Teaching and Learning, No. 11, pp. 71–83, 2008.
[2] M. Prince, "Does active learning work? A review of the research", Journal of Engineering Education, vol. 93(3), pp. 223–231, 2004.
[3] L. W., Anderson, B. S. Bloom, & D. R. Krathwohl, A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. New York: Longman, 2001.
[4] C. A. Shaffer, M. Cooper, & S. H. Edwards, "Algorithm Visualization: A Report on the State of the Field", In Proceedings of the 38th SIGCSE technical symposium on Computer science education, 2007.
[5] M. M. Fuad, D. Deb, J. Etim, and C. Gloster, "Using Interactive Exercise in Mobile Devices to Support Evidence-based Teaching and Learning", In Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE), pp.17-22, 2016.
[6] M. M. Fuad, D. Deb, and J. Etim, "An Evidence Based Learning and Teaching Strategy for Computer Science Classrooms and its Extension into a Mobile Classroom Response System", 14th IEEE International Conference on Advanced Learning Technologies (ICALT), Greece, pp. 149-153, 2014.
[7] http://compsci.wssu.edu/MRS
[8] L. Malmi, V. Karavirta, A. Korhonen, J. Nikander, O. Seppa¨la¨, and P. Silvasti. "Visual algorithm simulation exercise system with automatic assessment: TRAKLA2", Informatics in Education, 3(2), pp. 267–288, 2004.
[9] T. L. Naps, "Jhavé: Supporting algorithm visualization", Computer Graphics and Applications, IEEE, 25(5), pp. 49-55, 2005.
[10] S. Halim, Z. C. Koh, V. B. H. Loh, and F. Halim, "Learning algorithms with unified and interactive web-based visualization," Olympiads in Informatics, vol. 6, pp. 53–68, 2012.
[11] E. Fouh, D.A. Breakiron, S. Hamouda, M.F. Farghally, & C.A. Shaffer, "Exploring students learning behavior with an interactive eTextbook in Computer Science Courses", Computers in Human Behavior No. 41, pp. 478-485, 2014.
[12] V. Karavirta, C. A. Shaffer, "Creating Engaging Online Learning Material with the JSAV JavaScript Algorithm Visualization Library," in Learning Technologies, IEEE Transactions on, no.99, pp.1-1.

Simulation-Based Study on the Effect of Inversion Charge layer on Triple Material Double Gate Tunnel FET

Samia Safa,[1,*] Samantha Lubaba Noor,[2] and Md. Ziaur Rahman Khan[3]
[1−3]Department of Electrical and Electronic Engineering,
Bangladesh University of Engineering and Technology
Dhaka-1205, Bangladesh
*samia.safa@gmail.com

***Abstract*—The impact of inversion charge layer on triple material double gate (TMDG) Tunnel FET (TFET) has been inspected in this work by using 2D TCAD simulations. Simulation results show that band-to-band tunneling current saturates due to inversion layer formation. Inversion charges cause the pinning of surface potential and make drain current less responsive to gate voltage. The contribution of source and drain in forming the inversion layer is measured quantitatively. The influence of different device parameters on inversion layer formation has also been inspected.**

***Index Terms*—Triple material double gate (TMDG); tunnel field-effect transistor (TFET); inversion charge; inversion voltage.**

I. Introduction

Researches on novel device structures for VLSI applications are being done extensively in recent years to provide alternatives to conventional CMOS transistors due to the power scaling limitations faced by MOSFETs. Among all the alternative transistor designs, TFET is the most promising due to its low OFF-state leakage currents (I_{OFF}), better immunity to short channel effects (SCEs) because of a built-in tunnel barrier, steep subthreshold slope (SS) [1], [2] and resemblance to the conventional MOSFET. TFET is fundamentally a gated p-i-n diode which works on the principle of gate controlled band-to-band tunneling and this makes the device unique. Despite of its advantages, low on current (I_{ON}) resulting in low I_{ON}/I_{OFF} [3] and ambipolar conduction have made its application difficult in digital circuits. Numerous optimization techniques such as strain engineering [2], use of low bandgap material, heterostructures [4], gate underlap/overlap [5], use of high-*k* dielectric [6] have been proposed in recent years to boost the on current. Researches show that asymmetric source-drain doping [7], gate engineering [8], hetero gate dielectric structure [9] or short gate structures [10] are capable of suppressing the ambipolar device characteristics of TFETs. However, it has been found that inversion charge layer affects the performance of TFET devices [11], [12], [13]. In the early states of TFETs, it was claimed that formation of inversion charge layer is required to trigger band-to-band tunneling (BTBT) [14]. Recent researches show that BTBT between the source and channel can be induced without the formation of inversion charge layer [13], [15], [16]. Nevertheless, the inversion layer is formed with carriers coming from drain region at higher gate voltage [12]. Therefore, in order to optimize the performance of TFET, it is crucial to develop detail understanding on the influence of inversion charge layer on this device.

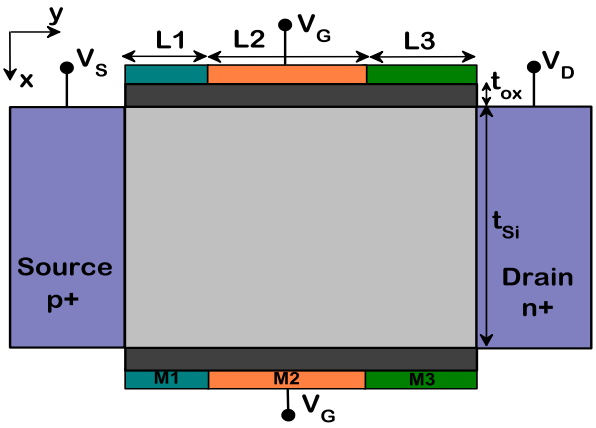


Fig. 1. Schematic representation of TMDG TFET Structure.

TMDG TFET is a recent addition to the TFETs which exhibits higher on current, lower SS, suppressed ambipolar behavior and better switching characteristics than the single metal and double metal gate structures [17], [18], [19], [20]. In this work, the impact of inversion charge layer on device performance is investigated by performing 2D TCAD simulations [21] for this structure. For quantitative analysis, the value of inversion voltage is extracted from the gate to drain capacitance and gate voltage relation. The controlling bias for inception of inversion layer has also been identified. Finally, the modulation of inversion voltage by varying gate dielectric material and the gate workfunction is studied.

II. Device Structure and Operating Principle

The schematic of n-type TMDG TFET with gate consists of three different metals M1, M2, and M3 is shown in Fig. 1. The workfunctions of three metal gates are $\phi_{M1} = 4.4eV$ (e.g. W, Ti), $\phi_{M2} = 4.6eV$ (e.g. Mo), and $\phi_{M3} = 4.3eV$ (e.g. Ag) respectively. When ϕ_{M2} is kept higher than ϕ_{M1} and ϕ_{M3}, it creates a suitable barrier and improves I_{ON}/I_{OFF} [22]. The total gate length is $L = 70nm$ with $L_1 = 20nm$, $L_2 = 30nm$, and $L_3 = 20nm$. The doping concentrations of the p-type source, lightly n-doped body and n-type drain regions are $10^{20}cm^{-3}$, $10^{16}cm^{-3}$, and $10^{19}cm^{-3}$ respectively. Ambipolar effect can be suppressed by using this type of asymmetrical doping profile [23]. The thickness of Si film is $t_{Si} = 10nm$ and oxide layer thickness is $t_{ox} = 2nm$.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

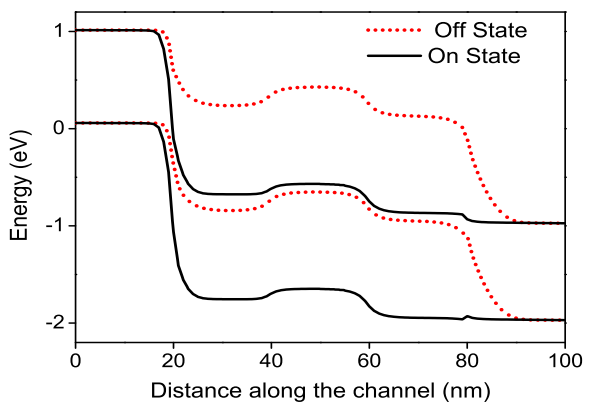


Fig. 2. Energy band diagram of TMDG TFET during off state ($V_{GS} = 0V$ and $V_{DS} = 1V$) and on state ($V_{GS} = 1V$ and $V_{DS} = 1V$).

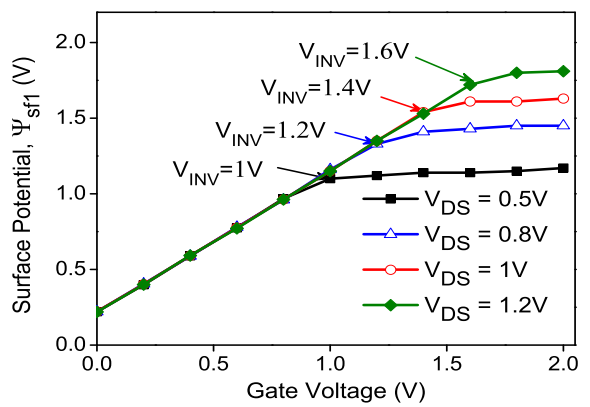


Fig. 3. Modulation of surface potential under M1 by gate voltage.

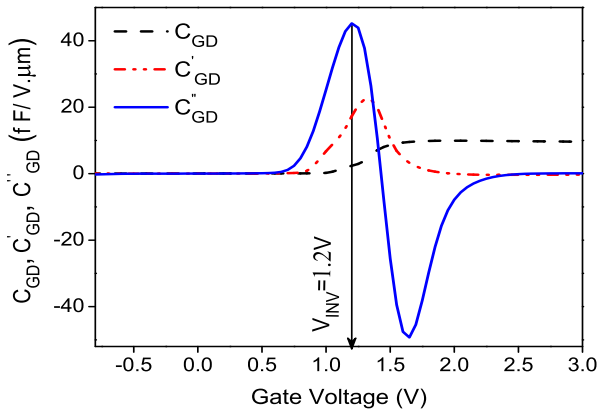


Fig. 4. Extraction of V_{INV} with $V_{DS} = 0.8V$ and $\epsilon_{ox} = 21$.

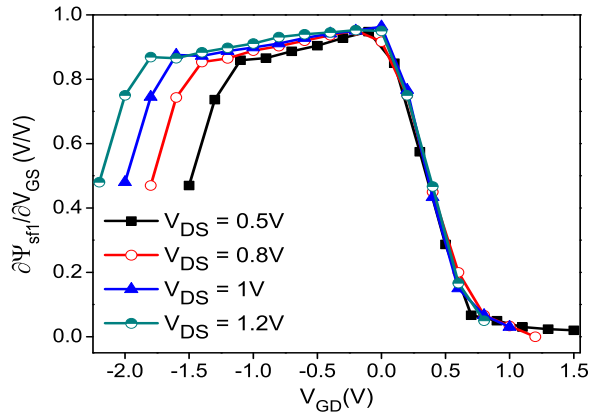


Fig. 5. $\partial\Psi_{sf1}/\partial V_{GS}$ as a function of gate to drain voltage.

The gate near the source is responsible for tunneling mechanism. Fig. 2 shows the energy band diagram of TMDG TFET during on-state and off-state. In the off-state, there is no tunneling path from the valence band of source to the conduction band of the channel and hence no band-to-band tunneling occurs. When sufficient gate voltage is applied, the conduction band of channel goes below the valence band of source, creates tunneling path and band-to-band tunneling current is induced. As gate voltage is increased upto some specific value, the tunneling barrier gets narrower and tunneling current keeps increasing.

III. Result and Discussion

The influence of inversion charge layer on TMDG TFET is studied by using Silvaco Atlas device simulator. Uniform doping profile is considered in the source, channel, and drain region and junctions were quasi-perfectly abrupt in all simulations. A fine mesh near the tunneling interface was created and carefully assembled. Kane's band-to-band tunneling is used for the simulations. The other models considered during simulation to incorporate the non-linearity of the device are: bandgap narrowing, SRH recombination, concentration dependent lifetime SRH, auger recombination, concentration dependent mobility, electric field dependent mobility.

The modulation of surface potential under gate M1 by gate voltage, V_{GS} is depicted in Fig. 3. Surface potentials, Ψ_{sf1} are taken at $1nm$ below the top oxide-silicon interface. Initially, Ψ_{sf1} increases as V_{GS} is increased. After a certain V_{GS}, the surface potential starts to saturate due to inversion charge layer formation. The inversion charges shield the channel surface from the gate electrode electronically which is the underlying reason of surface potential saturation. The saturation voltage is defined as inversion voltage, V_{INV}. From Fig. 3, it can be observed that the value of V_{INV} shifts to the right as V_{DS} is incremented. As V_{DS} is raised from $0.5V$ to $1.2V$, V_{INV} increases from $1V$ to $1.6V$. This linear increase happens due to quasi-Fermi level splitting at the surface channel. The extraction procedure of V_{INV} is shown in Fig. 4. Gate capacitance, C_{GD} increases as V_{GS} is increased and reaches oxide capacitance (C_{OX}) when inversion layer is formed. Therefore, the value of V_{GS} at which $\partial^2 C_{GD}/\partial V_{GS}^2$ is maximum corresponds to V_{INV} [13]. Fig. 4 shows that the extracted value of V_{INV} for $V_{DS} = 0.8V$ is $1.2V$ which matches with the value shown in Fig. 3.

The effect of drain voltage on inversion charge layer can be understood from Fig. 5. It can be noticed that when gate to drain voltage V_{GD} crosses $0.6V$, $\partial\Psi_{sf1}/\partial V_{GS}$ reaches almost zero which indicates the onset of Ψ_{sf1} saturation. From Fig. 5, it is evident that V_{GD} is the main biasing parameter which controls $\partial\Psi_{sf1}/\partial V_{GS}$ and eventually the formation of inversion charge layer. Therefore, it can be predicted that most of the inversion charges are thermally injected from the drain rather than band-to-band tunneling mechanism from source. Fig. 6 shows the concentration of total and drain injected inversion charges with and without considering band-to-band tunneling respectively for $V_{DS} = 0.5V$. The difference be-

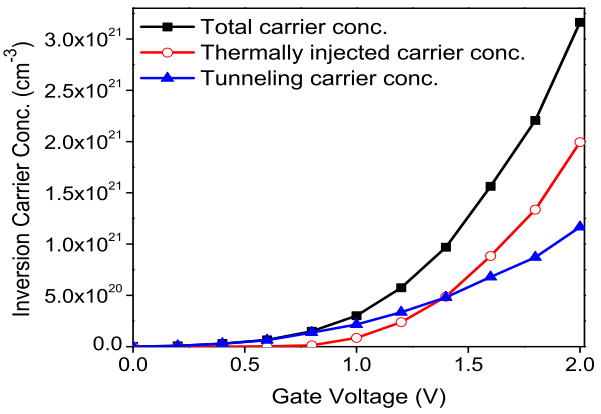


Fig. 6. Inversion carrier concentrations taken at the middle region under M1 with $V_{DS} = 0.5V$ and $\epsilon_{ox} = 21$.

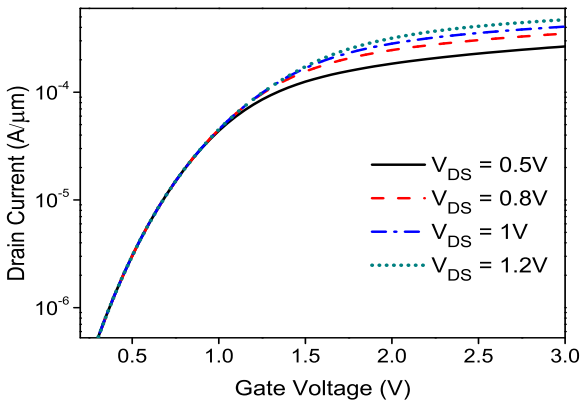


Fig. 7. Transfer characteristics of TMDG TFET for different drain bias and $\epsilon_{ox} = 21$.

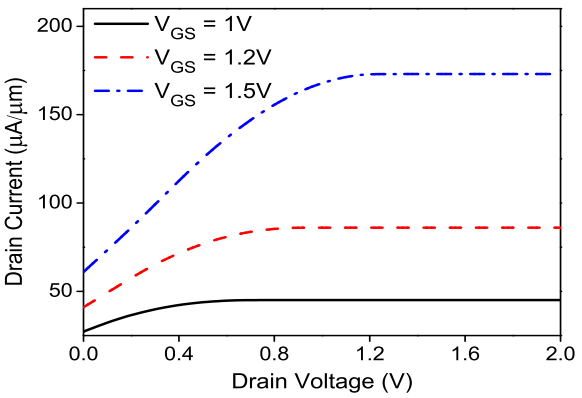


Fig. 8. Effect of drain biasing on band-to-band tunneling current under different V_{GS}.

tween these two concentrations provides the concentration of charges contributed from source by the tunneling mechanism. It can be observed that after inversion, the concentration of thermally injected charges is almost 3 times higher than the charges supplied by source. If V_{DS} is lowered, band-to-band tunneling occurs less frequently due to low Ψ_{sf1} and more charges will be thermally injected due to lower value of V_{INV}.

The influence of inversion charge layer on drain current, I_{DS} is illustrated in Fig. 7. When V_{GS} is below V_{INV}, I_{DS} increases with gate voltage and is independent of drain biasing.

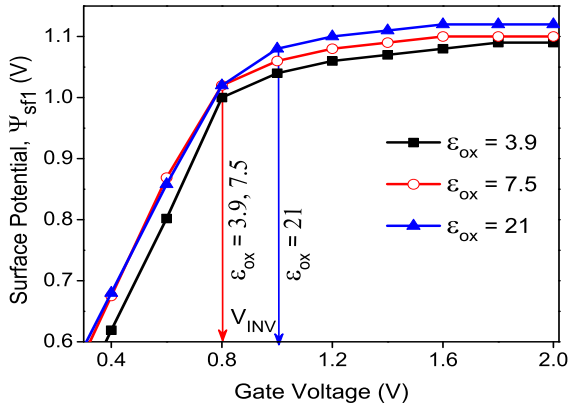


Fig. 9. Variation of Ψ_{sf1} with V_{GS} for $\epsilon_{ox} = 3.9, 7.5$, and 21 using $V_{DS} = 0.5V$.

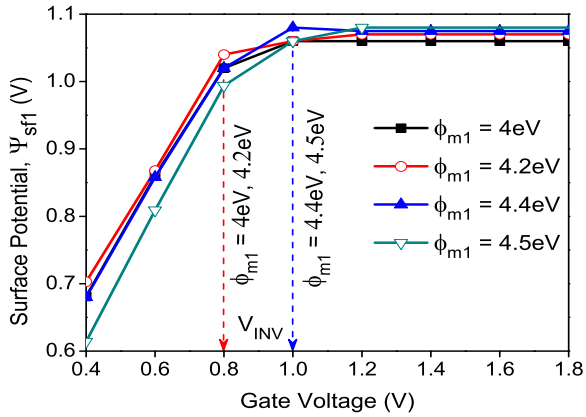


Fig. 10. Influence of workfunction variation of M1 on Ψ_{sf1} for $V_{DS} = 0.5V$ and $\epsilon_{ox} = 21$.

After inversion layer formation, I_{DS} saturates due to Ψ_{sf1} pinning and becomes dependent on V_{DS}: I_{DS} increases as V_{DS} is increased. Fig. 8 shows the dependence of I_{DS} on V_{DS}. At low V_{DS}, the current is dependent on drain voltage due to lower value of V_{INV}. When V_{DS} is high enough, V_{INV} shifts to a higher value and drain current becomes independent of drain biasing. It is also noticeable that when V_{GD} reaches $0.6V$, the drain current starts to saturate which is in agreement with previous observation. The drain current shifts upward as V_{GS} is increased due to higher Ψ_{sf1} under M1.

Fig. 9 shows the effect of different gate oxide material on V_{INV}. It can be observed that for $SiO_2(\epsilon_{ox} = 3.9)$ and $Si_3N_4(\epsilon_{ox} = 7.5)$ the value of inversion voltage is $0.8V$. When high-*k* material like $HfO_2(\epsilon_{ox} = 21)$ is used V_{INV} shifts to a higher value of $1V$. High-*k* dielectric improves the capacitive control of the gate and can provide higher surface potential than low-*k* dielectrics at same V_{GS} and thus slightly delays the saturation of Ψ_{sf1}.

The effect of different value of ϕ_{M1} on Ψ_{sf1} is depicted in Fig. 10. It can be observed that V_{INV} increases when high value of ϕ_{M1} is used. The effective gate voltage, V'_{GS} under M1 can be calculated as following

$$V'_{GS} = V_{GS} - (\phi_{M1} - \phi_s) \quad (1)$$

where ϕ_s is semiconductor workfunction. For low $\phi_{M1} = 4eV$ (e.g. Zn) and $4.2eV$ (e.g. Al), higher value of V'_{GS} can be obtained which facilitate the formation of inversion layer.

Therefore, saturation of Ψ_{sf1} is observed earlier for lower ϕ_{M1}. At the same time, lower value of ϕ_{M1} lowers the conduction band under M1. Hence, comparatively lower gate voltage is required to bring the conduction band of M1 below the valence band of source and start tunneling mechanism. For higher $\phi_{M1} = 4.4eV$ (e.g. W, Ti) and $4.5eV$ (e.g. Cu), although the value of V_{INV} increases, the on voltage, V_{ON} to start the band-to-band tunneling process also increases. Therefore, a trade-off exists between V_{INV} and V_{ON} when choosing the value of ϕ_{M1}.

IV. Conclusions

The influence of inversion charge layer on saturation of surface potential and thus I_{DS} limiting has been investigated by numerical simulation. The inversion voltage is extracted from the C_{GD} vs V_{GS} curve which can predict the onset of surface potential accurately. It is observed that after inversion, significant amount of inversion charge is contributed by the drain making V_{GD} one of the major control parameter for inversion layer formation. From simulation results, it has also been found that high-*k* dielectric material and higher metal workfunction near source can delay the inception of inversion layer formation.

Acknowledgment

The authors would like to express their gratitude to the Department of EEE, East West University, Bangladesh for the support with Atlas, Silvaco software.

References

[1] W. Y. Choi, B. -G. Park, J. D. Lee, and T. -J. K. Liu, "Tunneling field effect transistors (TFETs) with subthreshold swing (SS) less than 60mV/dec," *IEEE Electron Device Lett.*, vol. 28, pp. 743-745, Aug. 2007.

[2] T. Krishnamohan, D. Kim, S. Raghunathan, and K. Saraswat, "Double-gate strained-Ge heterostructure tunneling FET (TFET) with record high drive currents and $<<$ 60 mV/dec subthreshold slope," in *IEEE IEDM*, 2008, pp. 1-3.

[3] H. Zhao, Y. Chen, Y. Wang, F. Zhou, F. Xue, and Lee, J., "InGaAs Tunneling Field-Effect-Transistors with Atomic-Layer-Deposited Gate Oxides," *IEEE Trans. Electron Devices*, vol. 58, no. 9, pp. 2990-2995, Sept. 2011.

[4] O. M. Nayfeh, C. N. Chleirigh, J. Hennessy, L. Gomez, J. L. Hoyt, and D. A. Antoniadis, "Design of tunneling field effect transistors using strained silicon/strained germanium type II staggered heterojunctions," *IEEE Electron Device Lett.*, vol. 29, no. 9, pp. 1074-1077, Sept. 2008.

[5] A. Chattopadhyay and A. Mallik, "Impact of a Spacer Dielec-tric and a Gate Overlap/Underlap on the Device Performance of a Tunnel Field-Effect Transistor," *IEEE Trans. Electron Devices*, vol. 58, no. 3, pp.677-683, Mar. 2011.

[6] K. Boucart and A. M. Ionescu, "Double Gate Tunnel FET With High-Gate Dielectric," *IEEE Trans. Electron Devices*, vol. 54, no.7, pp. 1725-1733, July 2007.

[7] A. S. Verhulst, W. G. Vandenberghe, K. Maex, and G. Groeseneken, "Tunnel field effect transistor without gate drain overlap," *Appl. Phys. Lett.*, vol. 91, no. 5, pp. 053102-1–053102-3, 2007.

[8] N. Cui, R. Liang, J. Wang, and J. Xu, "Lateral energy band profile modulation in tunnel field effect transistors based on gate structure engineering," *AIP Adv.*, vol. 2, no.2, 022111, 2012.

[9] W. Y. Choi and W. Lee, "Hetero Gate Dielectric Tunneling Field Effect Transistors," *IEEE Trans. Electron Devices*, vol. 57, no. 9, pp. 2317-2319, Sep. 2010.

[10] J. Zhuge, A. S. Verhulst, W. G. Vandenberghe, W. Dehaene, R. Huang, Y. Wang, and G. Groeseneken, "Digital circuit analysis of short gate tunnel FETs for low voltage applications," *Semicond. Sci. Technol.*, vol. 26, no.8, pp. 085001-085008, Apr. 2011.

[11] C. Shen, S.-L. Ong, C.-H. Heng, G. S. Samudra, and Y.-C. Yeo, "A variational approach to the two-dimensional nonlinear Poissons equation for the modeling of tunneling transistors," *IEEE Electron Device Lett.*, vol. 29, no. 11, pp. 1252-1255, Nov. 2008.

[12] A. S. Verhulst, B. Soree, D. Leonelli, W. G. Vandenberghe, and G. Groeseneken, "Modeling the single-gate, double-gate, and gate-all-around tunnel field-effect transistor," *J. Appl. Phys.*, vol. 107, no. 2, pp. 024518-1–024518-8, Jan. 2010.

[13] W. Lee and W. Choi, "Influence of Inversion Layer on Tunneling Field-Effect Transistors," *IEEE Electron Device Lett.*, vol. 32, no. 9, pp. 1191-1193, Sept. 2011.

[14] K. Bhuwalka, J. Schulze, and I. Eisele, "A simulation approach to optimize the electrical parameters of a vertical tunnel FET" *IEEE Trans. Electron Devices*, vol. 52, no. 7, pp. 1541-1547, Jul. 2005.

[15] J. L. Padilla, F. Gmiz, and A. Godoy, "Impact of quantum confinement on gate threshold voltage and subthreshold swings in double-gate tunnel FETs," *IEEE Trans. Electron Devices*, vol. 59, no. 12, pp. 32053211, Dec. 2012.

[16] W. Vandenberghe, A. Verhulst, G. Groeseneken, B. Soree, and W. Magnus, "Analytical model for a tunnel field-effect transistor," in *Proc. MELECON*, May 2008, pp. 923928.

[17] R. Liang, N. Cui, M. Zhao, J. Wang, and J. Xu, "High performance Tunnel Field Effect Transistor with a Tri-Material-Gate structure," in *Proc. ISDRS*, Dec. 2011, pp. 1-2.

[18] P. Wang, Y. Zhuang, C. Li, and Z. Jiang, "Analytical modeling for double-gate TFET with tri-material gate," in *ICSICT*, 2014, pp. 1-3.

[19] D. Saraswathi, G. Balamurugan, G. L. Priya, and S. Manikandan, "A compact analytical model for 2D triple material surrounding gate nanowire tunnel field effect transistors," *Intelligent Computing and Applications*, vol. 343, pp. 325-332, Feb. 2015.

[20] B. Baral, A. K. Das, D. De, and A. Sarkar, "An analytical model of triple-material double-gate metaloxidesemiconductor field-effect transistor to suppress short-channel effects," *Int. J. Numer. Model.*, vol. 29, no. 1, pp. 4762, 2016.

[21] *ATLAS Device Simulation Software*, Silvaco Int., Santa Clara, CA, USA, 2012.

[22] N. Bagga and S. K. Sarkar, "An Analytical Model for Tunnel Barrier Modulation in Triple Metal Double Gate TFET" *IEEE Trans. Electron Devices*, vol. 62, no. 7, pp. 2136 - 2142, Jun. 2015.

[23] A. Singh, R. Narang, M. Saxena, and M. Gupta, "Ambipolar Behaviour of Tunnel Field Effect Transistor (TFET) as an Advantage for Biosensing Applications," in *Environmental Science and Engineering*, Springer Science + Business Media, 2013, pp. 171-174.

Characteristic Analysis of Triple Material Tri-Gate Junctionless Tunnel Field Effect Transistor

Monzurul Islam Dewan,[1,*] Md. Tashfiq Bin Kashem,[1] and Samia Subrina,[1] Member, *IEEE*

1. Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology
Dhaka – 1205, Bangladesh
*monzur.dewan@gmail.com

***Abstract*—We have designed a triple material tri-gate junctionless transistor (n-channel) which acts as a tunnel field effect transistor (TFET) by using the quantum tunneling phenomenon substituting thermionic emission. To turn the device ON, tunneling occurs by narrowing the barrier in between source to drain using middle gate voltage. This junctionless tunnel field effect transistor (JLTFET) operates similar to a regular TFET with superior performance without using any metallurgical junction. The device has been designed with high-k dielectric (HfO_2) and low-k air spacer to enhance better gate control. The effects of different device parameters have been studied. The device shows extremely low OFF-current (~10^{-14}A) whereas the drain current of the device rises sharply after crossing the threshold voltage, indicating a supreme ON-OFF current ratio (~10^{10}). This abrupt switching capability triggers an upgrade to digital VLSI and memory devices yet to come.**

***Index Terms*—Triple Material Tri-Gate JLTFET; TFET; ON-OFF Current Ratio; Standard BTBT; KANE Tunneling Model.**

I. Introduction

As the semiconductor industry has been forced to scale down the channel length to less than 20nm process, fabricating any types of regular DGMOSFET or TGMOSFET faces the same difficulties of fabricating super steep doping profile at any ends of the source or drain junctions. Also, these regular MOSFETs usually show highly steep subthreshold slope. Junctionless tunnel field effect transistor (JLTFET) in recent times shows most convincing performance for its similarity to gate engineered MOSFET and simplicity of fabrication [1-3].

In order to meet the demand of reduced gate length and minimum fabrication complexity & cost, JLTFET shows the most auspicious performance [4]. A highly doped JLTFET at zero gate voltage is completely depleted which results in a low OFF current. When there is a gate voltage slightly above zero (for n-channel silicon JLTFET), the channel starts entering into the flat band region producing a slightly higher ON current. But increasing the gate voltage above threshold pushes the channel into the accumulation region which produces much greater ON current. Thus the resistance of the device highly depends on the gate voltage resulting in a sufficiently huge ON-OFF current ratio [5]. JLTFET is the future of the digital VLSI i.e. it would be the ultimate future for the next generation memory devices like low power EEPROM, SRAM, DRAM or Flash drives [6-8]. Also, in case of ballistic transport at short channel lengths, JLTFET shows promising outcomes [9].

In this work, we explained the basic structure and operation of a triple material, tri-gate junctionless tunnel field effect transistor (JLTFET) with different gate bias voltages. We have used fixed bias voltages at two opposite gates and varied the mid gate voltage to turn ON the device. We introduced standard band-to-band tunneling effect initially and then incorporated other effects like bandgap narrowing, Auger recombination, etc. to examine its performance more accurately. We have also studied the effect of different parameters like channel length and thickness, oxide thickness etc. on the device performance.

II. Device Modeling

Holes and electrons follow the quantum mechanical phenomenon of transportation probability greater than zero over any energy barrier less than 10nm in thickness. This means there is always a non-zero chance for particles to tunnel from band to band if the thickness is less than 10nm. Apparently, this probability can be hastened by applying a non-zero gate voltage. This forces the valence band of the source to overlap with the conduction band of the channel, creating a passing tunnel window. Inside the TFET, electrons pass through source to channel and channel to drain specifically not using its own band but by tunneling to the other band. This actually differs from the conventional MOSFET particle transport mechanism where the thermionic emission causes particle to move in their own band. As tunnel FET uses band to band tunneling mechanism rather than thermionic emission, the device ON-OFF transition is extremely steep resulting in a lower subthreshold swing which has already been proved experimentally [10].

Fig. 1 shows the device structure and the corresponding parameters of the model are tabulated in Table 1.We have utilized lower-k dielectric, air (k=1) as spacer to separate the three gates to ensure proper isolation between them. In our device we have introduced three gates made of different materials with different gate potentials. The bias voltages used for gate 1 and gate 3 were -1V and 3V respectively. We varied the middle gate voltage V_{GS} from -1V to 3V which acted as the controlling gate to turn the device ON.

Our device is entirely N^+ doped ($N^+N^+N^+$) with doping concentration of $10^{20}cm^{-3}$.The two side gates are used to convert the JLTFET into a regular p-i-n device without any

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

physical doping by creating N^+ region and P^+ region near source and drain end by the gate voltages -1V and 3V

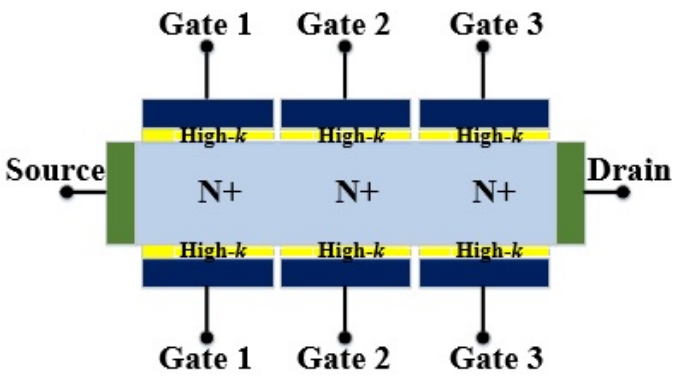


Fig. 1. A 2-D schematic structure of a triple material tri-gate JLTFET with high-k dielectric (HfO_2) gate oxide and low-k dielectric, air spacer.

Table 1. Model parameters used in this study

Parameter	Value
Gate length (Gate 1/Gate 2/Gate 3)	50nm
Channel length (L_{ch})	152nm
Channel thickness (T_{ch})	20nm
Oxide thickness (T_{ox})	3nm
Spacer gap	1nm
Channel/Source/Drain doping concentration	$10^{20}cm^{-3}$
Work function for gate 1 (copper)	4.5eV
Work function for gate 2 (silver, mid gate)	4.26eV
Work function for gate 3 (aluminum)	4.2eV

respectively. The middle gate voltage pushes the bandgap to be narrowed to a level less than 10nm so that the electrons from the source N^+ region in the valence band can pass through the narrow band into the conduction band of the P^+ drain region, resulting in a very high tunnel current (ON state).We have used high-k dielectric HfO_2 as gate oxide. OFF current has been taken at $V_{GS} = 0V$. Initially we examine the I_{DS}-V_{GS} of our device as shown in Fig. 2 by simply introducing standard band to band tunneling (BTBT) effect. Standard BTBT effect states that, if any significant electric field exists within the device , local band bending would be enough to allow electrons to tunnel by internal field emission resulting a generation of electron in the conduction band and a hole in the valence band. The tunneling generation rate [11] is reported as,

$$G_{BBT} = A.E^B.\exp\left(-C/_E\right) \quad (1)$$

Where E is the magnitude of electric field and A, B, C are device parameters. We have now introduced bandgap narrowing effect for the heavily doped region, Auger recombination effect for direct transition of carriers at high current densities, velocity saturation effect with Kane tunneling model [12] into the JLTFET to model the device more accurately and Fig. 3 shows the characteristics of the device including these effects.

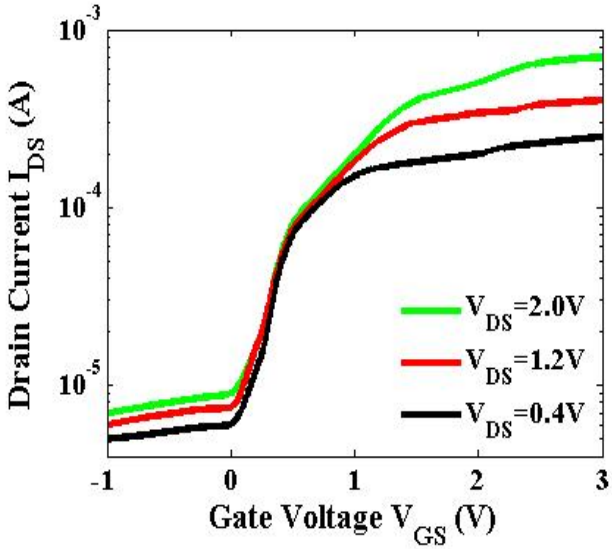


Fig. 2. Drain current variation against gate voltage for JLTFET with L_{ch}=152nm, T_{ch}=20nm and T_{ox}=3nm using standard band to band tunneling effect.

The Kane tunneling model considers the following generation rate,

$$G_{BBT_KANE} = \frac{A_{KANE}}{\sqrt{E_g}}.E^B.\exp\left(\frac{-C_{KANE}}{E}.E_g^{\,3/2}\right) \quad (2)$$

Where E_g denotes the position dependent bandgap and A_{KANE} and C_{KANE} are device parameters. For both the cases the tunnel current I_{DS} has been calculated numerically using the following expression [5],

$$I_{DS} = q.W_{Ch}\iint G_{BBT}\,dxdy \quad (3)$$

Where W_{ch} refers to the channel width. Solution to Fermi-Dirac statistics provides the intrinsic carrier concentration. For recombination modeling Auger recombination model has been included. We have also introduced Lombardi's mobility model with concentration dependent Shockley Read Hall model [13] and electric field dependent mobility in order to incorporate the effects for fixed minority carrier lifetime and temperature dependent effects respectively [14-15]. We have carried out the simulations using Silvaco TCAD device simulator [16].

III. Results and Discussion

Fig. 2 shows the variation of drain current against gate voltage for different values of V_{DS}. As the value of V_{DS} rises from 0.4V to 2V, drain current increases with the same voltage of middle gate. The middle gate voltage has been varied from $-1V$ to 3V. There is a sharp rise of drain current from 0V to 0.5V.

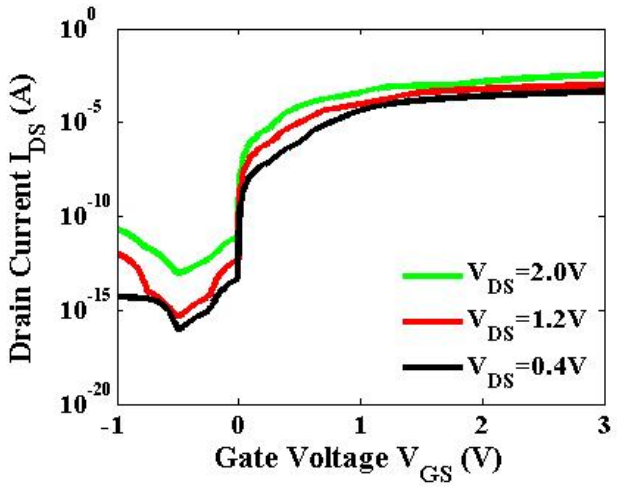


Fig. 3. Drain current variation against gate voltagefor JLTFET with L_{ch}=152nm, T_{ch}=20nm and T_{ox}=3nm using Kane's band to band tunneling effect.

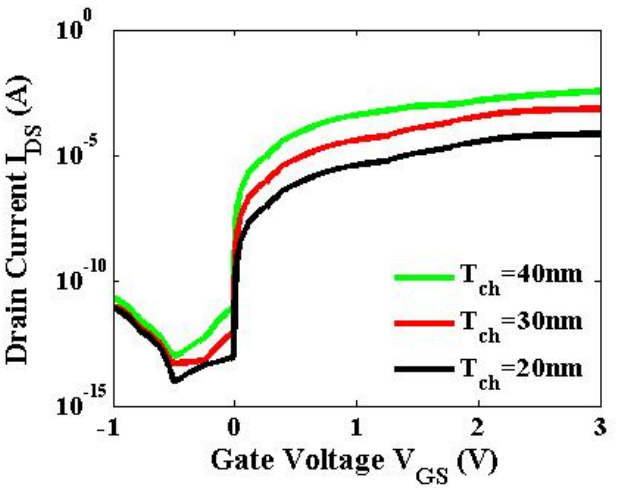


Fig. 4. Drain current as function of gate voltage for JLTFET with V_{DS}=1.5V, L_{ch}=152nm and T_{ox}=3nm for different values of channel thickness.

The drain current rises from $5x10^{-6}$A (I_{OFF}) to 10^{-4}A (I_{ON}) for $V_{DS} = 0.4V$ indicating a moderately high ON-OFF current ratio. The change in drain current is in the same region of V_{GS}(0V to 0.5V) which indicates that the band to band tunneling occurs for this transition in gate voltage, whereas the saturation current rises for rising V_{DS} indicates a much higher rate of carrier tunneling.

The characteristic of our more precise JLTFET model with Kane's tunneling effect is shown in Fig. 3. Keeping V_{DS} constant to 0.4V and varying V_{GS} from -1V to 3V, drain current changes abruptly from $8x10^{-14}$A (I_{OFF}) to 10^{-4}A (I_{ON}) which indicates a sharp rise in ON current. For the same structure, the ON-OFF current ratio is superbly higher than the previous value and even greater than any other regular MOSFETs. The reason behind this is that it uses direct and indirect transition of carriers which leads to sharp rise in generation rate, resulting in an increase in drain current. For extremely high ON-OFF current ratio, this enhanced model of the JLTFET is being used for further analysis.

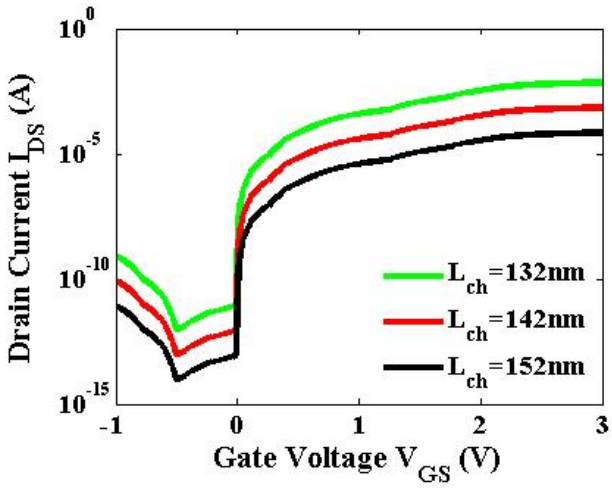


Fig. 5. Variation in Drain current against gate voltage withV_{DS}=1.5V,T_{ch}=20nm and T_{ox}=3nm showing the effect of channel length variation.

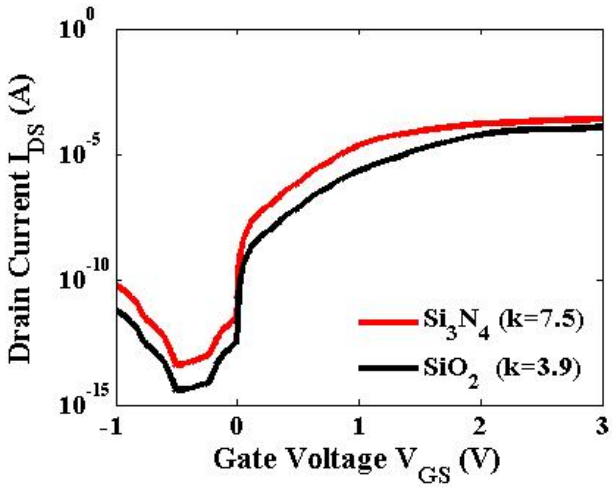


Fig. 6. Drain current as function of gate voltages withV_{DS}=1.5V,L_{ch}=152nm,T_{ch}=20nm and T_{ox}=3nm with different gate dielectrics.

Fig. 4 shows the effect of channel thickness variation. As the channel thickness increases keeping the voltage constant, both the ON and OFF current increases, indicating almost a constant ON-OFF current ratio. This is due to the thicker cross section of the channel allowing more carriers to tunnel simultaneously. Fig. 5 depicts the effect of channel length variation on device performance. Keeping the two side gate length at source and drain end constant to 50nm, we varied the middle gate length from 50nm to 30nm. As the channel length decreases, the drain current increases due to the gate dependent resistance decrement for the same V_{DS}.

Fig. 6 demonstrates the effect of gate oxide variation over the I_{DS} - V_{GS} characteristics. An improved and superior performance can be obtained using higher-k dielectric in terms of device ON current. We have used SiO_2 (k=3.9) and Si_3N_4 (k=7.5) to show the effects of using higher-k dielectric. Here the ON current increases with the increase of the dielectric constant (k). But the OFF current increases more than the ON current for greater-k dielectrics, resulting in a slightly lower ON-OFF current ratio. However a superbly low subthreshold

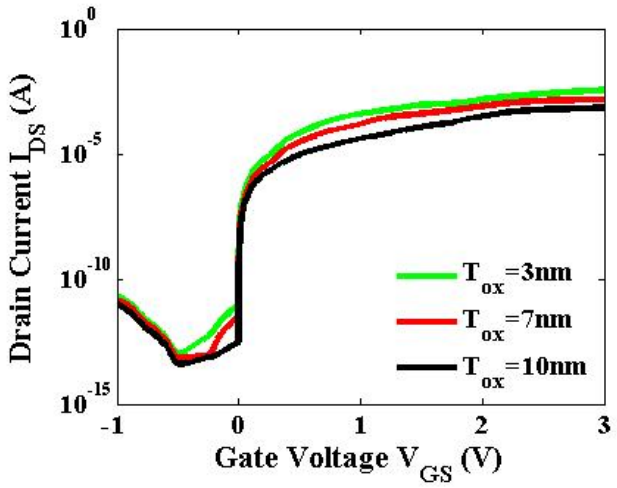


Fig. 7. Drain current variation against gate voltage for JLTFET withV_{DS}=1.5V, L_{ch}=152nm andT_{ch}=20nm using high-k dielectricHfO_2, varying oxide thickness.

slope (~47mV/decade) can be obtained using a higher-k dielectric which has already been reported [17]. Fig.7 shows the effect of gate oxide thickness on drain current. As the thickness increases, the drain ON current decreases gradually. The OFF current remains almost the same, resulting a lower ON-OFF current ratio.

IV. Conclusion

We proposed a model for triple material tri-gate junctionless tunnel field effect transistor (JLTFET) and examined the effect of band to band tunneling including the advanced tunneling effects. Also, we have studied device performance by changing the device parameters. In each case, we have calculated the ON-OFF current ratio to evaluate its performance as a digital switch. It has been found that the ON-OFF current ratio of our device is very large compared to any other regular MOSFETs. The device has very low band to band tunnel current in OFF state, resulting a small subthreshold leakage current. Detailed study on such device is thus necessary for the next generation low power memory devices.

References

[1] Q. Zhang, W. Zhao &A. Seabaugh, “Low-subthreshold-swing tunnel transistors.” *IEEE Electron Device Lett.,*27, 297-300 (2006).

[2] S. Gundapaneni, S. Ganguly & A. Kottantharayil, “Bulk planar junctionless transistor (BPJLT): an attractive device alternative for scaling”. *IEEE Electron Device Lett.,* 32(3), 261–263 (2011).

[3] S. Cho,K. R. Kim, B. G. Park, I. M. Kang, “RF performance and small-signal parameter extraction of junctionless silicon nanowire MOSFETs”. *IEEE Trans. Electron Devices,* 58(5), 1388–1396 (2011).

[4] J. P. Colinge, C. W. Lee, A. Afzalian, N. D. Akhavan, R. Yan, I. Ferain, P. Razavi, B. O’Neill, A. Blake,M. White, A. M. Kelleher, B. McCarthy,R. Murphy, ”Nanowire transistors without junctions”. Nanotechnol. 5(3), 225–229 (2010).

[5] S. Gundapaneni, M. Bajaj, R. K. Pandey, K. V. R. Murali, S. Ganguly, and A. Kottantharayil, “Effect of band-to-band tunneling on junctionless transistors,” *IEEE Trans. Electron Devices,* 59(4), 1023–1029 (2012)

[6] S. J. Choi, D. I. Moon, S. Kim, J. H. Ahn, J. S. Lee, J. Y. Kim, Y. K. Choi, “Nonvolatile memory by all-around-gate junctionless transistor composed of silicon nanowire on bulk substrate”. *IEEE Electron Device Lett.,* 32(5), 602–604 (2011).

[7] C. W. Lee, R. Yan, I. Ferain, A. Kranti, N. Akhvan, P. Razavi, R. Yu, J. Colinge, “Nanowire zero-capacitor DRAM transistors with and without junctions”. In: Proc. 10th IEEE-NANO, 242–245 (2010).

[8] A. Kranti, C. Lee, I. Ferain, R. Yan, N. Akhavan, P. Razavi, R. Yu, G. A. Armstrong, J. Colinge, “Junctionless 6T SRAM cell”. IET Electron. Lett. 46(22), 1491–1493 (2010).

[9] D. Sels, B. Sore, G. Groeseneken, “Quantum ballistic transport in the junctionless nanowire pinch-off field effect transistor”. J. Comput. Electron. 10(1/2), 216–221 (2011).

[10] S. H. Kim, H. Kam, C. Hu and T.-J. King-Liu, “Germanium-Source Tunnel Field Effect transistors with Record High ION/IOFF”, Symposium on VLSI Technology Digest of Technical Papers, 178-179 (2009)

[11] L. Gaul, S. Huber, J. Freyer, M. Claassen, "Determination of tunnel-generation rate from GaAs pin-structures", Solid-State Electronics, 34(7), 723–726 (1991)

[12] E. O. Kane, “Theory of Tunneling.”, Journal of Applied Physics, 32(1), 83-91 (1961).

[13] A. Schenk, “A model for the f eld and temperature dependence of SRH lifetimes in silicon,” Solid State Electron., 35(11), 1585–1596 (1992).

[14] Bahniman Ghosh ,Partha Mondal ,M. W. Akram ,Punyasloka Bal, and Akshay Kumar Salimath, “Hetero-gate-dielectric double gate junctionless transistor (HGJLT) with reduced band-to-band tunneling effects in subthreshold regime”, Journal of Semiconductors, 35(6), (2014).

[15] J. P. Duarte, S.-J. Choi, D.-I. Moon, and Y.-K. Choi, "Simple analyticalbulk current model for long-channel double-gate junctionless transistors",*IEEE Electron Device Lett.,* 32(6), 704—706 (2011).

[16] Atlas Silvaco Int., Santa Clara, CA.

[17] Kathy Boucart, Adrian Mihai Ionescu, “Double-Gate Tunnel FET With High-κ Gate Dielectric”, *IEEE Transactions on Electron Devices,* 54(7), (2007).

Implementation of a Three Phase Inverter for BLDC Motor Drive

Adnan Mohammad[1], Md. Anwarul Abedin[2], and Md. Ziaur Rahman Khan[1]

Department of Electrical and Electronic Engineering
[1]Bangladesh University of Engineering and Technology (BUET)
[2]Dhaka University of Engineering and Technology (DUET)
Dhaka, Bangladesh
E-mail: adnanmohammad403@gmail.com

Abstract — **In this paper a three phase inverter is designed for operating a brushless dc motor. Following the operating principle of a brushless dc motor, required switching is done by a microcontroller (arduino uno). Necessary gate driver circuit is also designed for proper operation of the inverter. The practical result matches with the simulation result and the motor runs successfully.**

Index Terms **– BLDC motor, Microcontroller and three phase inverter.**

I. Introduction

Nowadays, brushless dc motor is widely used in electric vehicles because of their efficiency, better speed versus torque characteristics, higher speed range and noiseless operation [1]. The motor has three phases and commutation is performed electronically. To run the motor a three phase inverter is necessary. Recently, many researches based on fuzzy logic, PID and FPGA control are done on BLDC motor operation [2-5].

Sensing rotor position is very important to generate proper switching sequences for the three phase inverter [6-9]. To do so, three hall sensors are provided with the BLDC motor for three phases of the motor.

Following the operation of BLDC motor a three phase inverter is simulated and designed in hardware in this paper. This drive system can change the speed of the motor as well. Microcontroller is very useful because of its several features such as faster speed, small chip size, low cost and most importantly easier troubleshooting [10]. So, in this design microcontroller is used to generate required switching for the three phase inverter.

Microcontroller senses the hall signals from motor and PWM signal which is generated from an electronic circuit. The generated pulse from microcontroller is send to the gate driver circuit which drives the three phase inverter. The three phase inverter is implemented for a 48 V and 500 W BLDC motor. Several output parameters of the inverter such as line to line voltages and phase voltages for varying switching pulse width are checked and they comply with the simulation result. The motor runs successfully by the inverter. Speeds at different output voltages of the inverter are noted from practical circuits.

II. Operating Principle of BLDC Motor and Inverter Design

Fig. 1 shows the equivalent circuit of a BLDC motor with three phase inverter. Here, L and R represent the motor inductance and resistance respectively. V_{AN}, V_{BN} and V_{CN} are back emf of BLDC motor. E is the battery and C is the input capacitor of the inverter. This inverter consists of six switches *Q1* to *Q6*. The three phase output is taken from *A, B* and *C*.

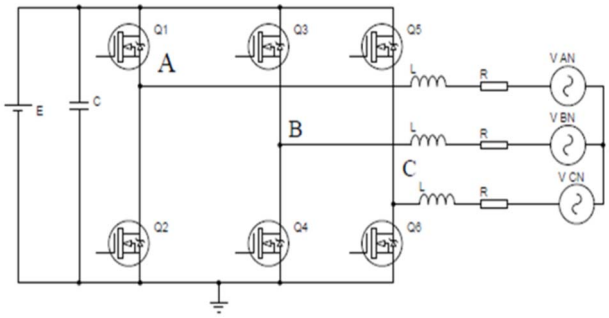


Fig. 1 Three phase inverter based BLDC motor drive.

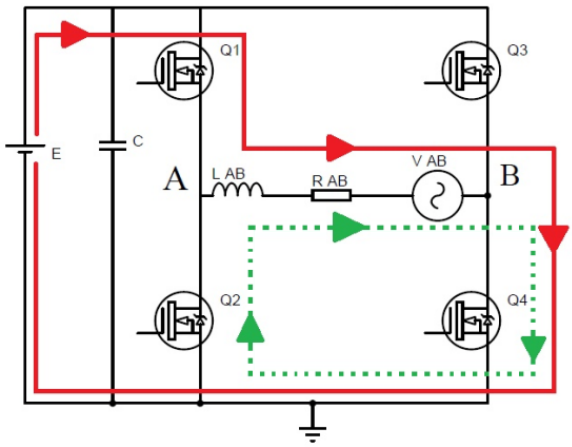


Fig. 2 The motor operation.

To run the motor the desired switching is shown in Fig. 2. Firstly, *Q1* and *Q4* are turned ON and the current flows from the battery to the motor (bold red line). During this switching the motor inductor stores energy and in the next cycle *Q1* is turned off and *Q4* is still ON. The inductor releases its energy via *Q4* and freewheeling diode of *Q2* (dotted green line). This operation is described for one state (60°). Similarly, there will be total six states (because one cycle is 360°). In the described state *Q4* is always ON and *Q1* is operated by PWM.

BLDC motor operation is dependent on the rotor position which is detected by hall sensors. Necessary switching is done based on the hall signal. Three hall sensors are provided in the motor for position sensing and they have 120° phase difference. Fig. 3 shows the response of hall sensor and back emf of BLDC motor in one electrical cycle.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

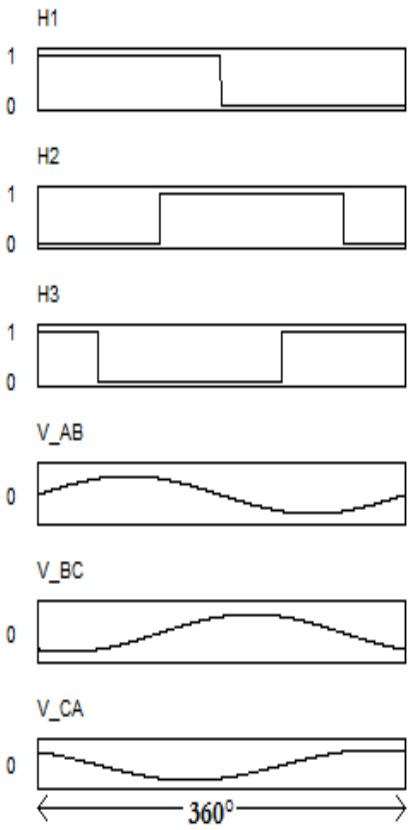


Fig. 3 Hall signal and Back emf.

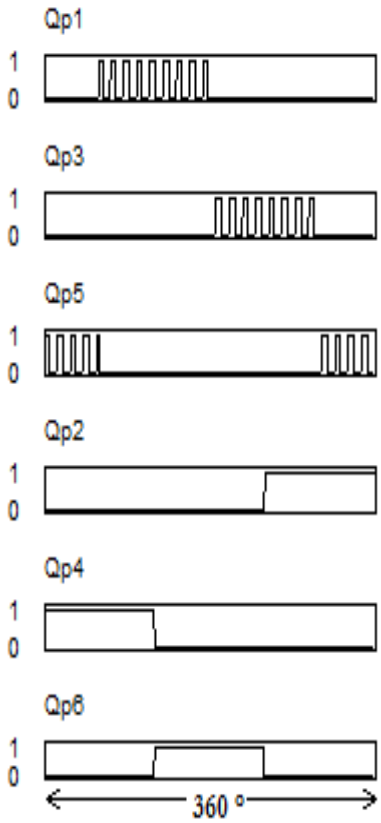


Fig. 4 The desired switching for motoring.

So, to run the motor, six switches of the three phase inverter should be operated according to the pattern shown in Fig. 4.

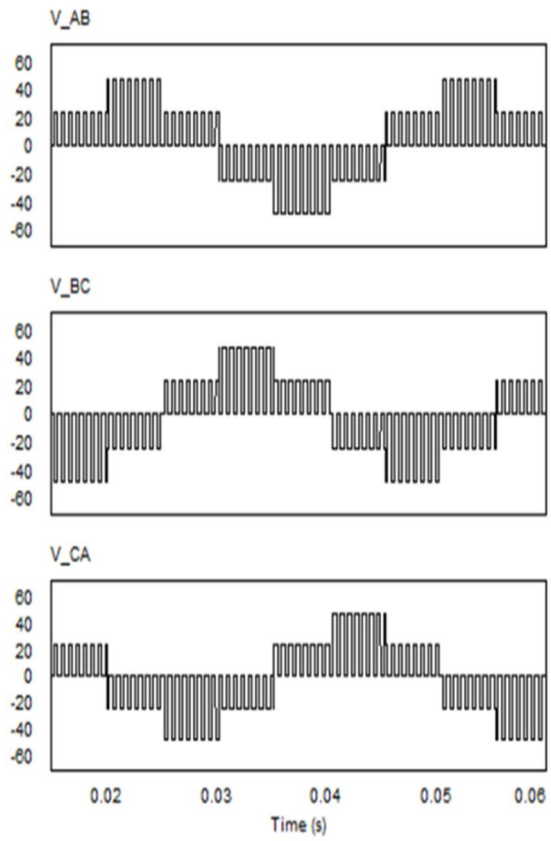


Fig. 5 Simulated line to line voltage of the inverter.

Based on the desired switching the simulation has been done in PSIM. The line to line voltages are shown in Fig 5. The output voltage of inverter can be changed varying the switching pulse width. In this simulation, the pulse frequency is set at 1.5 KHz and the duty cycle is at 50%. The output voltages have phase difference of 120°.

III. Hardware Implementation

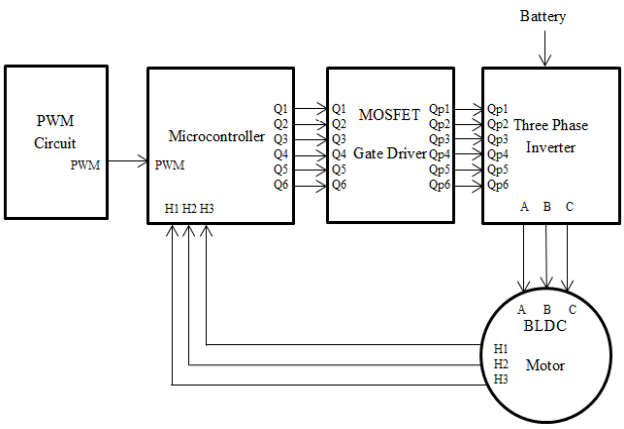
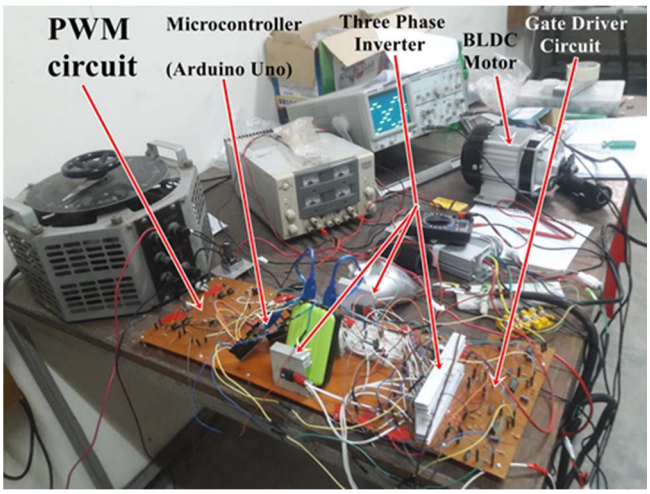
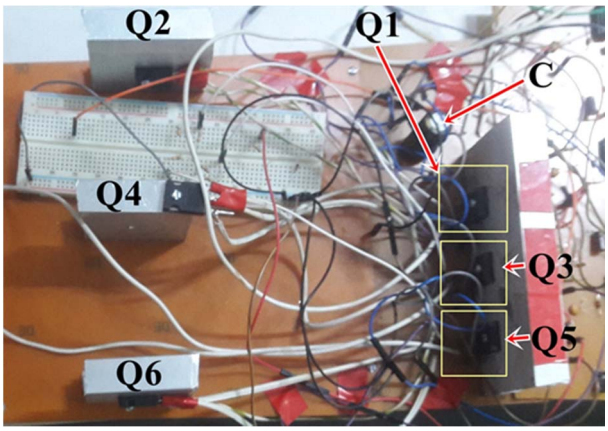


Fig. 6 Block diagram of the hardware set up.

The overall hardware set up is shown in Fig. 6. The microcontroller is coded to generate the required switching based on the inputs PWM, H1, H2 and H3. The output of the microcontroller is given to the gate driver circuit which is used to provide isolation between the control and main power circuit (inverter).



(a)



(b)
Fig. 7 The hardware setup.

Fig. 7 shows the hardware circuit which is implemented according to the Fig. 6. MOSFET (IRF250) is used as switch in this inverter and Arduino uno microcontroller is used to generate the desired switching pulses. IR2110 is used as gate driver circuit. The PWM circuit is implemented using IC741. To vary the pulse width a potentiometer is used.

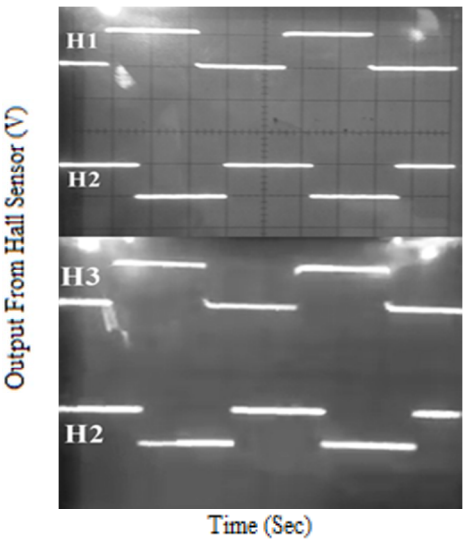


Fig. 8 Output from the hall sensors.

Response of hall sensor is observed in oscilloscope. Firstly, H1 and H2 are checked and the obtained output is according to the expected output described in Fig. 3. Here, the phase difference between H1 and H2 is 120° and the pulse voltage is 5 V. Again the channels are connected across H3 and H2. Overall, the outputs are achieved according to expected outputs.

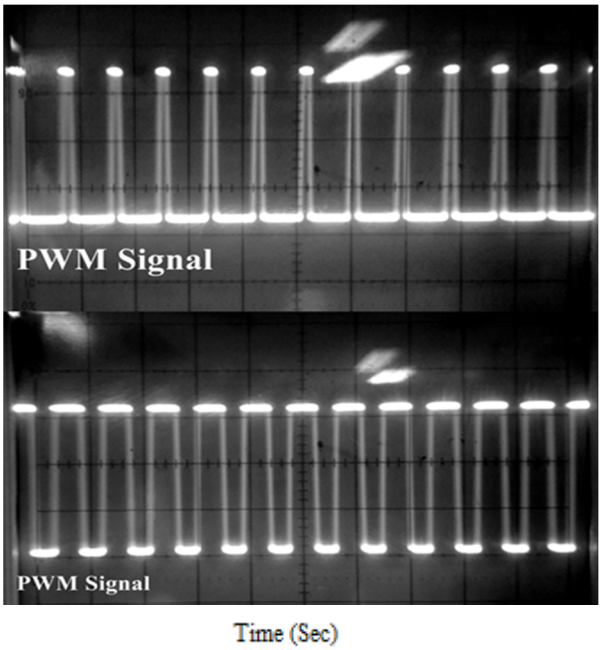


Fig. 9 Generated PWM signal.

Fig. 9 shows the generated PWM signal. Pulse of varying duty cycle (width) is obtained from the PWM circuit. The upper output has a duty cycle below 50% and the lower one has above 50%. This is done using a variable resistor.

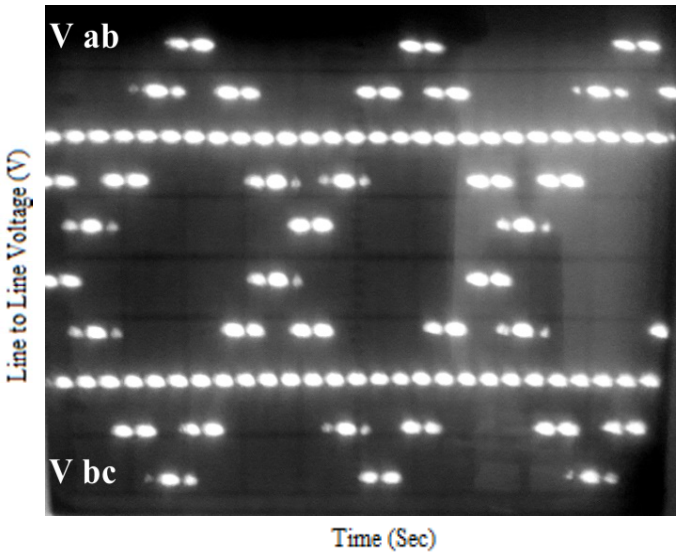


Fig. 10 Line to line output voltage of the three phase inverter (D = 50%).

The line to line voltage of the three phase inverter is shown in Fig. 10 and 11 respectively. The dc input to the inverter is 48 V and the inverter has peak voltage of 48 V. The output voltage of Fig.10 has been observed depending on a pulse signal of 50% duty cycle. The result exactly matches with the simulated result of Fig. 5. Another output voltage for 100% duty cycle is shown in Fig. 11. For all the outputs the phase difference is 120°.

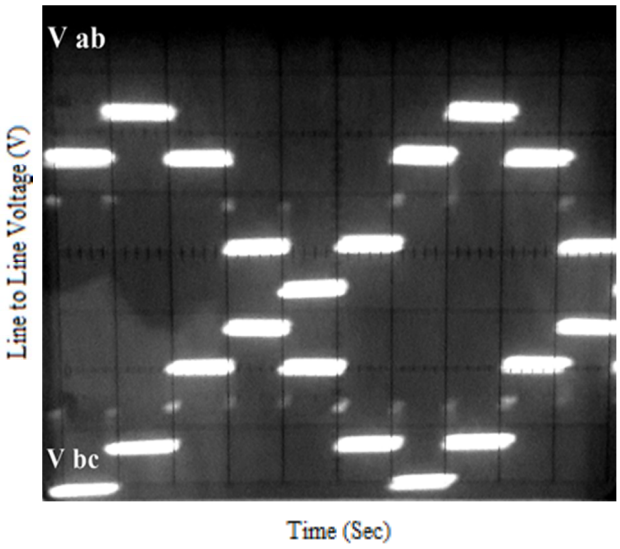


Fig. 11 Line to line output voltage of the three phase inverter (D = 100%).

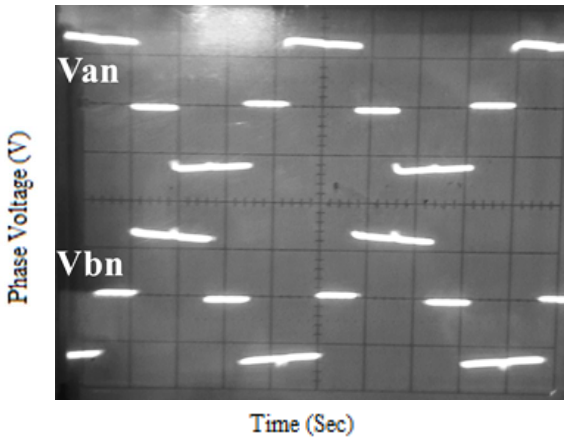


Fig. 12 Phase voltage of the three Phase inverter (D = 100%).

Fig. 12 shows the phase voltages of three phase inverter which have 120° difference and peak voltage of around 24 V. All the generated outputs are balanced.

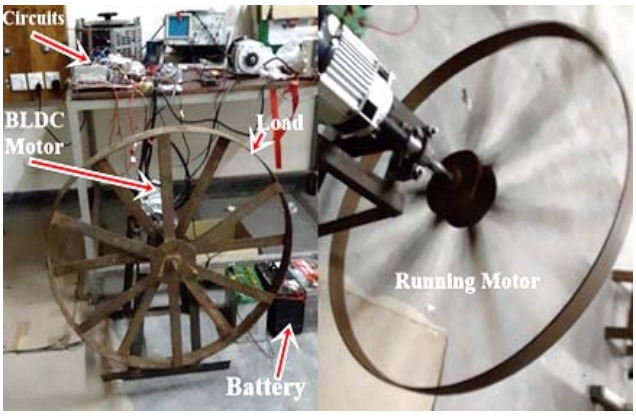


Fig. 13 Motor Running.

Fig. 13 shows the running motor. The ratings of the motor used in this research are 500 W and 48 V. The motor is connected with the inverter circuit and the pulse width is changed to vary the speed. The speed for different voltage has been noted.

TABLE I
INVERTER OUTPUT VOLTAGE AND RPM OF THE MOTOR

Duty Cycle	Inverter Output Voltage (V_{LL})	Speed
85%	31.2 V	448 rpm
70%	28.5 V	400 rpm
60%	26.2 V	367 rpm
40%	21.7 V	300 rpm
30%	18.1 V	250 rpm

Table I shows the rpm of motor for different inverter voltage. The frequency of the switching pulse was set at 1.5 KHz. The rpm has a proportional relation with the duty cycle or the applied voltage.

IV. CONCLUSIONS

In this paper a three phase inverter is designed to run a BLDC motor (500 W). The practical result matches with the simulation result and the motor runs successfully. All the line voltages & phase voltages are checked and they are like balanced three phase ac voltages having 120° phase difference. The necessary command is provided by a microcontroller (arduino uno) based on the PWM signal and hall sensor signal. The inverter is designed to run and control the speed of a BLDC motor. In future, command for different operations such as reverse rotation and regenerative braking can be implemented by providing proper instructions in microcontroller.

REFERENCES

[1] S. S. Dudhe and A. G. Thosar, "Mathematical Modelling and Simulation of Three Phase BLDC Motor using Matlab/Simulink," International Journal of Advances in Engineering & Technology, Vol. 7, Issue 5, pp. 1426-1433, Nov., 2014.

[2] M. J. Yang, H. L. Jhou, B. Y. Ma, and K. K. Shyu, "A Cost-Effective Method of Electric Brake With Energy Regeneration for Electric Vehicles," IEEE Transactions on Industrial Electronics, vol. 56, issue 6, pp. 2203 - 2212, June 2009.

[3] V. Sindhuja and G. Ranjitham, "Regenerative Braking System of Electric Vehicle Driven By BLDC Motor Using Neuro Fuzzy and PID," International Journal of Innovative Research in Science, Engineering and Technology, Vol. 3, Issue 12, pp. 17847-17854, Dec. 2014.

[4] G. Maruthaipandian, S. Ramkumar, and M. Muruganandam, "Design and Implementation of BLDC Motor Using Regenerative Braking for Electric Vehicle," International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 4, Issue 2, pp. 694-701, Feb. 2015.

[5] R. Elavarasi and P. K. SenthilKumar, "An FPGA based Regenerative Braking System of Electric Vehicle Driven by BLDC Motor," Indian Journal of Science and Technology, Vol 7(S7), pp. 1-5, Nov. 2014.

[6] A. Mohammad, M. A. Abedin, and M. Z. R. Khan, "Microcontroller Based Control System for Electric Vehicle," International Conference on Informatics, Electronics & Vision (ICIEV 2016), pp. 1-4, May 2016.

[7] A. Mohammad, and M. Z. R. Khan, "BLDC motor controller for Regenerative Braking," International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT 2015), pp. 1-6, May 2015.

[8] B. Long, S. T. Lim, J. H. Ryu, and K. T. Chong, "Energy Regenerative Braking Control of Electric Vehicles using Three-Phase Brushless Direct-Current Motors," Energies, vol. 7, pp. 99-114, January 2014.

[9] M. K. Yoong, Y. H. Gan, G. D. Gan, C. K. Leong, Z. Y. Phuan, B. K. Cheah, and K. W. Chew, "Studies of regenerative braking in electric vehicle," IEEE Conference on Sustainable Utilization and Development in Engineering and Technology (STUDENT 2010), pp. 40-45, Nov. 2010.

[10] Basics of Microcontrollers (2016, January 01). [Online]. Available: http://www.circuitstoday.com/basics-of-microcontrollers

Frequency Characterization of Bangladesh Power System

Md.Rifatul Haque, A. H. Chowdhury
Department of EEE, Bangladesh University of Engineering and Technology
Dhaka-1000, Bangladesh
E-mail: rifat8112@gmail.com

Abstract— **Due to lack of modern dynamic control over the system, the frequency of Bangladesh Power System varies significantly. This research work was started with an objective of finding common trends and behavior of the variation. Frequency data from national grid was collected for four months. They were analyzed in time domain, Fourier domain, with statistical tools and correlation tools. The results for different months were compared to identify any possible trend in the data. The observations of the study were interpreted such that suitable corrective actions could be recommended for stabilizing system frequency.**

Keywords— **Bangladesh power system (BPS), Frequency characterization, Rate of Change of Frequency (RoCoF), Distribution Function**

I. Introduction

Frequency control is vital in power system operation. One of the main concerns in power systems is that the amount of active power consumed plus losses should always be equal to the amount of the active power produced. If more power is produced than it is consumed the frequency would rise and vice versa. Even small deviations from the nominal frequency value would affect the performance of synchronous machines and other equipment. Making sure the frequency is constant, is usually the task of a transmission system operator (TSO).

Transients in power system cause serious disturbances in the operational factors of the system. The transient signals are transitory (short term duration) for which the frequency as well as varying time information are compulsory for the analytic studies. These disturbances occur for few cycles, which are difficult to be identified and classified by digital measuring and recording instrumentations.

In modern standard, a frequency tolerance of 0.05% has been accepted. For transient analysis of the time frequency signal Discrete Short Time Fourier Transform (DSTFT), Discrete Ambiguity Function (DAF), and the Discrete Wigner Distribution (DWD) are being used. P. Memon et al. [1] used Fast Fourier Transformation (FFT) to analyze the frequency contents of power system frequency variation using Simulink with wavelet analysis toolbox. J. Rivera et al. [2] demonstrated the use of two dimensional time frequency analysis as they can provide much information in the power quality analysis.

YongJune Shin et al. [3] showed that within the restriction of the uncertainty principle, time frequency analysis provides localized information of the disturbance signal as a function of time and frequency. Using simulation and actual power disturbance data, they showed that the so-called reduced 4 interference distribution (RID) appears to be most suitable for investigating the time-frequency characteristics of power disturbances.

This work identifies distinct observables, repetitions, measurement of statistical indices and performance parameters of the Bangladesh Power System (BPS) frequency data so that proper decision can be made from the analysis. The study is concerned with the analysis of frequency data in time domain, in frequency or Fourier domain, statistical analysis, different correlation and variation, and also interpretation of the results thus obtained.

II. System Description and Data Orientation

The Bangladesh power system is of moderate size with a generation capacity of about 12,365 MW [4] as of July 2016. The generation sector is coordinated by Bangladesh Power Development Board (BPDB) and its subsidiaries with dispersed generation sites located across the country. The generation is mainly dependent on natural gas as the main fuel with oil and coal playing increasingly greater parts. Only a very minute fraction of the total generation comes from renewable sources.

The National Load Dispatch Center (NLDC) operated by the national transmission system operator, the Power Grid Company of Bangladesh (PGCB), is responsible for demand-generation balancing and regulation of system frequency. Currently the governor controls of generators in the system have been disabled and there is no secondary or AGC control implemented. So NLDC attempts to maintain frequency by a very sharp demand forecasting and committing units. If frequency falls below specified limit NLDC operators cut off load automatically or command control posts to shed load. If there is steady rise in generation (hence frequency) then the control operators command a few chosen generators to decrease their production. This is a very primitive way of frequency control for a power system. Hence the frequency is very unstable. Most of the time it violates the nominal limits.

The frequency data for this study was obtained from the NLDC for the months of January, February, March and November of 2015. First three month's data were samples of 60 seconds apart and last month's data was 11 seconds apart. In this paper data from November was used for demonstration of the characterization study.

III. DATA Analysis

Data was characterized in time domain, frequency domain and with statistical and correlation tools using Matlab.

A. Analysis in Time Domain

Data of November contained 54,982 equally spaced samples of 11 seconds interval from November 18th to 24th.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

The data plotted along with its mean, 50 Hz baseline and boundaries of first and third standard deviation is shown in Fig. 1. Graphically the system is seen to have frequency above 50 Hz value most of the time. From numerical calculations, the mean and standard deviation were found out to be 50.470174 Hz and 0.459404 Hz, respectively.

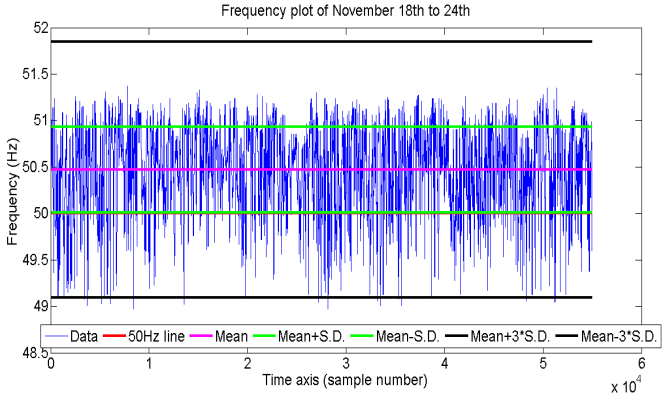


Fig. 1 Frequency data plot for Nov. 2015 with statistical parameters

Out of total 54,982 samples 45,639 samples were found to have a value greater than 50 Hz. Maximum Rate of change of frequency ($RoCoF_{Max.}$) was found out to be -0.0245 Hz/sec in the negative direction and 0.0186 Hz/sec in the positive direction respectively. This is much higher than the corresponding values of other months. The graph of first derivative presented in Fig. 2 shows the out of bound rate of change of frequencies. It is seen to have more upward deviations while fewer but prominent downward spikes.

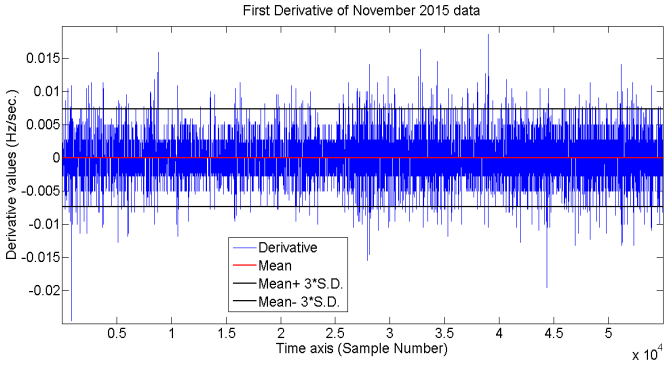


Fig. 2 First derivative of frequency with statistical limits

B. Fourier Analysis

Fast Fourier transformation of the data set, presented in Fig. 3, shows variational frequency components present within 2 Hz of the mean data. This is justified from the time domain observation of the components within specified limit.

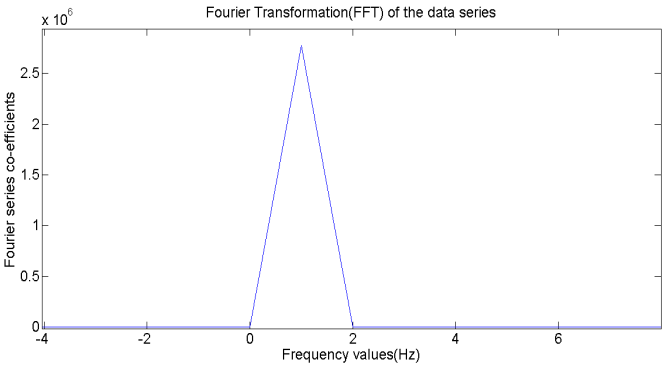


Fig. 3: Fourier transformation of the frequency data set of Nov. 2015

C. Statistical Analysis

The distribution of the data was compared to that of the standard normal distribution. While the data of first three months with lower resolution obeyed the distribution to closer extent, the data of November 2015 was found to be significantly out of boundary, as shown in Fig. 4. From the figure it can be seen that the data is positively skewed with respect to mean or base line. The data does not follow Gaussian approximation well enough.

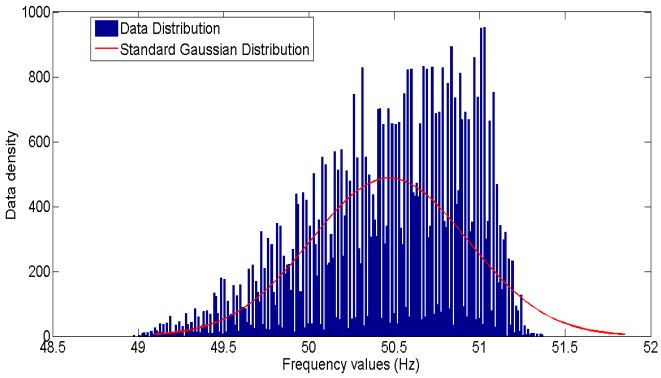


Fig. 4 Frequency data density of Nov. 2015 with Gaussian approximation

Probability density function of the frequency data was calculated and compared with that of standard normal distribution, as shown in Fig. 5, for further enhancing the anomaly. It can be seen that the data fluctuates significantly from standard normal distribution in higher frequency region but less deviation is observed in the lower frequency region, resulting in the interpretation that the data does not follow such distribution accurately.

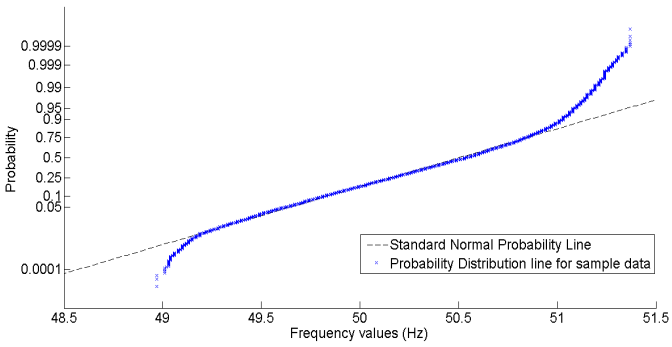


Fig. 5: Probability distribution of the frequency data with standard Normal probability line

The cumulative distribution function as shown in Fig. 6 reveals that the probability of frequency values staying at or below 50Hz is about 0.2.

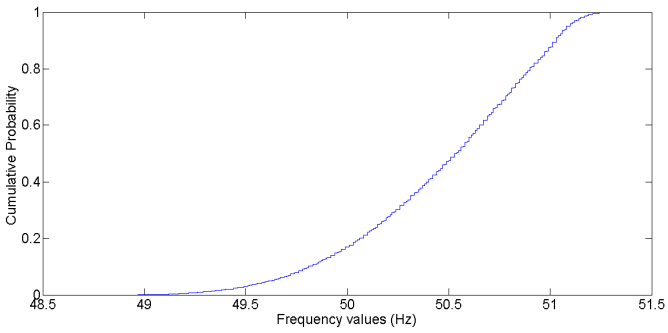


Fig.6: Cumulative probability distribution function of frequency data

D. Correlation Analysis

Correlation tools of signal analysis were used for checking if any limit violating trend was present in the data. Large data offshoots were traced by cross checking the data array with its first derivative array. Frequency samples outlying beyond three times the distance of standard deviation from mean values were identified. The common samples for a day were indexed as '1' and the rest being '0' in another matrix along a row. The matrix was summed up row wise to get another array which was plotted as shown in Fig. 7. Data of November had 7 days with approximately 7854 samples per day. The samples on which a large offshoot occurred or several smaller offshoots occurred can be identified from the vertical index of the figure. Larger index means a large frequency violation or several smaller violations on that same particular time through the data sampling period. Absence of very large summation indicates that not many correlated incidences have occurred.

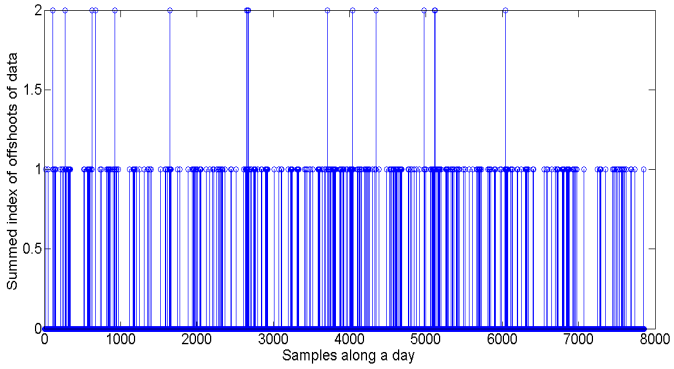


Fig. 7: Offshoot projections of frequency in Nov. 2015

The data of matrix was plotted using a three dimensional surface plot for estimating the offshoot trend in a particular month. As seen from Fig. 8, most of the normalized peaks have an index of '1'. This indicates to the presence offshoots of similar magnitude. Occurrences of similar peaks along the days indicate the possibility of a major switching event that is performed on a regular basis.

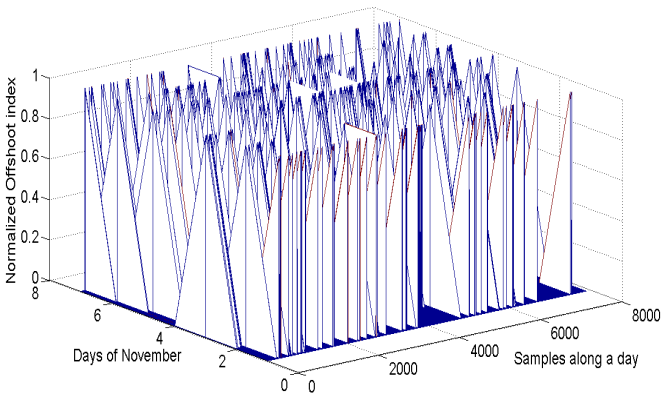


Fig. 8: Normalized peak offshoots projection of the data of Nov. 2015

Since the data resolution of November is higher than that of previous three months, it contains in-depth information. The number of sudden frequency jumps is computed with higher accuracy from data sampled at higher rate. As shown in Fig. 9, with lower data sampling rates on January, February and March, the offshoots indices peaked to 5, 4, and 3 respectively whereas it peaked to 2 for November. For an application oriented characterization study of trend analysis, a sampling rate of multiple times higher than present data is required.

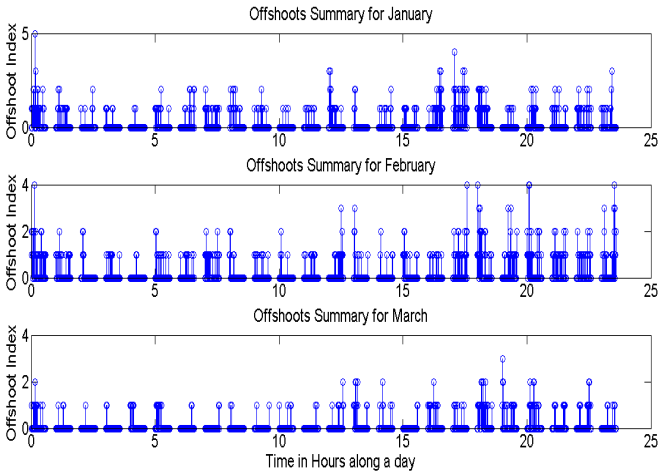


Fig. 9: Offshoot projection comparisons of three months

IV. Decision Criteria obtained from study

The observations from previous plots and summarized results can be put into defining the set points for security and protective devices. Relay trips that would occur for different cases were identified and summarized.

A. Number Of Under Frequency Trips

The number of trips an under frequency relay would make when it encounters frequency below a threshold value was calculated. It is assumed that the frequency data represents average values over an 11 second period for November 2015. Fig. 10 presents the frequency threshold vs. numbers of under frequency relay trips. It is seen that zero relay trips would occur if relay is set to trip on 49 Hz, which is much lower than standard limit. The number of trips that would occur is seen to rise exponentially with a rise in trip frequency setting.

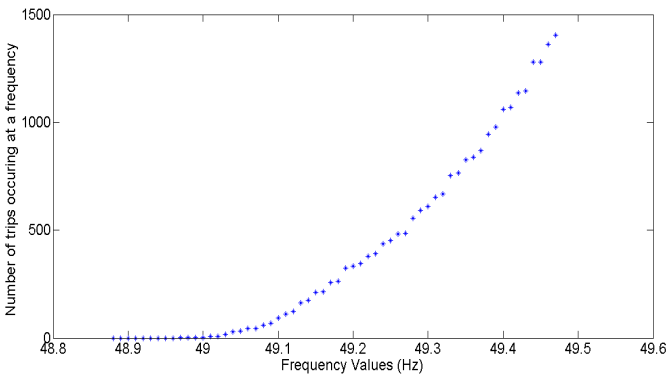


Fig. 10: Number of trips of under frequency relay with change in frequency threshold

B. Number of Trips Associated with Rate of Change of Frequency

Rate of change of frequency (RoCoF) denotes the system inertia. Sometimes due to large faults, frequency changes at a faster rate which induces oscillatory phenomena in rotor angle of the alternators connected to system grid. So relays are often set at RoCoF. The number of trips were counted if RoCoF exceeded a particular value and plotted, as shown in Fig. 11, for the November frequency data. Zero trips occur at a rate of about 0.049 Hz/sec., which is very high. The number of trips rises linearly with a low rate down to 0.01 Hz/sec.

and then rises again at a much steeper rate. The graphs of other month were verified to be of similar characteristics.

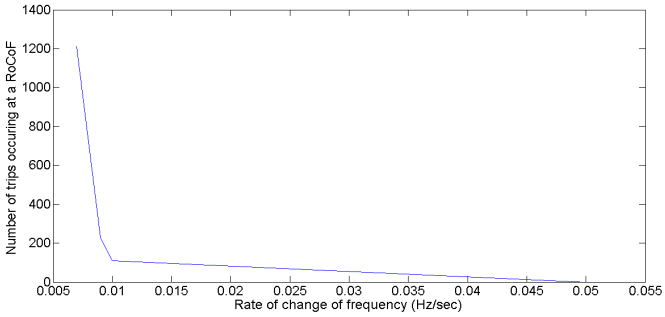


Fig. 11: Number of trips for particular RoCoF

The trips for different months were compared and presented in Table I to identify the effect of sampling rate on the RoCoF. It is seen that the calculated RoCoF and the number of trips both rise tremendously with higher data resolution. It is to be noted that the frequency data available had a very low sampling rate; hence the RoCoF and number of trips obtained would be different in reality. Moreover, different relay may calculate the RoCoF differently, and their trip logic may also differ.

It is also seen from Table I that the number of trips increased for decreasing RoCoF for every month as a verification of results displayed in Fig. 11.

TABLE I
COMPARISON OF NUMBER OF TRIPS FOR DIFFERENT ROCOF

RoCoF (Hz/sec)	Number of trips			
	January	February	March	November
0.01	3	5	6	110
0.009	6	8	7	229
0.007	90	45	20	1210

C. Number of Trips Associated with Both Under frequency and Rate of Change of Frequency

Analysis was conducted varying both the RoCoF and frequency threshold. At first the rates were changed for a particular frequency and the rate at which no trip encountered was identified and recorded. Then the frequency was decreased by a prefixed amount to a new set point and the calculations were repeated. Results are summarized in Table II. It is seen from the table that rates become constant for any particular month. For February, the number of zero trips was first encountered at a very low frequency value of 48.78 Hz. The first zero trips for March occurred at a RoCoF of 0.009 Hz/sec. on any low frequency. So if relay set point is to be decided based on the data supplied then it is to be set at 48.78 Hz at a RoCoF value greater than 0.009 Hz/sec.

TABLE II
COMPARISON OF NUMBER OF FIRST ZERO TRIP FOR DIFFERENT MONTHS BY VARYING BOTH FREQUENCY AND ROCOF

Zero Trip Frequency (Hz)	January	February	March	November
	RoCoF (Hz/sec)	RoCoF (Hz/sec)	RoCoF (Hz/sec)	RoCoF (Hz/sec)
48.90	0.0025	N/A	0.009	0.0025
48.89	0.0025	N/A	0.009	0.0025
48.88	0.0025	N/A	0.009	0.0025
48.87	0.0025	N/A	0.009	0.0025
48.86	0.0025	N/A	0.009	0.0025
48.85	0.0025	N/A	0.009	0.0025
48.84	0.0025	N/A	0.009	0.0025
48.83	0.0025	N/A	0.009	0.0025
48.82	0.0025	N/A	0.009	0.0025
48.81	0.0025	N/A	0.009	0.0025
48.80	0.0025	N/A	0.009	0.0025
48.79	0.0025	N/A	0.009	0.0025
48.78	0.0025	0.0025	0.009	0.0025

V. CONCLUSION

This paper presents frequency characterization of the Bangladesh Power System (BPS). Frequency data for four different months were used for the purpose. Despite being a system of moderate inertia, BPS has its supply frequency varying at a high rate. The data used was seen to vary mostly within ±1.5 Hz from the median, with little correlation and the mean value was considerably above 50 Hz. RoCoF of data was above acceptable standard and density of data was seen to be positively skewing for higher sampling rate. The decision of operating the system with a relay trip frequency that is lower than the standard and setting contingent value of RoCoF to be high was justified from the observations. Similar study with frequency data having higher sampling rate will yield information that can be used to avert any unexpected system contingency and make the system more stable.

REFERENCES

[1] Aslam P. Memon, M. Aslam Uqaili, Zubair A. Memon, Asif A. Akhund, "*Time-Frequency Analysis Techniques for Detection of Power System Transient Disturbances*", *International Journal of Emerging Trends in Electrical and Electronics* (IJETEE – ISSN: 2320-9569) Vol. 9, Issue. 1, Nov. 2013.

[2] Ivan J. Rivera, Ana B. Ramirez, Domingo Rodriguez, "*A Time-Frequency Signal Analysis System for Power Quality assessment*", 48th Midwest Symposium on Circuits and Systems, 2005, pp 1677 - 1680 Vol. 2, Aug 2005.

[3] YongJune Shin, Antony C. Parsons, Edward J. Powers, W.M. Grady, "*Time-Frequency Analysis of Power System Disturbance Signals for Power Quality*", Power Engineering Society Summer Meeting, IEEE, Volume 1, pp 401-407, Jul 1999.

[4] Website of Bangladesh power development board : http://www.bpdb.gov.bd /bpdb/

[5] Website of Power Grid Company of Bangladesh limited : http://www.pgcb.org.bd/PGCB/?a=pages/electricity_transmission

Wind Farm Transient Stability improvement by Fuzzy Logic Controlled Series Variable Resistor

A B M Samsuzzaman[1,*], Rabayet Sadnan[2] and A.S.M. Jahid Hasan[2]
[1]LHP Engineering Solution, Columbus, Indiana, USA
[2]Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology (BUET). Dhaka, Bangladesh
*samsuzzaman.abm.eee@gmail.com

***Abstract*—Wind farms with doubly fed induction generators (DFIG) are susceptible to the grid faults. Without any auxiliary device, they are subjected to high disturbance at the event of grid fault. The wind farm may be forced to unstable condition which may eventually lead to instability of the whole grid as the wind power penetration is higher these days. So it is very important to ensure stable transient performance of the wind farms. Application of a fuzzy logic controlled series variable resistor (FLC-SVR) is proposed in this work for transient stability enhancement of the grid connected wind farms. Transient performance is analyzed for the most severe three-line-to-ground (3LG) fault applied to a test system through simulations in Matlab/Simulink software. To emphasis the usefulness of the proposed FLC-SVR, its performance is compared against series dynamic braking resistor (SDBR). It was found that the FLC-SVR is very skillful in improving wind farm transient stability and shows superior performance compared to SDBR.**

***Index Terms*—Doubly fed induction generator (DFIG), fuzzy logic controller (FLC), series dynamic braking resistor (SDBR), series variable resistor (SVR), transient stability.**

I. Introduction

Electricity generation from wind farms are getting popular around the world. Most of the present days wind farms are equipped with doubly fed induction generator (DFIG) based variable speed wind farms. DFIGs account for 50% of the wind energy market and this trend will continue to expand in future [1]. Because more and more countries are getting committed to make the world more sustainable by producing some portion of their total electricity from the wind [2]. But these grid connected DFIG based wind farms are susceptible to the grid faults and some auxiliary devices are required to ensure their stable operation and overall grid stability [1].

To resolve this issue some solutions are available in literature. Some propose rotor side converter (RSC) and grid side converter (GSC) control technique for example in [3]. But they cannot handle severe fault as the RSC and GSC have partial rating. Others propose the use of some auxiliary devices like crowbar [4], series dynamic braking resistor(SDBR) [5], fault current limiters [6], [7], static synchronous compensator (STATCOM) [8], flywheel energy storage (FES) [9], superconducting magnetic energy storage (SMES) [10], superconducting fault current limiter (SFCL) [11].

Due to the nonlinear nature of the power system, a fuzzy logic controller (FLC) is proposed to control a series variable resistor (SVR) for the transient stability enhance application of grid connected wind farm. The performance of the FLC-SVR is investigated in a test system and its performance is compared with SDBR. From the simulations executed in Matlab/simulink, it was found that the FLC-SVR is very useful device for transient stability improvement of the DFIG based variable speed wind farm and performs superior to the SDBR.

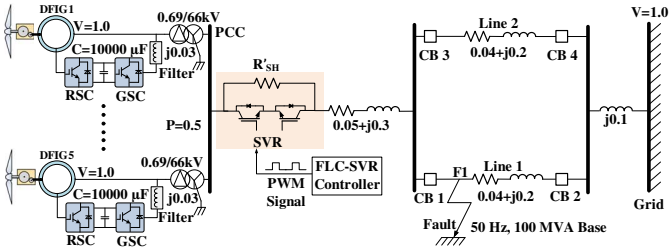


Fig. 1: Model system with the DFIG based wind farm

II. Test Power System

The power system model shown in Fig. 1 has been used for the simulation analysis. The system consists of a 50 MVA DFIG based wind farm, which is delivering power to an infinite bus through transformers and transmission line with two circuits, respectively. There is also a single circuit local transmission line in between the wind farm and the two circuit transmission line. As the wind farms, in reality are composed of many generator turbine systems, the farm is built form five individual turbines each representing six 1.67 MVA wind turbines. This arrangement of wind farm is considered as it captures much more dynamic situation compared to a wind farm made of an single individual turbine. The power system model data is given in Fig. 1.

III. Modeling of Wind farm

Wind turbine and wind generator modeling is given below,

A. Wind Turbine Modeling

The wind power captured by a wind turbine as the mechanical power can be calculated as follows [12],

$$P_w = \frac{1}{2}\pi\rho R^2 V_w^3 C_p(\lambda,\beta) \tag{1}$$

where P_w is the extracted power from the wind, ρ is the air density, R is the blade radius, V_w is the wind velocity, and C_p is the power coefficient which is a function of both the tip speed ratio λ and the blade pitch angle β. C_p is given by [13],

$$C_p(\lambda,\beta) = 0.5176(\frac{116}{\lambda_i} - 0.4\beta - 5)e^{-\frac{21}{\lambda_i}} + 0.0068\lambda \tag{2}$$

$$\lambda = \frac{\omega_r R}{V_w} \tag{3}$$

$$\frac{1}{\lambda_i} = \frac{1}{\lambda + 0.08\beta} - \frac{0.035}{\beta^3 + 1} \tag{4}$$

where ω_r is the angular mechanical speed. The two mass reduced order equivalent mass-spring-damper model of the wind turbine attached to the rotor shaft of the DFIG through an equivalent gearbox is considered. To see the related equations of motion, the reader is referred to [13].

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

TABLE I: WIND GENERATOR AND TURBINE DATA

Characteristic	Value
Nominal power	1.67 MVA
Rated voltage	690 V
Stator to rotor turns ratio	0.3
Rated frequency	50 Hz
Stator resistance (Rs)	0.012 pu
Stator inductance (Ls)	0.15 pu (referred to stator)
Rotor resistance (Rr)	0.012 pu
Rotor reactance (Lr)	0.15 pu (referred to stator)
Mutual inductance (Lm)	4 pu
Inertia Constant (H)	0.0685 pu
DC link rated voltage (E_{dc})	1200 V
Turbine inertia constant	4.32 s
Shaft spring constant	1.11 pu
Shaft mutual damping	1.5 pu

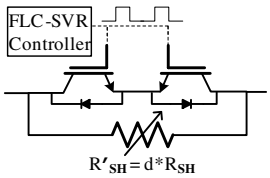
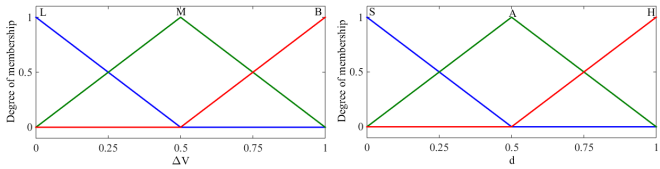


Fig. 2: Per phase series variable resistor



(a) (b)

Fig. 3: Membership functions (a) input ΔV, and (b) output d

TABLE II: FUZZY RULE TABLE

PCC Voltage Deviation ΔV	Duty Cycle d
L	S
M	A
B	H

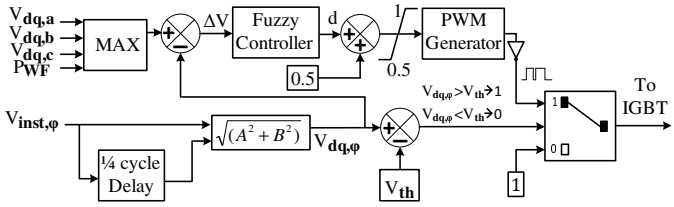


Fig. 4: Per phase SVR controller

B. DFIG Modeling

The modeling technique of DFIG is adopted and utilized form [14] in this work. It is the Park's transformation model. A synchronously rotating d-q reference frame is used. The d-axis aligned with the stator flux orientation. A decoupled control between the rotor excitation current and the electrical torque is thus achieved. Hence, the reference frame is rotating with the equal speed as the stator flux. The DFIG parameters used in this work are listed in Table I [15]. Some linear controllers are used to control the rotor side converter (RSC) and the grid side converter (GSC) which is available in [16].

IV. FLC-SVR

The proposed FLC-SVR is illustrated below.

A. SVR

The SVR topology is presented in Fig. 2. A shunt resistor R_{SH} is placed in parallel with the IGBT switches. During normal operation, the IGBT switches remain closed and the full line current flows through them. When fault is detected in the system, the current is bypassed to the resistor by turning on and off the IGBT. The IGBT switching facilitated by pulse width modulation (PWM) block. The PWM duty ratio d is used to control the IGBT turn off period. d=T_{off}/T_c, T_{off} is the time IGBT is in off position within the period of the PWM carrier frequency T_c. A dynamic variable resistance value $R'_{sh} = d * R_{SH}$ is offered to the fault to limit the fault current as per the voltage sag ΔV. Hence the SVR maintains the voltage profile. When the normal operating scenario is restored followed by the fault recovery, the IGBT switches are forcefully closed to carry the normal operating line current. Value of d is determined by FLC which is described below.

B. FLC

1) Fuzzyfication: Deviation of the wind farm terminal voltage, ΔV, and the duty ratio d of the PWM block, are selected as the input and output, respectively, for the design of the proposed fuzzy logic controller. Triangular membership functions for ΔV and d are shown in Fig. 3a and Fig. 3b, in which the linguistic variables L, M, B, S, A, and H stand for Low, Medium, Big, Small, Average and High, respectively. Adjustment has been made to the membership functions to obtain the best system performance. The equation of the triangular membership function used to determine the grade of membership values is as follows [17],

$$\mu_i(\Delta V) = \frac{1}{b}(b - |\Delta V - a|) \tag{5}$$

where $\mu_i(\Delta V)$ is the value of grade of membership, 'b' is the width, 'a' is the coordinate of the point at which the grade of membership is 1.

2) Fuzzy Rule Base: The control rules of the FLC controller are determined from the viewpoint of practical system operation and best system performance and are shown in Table II.

3) Fuzzy Inference: Using Mamdani's method [17] the degree of conformity, W_i, of each fuzzy rule is as follows:

$$W_i = \mu_i(\Delta V_{PCC}) \tag{6}$$

where i is rule number.

4) Defuzzification: The center-of-area method is used to determine the output crispy value (i.e., the chopper duty ratio, d). This is given by the following expression:

$$d = \sum W_i C_i / \sum W_i \tag{7}$$

where C_i is the value of d expressed in terms of linguistic variables in the fuzzy rule table.

C. SVR Controller

The SVR controller used per phase is shown in Fig. 4. $V_{dq,\varphi}$ is compared with a threshold voltage level V_{th}=0.9 pu, to detect the fault. During normal operation $V_{dq,\varphi}$>V_{th} and the IGBT is forcefully kept closed and when $V_{dq,\varphi}$<V_{th}, that is when a fault is detected in the system, the switching signal

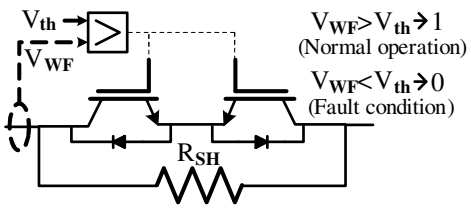


Fig. 5: Per phase SDBR with controller

from the PWM is sent to the IGBT. Switching duty ratio d is governed by the fuzzy logic controller. This duty ratio d makes the SVR dynamically changing the parallel resistance according to the voltage sag severity.

D. Design Considerations

The SVR contributes to keeping a better transient stability by consuming the excess active power during fault. The voltage at the wind farm terminal is increased and the oscillating electrical torque is subsided. The maximum allowable resistance R_{TH}, offered to the wind farm during the fault can be approximated from R_{TH}=V_{PCC}^2/P_{WF} where R_{TH} is given as R_{TH}=R_{SH}+R_L+R_T. Here, R_{SH}, R_L, R_T are the value of the shunt bypass resistance used in the SVR, the line resistance and the resistance of transformers, respectively. The effective resistance R'_{SH} of the SVR bypass path is governed by R'_{SH}=d*R_{SH} that dynamically varies with the severity of the fault.

E. Series Dynamic Braking Resistor

The SDBR and its conventional controller [7] used here is shown in Fig. 5. The wind farm terminal voltage V_{WF} is compared with the threshold voltage level V_{th}=0.9 pu. If V_{WF} is greater than V_{th}, there is no fault in the system and the IGBT switch are remained closed. At fault, V_{WF} becomes less than V_{th} and the the IGBT switch are opened to bypass the current through the resistor R_{SH} till the fault is recovered. When the fault is recovered and V_{WF} crosses V_{th}, the IGBT switches are closed and normal operation is resumed. To keep the comparison between the FLC-SVR and the SDBR, same value of R_{TH} is used in both of them.

V. Results

To evaluate the FLC-SVR performance in transient stability enhancement of wind farm, its performance is compared to the SDBR. The simulation analysis is carried out on the system showed in Fig. 1. Simulation results are given with discussion below.

A. Simulation Considerations

For the transient analysis purpose, a fixed wind speed of 15 m/s is considered. The wind speed is assumed constant as the duration of the fault is too short for the wind speed to influence the transient performance. The most severe three-line-to-ground (3LG) temporary fault is applied. At point F1, fault is applied at 0.1 s. Circuit breakers on the faulted line open at 0.2 s and reclose successfully at 1.2 s. The simulation time step used for the study is 20 μs. In the following sections, system responses are shown for three different cases considered in the simulations. They are as follows:

1) Case A : with no auxiliary device
2) Case B : with SDBR
3) Case C : with FLC-SVR

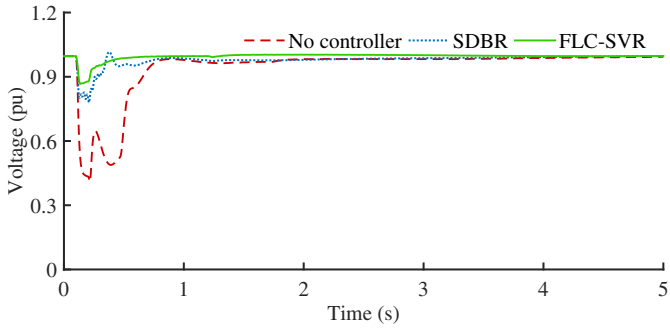


Fig. 6: Wind farm voltage response for 3LG fault.

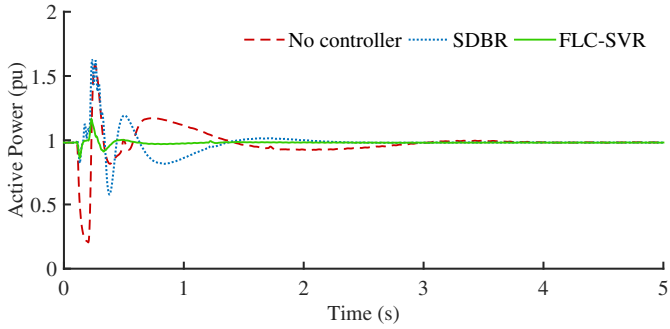


Fig. 7: Wind farm active power response for 3LG fault.

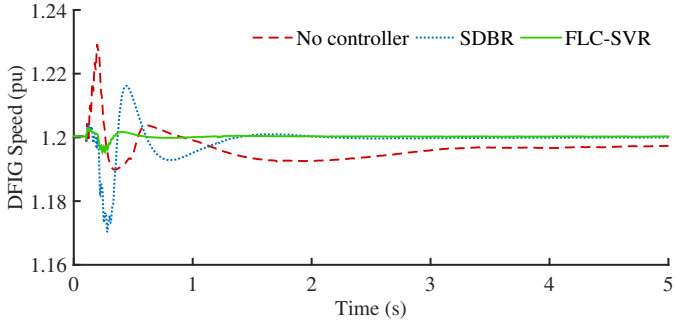


Fig. 8: DFIG speed response for 3LG fault.

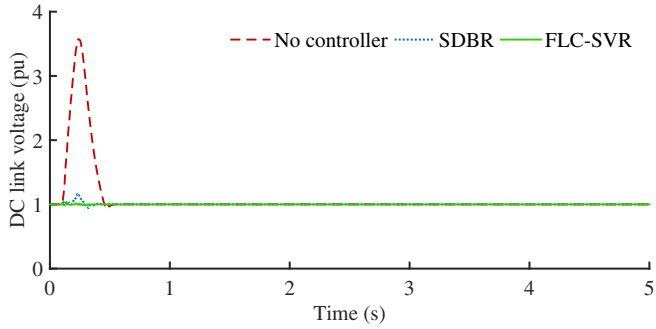


Fig. 9: DFIG DC link voltage response for 3LG fault.

B. Transient Stability Improvement for 3LG fault

Fig. 6 shows the wind farm terminal voltage response. Without any auxiliary device the voltage goes very low and takes a longer time to come back to nominal value after circuit breaker reclosing. With the application of the FLC-SVR, the voltage profile is kept to the highest level and closest to the prefault nominal voltage. The conventionally controlled SDBR can contribute in improving the voltage profile but the performance of the FLC-SVR is noticeably better compared to that of the SDBR.

High fluctuation of output active power is seen in Fig. 7 if no auxiliary device is used for transient performance improvement. The SDBR can limit this fluctuation till breaker opening. But after the breaker opening, the SDBR can not control the fluctuation. It is apparent that the FLC-SVR is the most skillful in minimizing the fluctuation at fault initiation and after the breaker opening.

Fig. 8 provides the DFIG speed response. There is a high oscillation of the rotor after the fault without the help from any auxiliary device and it takes long time for the speed to come back to nominal value. The SDBR can bring back the speed after the breaker opening in a while but can not prevent the speed oscillation after breaker opening. This mainly happens because the SDBR cannot control the active power control as opposed to the FLC-SVR. With the FLC-SVR, the speed oscillation is the lowest and the generator comes back to nominal speed at the fastest time. This really reduces the electromechanical stress on the whole turbine generator system and enhances the transient performance.

DFIG dc-link voltage response is shown in Fig. 9. At fault, the dc-link voltage goes very high if no auxiliary controller is applied. This high level of voltage may destroy the dc link capacitor and the capacitor protective circuit. The FLC-SVR helps the DFIG to keep the dc-link voltage fluctuation to the lowest level. If the FLC-SVR is in the system, protection device activation of the dc-link capacitor can be avoided. The SDBR can also help keeping dc-link voltage lower but its performance is inferior to the FLC-SVR.

C. Index Based Comparison

Some performance indices, namely, $vlt(pu.s)$, $spd(pu.s)$, and $pow(pu.s)$ are considered to provide a numeric representation of the performance comparison. A better transient performance is indicated by the lower value of the indices. Their definitions are given below in the following equations.

$$vlt(pu.s) = \int_0^T \Delta V dt \tag{8}$$

$$spd(pu.s) = \int_0^T \Delta \omega dt \tag{9}$$

$$pow(pu.s) = \int_0^T \Delta P dt \tag{10}$$

where ΔV, $\Delta\omega$ and ΔP are the deviation of the wind farm terminal voltage, the DFIG speed, and the active power output of the wind farm, respectively. T=5 s refers to the end point of the time duration of interest starting from 0 s. The values of the indices are presented in Table III for the 3LG fault. It shows that the system performance is the worst without any auxiliary device. A significant improvement is observed using the FLC-SVR as it gives the lowest index values. Compared to the SDBR, the FLC-SVR gives lower value of the indices and hence performs better than the SDBR.

TABLE III: PERFORMANCE INDEX VALUES FOR 3LG FAULT

Index parameters(%)	Values of indices		
	No auxiliary controller	SDBR	FLC-SVR
vlt(pu.s)	28.28	9.978	3.57
pow(pu.s)	29.54	19.18	2.59
spd(pu.s)	2.42	1.01	0.11

VI. CONCLUSION

A fuzzy logic controller is proposed to facilitate a series variable resistor for wind farm transient stability enhancement purpose. From the simulation results the below points can be noted.

1) The FLC-SVR is an effective method to augment the wind farm transient stability.
2) The wind farm electromechanical turbine-generator systems encounter lower stress with the FLC-SVR actions.
3) The FLC-SVR is more adaptive to the faults and offers a superior performance than the SDBR in every aspect as evident by graphical and numerical results.

Application of FLC-SVR can be investigated for smart grid and microgrid transient stability enhancement. Also other nonlinear controller can be applied to control the SVR other than the fuzzy logic controller.

REFERENCES

[1] R. Cárdenas, R. Peña, S. Alepuz, and G. Asher, "Overview of control systems for the operation of DFIGs in wind energy applications," *IEEE Trans. Ind. Electron.*, vol. 7, no. 60, pp. 2776–2798, 2013.

[2] G. Rashid and M. Hasan Ali, "Bridge-type fault current limiter for asymmetric fault ride-through capacity enhancement of doubly fed induction machine based wind generator," in *Proc. IEEE Energy Convers. Congress and Exp.(ECCE)*, pp. 1903–1910, Sept 2014.

[3] F. K. A. Lima, A. Luna, P. Rodriguez, E. H. Watanabe, and F. Blaabjerg, "Rotor voltage dynamics in the doubly fed induction generator during grid faults," *IEEE Trans. Power Electron.*, vol. 25, pp. 118–130, Jan. 2010.

[4] L. Meegahapola, T. Littler, and D. Flynn, "Decoupled-dfig fault ride-through strategy for enhanced stability performance during grid faults," *IEEE Trans. Sustain. Energy*, vol. 1, pp. 152–162, Oct 2010.

[5] J. Yang, J. Fletcher, and J. O'Reilly, "A series-dynamic-resistor-based converter protection scheme for doubly-fed induction generator during various fault conditions," *IEEE Trans. Energy Convers.*, vol. 25, no. 2, pp. 422–432, 2010.

[6] G. Rashid and M. Ali, "Transient stability enhancement of doubly fed induction machine-based wind generator by bridge-type fault current limiter," *IEEE Trans. Energy Convers.*, vol. 30, pp. 939–947, Sept 2015.

[7] G. Rashid and M. H. Ali, "A modified bridge-type fault current limiter for fault ride-through capacity enhancement of fixed speed wind generator," *IEEE Trans. Energy Convers.*, vol. 29, no. 2, pp. 527–534, 2014.

[8] K. S. Latha and M. V. Kumar, "STATCOM for enhancement of voltage stability of a DFIG driven wind turbine," in *Proc. Power Energy Syst. Conf. Towards Sustain. Energy*, pp. 1–5, Mar. 2014.

[9] L. Wang, J.-Y. Yu, and Y.-T. Chen, "Dynamic stability improvement of an integrated offshore wind and marine-current farm using a flywheel energy-storage system," *IET Renew. Power Gener.*, vol. 5, pp. 387–396, Sept. 2011.

[10] L. Wang and S.-S. Chen, "Stability improvement of a grid-connected offshore wind farm using a superconducting magnetic energy storage," in *Proc. IEEE Ind. Appl. Soc. Annu. Meet.*, pp. 1–8, Oct. 2012.

[11] Y. Zhao, O. Krause, T. K. Saha, and Y. Li, "Stability enhancement in distribution systems with DFIG-based wind turbine by use of SFCL," in *Proc. Australas. Univ. Power Eng. Conf. (AUPEC)*, pp. 1–6, Sept. 2013.

[12] G. Rashid and M. Hasan Ali, "Asymmetrical fault ride through capacity augmentation of DFIG based wind farms by parallel resonance fault current limiter," in *Accepted for presentation at IEEE Power and Energy Society General Meeting*, pp. 1–5, paper ID:16PESGM0946, July 2016.

[13] T. Ackermann, *Wind power in power systems*. John Wiley & Sons, 2005.

[14] A. Petersson, *Analysis, modeling and control of doubly-fed induction generators for wind turbines*. PhD thesis, Göteborg, Sweden, 2003.

[15] L. Xu and P. Cartwright, "Direct active and reactive power control of DFIG for wind energy generation," *IEEE Trans. Energy Convers.*, vol. 21, pp. 750–758, Sept. 2006.

[16] G. Rashid and M. Hasan Ali, "Application of parallel resonance fault current limiter for fault ride through capability augmentation of DFIG based wind farm," in *Proc. IEEE PES Transmission & Distribution (T&D) Conference & Exposition*, pp. 1–5, paper ID:TD0149, May 2016.

[17] M. H. Ali, M. Park, I.-K. Yu, T. Murata, J. Tamura, and B. Wu, "Enhancement of transient stability by fuzzy logic-controlled smes considering communication delay," *Int. Journal of Electrical Power & Energy Syst.*, vol. 31, no. 7, pp. 402–408, 2009.

Analytical Modeling of VFET

Nishith Nirjhar Chakraborty, Imam Al Razi, Mehedy Hasan Bappy and Md. Ziaur Rahman Khan

Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology
Dhaka, Bangladesh
oishi.nnc@gmail.com

Abstract—**In this paper, an analytical model has been developed for the vacuum field-effect transistor (VFET). Mathematical expressions for drive in voltage and current has been derived. A drive in voltage has been modeled based on minimum current density for space charge limited emission. Analytical model of the channel current incorporates the device operating physics such as space charge limited emission and Fowler-Nordheim tunneling. The analytical model has been numerically verified with the practical data available. Proposed model provides the opportunity to evaluate the vacuum device characteristics and determine the future projections as well.**

***Index Terms*—*VFET, vacuum microelectronics, Fowler-Nordheim tunneling*.**

I. Introduction

High speed electronic devices rely on short carrier transport times, which are usually achieved by decreasing channel length and increasing carrier velocity [1]. But it is difficult to achieve ballistic transport in a solid-state medium because high electric fields used to increase carrier velocity also increase scattering [2-3]. The vacuum is intrinsically superior to solid as carrier transport medium since it allows ballistic transport. Early electronics centered around the vacuum tube used to amplify, switch or modulate electrical signals [4-5]. The transition from vacuum tube to solid-state device was driven by the ease of fabrication, low-cost, low-power consumption, lightness, long lifetime, and ideal form factor for integrated circuits. But transistors were not better in all respects. Besides being slower than vacuum tubes, another disadvantage of the transistors is that semiconductors are susceptible to strong radiation, which can disrupt the atomic structure of the silicon such that charges no longer move properly.

Advantages of both devices can be achieved together if macro-scale vacuum tube is miniaturized to nanometer scale [8]. The nano vacuum tubes can provide high frequency/power output and provide stability at harsh conditions [7]. Further downscaling can allow a cold cathode because the electric field itself is strong enough to emit electrons [5]. Also, downscaling in conjunction with low work function materials may decrease the turn-on gate and drain voltages, thus enabling devices to be competitive with modern semiconductor technology [8]. Thus the new device, vacuum field-effect transistor (VFET) is a cross between today's transistors and vacuum tubes of yesteryear. It's small and easily manufactured, but also fast and radiation-proof.

In this paper, an analytical model for VFET has been developed. First, the device structure and modes of operation is specified. It also includes device operating physics of space charge limited emission and mechanisms of quantum tunneling, as shown in section II. The mathematical expression for drive-in voltage along with the on current equation of the device is derived in section III. It also depicts the process of obtaining effective electric field and emission area of cathode. Section IV includes the results and discussions of the analytical model. The proposed analytical model for channel current is numerically verified here by the simulation data available.

II. Vacuum Microelectronics

A. Device Structure and Physics

The most common design of VFET is a vertical field emitter consisting of the emitter, gate, and collector as shown in Fig. 1(a). Unfortunately, the vertical structure may be undesirable for circuit implementation due to the difficulties in achieving identical geometrical dimensions over all devices on the substrate. In contrast, the geometry of the planar structure (Fig. 1(b)) is defined by photolithography enabling practical integration [10]. Herein, we have implemented a gate-insulated planar lateral triode to reduce the gate leakage similar to silicon dioxide in MOSFET.

In this device, electrons are emitted from the emitter due to a voltage applied across it and the collector, while the gate controls the electron flow across the cavity. The electron flow through a metal nano-gap structure usually involves two step process : field emission from the metal surface, commonly described by Fowler-Nordheim theory, and subsequent transport through the gap, governed by Child-Langmuir SCL current flow..

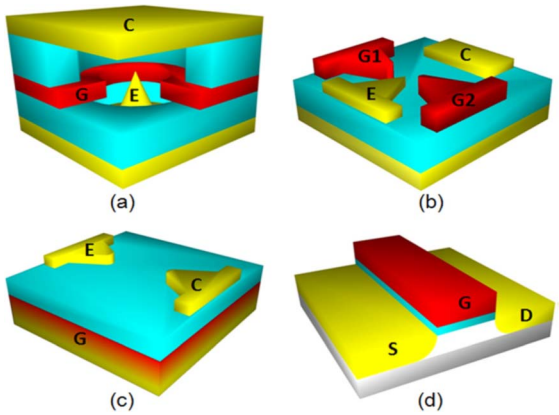


Fig. 1 Structures of Vacuum Devices: (a) Vertical Field Emitter (b) Planar Lateral Field Emitter (c) MOSFET (d) Gate insulated air channel transistor.[6]

B. Physics of Space-Charge Limited Current Flow

Following is a qualitative description of the space-charge limit, illustrated in Figure 2. The first electrons in a beam leaving an emitter move freely across the emission gap to their target: i.e., across a vacuum gap to an anode. Subsequent electrons, however, are influenced by the

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

negative electric field generated by each electron already in the beam, causing them to decelerate. This is referred to as a space-charge effect. The impact of each emitted electron is cumulative such that, as the emission current increases, so does the deceleration force. The deceleration of electrons creates volume of increased beam density in the emission gap where electrons reach their minimum velocity. At sufficiently high currents, the velocity of the slowest electrons drops to zero and below. Once this occurs, the charge at the high density point in the beam, referred to as the virtual cathode, builds up rapidly. Next, more and more subsequent electrons approaching the virtual cathode are reflected back to the emitter. Thereafter, the virtual cathode forms and dissipates in a time varying pattern. The current above which this occurs is referred to as the space-charge limit.

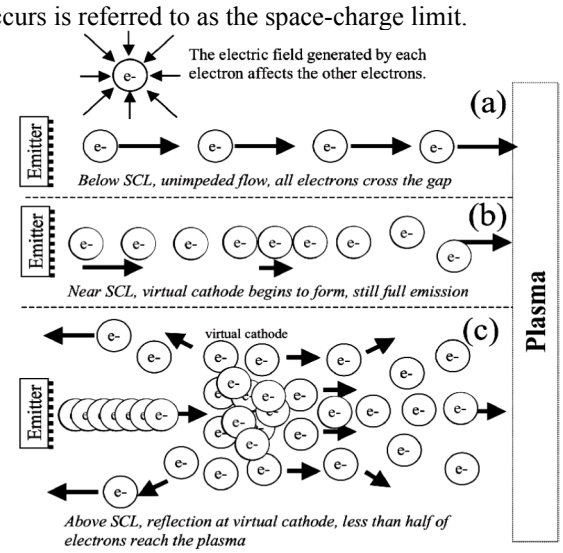


Fig. 2 Graphical depiction of the space-charge limit for electron emission from a space-craft cathode. [9]

Space-charge current density can be given by,

$$J(z) = \rho(z)v(z) = -J_{SCL} \tag{1}$$

For the motion of an electron, we can write,

$$\frac{1}{2}mv^2 - e\Phi(z) = 0 \tag{2}$$

Poisson's equation,

$$\frac{d^2\Phi}{dz^2} = -\frac{\rho(z)}{\epsilon_0} = \frac{J_{SCL}}{\epsilon_0}\sqrt{\frac{m}{2e}}\frac{1}{\sqrt{\Phi(z)}} \tag{3}$$

Multiplying equation (3) by $(\frac{d\Phi}{dz})$ on both sides, integrating and using boundary conditions we get,

$$J_{SCL} = \frac{4}{9}\epsilon_0\sqrt{\frac{2e}{m}}\frac{V^{\frac{3}{2}}}{D^2} \tag{4}$$

Here, the boundary conditions used are:

For, z=0, $\Phi(z) = 0$

For, z=D, $\Phi(z) = V$

C. Fowler-Nordheim Equation

The tunneling current density for an electron, passing through a general potential is derived here. The E label indicates that the vertical axis refers to the energy in the band diagram $q.\Phi(x)$ is the shape of the potential barrier, E_x is the electron energy along the x-direction and T_{tun} is the tunneling distance. The expression of the current density (J) induced by electrons tunneling in the x-direction through a generic potential barrier is

$$J = q\int_0^{\alpha} N(v_x)T(E_x)v_x dv_x \tag{5}$$

Where, q is the charge of an individual electron, $T(E_x)$ is the tunneling probability, i.e. the probability that one electron having energy Ex along the x-axis goes through the potential barrier. $N(v_x)dv_x$ is the density of electrons with velocity between vx and v_x+dv_x along the x-axis. The integral is taken from 0 to infinity since the electrons are trapped in the metal for x<0.

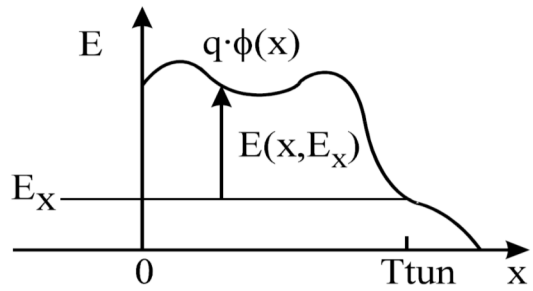


Fig. 3 Potential barrier of a general shape [11].

$$N(v_x) = \int_{-\alpha}^{\alpha} dv_y \int_{-\alpha}^{\alpha} dv_z f(\frac{1}{2}mv^2)$$

$$N(v_x) = \frac{2\pi}{m}\int_{-\alpha}^{\alpha} dE_r f(E_r + E_x) \tag{6}$$

Assuming that the electrons in the metal follow the free electron gas model, Pauli Exclusion Principle applies and Fermi-Dirac distribution function is introduced

$$f\left(\frac{1}{2}mv^2\right) = \frac{1}{1+e^{\frac{\frac{1}{2}mv^2-E_f}{k_BT}}} \tag{7}$$

The current density becomes, substituting the previous expression for $N(v_x)$ and considering the variable change $E_x = mv_x^2/2$, so that $v_x dv_x = dEx/m$,

$$J = \frac{4\pi q m k_B T}{h^3}\ln(e^{\frac{E_x-E_f}{k_BT}} - 1)\int_0^{\alpha} T(E_x)dE_x \tag{8}$$

The tunneling probability *T(Ex)* depends on the potential barrier shape and it can be determined by the Wentzel-Kramers-Brillouin (WKB) approximation and some calculations:

$$T(E_x) = T(E_f)e^{\sqrt{\frac{8\pi^2 m}{h^2}}(E_f-E_x)\int_0^{Ttun}\frac{1}{\sqrt{q\Phi(x)-E_x}}dx} \tag{9}$$

Where, $(q\Phi(x) - E_x)$ is the difference between the energy of the potential barrier at position x *(q·Φ(x))* and the electron energy in the x-direction (E_x).

$$T(E_f) = e^{-2\sqrt{\frac{8\pi^2 m}{h^2}}\int_0^{Ttun}\sqrt{,(q\Phi(x)-E_f)}\,dx} \tag{10}$$

For a triangular potential barrier,

$$q.\Phi(x) = q.\Phi + E_f - q.F.x$$

$$q.\Phi - q.F.T_{tun} = 0 \tag{11}$$

The tunneling probability becomes,

$$T(E_f) = e^{-\frac{4}{3}\sqrt{\frac{8\pi^2 m}{h^2}}\frac{(q\Phi)^{\frac{3}{2}}}{qF}} \tag{12}$$

Thus,

$$J = \frac{4\pi q m}{h^3}\frac{(k_BT)^2}{C^2}T(E_f)\frac{C\pi}{\sin C\pi} \tag{13}$$

Where the constant C becomes

$$C = k_BT\sqrt{\frac{8\pi^2 m}{h^2}}\int_0^{T_{tun}}\frac{1}{\sqrt{q\Phi(x)-E_f}}dx \tag{14}$$

Substituting the expressions for $T(E_f)$ and *C* into the current density equation,

$$J = \frac{q^2}{8\pi h\Phi} F^2 e^{-\frac{4}{3}\sqrt{\frac{8\pi^2 m}{h^2}}\frac{(q\Phi)^{\frac{3}{2}}}{qF}} \frac{C\pi}{\sin C\pi}$$

$$\approx \frac{q^2}{8\pi h\Phi} F^2 e^{-\frac{4}{3}\sqrt{\frac{8\pi^2 m}{h^2}}\frac{(q\Phi)^{\frac{3}{2}}}{qF}} \quad (15)$$

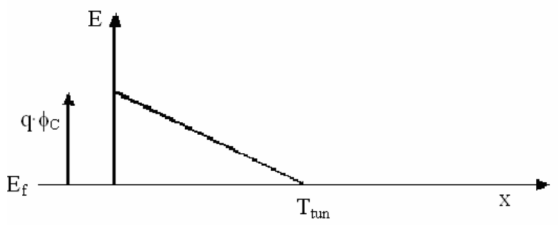


Fig. 4 Triangular shape barrier for FN tunnelling.

We can write,

$$J \approx A\frac{F^2}{\Phi} e^{-\frac{B\Phi^{\frac{3}{2}}}{F}} \quad (16)$$

Where, A=1.54X10^{-6} and B= 6.82X10^{9}.

III. Analytical Model

A. Drive in Voltage

In VFET, SCL regime seems to appear at very low drain voltages (~0.5V). The minimum voltage at drain that starts the SCL regime (a current density of 105 Acm-2) can be defined as drive in voltage of VFET.

Current density for space charge limited emission,

$$J_{SCL} = \frac{4}{9}\epsilon_0\sqrt{\frac{2e}{m}}\left(\frac{V}{D}\right)^2\frac{1}{\sqrt{V}}$$

$$J_{SCL} = \frac{4}{9}\epsilon_0\sqrt{\frac{2e}{m}}\,(E)^2\frac{1}{\sqrt{V}} \quad (17)$$

Effective electric field, $E_{eff} = \beta V_A$ (18)

Here, V_A = Applied anode voltage.

Replacing in equation (17)

$$J_{SCL} = \frac{4}{9}\epsilon_0\sqrt{\frac{2e}{m}}\,(\beta V_A)^2\frac{1}{\sqrt{V_A}} \quad (19)$$

Drive in voltage (V_{AD}) is the minimum anode voltage while device has the off current (J_{min}). Placing these in equation (19) and after some calculation,

$$V_{AD} = \left[\frac{4}{9}\epsilon_0\sqrt{\frac{2e}{m}}\,(\beta)^2\frac{1}{J_{min}}\right]^{-\frac{2}{3}} \quad (20)$$

B. Effective Electric Field

Electric field at the cathode tip of VFET has contribution from both gate and anode voltages. Both components referred by figure 5 can be summed up and written as,

$$F_0 = \frac{V_{ds}}{d} + \frac{2V_{gs}}{\sqrt{r_c^2+d^2}}\cos\Theta \quad (21)$$

Where, $\Theta = \tan^{-1}\frac{r_c}{d}$ (22)

For, a higher anode-cathode distance and channel radius i.e. cathode radius field component due to gate voltage gets negligible. To find the effective field we have to multiply potential distribution due to electric field (equation (21)) by field enhancement factor.

The effective electric field, $F = \beta V(F_0)$ (23)

Where, field enhancement factor, $\beta = \frac{1}{kr}$

r = microtip radius

And $\begin{cases} \frac{1}{2}\ln\left(\frac{4d}{r}\right) \ldots\ldots r \ll d \\ \frac{d}{r} \quad \ldots\ldots\ r > d \end{cases}$

C. Effective Emission Area

The emission area is supposed to be the tip area of the cathode. But in practical, the whole tip area is not engaged in emission. The effective emission area is reduced due to the angular dependence of the regular fields [12].

Effective emitting area, $S = 2\pi {r_{in}}^2 \frac{F^2}{\gamma_F \beta F_0}$ (24)

Where, the field can be given by,

$$F \approx \beta F_0$$

So, $S = 2\pi {r_{in}}^2 \frac{\beta F_0}{\gamma_F}$ (25)

Here, F_0= Applied electric field,m=Electron mass, Φ= Work function of emitting material, β= Field enhancement factor, γ=0.66 and

$$F = \frac{2\pi}{3eh} 4\sqrt{2m}\Phi^{\frac{3}{2}} \quad (26)$$

The maximum emission angle ranges from 10 to 25° for a gate voltage of 10V gives an energy spread of 1eV [13]. A successful emission and transmission both are required for anode to cathode current. Thus, we can conclude about the effective emission area,

$$S = 2\pi {r_{in}}^2 \frac{\beta F_0}{\gamma_F}\frac{2\Theta}{180} \quad (27)$$

Where , $\Theta = 10°$ to 25°.

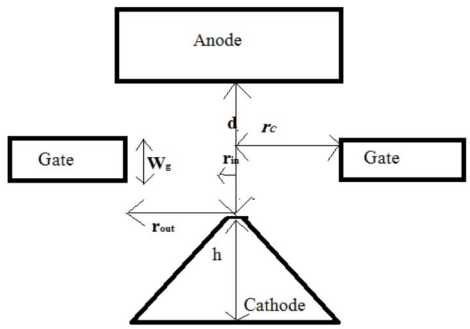


Fig. 5 Cross section of a VFET showing different structural parameters.

D. Overall Anode to Cathode Current Equation

Usually, SCL regime begins to appear at very low voltages (<0.5 V), but the Fowler– Nordheim regime is absent in the voltage range < 2 V. The different voltage dependences of the two regimes (Fowler–Nordheim versus 2DES-enabled SCL) were confirmed in a different experiment [14]. So the total current density,

$$J = J_t + J_{SCL}$$

So, channel current $I = JS$

So, the final expression for channel current,

$$I = \left\{A\frac{F^2}{\Phi}e^{-\frac{B\Phi^{\frac{3}{2}}}{F}} + \frac{4}{9}\epsilon_0\sqrt{\frac{2e}{m}}\frac{(V_A-V_{AD})^{\frac{3}{2}}}{D^2}\right\}\left(2\pi {r_{in}}^2\frac{\beta F_0}{\gamma_F}\frac{2\Theta}{180}\right) \quad (28)$$

IV. Results And Discussions

This section depicts the numerical verification of our proposed model for VFET. The numerical analysis i.e. the simulation of the analytical model has been done in

MATLAB and verified with the practical work i.e. fabricated device data of different parameters [1,8] available.

Figure 6 & 7 shows channel current as a function of drain voltage for a gate voltage of 2V and 6V respectively. Both the analytical model and simulation data obtained from [8] have been plotted here. There is mismatch between the plots and the deviation of the analytical model from the simulated one is increasing with the increase of drain voltage. The explanation of this deviation is that, in the simulation data or the real time data the cathode efficiency has a big deal. With the increase of drain voltage and current cathode tips get hotter and cathode tip may not be as sharp as it was at start. This might effect parameters like threshold and drive in voltage. The temperature issue has viable significance on the performance of the cathode. Thus it can't be said that, the efficiency of the cathode was same all along the simulation voltage range from 1V to 10V.

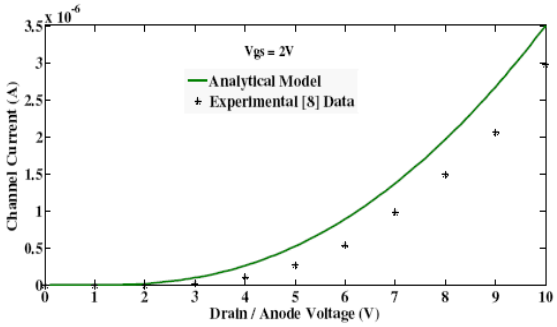


Fig. 6 Channel Current vs. Drain/Anode Voltage with Vgs at 2V. For Anode to Cathode Gap of 100 nm and Tip Radius of 5 nm.

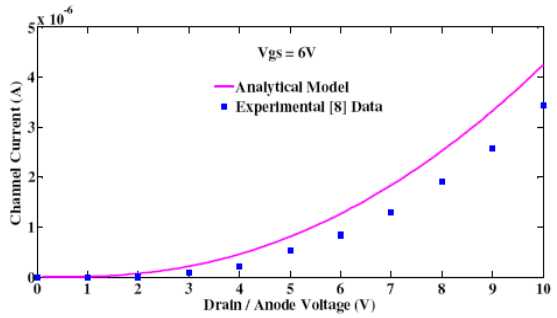


Fig. 7 Channel Current vs. Drain/Anode Voltage with Vgs at 6V. For Anode to Cathode Gap of 100 nm and Tip Radius of 5 nm.

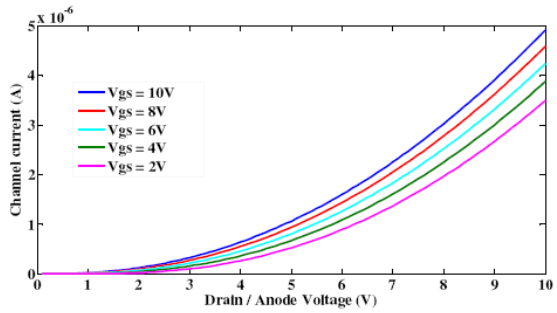


Fig. 8 Current-Voltage Characteristics of VFET. For Anode to Cathode Gap of 100 nm and Tip Radius of 5 nm.

Besides this deviation, our proposed model shows enough conformity with the actual data. The solution of the deviation can be including a temperature dependent efficiency factor in our model which will incorporate the efficiency degradation with increase of drain current and voltage.

The analytical model demonstrates a clear well defined FET characteristics as shown in Figure 8. We can see the increase in slope of the channel current with the increase of gate to source voltage.

V. Conclusions

A number of experimental works has been done for VFET, but very few analytical works exist. A concise analytical model of this device gives the scope to explore various characteristics of the device. Proposed model includes drive in voltage and channel current. It incorporates the emission physics of Child Langmuir and Fowler-Nordheim. Channel current components, such as space charge limited and tunnel current have dominance in different region of device operation. Inclusion of effective electric field and emission area makes the model more accurate. The model shows good enough conformity with the available actual simulation data.

References

[1] Srisonphan, S.; Jung, Y. S.; Kim, H. K., "Metal-oxide-semiconductor field-effect transistor with a vacuum channel", Nature nanotechnology, 7(8), 504-508, 2012.J. Breckling, Ed., *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*, ser. Lecture Notes in Statistics. Berlin, Germany: Springer, 1989, vol. 61.

[2] Leitenstorfer, A.; Hunsche, S.; Shah, J.; Nuss, M. C.; Knox, W. H., "Femtosecond charge transport in polar semiconductors", Physical review letters, 82(25), 5140, 1999. (2002) The IEEE website. [Online]. Available: http://www.ieee.org/

[3]] Han, J. W.; Meyyappan, M., "The device made of nothing", Spectrum,IEEE, 51(7), 30-35, 2014.

[4] Kesling, W. D.; Hunt, C. E., "Beam focusing for field-emission flat-panel displays", Electron Devices, IEEE Transactions on, 42(2), 340-347,1995. *FLEXChip Signal Processor (MC68175/D)*, Motorola, 1996.

[5] Sze, S. M., "Semiconductor devices: physics and technology", John Wiley & Sons, 2008. J. Breckling, Ed., *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*, ser. Lecture Notes in Statistics. Berlin, Germany: Springer, 1989, vol. 61.

[6] Han, J. W.; Oh, J. S.; Meyyappan, M., "Vacuum nanoelectronics: Back to the future?—Gate insulated nanoscale vacuum channel transistor",Applied Physics Letters 100(21),213505,2012.

[7] *Symons, R. S., "Tubes: Still vital after all these years", Spectrum, IEEE,35(4),52-63,1998* .

[8] Han, J. W.; Oh, J. S.; Meyyappan, M.; "Co-fabrication of Nanoscale Vacuum Field Emission Transistor (VFET) and MOSFET", 2014. R. E. Sorace, V. S. Reinhardt, and S. A. Vaughn, "High-speed digital-to-RF converter," U.S. Patent 5 668 842, Sept. 16, 1997.

[9] Bardeen, J., & Brattain, W. H. (1949). Physical principles involved in transistor action. *Physical Review*, *75*(8), 1208. J. Padhye, V. Firoiu, and D. Towsley, "A stochastic model of TCP Renocongestion avoidance and control," Univ. of Massachusetts, Amherst,MA, CMPSCI Tech. Rep. 99-02, 1999.

[10] *Park, S.S.; Park, D.I.; Hahm, S.H.; Lee, J.H.; Choi, H.C.; Lee, J.H.; "IEEE Trans. Electron Devices" 46, 1283 (1999).*

[11] Yazdani, A.; "Watching an atom tunnel", Nature, 409, 471 (2001).

[12] Raichev, O. E.; "Coulomb blockade of field emission from nanoscaleconductors." Physical Review B73.19(2006): 195328 .

[13] *Nation, J. A., Schächter, L., Mako, F. M., Len, L. K., Peter, W., Tang, C. M., & Srinivasan-Rao, T. (1999). Advances in cold cathode physics and technology.Proceedings of the IEEE, 87(5), 865-889. Park, S.S.; Park, D.I.; Hahm, S.H.; Lee, J.H.; Choi, H.C.; Lee, J.H.; "IEEE Trans. Electron Devices" 46, 1283 (1999).*

[14] Lau, Y. Y.; Liu, Y. & Parker, R. K.; "Electron emission: from the Fowler–Nordheim relation to the Child–Langmuir law". Phys. Plasmas 1, 2082–2085 (1994).

Analysis of ECG Signal and Classification of Heart Abnormalities Using Artificial Neural Network

Tanoy Debnath, Md. Mehedi Hasan, Tanwi Biswas
Department of Electronics and Communication Engineering
Khulna University of Engineering & Technology
Khulna 9203, Bangladesh
kuettanoydebnath@yahoo.com, mehedi_ece@yahoo.com, tanwibiswas.tb@gmail.com

Abstract—**Cardiac arrhythmia indicates abnormal electrical activity of heart that can be a great threat to human. So it needs to be identified for clinical diagnosis and treatment. Analysis of ECG signal plays an important role in diagnosing cardiac diseases. An efficient method of analysing ECG signal and predicting heart abnormalities have been proposed in this paper. In the proposed scheme, at first the QRS components have been extracted from the noisy ECG signal by rejecting the background noise. This is done by using the Pan Tompkins algorithm. The second task involves calculation of heart rate and detection of tachycardia, bradycardia, asystole and second degree AV block from detected QRS peaks using MATLAB. The results show that from detected QRS peaks, arrhythmias which are based on increase or decrease in the number of QRS peak, absence of QRS peak can be diagnosed. The final task is to classify the heart abnormalities according to previous extracted features. The back propagation (BP) trained feed-forward neural network has been selected for this research. Here, data used for the analysis of ECG signal are from MIT database**

Index Terms*— *ECG; Tachycardia; Bradycardia; Asystole; Second degree AV block; Backpropagation.

I. Introduction

Electrocardiogram (ECG) is the graph of voltage versus time which is produced due to electrical activities of human heart. An ECG signal consists of various parts such as P wave, QRS complex and T wave which are the waveform of corresponding electrical activities of various parts of the human heart. During the analysis of ECG signal, data including the positions or magnitudes of the QRS, PR, QT and ST intervals, the PR and ST segments etc. are considered for accurately diagnosing the heart diseases [1]. With the features present in the ECG signal, it is possible to predict various cardiac arrhythmias. There are many established algorithms for extracting features from ECG signal to predict heart abnormalities, for example, algorithms from the field of artificial neural network [2], genetic algorithm [3], Wavelet Transform [4] etc.

For analyzing the ECG signal and before the steps of feature extraction, preprocessing of ECG signal is necessary because signal is noisy and so many artifacts. For the analysis of ECG signal MIT-BIH Arrhythmia Database [5] with patients Information has been used. This database is a large collection of recorded physiologic signals and it has free web access for researchers for their experimental purpose. For the preprocessing purpose a widely popular technique called Pan-Tompkins Algorithm has been adapted. This is for QRS complex detection for the measurement of Heart beat (BPM) [6]. The proposed algorithm automatically calculate heart rate and finds the Heart condition according to the measured Heart Beat (BPM). Extracted features are based on patient's age, R-R intervals, Heart rate and Heart condition. The purpose of classification is to identify various heart conditions with patients age/gender, R-R intervals and Heart Rate which helps clinical cardiac diagnostic system. For this reason, artificial neural network has chosen for classification because of its vast use in machine learning areas and biomedical fields. For the classification purposes Back propagation has chosen which is easy to utilize for proper classification. It is a feed forward neural network which is used by training and testing data. Back propagation refers to backward propagation of errors so that there is a minimum difference between estimated data and output data. In back propagation, an iterative way is used mainly for this purpose so that all the weight can be updated to a new value after each iteration and a feed forward network is chosen because here all the input, weight and hidden neuron are interconnected to each other.

II. Proposed Methodology

The Methodology of this proposed work can be predicted from figure (1) in Block Diagram.

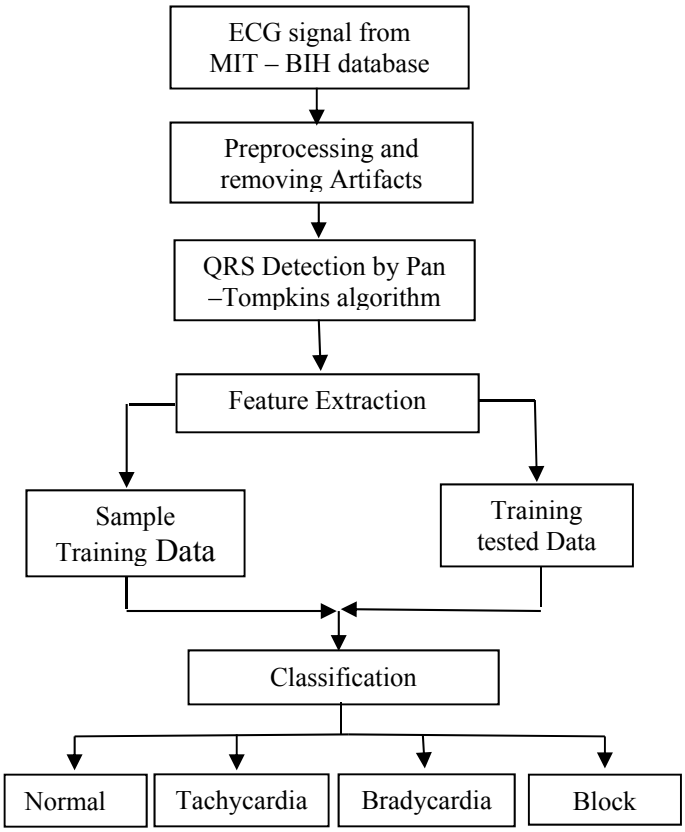


Figure 1. Block Diagram of Proposed Methodology

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

Firstly ECG signals are collected from database and preprocessed for removing artifacts and using Pan – Tompkins algorithm Q,R,S are identified for the heart rate calculation. This deflection of ECG signal is very important features for identify various hearts conditions. A neural network is trained with 26 datasets containing features of QRS complex with patients age/gender, R-R intervals and Heart Rate with condition of hearts. After the training, the network is tested with 12 more datasets. Tested results have gone through the same as by training datasets.

III. METHODOLOGY WORK

A. Preprocessing of ECG signal

ECG signal collected from database is recorded signal that is often contaminated with electrical interference or noise that is recorded from sources other than the electronic signals of the heart. In order to eradicate noisy ECG signal we have to preprocess the basic ECG signal. This preprocessing can be done by two ways.

1) Preprocessing and
2) Feature Extraction.

Preprocessing step eliminates electrical interference of ECG signal like power line artifacts, Electromyographic (EMG) noise, baseline wandering which are more significant among artifacts and they can impact on ECG signal analysis. Feature extraction step refers to obtain the most relevant information from the original ECG signal. For this preprocessing and feature extraction stage Pan – Tompkins algorithm has been applied.

B. Pan-Tompkins Algorithm

Compared to other QRS detection technique the utilization of Pan-Tompkins algorithm is more vulnerable. The literature survey of this approach makes it as one of the most important algorithm in detection QRS peak [7]. For better diagnosing the heart abnormalities the accuracy of any ECG waveform extraction plays a vital role.

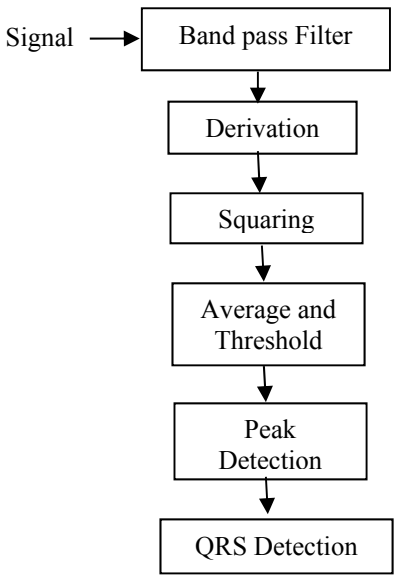


Fig 2. Pan-Tompkins Structure

Fig 2. Shows the graphical representation of the Pan-Tompkins algorithm. This algorithm is one of the most popular QRS detection process in biomedical signal processing textbooks [8]. Initially the contaminated signal is processed by filtering, derivation, squaring, thresholds detecting and finally detecting the QRS complex [9].

1) Band Pass Flitering

Signal passes through a band pass filter which is a cascaded low pass filter and high pass filter. This filtering process reduces noise in the ECG signal and an average QRS complex is found by matching the spectrum of the signal. Noises like Muscle noise, power line interference, baseline wander, T wave interference are attenuated by band pass filtering. Pass band maximizes the energy of QRS complex in the range of 5 Hz- 15 Hz. The cascaded low pass filter and high pass filter are described by the formula. For low-pass filter,

$$y(n)= 2y(n-1)-y(n-2)-x(n)-2x(n-6)-x(n-12) \quad (1)$$

For High –pass filter,

$$y(n)= y(n-1)- 1/32x(n)-x(n-16)-x(n-17)+1/32x(n-32) \quad (2)$$

2) Derivation

Signal then processed with derivative operator which is a standard technique to identify a peak in the signal. Derivative filter finds out the high slopes. It identifies the direction in the slope of the signal and distinguishes the QRS complexes from other ECG waves. It suppresses the low frequency components and a large gain to the high frequency components arises. The signal is differentiated to find out QRS slope using the following formula:

$$y(n)= 1/8[2x(n)+ x(n-1)-x(n-3)-2x(n-4)] \quad (3)$$

3) Squaring

The squaring function is a nonlinear transformation that squaring the signal samples point by point. It makes the all signal values positive and amplifies the output of the previous stage. It suppresses the P and T wave and high frequency components of the QRS complex are increased. Thus emphasizing the R peaks in the signal. After differentiating the signal is then squared making all the values positive and give importance to the higher frequencies.

$$y(n)=x^2 (n) \quad (4)$$

4) Averaging and Threshold Process

For visual purposes and better identification of R peaks, averaging is utilized that maximizes the signal compared with squaring output. After the averaging process the threshold of the obtained signal is done that detects peaks in the ECG signal and identifies a threshold value. Then it is classified as QRS peak if peak crosses threshold. The signal now passes through a moving window integrator that performs as suitable interval under the squared waveform, sums the area and advances the new window. It performs as a smoothing of output of the previous operations. N=40 was found as a suitable window width for frequency, f_S=360 Hz.

5) *QRS peak detection*

For measuring Heart Beat in BPM detection of QRS complex is necessary. By calculating the time period of the R-R interval it can be measured. Different heart condition called arrhythmia

maximum and minimum amplitude is also found by analyzing the ECG signal. Below are several steps for real time ECG signal processing and QRS detection.

- Data Acquisition from MIT-BIH database.
- Filtering the signal by using low pass filter and high pass filter.
- Differentiation of the signal that's passes through band pass filter.
- Squaring the differentiated signal for getting all positive value.
- Averaging, Threshold and integrating the squared signal and clear R peak detection.
- Determination of R-R interval.
- Calculation of heart rate by R-R interval.

For Heart Rate calculation,

Heart Rate = 60/ R-R interval (m sec).

C. Feature Extraction

By using the Pan–Tompkins algorithm, ECG signal is preprocessed and we obtained peak values. Now for the classification purposes, features are selected by four ways. Age/gender (from MIT database), R-R intervals, Heart Rate are chosen for the features for classification for different heart conditions. Firstly age is considered as 95 is the highest age, secondly gender is chosen as male=1; Female=0; other=0.5, thirdly R-R interval/max. of R-R inerval value is chosen and finally heart rate is guessed (160 is the maximum value).

D. Classification

In biomedical engineering and machine learning, Artificial Neural Network (ANNs) are a family of statistical learning algorithm inspired by biological neural network (the central system of animals) that are utilized to estimate or approximate functions that depend on a large number of inputs that are usually unknown [10]. For classification for different heart conditions the ANN used in this work is based on back propagation (BP).

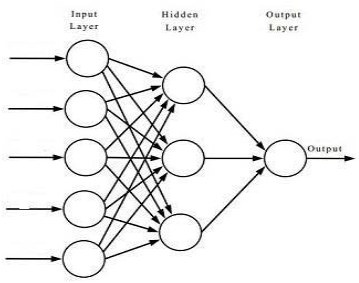


Fig. 3: Three layer feed forward network

Back propagation refers to backward propagation of errors. It is a three layer feed forward network and a common method of training artificial neural network that is efficient. For obtaining desired classification performance, the selection of hidden nodes is very important. Number of hidden nodes is commensurate to the complexity and memory requirement. The network cannot adjust the weight efficiently when the number of hidden nodes decrease. Sometimes an overfitting can be occurred. That's why, the network is generalized to nullify overfitting [11]. The back propagation algorithm can be divided into two phases:1) Propagation of errors and 2) Weight update. When the errors propagate in the backward direction, the weight in the layers will change. It is an iterative algorithm to minimize the mean square error between the desired output and the actual network output. This algorithm has limitation like it always changes weights in such a way to cause the error to fall.

TABLE I. CONDITIONS FOR HEART ABNORMALITIES

Serial No	Heart Abnormalities	Conditions
1	Tachycardia	Heart Rate>100BPM
2	Bradycardia	Heart Rate<60BPM
3	Second Degree AV Block	QRS dropped
4	First Degree AV Block	Long PR interval
5	Complete Heart Block(Asystole)	Complete drop out of cardiac cycle

Table 1 shows the conditions for the different heart conditions. For the classification purposes second degree heart block, first degree heart block and complete heart block are considered as block condition. For scarcity of above mentioned conditions data we considered these for block.

IV. RESULT ANALYSIS AND DISCUSSION

Table 2 shows the extracted features for the sample signals for different heart conditions and features are divided by four ways. Table 3 shows the confusion matrix for heart abnormalities classification performance that shows the correct and incorrect predictions made by BPNN. Firstly 26 datasets are trained and tested with 12 data. Three conditions for heart such as normal, bradycardia, tachycardia are classified perfectly but block condition is not correctly identified. Fig. 5 shows the iteration number to train error graph for training the data set where mean square error decreases in the increasing of iteration. It will be tends to zero of the higher value of iteration number. The mean square error will not decrease sharply for a value of iteration number. An approximate value of 0.056 after which the value of mean square have not decreased sharply.

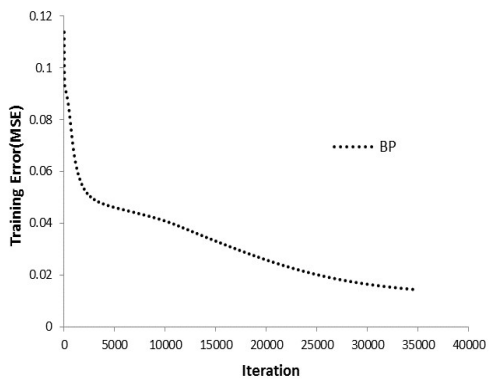


Fig. 5: Training error vs. number of iteration curve for training data set.

TABLE II. EXTRACTED FEATURES FOR SAMPLE SIGNALS

Test	Age/Gender	R-R interval	Heart rate	Identification
103	84/F	0.83	72	NORMAL
104	66/F	0.772	77.2	NORMAL
106	24/F	0.90	66	NORMAL
16272	20/F	1	60	NORMAL
16273	28/F	0.714	84	NORMAL
16795	20/F	0.711	84.375	NORMAL
16786	32/F	0.833	72	NORMAL
16773	26/M	0.833	72	NORMAL
16439	35/F	0.90	66	NORMAL
101	75/F	1.2	49	BRADYCARDIA
232	76/F	1.25	1.25	BRADYCARDIA
800	-	1.25	48	BRADYCARDIA
801	-	1.66	36	BRADYCARDIA
802	-	1.42	42	BRADYCARDIA
16483	42/M	1.66	36	BRADYCARDIA
16420	38/F	1.5	40	BRADYCARDIA
18177	26/F	2.5	24	BRADYCARDIA
PERSON30		1.11	54.05	BRADYCARDIA
PERSON40		1.06	56.4	BRADYCARDIA
212	32/F	0.659	91	TACHYCARDIA
215	81/M	0.56	106.3	TACHYCARDIA
16265	-	0.51	118.25	TACHYCARDIA
Chf01	71/M	0.430	139.3	TACHYCARDIA
CU03	-	0.69	86.4	VENTRICULAR TACHYCARDIA
30		0.6	106.4	TACHYCARDIA
31		0.44	`136.7	TACHYCARDIA
32		0.427	140.2	TACHYCARDIA
35		0.5	130.2	TACHYCARDIA
36		0.54	110.7	TACHYCARDIA
45		0.4	143.1	TACHYCARDIA
107	63/F	0.77	78.3	SECOND DEGREE HEART BLOCK
108	87/F	0.7	85.71	FIRST DEGREE HEART BLOCK
111	47/F	0.83	72	FIRST DEGREE HEART BLOCK
207	69/F	0.7	82.1	HEART BLOCK(DIFFICULT RECORD)
Chf04	54/M	1.66	36	HEART BLOCK
Chf15	53/M	1.2	48.2	HEART BLOCK
228	80/F	0.83	72	FIRST DEGREE HEART BLOCK
37		1.9	31.5	COMPLETE HEART BLOCK

Table 3. Confusion matrix for heart abnormalities classification by BPNN

Classes	Normal	Bradycardia	Tachycardia	Block
Normal	0.981165	0.063549	0.002401	0.011074
Normal	0.974685	0.276393	0.000745	0.003432
Bradycardia	0.061554	0.813611	0.000231	0.209684
Bradycardia	0.542276	0.872049	0.000107	0.004602
Tachycardia	0.008373	2.95E-05	0.994776	0.000108
Tachycardia	0.000288	1.64E-07	0.999789	0.08414
Block	0.280936	0.006238	0.048266	0.465988
Block	0.359838	0.014197	0.024578	0.66721

V. CONCLUSION

In this paper we established a network to classify different heart abnormalities e.g. normal, bradycardia, tachycardia and block as discussed earlier. A back propagation algorithm with feed forward neural network is used for classification purpose that is efficient. Though in our proposed work block condition is not classified due to some unintentional errors but the overall performance is excellent. The correct and incorrect relationship has shown in confusion matrix that is satisfactory. For future work, we will try to find out more features and utilize other classifier like SVM, kNN classifier and will make comparison with ANN based on their classification performance.

REFERENCES

[1] Yun-Chi Yeh, Wen-June Wang, Che Wun Chiou, "Cardiac arrhythmia diagnosis method using linear discriminant analysis on ECG signals", ELSEVIER, Measurement vol. 42, pp. 778-789, January 2009.

[2] G. Karraz, G.M., " Automatic classification of heartbeats using neural network classifier based on a bayesian framework", 28*th Annual International Conference of the IEEE Publication*, pp. 4016-4019, 2006.

[3] Poli, R., Cagnoni, S. and Valli, G., "Genetic design of optimum linear and nonlinear QRS detectors", *IEEE Transaction Biomedical Engineering*, vol. 42, pp. 1137-1141, 1995.

[4] Prasad, G.K. and Sahambi, J.S., "Classification of ECG arrhythmias using multi-resolution analysis and neural networks", *Proceedings of IEEE Conference on Convergent Technologies*, vol. 1, pp. 227-231, 2003.

[5] https://physionet.org/cgi-bin/atm/ATM

[6] J. Pan and W. Tompkins, "Real Time Algorithm detection for QRS", IEEE Trans. Eng. Biomed Eng, 32(3), pp.230-236, 1985.

[7] Friesen, G. M., Jannett, T. C., Jadallah, M. A., Yates, S. L., Quint, S.R., Nagle, H. T, "A comparison of the noise sensitivity of nine QRS detection algorithms", IEEE Trans. Biomed.Eng., BME-37: pp. 85–97, 1990.

[8] L. Ahmad, M. Mohamed, N. A. Ab. Ghani," An Improved QRS Complex Detection for Online Medical Diagnosis", World Academy of Science, Engineering and Technology, 68, 2012.

[9] Shital L. Pingale, Nivedita Daimiwal, "Detection of Various Diseases Using ECG Signal in MATLAB", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-3, Issue-1, March 2014.

[10] https://en.wikipedia.org/wiki/Artificial_neural_network

[11] M.R.Ahsan, M.I.Ibrahimy, and O.O.Khalifa, "EMG Signal classification for Human Computer Interaction:A Review,"European Journal of Scientific Research,Vol.33,no.3,pp.480-501,2009

Effect of Response Time and Starting Point of Duty Cycle in Maximum Power Point Tracking Operation of Solar Panel Using PO Algorithm

Ratul Das[1,*] and K. M. Rahman[1]

[1]Dept. of EEE, Bangladesh University of Engineering and Technology
Dhaka, Bangladesh
*ratuldas@live.com

Abstract— **This paper presents a detailed view of Perturb and Observe (PO) method for maximum power point (MPP) tracking (MPPT) of a solar panel. After development of Simulink model of a PV array, a load was connected via a DC-DC converter and PO method was applied to adjust the duty cycle of the converter. The change required in duty cycle with respect to load and insolation have been critically observed. Observations shows that MPPT algorithms work erratically due to insufficient delay between two consecutive feedbacks and improper starting point of duty cycle. These points are discussed in details in this paper. An optimum delay between two feedbacks and starting point for duty cycle have been proposed to track the maximum power for a wide range of loads.**

Index Terms— **solar panel, MPPT, PO method, response time, duty cycle.**

I. Introduction

The efficiency of a solar panel is highly dependent on solar insolation, temperature as well as output voltage of the panel. So, maximization of power from a solar photovoltaic module is of special interest. A maximum power point tracker can increase the efficiency of a PV solar module. There are a lot of MPPT techniques [1] proposed in the literature. Perturb and Observe (PO) method [2] is one of the most popular ones. Incremental conductance method [3], constant voltage method [4], Fuzzy logic control [5], neural network method [6] etc. have also uses. Researches are going on to improve these methods and introduce new methods to maximize the efficiency of MPPT [7-9]. These algorithms are involved with controlling the duty cycle of the intermediate DC-DC converter between the panel and load.

The reason behind researches and improvement on existing and novel algorithms for MPPT is, in practical cases, almost all of them do not work properly. The algorithms are implemented in a close loop manner. Control mechanism depends on the feedback. After introducing a change in the control based on the feedback, the total system takes some time to follow that change or to settle. This is called response time delay. If the control mechanism does not wait this response time and takes feedback of the unsettled system and introduces more change, the system will be oscillating and fail to track the maximum power. Another big constraint is starting point of the duty cycle of the PWM controlled DC-DC converter. Duty cycle can be varied 0 to 100 percent and starting point can also be anywhere between this. But not all starting point will successfully track the maximum power. Many complete systems fail just only for this reason.

This work concentrates on determining the response time of a solar panel system and proper staring point of duty cycle and synchronize the feedback in order to maximize the power. In section II, the basic characteristics of a 75 watt solar panel [10] developed by Siemens has been presented. In section III and IV, maximum power transfer theorem in the case of a solar panel system has been analysed in terms of impedance matching and duty cycle of DC-DC converter. After that, PO method was stated in section V. Effect of response time delay and starting point of duty cycle has been presented in section VI and VII respectively. Section VIII includes the results.

II. Solar Panel Model

The equations and figures stated in previous works [11-12] were used to study the I-V and P-V characteristics of a solar panel (e.g. Fig. 1 and Fig. 2). The specification of a 75 W solar panel [10] by Siemens is used to generate the I-V and P-V curves. At first, the expected power at optimum condition was matched selecting resistances and ideal factor by trial and error. Then those values were used in the model. From these curves it is seen that, at different solar insolation, the current for a particular voltage is different and as a result, the power also varies. And the magnitudes of maximum power point (MPP) for different insolation are also different. Current and power also varies because of temperature which is not shown here.

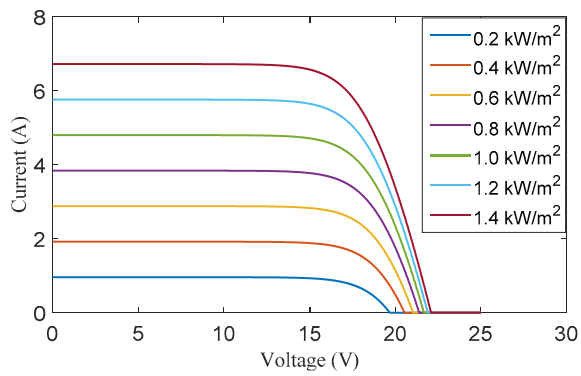


Fig. 1: I-V Characteristics

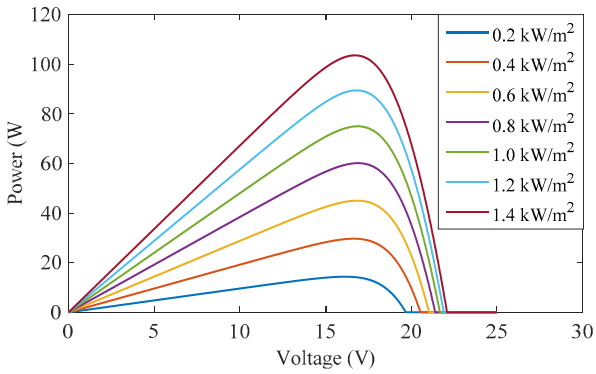


Fig. 2: P-V Characteristics

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

III. Maximum power transfer theorem for solar panel

The maximum power transfer theorem states that, to obtain maximum external power from a source with a finite internal resistance, the resistance of the load must equal the resistance of the source as viewed from its output terminals. But this theorem can not be applied directly if the source is a solar panel. It is obvious from Fig. 1-2 that the voltage and current for different insolation are different and as a result, the resistance of solar panel also varies with insolation. The resistances of the panel are calculated at different insolation by applying this formula,

$$Rin = \frac{V}{I} \tag{1}$$

Where, V and I are the voltage and current at maximum power point.

The calculated values are listed in Table 1. To extract maximum power the load should be varied with changed solar panel resistance. But it is not a practical approach.

TABLE I
INPUT RESISTANCES OF A SOLAR PANEL AT DIFFERENT SOLAR INSOLATION

Solar irradiation (kW/m2)	Maximum power available (W)	Voltage at maximum power point(V)	Current at maximum power point (A)	Input Resistance (Ω)
0.2	14.3646	16.14	0.8900	18.1348
0.4	29.6846	16.63	1.7850	9.3165
0.6	45.0103	16.82	2.6760	6.2855
0.8	60.1356	16.84	3.5710	4.7158
1.0	74.9764	16.86	4.4470	3.7913
1.2	89.4683	16.83	5.3160	3.1659

IV. DC-DC Converter for Maximum Power Transfer

To match the solar panel resistance and the load resistance a DC-DC converter can be used between solar panel and the load. DC-DC converter is an electronic circuit which converts a source of direct current (DC) from one voltage level to another. There are a lot of classes for DC-DC converters. In this work, buck converter topology is used for software simulation (i.e. Fig. 3).

For a Buck converter, voltage transfer ratio is

$$\frac{Vout}{Vin} = \frac{Ton}{T} = D \tag{2}$$

Where, D is the duty cycle. If we consider no power loss between input and output, then,

$$Pin = Pout$$

$$\frac{Vin^2}{Rin} = \frac{Vout^2}{Rout}$$

$$Rin = \frac{Vin^2}{Vout^2}$$

Replacing $\frac{Vin}{Vout} = \frac{1}{D}$ from equation (2), we get

$$Rin = \frac{1}{D^2} * Rout \tag{3}$$

So, if the input resistance of a source does not match the load then, a converter can be used to match them by varying its duty cycle by this formula which is derived from eqn. (3).

$$D = \sqrt{\frac{Rout}{Rin}} \tag{4}$$

At 1kW/m^2 solar insolation, solar panel resistance is about 3.8 ohm. From eqn. (4), it is clear that load resistance must the less than solar panel resistance if the intermediate DC-DC converter is of Buck topology. So, a bunch of resistances value less than 3.8 ohm has been used as load and duty cycles required to track the maximum power for them are enlisted in table II. But calculated duty cycle for one insolation will not guarantee maximum power at other insolation. As, weather conditions and other factors related to insolation always changes, it is nearly impossible to get always a constant insolation. So, a fixed duty cycle for a fixed load will not provide maximum power.

TABLE II
DUTY CYCLES OF A BUCK CONVERTER REQUIRED TO EXTRACT MAXIMUM POWER FROM A SOLAR PANEL AT A CONSTANT INSOLATION

1 kW/m^2 Solar irradiation			
Load (Ω)	Duty Cycle (%)	Load (Ω)	Duty Cycle (%)
0.01	5.1	1.0	51.3
0.05	11.5	1.5	62.9
0.10	16.2	2.0	72.6
0.30	28.1	2.5	81.2
0.50	36.3	3.0	88.9

V. PO method

In PO method (e.g. Fig. 3), voltage and current from solar panel are sampled continuously. Corresponding power is calculated by multiplying them. The power of a particular sample is compared with the power of the previous sample. If the present sample of power is greater, then if the present voltage is greater than the previous, the voltage is further increased. Otherwise, it is decreased. And if the present sample of power is smaller, it should be checked again which of the previous and current sample of voltage is greater. If the current one is greater, voltage is decreased or it is increased in the opposite case. Voltage is changed by changing duty cycle. The process is repeated periodically. The system then oscillates around the MPP. A solar panel system with a DC-DC converter and resistive load has been designed in MATAB-Simulink environment (Fig. 4) and further analysis has been carried out.

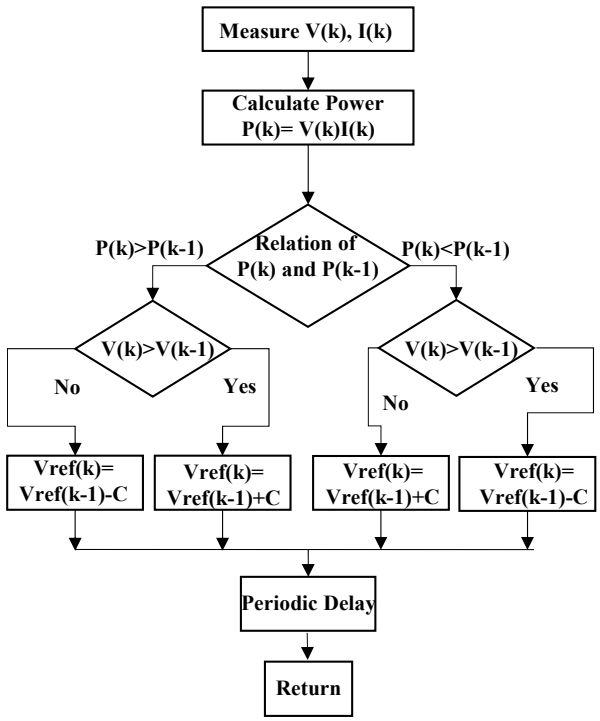


Fig. 3: Flowchart of PO method

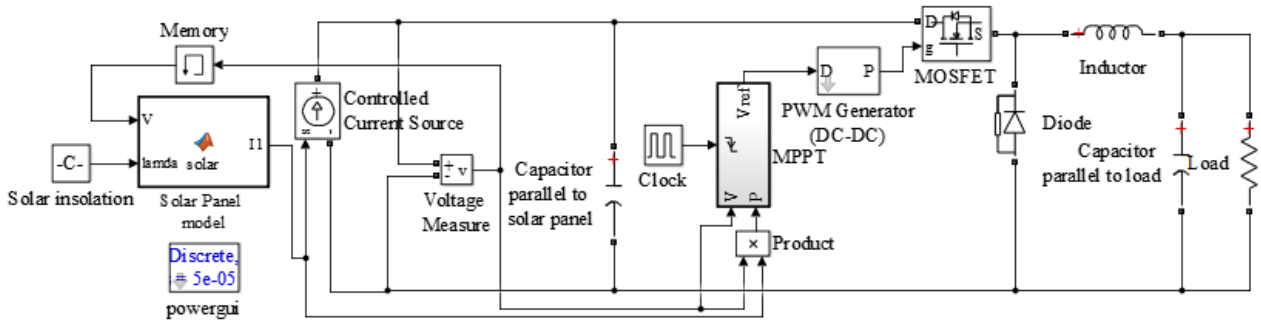


Fig. 4: Simulink model of solar panel system

VI. Effect of Response Time

Before starting the tracking, the system acts as a no load system because the MOSFET remains off. So, the parallel capacitor across solar panel becomes fully charged up to the open circuit voltage of the solar panel and solar panel current becomes zero. At the beginning, a certain duty cycle is applied as the starting duty cycle. And successful tracking can reach up to the desired duty cycle. To calculate the response time, a step change of 0 to 90 percent was introduced and response time was calculated by observing the time needed for the output voltage to be stable. All the loads enlisted in table II are also used here to observe the response time. The output voltage for maximum and minimum load has been shown in Fig. 5. All the response time calculated are enlisted in table III.

TABLE III
Response time for different loads

Load (Ω)	Response time (sec)	Load (Ω)	Response time (sec)
0.01	7.0	1.0	2.0
0.05	6.0	1.5	1.0
0.10	4.5	2.0	0.7
0.30	3.0	2.5	0.5
0.50	2.5	3.0	0.2

When the periodic delay is less than the response time, the system remain oscillatory but shows quite stability in the extracted power if the periodic delay is greater than response time (i.e. Fig. 6). The large spikes are due to discrete selection of duty cycle.

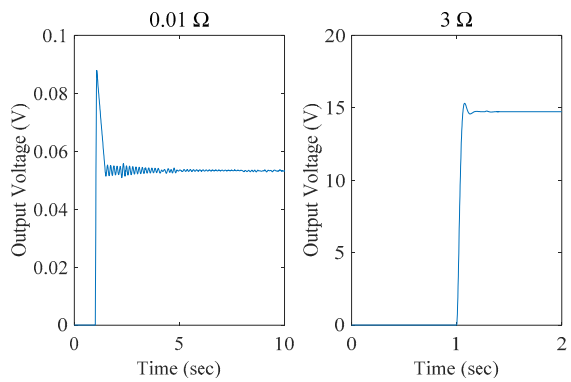


Fig. 5: Response time calculation for a step change in duty cycle from 0 to 90% at time 1s

VII. Effect of Starting Point of Duty Cycle

A load of 2 ohm was attached with the system and MPPT algorithm was applied starting from two different point of duty cycle 50 and 90 percent (i.e. Fig. 7). Estimated duty cycle at MPP was 72.6 percent. Other conditions remain unchanged, starting duty cycle from 90 percent, the system reached the MPP. But in the other case, it failed. It can be explained by the solar panel voltage at these corresponding duty cycles and P-V and I-V characteristics of the panel (i.e. Fig. 1 and 2). At 90 and 50 percent duty cycle, the solar panel voltages are around 12 V and 20 V respectively. That means, starting point for these duty cycles are left and right side of the MPP. In the left side of MPP, slightly decreased or almost constant current is needed from the solar panel to go towards MPP at every step. But, as the current remains almost constant, it is possible for the system to reach the MPP after some steps. In the right side, huge amount of increased current is required to increase the power in each step. This increased current can not go to the output side because of the current restriction of the inductor and high switching frequency of the MOSFET. So the increased current feds into the capacitor which prevents any voltage decrease. This is the reason behind the failure of MPPT starting at 50% duty cycle.

VIII. Results

Considering the response time and starting point constraint of duty cycle, MPPT was conducted for a number of loads. In every single case, the system successfully tracked the desired duty cycle for which the power maximizes. For higher loads, response time is faster and the desired duty cycles are also closer to the starting point. But lower loads have not only slower response time but also the duty cycle at MPP is much less than the starting point. So, lower loads required longer time to track MPP. However, if prior knowledge of the load is unavailable, then worst case possible response time for the minimum possible load should be considered as periodic delay and the starting point should be selected more than the highest possible duty cycle at MPP. Such as, if we consider the possible loads ranges from 0.01 ohm to 3 ohm, periodic delay should be at least 7 s and starting point should be greater than 88 percent duty cycle. Considering this, MPPT was conducted for all the loads. Longer periodic delay resulted in delayed tracking but it was as successful as the shorter one (i.e. Fig.8).

IX. Conclusions

In this paper, detailed analysis of two constraints behind the failure of PO algorithm has been presented. Insufficient periodic delay results in oscillations in the extracted power whereas improper starting point of duty cycle fails to change duty cycle beyond a particular limit. The reasons behind the failure for both have been discussed. A method has been proposed to select optimum periodic and optimum duty cycle starting point for a particular load. And at last, a general strategy for overcoming these constraint for an unknown load and known solar panel system has been presented.

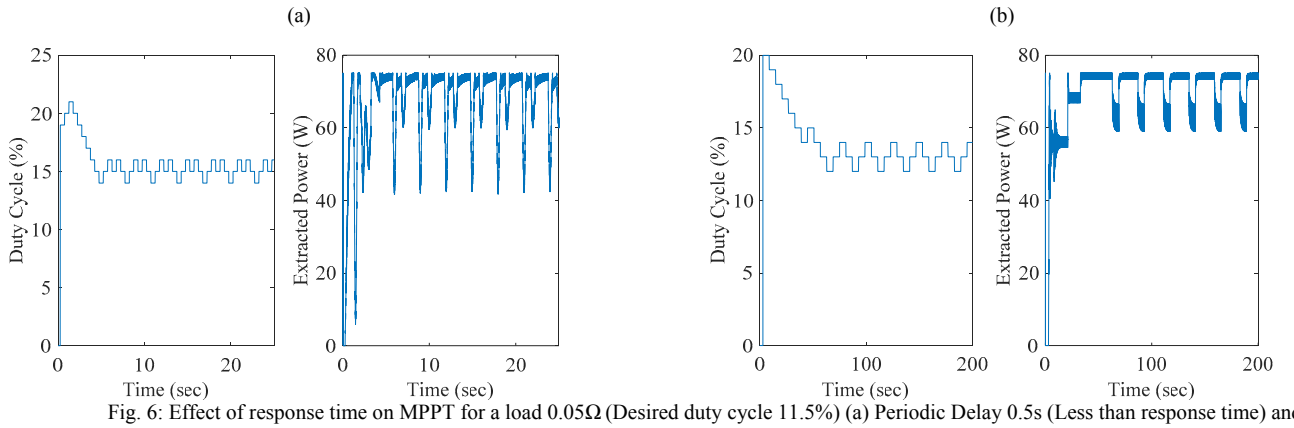


Fig. 6: Effect of response time on MPPT for a load 0.05Ω (Desired duty cycle 11.5%) (a) Periodic Delay 0.5s (Less than response time) and (b) Periodic delay 6.0 s (Equal to response time)

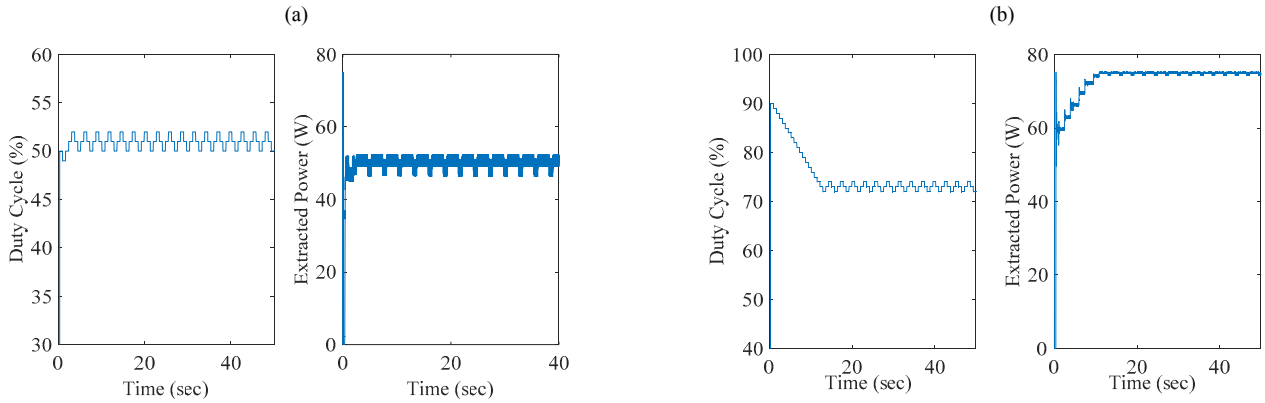


Fig. 7: Effect of starting point on MPPT for a load 2 Ω at periodic delay 0.7s (Desired duty cycle 72.9%) (a) Starting Point 50% duty cycle and (b) Starting point 90% duty cycle

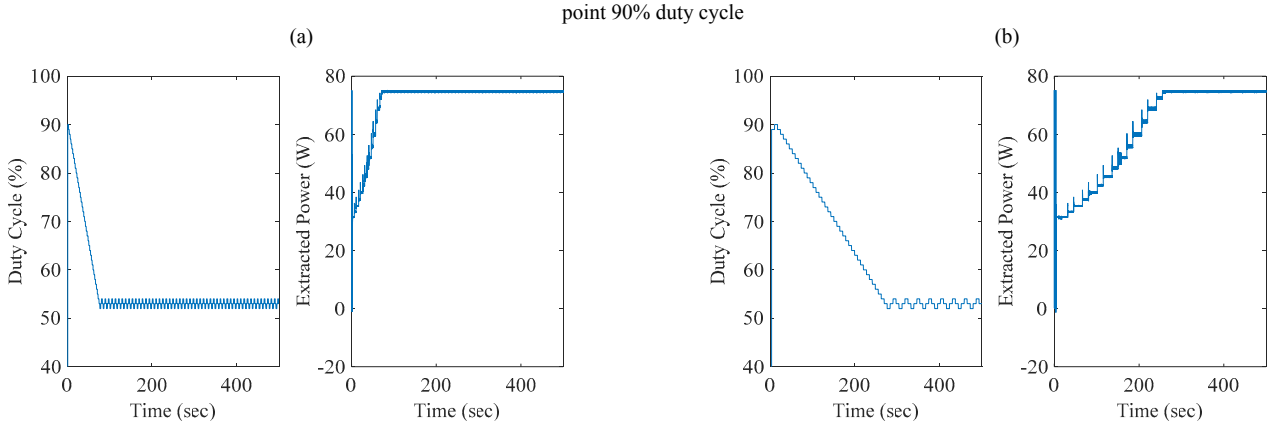


Fig. 8: Effect of longer periodic delay on MPPT for a load 1Ω (Desired duty cycle 51.3%) (a) Periodic delay 2s and (b) Periodic delay 7s

REFERENCES

[1] B. Subudhi and R. Pradhan, "A comparative study on maximum power point tracking techniques for Photovoltaic power systems," *IEEE Transactions on Sustainable Energy*, vol. 4, pp. 89–98, Jan. 2013.

[2] N. Femia, G. Petrone, G. Spagnuolo, and M. Vitelli, "Optimization of Perturb and observe maximum power point tracking method," *IEEE Transactions on Power Electronics*, vol. 20, pp. 963–973, Jul. 2005.

[3] A. Morales-Acevedo, J. L. Diaz-Bernabe, and R. Garrido-Moctezuma, "Improved MPPT adaptive incremental conductance algorithm," *IECON 2014 - 40th Annual Conference of the IEEE Industrial Electronics Society*, pp. 5540–5545.

[4] C. Dorofte, U. Borup, and F. Blaabjerg, "A combined two-method MPPT control scheme for grid-connected photovoltaic systems," *2005 European Conference on Power Electronics and Applications*, p. 10.

[5] L. Bouselham, M. Hajji, B. Hajji, A. E. Mehdi, and H. Hajji, "Hardware implementation of fuzzy logic MPPT controller on a FPGA platform," *2015 3rd International Renewable and Sustainable Energy Conference (IRSEC)*, pp. 1-6.

[6] X. Li, Y. Li, and J. E. Seem, "Maximum power point tracking for Photovoltaic system using Adaptive Extremum seeking control," *IEEE Transactions on Control Systems Technology*, vol. 21, pp. 2315–2322, Nov. 2013.

[7] M. E. Basoglu and B. Cakir, "A new maximum power point tracking approach for decreasing convergence time," *2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC)*, pp. 172-176.

[8] K. Gaikwad and S. Lokhande, "Novel maximum power point tracking (MPPT) algorithm for solar tree application," *2015 International Conference on Energy Systems and Applications*, pp. 189-93.

[9] N. E. Zakzouk, B. W. Williams, A. A. Helal, M. A. Elsaharty, and A. K. Abdelsalam, "Improved performance low-cost incremental conductance PV MPPT technique," *IET Renewable Power Generation*, no. 4, pp. 561–574, Apr. 2016.

[10] [Online]. Available: http://www.abcsolar.com/pdf/sp75.pdf.

[11] K. H. Hussein, "Maximum photovoltaic power tracking: An algorithm for rapidly changing atmospheric conditions," *IEE Proceedings - Generation, Transmission and Distribution*, vol. 142, p. 59, 1995.

[12] K. Kachhiya, M. Lokhande, and M. Patel, "Matlab/simulink model of solar PV module and MPPT algorithm," *2011 National Conference on Recent Trends in Engineering & Technology*, pp. 13-14.

Tunable Fabry Perot Filter Using All-Dielectric Multilayer Anisotropic Metamaterial In The Mid-Infrared Frequency Range

Ayed Al Sayem[1*] and Md. Saifur Rahman[2]

[1]Dept. of EEE, Daffodil International University
Dhaka, Bangladesh

[2]Dept. of EEE, Bangladesh University of Engineering and Technology
Dhaka-1205, Bangladesh

e-mail: ayedalsayem143@gmail.com[1], saifur@eee.buet.ac.bd[2]

Abstract—**In this article, we have theoretically proposed and numerically analysed a resonant wavelength tunable Fabry Perot (FP) filter without changing the physical length of the cavity in the mid-infrared range. The proposed FP filter consists of a multilayer cavity, which is constructed by alternately repeating two different sub-wavelength dielectric or semiconductor materials. Such a multilayer structure can behave as an anisotropic metamaterial. By varying the fill factor of the metamaerial, the effective refractive index of the cavity can be easily tuned. As the resonant wavelength of such a filter strongly depends on the cavity refractive index, the resonant wavelength of the filter can be easily tuned between a wide wavelength range of approximately 400nm by changing the fill factor. Distributed Bragg reflectors (DBRs) have been used as the high reflectivity mirrors. Both the DBRs and the cavity can be constructed by using only two materials, such as Silicon (Si) and SiO_2. This device may find substantial applications in the fields of spectroscopy, color filtering, sensing etc.**

Index Terms—**Cavity, Filter, Distributed Bragg Reflectors, Metamaterial.**

I. Introduction

Fabry Perot (FP) resonator is one of the most common and widely used high resolution optical filters used for various photonic applications such as spectroscopy [1]. A FP filter is a resonant transmission filter where only waves with a narrow bandwidth are transmitted through the filter and all other waves with different wavelengths are reflected. Typical FP resonators are made by using a pair of broadband high reflectivity mirrors such as metals or Distributed Bragg Reflectors (DBRs) and a cavity filled with a dielectric material [2]. The resonant frequency of a typical FP filter can be varied by varying the cavity thickness [3-5] and also the dielectric material used in the cavity. The resonant frequency of FP filters can also be altered without changing the physical thickness of the filter. Such a concept has been demonstrated where FP filter arrays have been constructed by metallic mirrors and an effective medium created by sub-wavelength 2d patterning [6]. 1d sub-wavelength gratings along with dielectric mirrors have also been theoretically proposed to construct tunable FP filters [7]. Double sub-wavelength grating reflectors [8] and metasurfaces [9] have also been recently used to demonstrate a tunable FP filter both theoretically and experimentally.

In this article, we have used an anisotropic metamaterial as the cavity to construct a tunable (having various resonant wavelengths) FP filter without changing the physical length of the cavity. This anisotropic metamaterial has been constructed by alternately repeating two deeply sub-wavelength dielectric (SiO_2) or semiconductor (Silicon) materials. Recently, such all–dielectric or semiconductor based sub-wavelength multilayer structurehas been proposed as anisotropic metamaterial [10, 11]. Better wave-guiding performance is expected by using such metamaterials as the cladding material in optical waveguides [11]. Such anisotropic metamaterials have also been used in constructing high Q resonators [12] and compact air-cavity resonator [13]. Intriguing phenomenon, such as negative refraction, has also been demonstrated by such metamaterials [14]. However, such an anisotropic metamaterial has not yet been considered for constructing a tunable FP filter to the best of our knowledge. Without changing the physical length of the proposed FP filter, the resonant wavelength can be tuned easily by changing the fill factor of the metamaterial for a wide range of wavelength (up to ~400nm). Also the proposed FP filter can be constructed by using two different materials only. The proposed theory has been validated by rigorous 4×4 transfer matrix method (TMM) [15-17]. This device may find fundamental applications in the fields of spectroscopy, color filtering, sensing etc.

II. Theory

A schematic view of the proposed structure is shown in Fig. 1. The proposed FP consists of a multilayer cavity constructed by alternately repeating two different sub-wavelength dielectric (semiconductor) materials with permittivities, ε_{t1} and ε_{t2}, and thicknesses, t_1 and t_2, respectively. DBRs have been used as the highly reflective mirrors above and below the cavity. Silicon and SiO_2 were the chosen materials for constructing both the anisotropic multilayer metamaterial and also for the DBRs.

The Bloch equation for the 1d multilayer structure (cavity) can be derived as,

$$\cos(k_z t) = \cos(k_{z2} t_2)\cos(k_{z1} t_1) - \frac{1}{2}\left(\frac{\varepsilon_{t2}}{k_{z2}}\frac{k_{z1}}{\varepsilon_{t1}} + \frac{k_{z2}}{\varepsilon_{t2}}\frac{\varepsilon_{t1}}{k_{z1}}\right) sin(k_{z2} t_2)\, sin(k_{z1} t_1) \quad (1)$$

Where, $k_{zi}^2 = \varepsilon_{ti} k_o^2 - k_x^2$; i=1, 2 and $t = t_1 + t_2$ is the thickness of the unit cell. k_x and k_z are the x and z components of the wave vectors and $k_o = \frac{2\pi}{\lambda}$, where λ is the free space wavelength.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

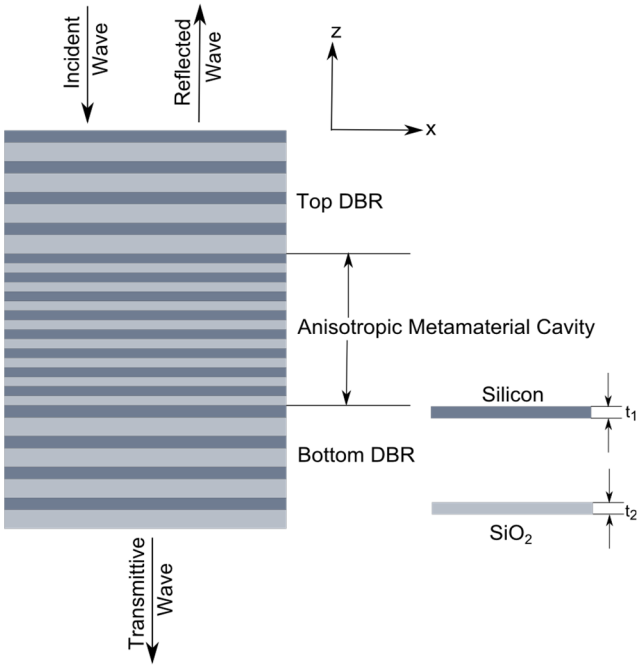


Fig. 1: Schematic of the proposed FP Filter. The structure consists of top and bottom DBRs where each DBR consists of 4 pairs of Si and SiO_2 with thicknesses of 110nm and 260nm, respectively. The cavity also consists of alternate layers of Si and SiO_2 but with deeply sub-wavelength thicknesses, t_1 and t_2, respectively, where the unit cell thickness is 20nm.

In the deeply sub-wavelength limit ($k_0 t \ll 1$), multilayer structure (cavity), can be considered as anisotropic metamaterial whose permittivity tensor can be calculated from equation (1) as,

$$\varepsilon_{\parallel} = \varepsilon_x = \varepsilon_y = f\varepsilon_{t1} + (1-f)\varepsilon_{t2} \tag{2a}$$

$$\varepsilon_{\perp} = \varepsilon_z = \left(\frac{f}{\varepsilon_{t1}} + \frac{1-f}{\varepsilon_{t2}}\right)^{-1} \tag{2b}$$

where, f is fill factor, given by, $f = \frac{t_1}{t_1+t_2}$

Equations (2a) and (2b) are commonly known as effective medium theory (EMT) and has been regularly used in literature [10, 11, 18-20].

For a FP filter, the relation between the FP resonant wavelength, λ, the cavity thickness, d and the refractive index, n can be given by,

$$\lambda \propto nd \tag{3}$$

The FP resonance follows a single peak in the transmission spectrum. As can be seen from equation (3), the resonant wavelength of the filter can be tuned by varying either the refractive index of the cavity or the thickness of the cavity or both. Tuning the refractive index is in general a difficult task as a result; the cavity thickness is varied to tune the resonant wavelength. But for our proposed structure, the cavity refractive index n is the effective refractive index of the multilayer metamaterial given by,

$$n = n_{eff} = \sqrt{\varepsilon_{\parallel}} \tag{4}$$

With the change of fill factor (the thickness of Si and SiO_2 in the unit cell), the effective permittivity $\varepsilon_{\parallel}$ can be easily varied between the values of Si and SiO_2 and thus the cavity resonant wavelength can be easily tuned without altering the physical length of the cavity.

III. Results and Discussion

Fig. 2 shows the effective permittivity; parallel permittivity $\varepsilon_{||}$ and perpendicular permittivity $\varepsilon_{\perp}$ of the Si-SiO_2 multilayer metamaterial as a function of fill factor. Here the refractive indices of Si and SiO_2, have been considered as 3.40 and 1.47, respectively. As can be seen from Fig. 2, effective permittivity of the structure can be varied between the values of Silicon and SiO_2 by changing the fill factor.

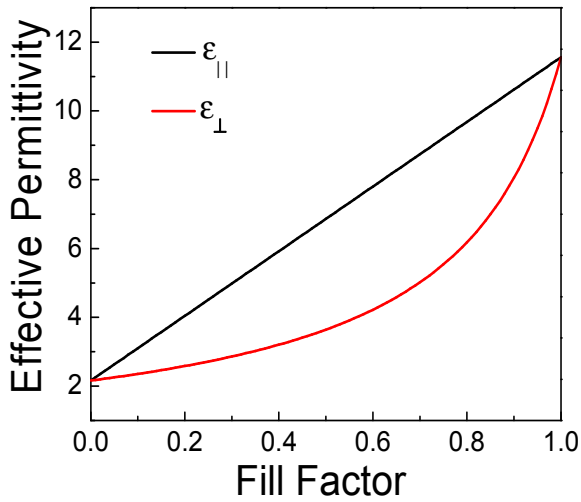


Fig. 2: Effective permittivity of the multilayer anisotropic metamaterial as a function of fill factor.

Fig. 3 shows the reflection spectrum of the identical top and bottom DBRs for the normally incident TM polarized wave as a function of wavelength calculated by TMM. Both DBRs are identical and have been constructed by alternate layers of Si and SiO_2. 4 pairs of quarter-wavelength stacks have been used to construct the DBRs. The thicknesses of the Si and SiO_2, in each pair of the DBRs was 110nm and 260nm, respectively. As can be seen from Fig. 3 more than ~99% reflection is observed from the DBRs for the wavelength range lying between 1300nm to 1800nm.

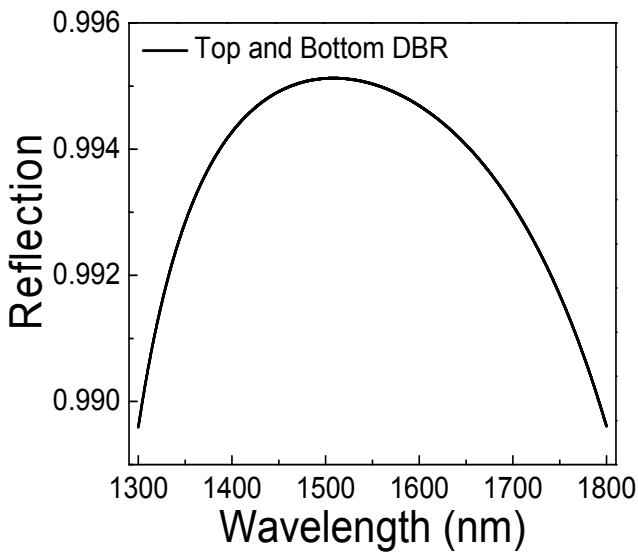


Fig. 3: Reflection coefficient for the normally incident TM polarized wave as a function of wavelength for the DBRs. The DBR consists of 4 pairs of alternate Si and SiO_2 layers with thicknesses 110nm and 260nm, and refractive indices 3.40 and 1.47, respectively.

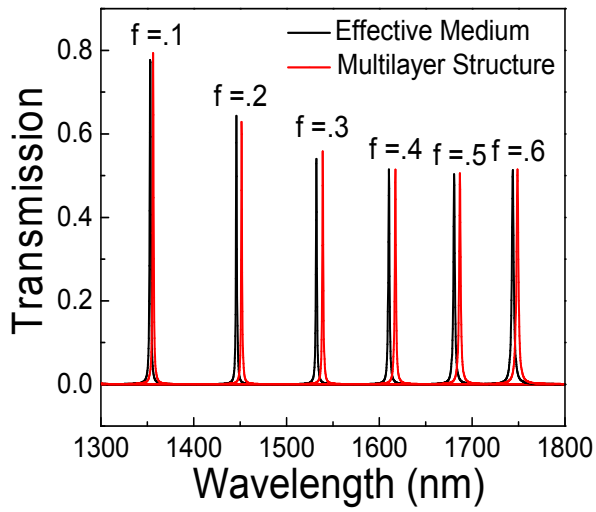


Fig. 4: Transmission coefficient for the normally incident TM polarized wave as a function of wavelength for the proposed FP filter. The DBRs have same physical parameters as mentioned in Fig. 3. Length of the cavity is 520nm. For the multilayer cavity 26 unit cells have been considered, where each unit cell thickness is 20nm.

Fig. 4 shows the simulated transmission coefficients as a function of wavelength for normal incidence for TM polarized wave calculated by 4×4 TMM with different fill factors (0.1 to 0.6). The red and black lines represent the data for the real multilayer structure and effective medium, respectively. Top and bottom DBRs are same for both the multilayer structure and effective medium. Height of both the cavities (real structure and effective medium) was set to 520nm. For the multilayer structure, 26 unit cells have been used, where each unit cell contains one pair of Silicon and SiO_2. Total thickness of the unit cell is 20nm. As can be observed from Fig. 4, the resonant transmission wavelength of the proposed FP filter is a strong function of fill factor and can be varied from ~1350nm to ~1750nm by changing the fill factor from 0.1 to 0.6. The real multilayer structure and the effective medium show very similar characteristics. Here, we have not used the full range of fill factor form 0 to1 because the DBRs show high reflection (~99%) only between the range of 1300nm to 1800nm. And fill factors 0.1 to 0.6 are enough to tune the resonant wavelength in this spectral range.

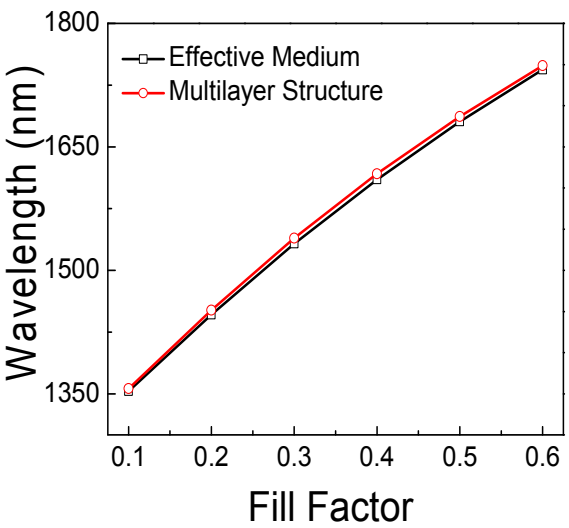


Fig. 5: Resonant transmission wavelength of the proposed FP filter as a function of fill factor. All the parameters used are same as Fig. 4.

Fig. 5 shows the resonant wavelength of the proposed FP filter as a function of fill factor for both the real structure and effective medium. With the increase of fill factor, the effective index of the filter increases. As a result, the resonant wavelength of the filter moves to higher wavelength which is expected from equation (3). Also a clear validation of the effective medium theory can be observed from the figure.

Fig. 6 shows the simulated transmission coefficients for TM polarized wave as a function of wavelength for different incidence angles. The proposed FP filter exhibit strong angular dependence.

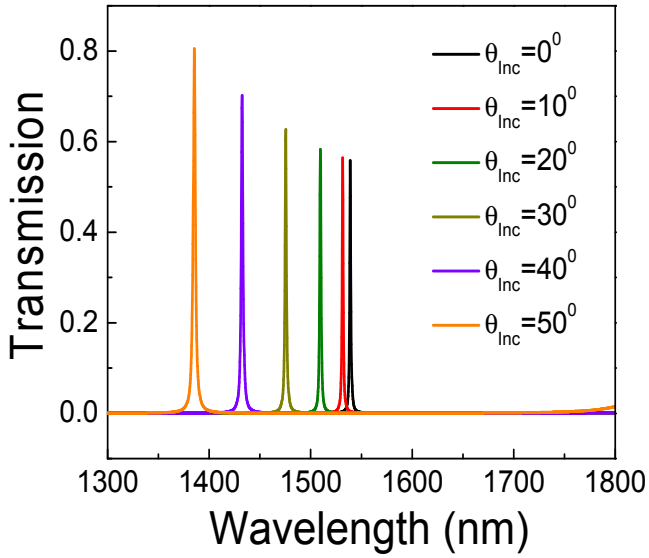


Fig. 6: Transmission coefficient for TM polarized wave as a function of wavelength for the proposed FP filter with different incidence angles. The DBRs and the cavity have the same physical parameters as mentioned in Fig. 3 and 4 but with fill factor 0.3.

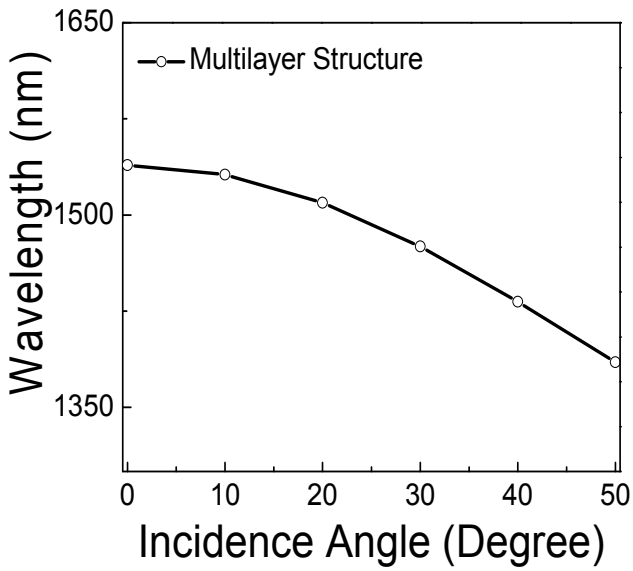


Fig. 7: Resonant transmission wavelength of the proposed FP filter as a function of incidence angle. All the parameters used are same as Fig. 4.

Fig. 7 shows the resonant wavelength of the proposed FP filter as a function of incidence for TM polarized wave. With the increase of incidence angle, the resonant wavelength moves to shorter wavelengths.

IV. Conclusions

In this article, we have theoretically proposed a simple way to construct a tunable Fabry Perot (FP) filter without changing the physical length of the cavity. The resonant wavelength of the proposed cavity can be tuned to a wide wavelength range of about 400nm. The proposed FP filter can be constructed by using only two different materials such as Silicon and SiO_2. Widely used thin film deposition can be used to construct such a tunable filter. Though the proposed device works in the mid infrared region, such concept can also be tailored to construct tunable filter in the visible region. This device may find fundamental applications in the fields of spectroscopy, angular color filtering, biological sensing etc.

References

[1] G. Minas, R. F. Wolffenbuttel and J. H. Correia, "A lab-ona-chip for spectrophotometric analysis of biological fluids," *Lab Chip,*vol. 5, pp. 1303–1309, 2005.

[2] J. H. Correia, M. Bartek and R. F. Wolffenbuttel, "High selectivity single-chip spectrometer in silicon for operation in visible part of the spectrum," *IEEE Trans. Electron Dev.,*vol. 47, pp. 553–559, 2000.

[3] A. Emadi, H. Wu, G. Graaf and R. Wolffenbuttel, "Design and implementation of a sub-nm resolution microspectrometer based on a linear-variable optical filter," *Opt. Express*, vol. 20, pp. 489–507, 2012.

[4] N. A. O'Brien, C. A. Hulse, D. M. Friedrich, F. J. V. Milligen, M. K. Gunten, F. Pfeifer and H. W. Siesler, "Miniature near-infrared (NIR) spectrometer engine for handheld applications," in *Proc. SPIE*, 2012, Vol. 8374, p. 837404.

[5] J. Xiao, F. Song, K. Han and S. Seo, "Fabrication of CMOS-compatible optical filter arrays using gray-scale lithography," *J. Micromech. Microeng.,* vol. 22, pp. 025006, 2012.

[6] K. Walls, Q. Chen, J. Grant, S. Collins, D. R. S. Cumming and T. D. Drysdale, "Narrowband multispectral filter set for visible band," *Opt. Express,* vol. 20, pp. 21917–21923, 2012.

[7] S. Kaushik and B. R. Stallard, "Twodimensional array of optical interference filters produced by lithographic alterations of the index of refraction," in *Proc. SPIE*, 1995, Vol. 2532, pp. 276–281

[8] Y. Horie, A. Arbabi, S. Han and A. Faraon, "High resolution on-chip optical filter array based on double subwavelength grating reflectors," *Opt. Express,* vol. 23, pp. 29848-29854, 2015.

[9] Y. Horie, A. Arbabi, E. Arbabi, S. M. Kamali and A. Faraon, "Wide bandwidth and high resolution planar filter array based on DBR-metasurface-DBR structures," *Opt. Express*, vol. 24, pp. 11677-11682, 2016.

[10] S. Jahani and Z. Jacob, "Transparent subdiffraction optics: nanoscale light confinement without metal," *Optica*m vol. 1, pp. 96-100, 2014.

[11] S. Jahani and Z. Jacob, "Photonic skin-depth engineering," *J. Opt. Soc. Am. B,* vol. 32, pp. 1346–53, 2015.

[12] W. Liu, A. E. Miroshnichenko and Y. S. Kivshar, "Q-factor enhancement in all-dielectric anisotropic nanoresonators," arXiv:1603.02111, 2016.

[13] S. Atakaramians and B. T. Kuhlmey, "Compact air-cavity resonators within a metamaterial waveguide," *Opt. Lett.*, vol. 41, pp. 3379-3382, 2016.

[14] A. A. Sayem, M. R. C. Mahdy and M. S. Rahman, "Broad angle negative refraction in lossless all dielectric or semiconductor based asymmetric anisotropic metamaterial," *Journal of Optics,* vol. 18 (1), 2016.

[15] B Dwight, "Optics in stratified and anisotropic media: 4×4-matrix formulation," *JOSA A,* vol. 62, pp. 502–510, 1972.

[16] P. Yeh, "Optics of anisotropic layered media: A new 4×4 matrix algebra," *Surface Science,* vol. 96, pp. 41, 1980.

[17] M. Schubert, "Polarization-dependent optical parameters of arbitrarily anisotropic homogeneous layered systems," *Phys. Rev. B,* vol. 53, pp. 4265, 1996.

[18] P. Belov and Y. Hao, "Subwavelength imaging at optical frequencies using a transmission device formed by a periodic layered metal–dielectric structure operating in the canalization regime," *Phys. Rev. B*, vol. 73, pp. 113110, 2006.

[19] V. M. Agranovich and V. L. Ginzburg, *Crystal Optics with Spatial Dispersion, and Excitons,* 2nd ed., Berlin, Germany: Springer, 1984.

[20] W. Cai and V. Shalaev, *Optical Metamaterials: Fundamentals and Applications,* 1st ed., New York, USA: Springer-Verlag, 2010.

Roboticon: A Realistic Platform to Imitate Facial Expressions

Amit Kumar Kundu*, Md. Didarul Islam* and S. M. Mahbubur Rahman †
*Department of Electrical and Electronic Engineering, Uttara University, Bangladesh
†Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology, Bangladesh
Email: amit31416@gmail.com, s102bueeet@yahoo.com, mahbubur@eee.buet.ac.bd

***Abstract*—This paper presents a concept that quickly detects the facial expressions of humans with a sufficient accuracy and imitates the expressions remotely in terms of robot-based mimics. In traditional online communication of expressions, it is not certain that the sender remains in the same state of expression, when the state is actually received in the remote side. In addition, the expression received in the remote side is mimicked by special visual characters commonly known as emoticons. The approach presented in this paper addresses the issues of real-time detection, communication, and imitation of expressions in order to solve the existing shortcomings. In particular, a computationally efficient intensity threshold-based algorithm is adopted to detect facial expressions from video signals. Because of such a simplicity, the proposed system ensures fast detection of an expression, which is quite convenient for real-time communication. At the same time, the system copies the expression on the face of a humanoid robot at the receiving end. It is expected that this novel approach, which is referred to as Roboticon will play a key role to imitate facial expressions in real-time.**

***Keywords*—Action units, expression detection, expression imitation, interfacing, Viola-Jones algorithm.**

I. Introduction

Facial expression is considered to be the most effective nonverbal way for human beings to express their emotions and interact with others. Emoticons and video calls are already being used for expressive communications. But a more effective representation would be a holographic or robotic face that can actually mimic someone's expressions. In this research work, focus has been given on copying sender's expression on a robotic face instead of emulating expression on a holographic one. Therefore, it is necessary for an expression copying robot to recognize and imitate human facial expressions as close as possible [1].

Many researchers have attempted to automatically recognize facial expressions of human beings. But understanding and imitating the human emotions appropriately have been remained a challenge in today's robotics while researchers are trying to make the robots with high level of sophistications [2]. For example, in [3], facial expression is detected utilizing twenty two facial muscle tensions. The 2D Gaussian-Hermite moments of facial distortions from the neutral images are used for recognizing human expressions in [4]. In order to express various emotions on a robot head; arm type antennae, emoticon eyes and exaggerating jaws and checks are used in [5].

In this paper, a simple, fast but efficient algorithm is proposed for automatic expression recognition which can be easily implemented in real time communication. The research scope of the paper consists of three expressions: happiness, sadness and surprise. Simple algorithm is used here for fast and real time implementation. At the receiving end a realistic approach of human expression imitation is attempted.

The organization of the rest of the paper can be summarized as follows. In Section II, methodology of the system is described. Section III presents the experimentation part where the whole idea is implemented practically. Finally, in Section III, conclusion is drawn and some future directions of works are proposed.

II. Methodology

The proposed facial emotion detection method comprises of three major parts which are expression detection, expression transmission and expression imitation as depicted in Fig. 1. In the first segment, images are acquired continuously. Then, a simple threshold based scheme is proposed to detect facial emotions such as: happiness, sadness or surprise. These expressions are detected based on RGB values in the informative portion of test images and after that by comparing them with their corresponding RGB values of neutral images of the same test subject. Later, depending on the expression of consecutive frames a decision making algorithm is proposed to confirm the facial expression in case of a falsely detected expression. The second step involves the transmission part from one computer to another computer. The detected expression is sent to another computer for the purpose of imitating that expression at the receiving end. The last part is the expression imitation process in which corresponding expression is imitated in a robot face with the help of LEDs and Arduino interfacing.

A. Image Acquisition

In order to detect facial expression, real time image of a human face is needed to be acquired. To serve that purpose, CCD camera is used to acquire continuous images. Fig. 2 shows an example of continuous image acquisition of a test subject. Then, individual frames are extracted and fed to the next step to recognize expression for each one separately.

B. Expression Detection for Individual Frame

To avoid false expression detection, it is required to detect expression of individual frames and ensure that this is the correct expression. At first, Viola-Jones algorithm is implemented to detect face parts such as: mouth, eye-pair and nose of a subject. This algorithm is widely used for many cascade

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

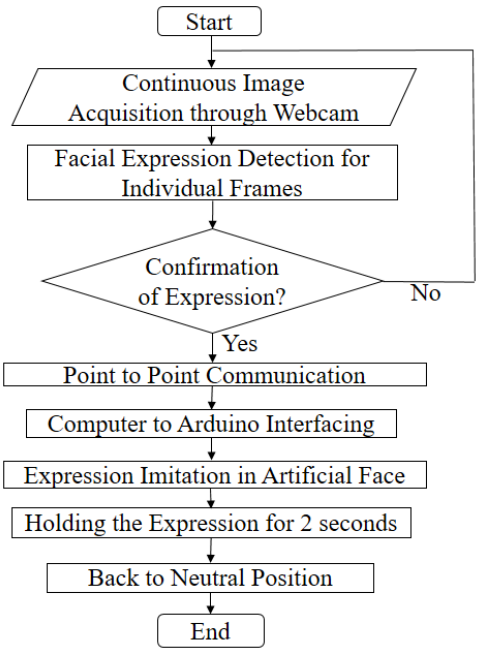


Fig. 1: Overview of system architecture.

object detection applications [6]. It is seen that for some subjects, multiple mouths or eye-pairs are detected considering environmental conditions (e.g., lighting, background, photo quality). Even though the number of multiple detection is too few, this multiple detected mouths or eye-pairs are on the opposite sides (up and down) of the nose. Hence, to remove this redundancy, coordinates of detected mouths or eye-pairs are compared with respect to coordinate of nose and falsely detected mouth or eye-pair is removed.

Next, from separate face parts, expressions are detected based on simple RGB values. For instance, smiling varies from person to person. But for simplicity, it is assumed that a person will reveal his teeth at the time of smiling. So, based on the RGB values of teeth, the areas of teeth are calculated and compared with his neutral image. If the teeth area of an image exceeds a certain threshold value (teeth area of neutral face), then the expression can be considered as happy. Similarly, it is expected that for a surprised face, the eye-ball area will increase with respect to neutral eye. Utilizing this technique, a surprised expression is detected. Last of all, for sad expression, it is assumed that the corner of the lips will get lower than the neutral lips corner. Therefore, distance of lips corner from nose is measured and compared with respect to neutral expression. If this distance gets increased by a certain amount then the image is considered as sad expression.

C. Decision Making

After detecting expression for an individual frame, decision of consecutive frames are compared to confirm which expression is shown by the test subject. If the detected expression is neutral, the system will continue to check the next frame. But in case of any other expression (except neutral), if that expression is detected for five or more consecutive frames, that expression is saved temporarily. Otherwise (non-neutral expression's number of frame is less than five), the system does not save any expression temporarily rather searches for next expression. This step helps to remove falsely detected expressions as any expression cannot stay in progress only for one or two image frames. After that, when expression is over, it is expected that, neutral expression will be back in the next frames. Therefore, this feature is incorporated in the algorithm that if three consecutive frames are detected as neutral (after a temporarily saved non-neutral expression), the temporarily saved expression will be confirmed and will indicate the end of expression. Later, this expression is sent to remote end for imitation. Once the whole process is completed the detection algorithm will search for new expressions.

TABLE I: Action Unit Number of Different Expressions

Emotion	Action Units	LED Color
Happiness	6+12	Green
Sadness	1+4+15	Red
Surprise	1+2+5+26	Yellow

TABLE II: Name of Each Action Unit

AU Number	FACS Name
1	Inner Brow Raiser
2	Outer Brow Raiser
4	Brow Lowerer
5	Upper Lid Raiser
6	Cheek Raiser
12	Lip Corner Puller
15	Lip Corner Depressor
26	Jaw Drop

D. Point to Point Communication

In this research work, sending and receiving points (computers in our case) are connected via physical links and communicate through a suitable program (MATLAB is used for implementation). The detected expressions from sending end computer is transmitted to the receiving end computer online.

E. Expression Imitation

Once an expression is received by the receiving end computer, that expression is required to be emulated on a humanoid robotic face. Suitably placed bidirectional motors (in the robotic face) will rotate in one direction to activate certain action units and after a specific time, will rotate in opposite direction to show the neutral expression again. Facial action unit system (FACS) is a system to express the facial muscle movement during any expression given by a person [7]. Table I shows the action unit numbers that get activated during any expression and Table II shows the name of action units and their corresponding numbers.

III. Experimentation

For experimentation purpose, 3 test subjects (of 3 different ages) are chosen for expression detection. Although there are six basic classes of expressions, this paper deals with these three expressions: happiness, sadness and surprise [8]. These particular three emotions were chosen as their clusters are most distant with respect to the neutral expression (see for example in [4]). The CCD camera connected with the sender's

Fig. 2: Continuous image acquisition (oldest to newest frame from left to right).

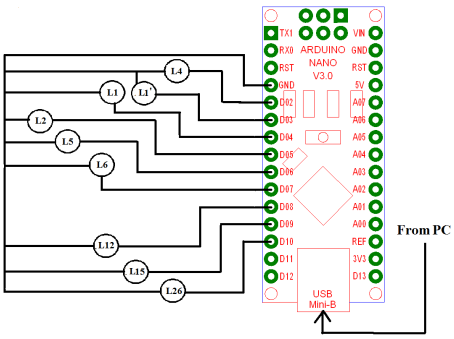


Fig. 3: Arduino connection diagram.

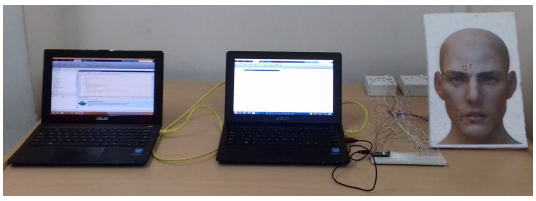


Fig. 4: Overall setup (from left to right: sender PC, physical link, receiver PC, Arduino, Robotic face).

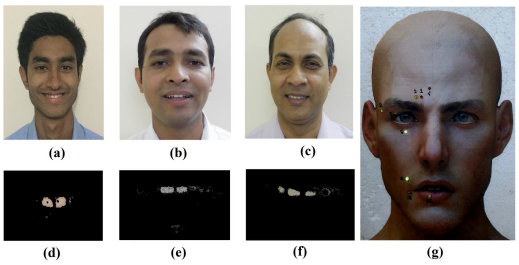


Fig. 5: Happy expression detection; (a), (b), (c) Happy expressions of the subjects; (d), (e), (f) Teeth areas of happy expressions; (g) Happy expression imitation on robotic face.

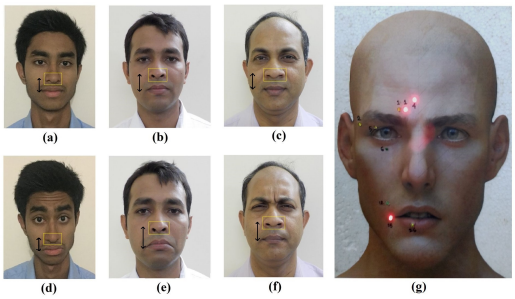


Fig. 6: Sad expression detection; (a), (b), (c) Sad expressions of the subjects (Nose to lip corner distance is shown); (d), (e), (f) Lip corner depression in sad expressions (Nose to lip corner distance increased); (g) Sad expression imitation on robotic face.

computer is used for continuous image acquisition at a rate of 5fps. After that, test subject's expressions are detected from continuous images, transmitted to receiving end computer and implemented there on a robotic face with the help of Arduino interfacing. Fig. 3 and Fig. 4, respectively, show the Arduino LED connection diagram and the setup of overall system. To detect expression from individual images, at first, Viola-Jones algorithm is used to detect face parts such as: mouth, eye-pair and nose. Then, individual expressions from images are detected using simple threshold based algorithm utilizing the RGB values of different face parts.

After detecting mouth of a subject, teeth is distinguished from its RGB values. Measurement shows that RGB values lie between 80 to 160 for teeth pixels for the considered subjects. Fig. 5((a)-(f)) shows the happy faces of the subjects and their corresponding detected teeth areas. These teeth areas are compared with the teeth areas of neutral faces. It is observed that in happy expression the teeth area is around 7 times greater. So, if the teeth area in an image is found 7 times or more than the teeth area in the neutral expression, then the expression is considered as happy. For a happy expression, action units 6 and 12 get activated which is shown with the help of green LEDs in Fig. 5(g). However, it is to be noted that teeth area does not increase that much for other expressions.

It is obvious that, for sad expression, the lips corner depresses. Experimentation shows that RGB values of lip corner lies between 25 to 35. Distance between nose level to the lip corner is measured. If the distance increases compared to their neutral value then, the subject is considered to be a sad. Fig. 6((a)-(f)) shows the neutral faces and respective changes in sad faces of all subjects. For a sad expression, action units 1, 4 and 15 get activated which is shown with the help of red LEDs in Fig. 6(g).

Similarly to detect surprised expressions, it is assumed that a persons eyeball will get bigger when he/she is surprised. Therefore, after properly detecting the eye-pair, the area of eyeball is detected using RGB values. Experimentation shows that RGB values lie between 90 to 150 for eyeball pixels. Fig 7((a)-(f)) shows the surprised faces of the subjects and their corresponding detected eyeball areas. After detecting the surprised eyeball areas, they are compared with the neutral eyeball areas and it is seen that in surprised images the eyeball areas is double or more than the eyeball area in the neutral image. This technique is used to detect surprised expressions. Action units 1, 2, 5 and 26 get activated for surprised expression, which is imitated in robotic face by

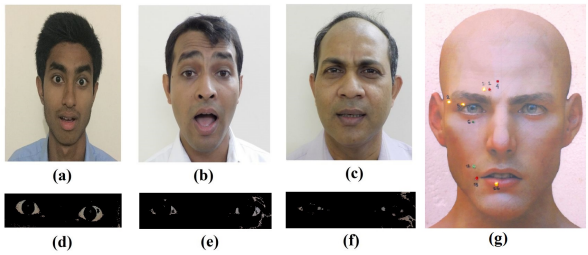


Fig. 7: Surprised expression detection; (a), (b), (c) Surprised expressions of the subjects; (d), (e), (f) Eyeball areas of surprised expressions; (g) Surprised expression imitation on robotic face.

yellow LEDs in Fig. 7(g). However, it is to be noted that for other considered expressions eyeball area remains same as neutral expression.

Once an expression is confirmed by the decision making algorithm, it is sent to the receiving end computer for imitation purpose. This transmission of expression is done via TCP/IP based socket communication with MATLAB [9]. The IP address of both host and server computer is required for this purpose.

After receiving the expression from the sending computer, the expression is intended to perform by bidirectional motors. But for better visual representation, LEDs are used instead of bidirectional motors. Three different color LEDs are lit for 2 seconds to show the activation of certain action units and thus imitating different expressions as explained in Table I. The LEDs are operated using arduino interfacing [10].

IV. Conclusion

In this paper, a simple but effective expression detection scheme in RGB color space has been proposed. The RGB values of lips, teeth, and eyeball have been found to be very effective for differentiating among different expressions. The simple threshold based scheme makes the process of expression detection and communication faster and involves less computation. In the experimentations on test subjects of different ages it is observed that the proposed method is capable of detecting almost all the expressions correctly with a rare exception. Therefore, it can be concluded that the proposed system can serve as an effective way to transmit real-time expressions on-line. The Arduino-based LED outputs on the FACS open the door of robotic imitation of expressions in the receiving end. In future, the system can be extended to communicate wirelessly using Bluetooth module (e.g., BLE v4.0). Moreover, other emotional expressions such as anger and disgust can be incorporated to the system along with happiness, sadness, and surprise. At the same time, the imitation of expressions can be confirmed as 'seen', when the receiving end uses motorized units on facial parts of robots.

References

[1] Y. Yang, S. S. Ge, T. H. Lee, and C. Wang, "Facial expression recognition and tracking for intelligent human-robot interaction," *Intelligent Service Robotics*, vol. 1, no. 2, pp. 143–157, 2008.

[2] S. S. Ge, "Social robotics: Integrating advances in engineering and computer science," in *Proceedings of Electrical Engineering/ Electronics, Computer, Telecommunications and Information Technology International Conference*, 2007, pp. xvii–xxvi.

[3] S. S. Ge, C. Wang, and C. C. Hang, "Facial expression imitation in human robot interaction," in *Proceedings of 17th IEEE International Symposium on Robot and Human Interactive Communication*, Munich, Germany, 2008, pp. 213–218.

[4] S. M. Imran, S. M. Mahbubur Rahman, and D. Hatzinakos, "Differential components of discriminative 2D Gaussian-Hermite moments for recognition of facial expressions," *Pattern Recognition*, vol. 56, pp. 100–115, 2016.

[5] H. Song, Y. M. Kim, J. C. Park, C. H. Kim, and D. S. Kwon, "Design of a robot head for emotional expression: Eeex," in *Proceedings of 17th IEEE International Symposium on Robot and Human Interactive Communication*, Munich, Germany, 2008, pp. 207–212.

[6] Y. Muttu and H. G. Virani, "Effective face detection, feature extraction neural network based approaches for facial expression recognition," in *Proceedings of International Conference on Information Processing (ICIP)*, Pune, India, 2015, pp. 102–107.

[7] FACS- Facial action coding system. [Online]. Available: https://www.cs.cmu.edu/~face/facs.htm [Accessed on August 30, 2016].

[8] P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion," *Journal of Personality and Social Psychology*, vol. 17, no. 2, pp. 124–129, 1971.

[9] Robotics, teaching and learning. [Online]. Available: http://www.lucidarme.me/?p=4087 [Accessed on August 30, 2016].

[10] Arduino. [Online]. Available: https://www.arduino.cc [Accessed on August 30, 2016].

Design of a Tunable Ring Resonator with Enhanced Quality Factor

Riazul Arefin*, Md. Omar Faruque, Rabiul Al Mahmud, Rakibul Hasan Sagor

Department of EEE, Islamic University of Technology
Board Bazar, Gazipur-1704, Bangladesh
*riazularefin@iut-dhaka.edu

Abstract—**A two dimensional structure consisting of a straight waveguide and a ring resonator is presented here. Both the ring and straight waveguide consist of two parts: core and cladding. Herein, a microring resonator based notch filter was simulated. By varying the core width of the ring resonator it was made tunable and thus the notch frequency of the filter was controlled. Material to build the ring resonator was selected as Quartz and $AgGaS_2$. An improvement in the quality factor was observed for the frequencies in the infrared and visible range. Variation of the parameters provided achievement of a quality factor of maximum 1325.224 and also its notch frequency was made tunable for a range of 0.0194 micrometre. The response of the ring resonator due to parameter variation was clearly marked during the simulation. This gives a good idea of the sensitivity of the device. Simulation software COMSOL Multiphysics was employed for this purpose.**

Index Terms—**waveguide, notch filter, ring resonator, quality factor.**

I. Introduction

The field of photonics is getting much more attraction due to its nanoscale electronics and broadband photonics [1]. Microwave photonic filters are attracting increasing interest than the traditional microwave filters because of their advantages like wider bandwidth, reconfiguration and tunability [2]. Among the photonic devices microring resonator is one of the most versatile device because it can serve as building block for massive integration of frequency selective components [3] which can be used as tunable bandwidth filters, laser cavities, and switches [4]. The microring resonator has also attracted tremendous interest in the past decade for optical detection of small biological or chemical entities in liquids [5] which can serve as a label free optical biosensor in specific fields like development of drug, medical diagnostic, environmental monitoring etc. [6].

Besides these applications, it is a good candidate for fabricating very large scale integrated (VLSI) photonic circuits as it provides a large range of optical signal processing capability [7].

One of the most useful applications of ring resonator is that it can act as a frequency selective filter [3]. Among the ring resonator filters, one notable device is the notch filter because of its ability to block undesired spectral content [8]. Optical filters are getting more consideration as they can act as demultiplexer in dense wavelength-division-multiplexed (DWDM) optical communication systems [9]. To configure the optical functions of a ring resonator, the most important property is tunability [10]. Resonant frequency can be tuned by several approaches like variation of internal radius of the ring, width of the ring, length of the gap [11]. Since it is considered as an important device to contribute to further integration of optical devices, continuous researches are going on to improve its efficiency and frequency selectivity [12, 13]. Some novel approaches have brought a path to show blue and red shift of the frequency selectivity of notch filter [14, 15]. Quality factor is also a big concern in this regard as more the variation of the quality factor can be achieved, it will provide a larger scope to integrate devices [14].

In this paper an approach has been proposed to have control over the notch frequency of a ring resonator based bandstop filter. The controlling of the quality factor is also studied here. A two dimensional microring resonator has been used where the width of the core was taken as the controlling parameter of the structure. The core width has been increased from 0.18 micrometre to 0.30 micrometre and the corresponding transmittance has been calculated.

II. Structure And Theoretical Overview

A. Basic Configuration and Theory

A simple ring resonator follows the principal of whispering gallery mode [16]. The basic configuration of a ring resonator is shown in the Fig. 1. It consists of a straight waveguide and a ring waveguide coupled together.

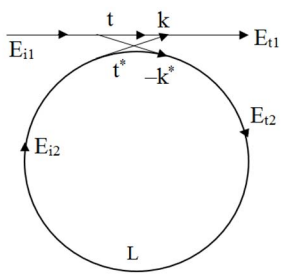


Fig. 1: Basic configuration of a ring resonator.

Excitation signal (E_{i1}) is provided in the input port and its coupling with the ring resonator can be described by the matrix relation:

$$\begin{bmatrix} E_{t1} \\ E_{t2} \end{bmatrix} = \begin{bmatrix} t & k \\ -k^* & t^* \end{bmatrix} \begin{bmatrix} E_{i1} \\ E_{i2} \end{bmatrix} \quad (1)$$

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

Here E_{t1} and E_{t2} are normalized complex mode amplitudes, couplers transmission coefficient and couplers coupling coefficient are denoted by t and k respectively, related by:

$$|t^2| + |k^2| = 1 \quad (2)$$

In the ring waveguide wave propagation can be described by:

$$E_{i2} = E_{t2} L\, e^{-j\Phi} \quad (3)$$

Here, L is the loss coefficient for the propagation around the ring and Φ is the accumulated phase.
Coefficient of transmission and its phase can be written as;

$$t = |t| e^{-j\Phi_t} \quad (4)$$

From equation (1), (2), (3) and (4) the expression of transmitted field can be written as;

$$E_{t1} = \frac{|t| - Le^{-j(\Phi-\Phi_t)}}{1 - |t| Le^{-j(\Phi-\Phi_t)}} E_{i1}\, e^{\Phi_t} \quad (5)$$

Here, the transmission coefficient is separated into the transmission loss $|t|$ and the corresponding phase Φ_t.

From the above equations it can be observed that when the phase difference $\Phi - \Phi_t$ is an integer multiple of 2π and $|t| = L$ the field is zero and thus the optimum condition for a notch filter is obtained [17] [18-20].

The quality factor Q can be calculated by using the following formula [21] :

$$Q = \frac{f_o}{f_H - f_L} \quad (6)$$

where, f_H is the higher cut off frequency and f_L is the lower cutoff frequency.

B. Proposed Structure

The proposed structure is almost similar to the basic ring resonator where a straight waveguide is coupled with ring waveguide. Both the straight and ring waveguide composed of two parts: the cladding and the core [17].

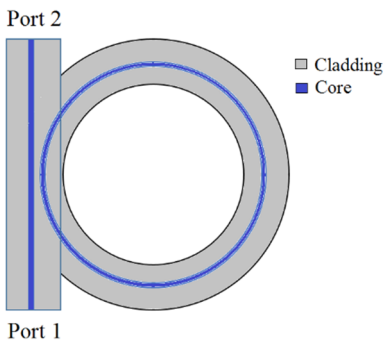


Fig. 2: Proposed structure of notch filter.

The core is composed of $AgGaS_2$ which is sandwiched by the Quartz cladding. The incident wave propagates through the core. The core and cladding are chosen in such a way that the phenomenon of total internal reflection of incident wave takes place at the interface. The refractive index of Quartz and $AgGaS_2$ are around 1.5 and 2.5 respectively [22].

In the proposed model, a ring was used having a radius of 6.2 micrometre with three inner layers. The structure is presented in Fig. 2. Blue and grey part represents core and cladding respectively. The middle layer was taken as the core of the structure. Width of that layer varied from 0.18 to 0.30 micrometre and thickness of the other two layers was 0.9 micrometre. The straight waveguide is also layered into three parts just like the ring waveguide. Excitation signal was given by lights having various wavelengths into port 1 and the output was recorded from port 2 depicted in Fig. 2.

III. Simulation Results And Discussion

For the simulation purpose COMSOL Multiphysics was used. The simulation results show the transmittance for different wavelengths. The core width has been varied and the corresponding changes in transmittance have been investigated. Fig. 3 shows the transmittance vs wavelength plot for core width of 0.18 micrometre. Here, three modes of the notch filter are presented. If we consider mode 3, the notch frequency was taken as 1.5816 micrometre and quality factor for this mode is 453.422.

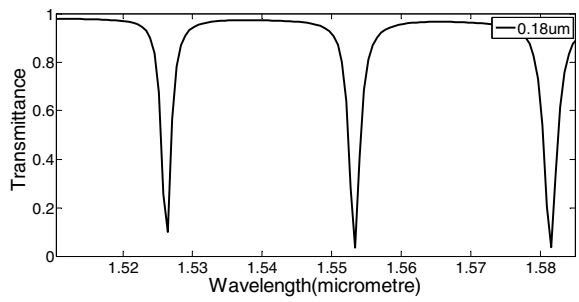


Fig. 3: Filter response for 0.18 micrometre core width.

A. Frequency Shifting Property

The frequency shifting property has been investigated with the simulation results. The notch frequency as well as the wavelength is shifting with varying core width which is shown in the Fig. 4(a) and 4(b). In Fig. 4(a) three modes for core width 0.18 micrometre and 0.22 micrometre has been plotted. It is found that, as the core width increases, the notch wavelength shows left shift.

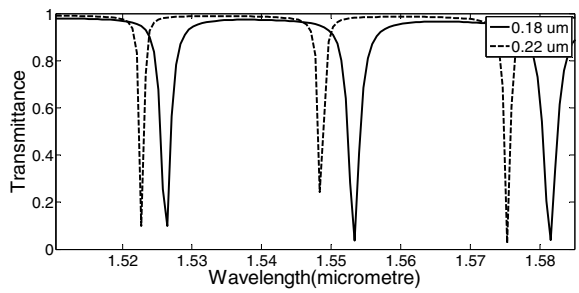


Fig. 4(a): Comparison of transmittance vs wavelength plot for 0.18 micrometre and 0.22 micrometre for three modes.

In Fig. 4(a) the solid and dashed lines represents the transmittance for 0.18 micrometre and 0.22 micrometre core width. It shows the amount of shift in notch wavelength due to increase in the core width. In Table 1 notch wavelength shifting characteristics is shown. It is seen that, up to core width 0.26

micrometre notch wavelength shifts toward left and after that it exhibits non-linearity.

Frequency shifting characteristics is also depicted in Fig. 4(b). We see that different core width shows different transmittance characteristics.

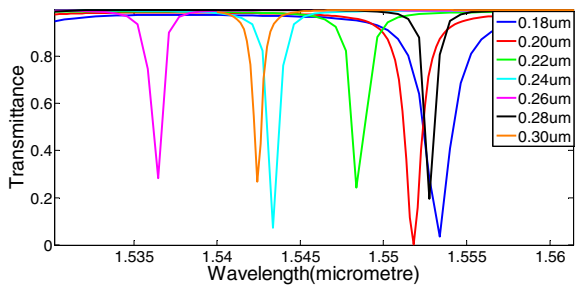


Fig. 4(b): Comparison of transmittance plot for 0.18 micrometre - 0.30 micrometre core width showing one mode with interval 0.02 micrometre.

If the core width increases more than 0.26 micrometre, figured out in Fig. 4(c), linearity property does not exist. We find a negative slope from core width 0.18 micromete to 0.26 micrometre.

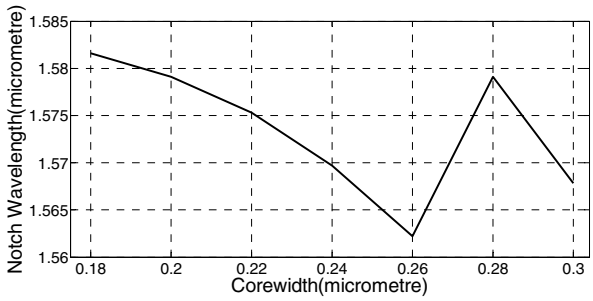


Fig. 4(c): Notch wavelength vs core width.

The numerical values of notch wavelengths are shown in Table 1.

Table 1: Notch wavelength shifting property

Core width (micrometre)	Notch wavelength (micrometre)
0.18	1.5816
0.20	1.5792
0.22	1.5753
0.24	1.5697
0.26	1.5622
0.28	1.5791
0.30	1.5679

B. Observation of Quality Factor

From the simulation results, the quality factors of the notch filters were calculated using equation (6). The quality factor has been found to be linearly increasing with the increase of core width.

Obtained quality factors for various core width have been shown in Table 2. It is observed that for core width more than 0.22 micrometre the quality factor is as high as 1000 which can be considered as a reliable one in most of the applications [21].

Table 2: Quality factor variation

Core width(micrometre)	Quality factor
0.18	453.422
0.20	678.150
0.22	985.034
0.24	1266.932
0.26	1212.276
0.28	1325.224
0.30	1274.792

From the simulation results, linearity in the improvement of the quality factor was observed upto core width 0.24 micrometre shown in Fig. 5. Nonlinearity appears from core width 0.26 micrometre with high values of quality factors which are still reliable.

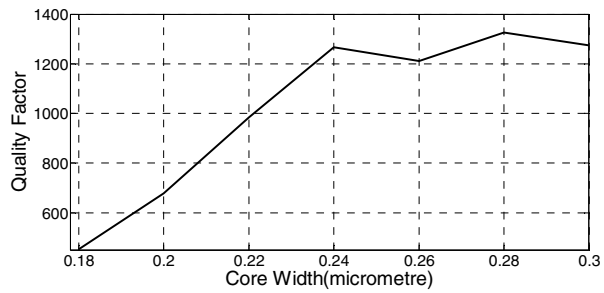


Fig. 5: Variation of quality factor for different core widths.

IV. Conclusions

In this work, a new notch filter based on a ring resonator has been proposed. In micro scale optics, to design new devices, it is important to have control over the notch frequency of a filter within a bandwidth to reject one channel while allow other channels to pass. This paper proposes a new technique to control this by varying the core width. At the same time, a new phenomenon was observed. Varying the core width also controls the quality factor within a range. So in particular, it can be said that, this paper has the potential to serve as a building block to fabricate new optical devices.

References

[1] R. H. Sagor, *et al.*, "Extraction of modeling parameters for low-loss alternative plasmonic material," *Procedia-Social and Behavioral Sciences,* vol. 195, pp. 2061-2066, 2015.

[2] J. Palací, *et al.*, "Single bandpass photonic microwave filter based on a notch ring resonator," *IEEE Photonics Technology Letters,* vol. 22, pp. 1276-1278, 2010.

[3] P. Absil, *et al.*, "Compact microring notch filters," *IEEE Photonics Technology Letters,* vol. 12, pp. 398-400, 2000.

[4] X.-Y. Zhang, *et al.*, "Tunable microring resonator based on dielectric-loaded surface plasmon polariton waveguides," *Journal of nanoscience and nanotechnology,* vol. 11, pp. 10520-10524, 2011.

[5] K. Q. Le and P. Bienstman, "Nanoplasmonic ring resonator for biosensing applications," in *15th Annual Symposium of the IEEE Photonics Benelux*.

[6] T. Claes, *et al.*, "Label-free biosensing with a slot-waveguide-based ring resonator in silicon on insulator," *IEEE Photonics journal,* vol. 1, pp. 197-204, 2009.

[7] B. Little, *et al.*, "Microring resonator arrays for VLSI photonics," *IEEE Photonics Technology Letters,* vol. 12, pp. 323-325, 2000.

[8] M. S. Rasras, *et al.*, "Demonstration of a tunable microwave-photonic notch filter using low-loss silicon ring resonators," *Journal of Lightwave Technology,* vol. 27, pp. 2105-2110, 2009.

[9] C. Chen, *et al.*, "Bandpass filters based on phase-shifted photonic crystal waveguide gratings," *Optics express,* vol. 15, pp. 11278-11284, 2007.

[10] M.-C. Lee and M. C. Wu, "MEMS-actuated microdisk resonators with variable power coupling ratios," *IEEE Photonics Technology Letters,* vol. 17, pp. 1034-1036, 2005.

[11] A.-N. A. Jabita, "Design of singly split single ring resonator for measurement of dielectric constant of materials using resonant method," 2013.

[12] W.-j. Lui, *et al.*, "Improved frequency notched ultrawideband slot antenna using square ring resonator," *IEEE Transactions on Antennas and Propagation,* vol. 55, pp. 2445-2450, 2007.

[13] F. Vollmer and S. Arnold, "Whispering-gallery-mode biosensing: label-free detection down to single molecules," *Nature methods,* vol. 5, pp. 591-596, 2008.

[14] K. Djordjev, *et al.*, "Microdisk tunable resonant filters and switches," *IEEE Photonics Technology Letters,* vol. 14, pp. 828-830, 2002.

[15] T.-B. Wang, *et al.*, "The transmission characteristics of surface plasmon polaritons in ring resonator," *Optics express,* vol. 17, pp. 24096-24101, 2009.

[16] Y.-F. Xiao, *et al.*, "Coupling whispering-gallery-mode microcavities with modal coupling mechanism," *IEEE Journal of Quantum Electronics,* vol. 44, pp. 1065-1070, 2008.

[17] Comsol, *COMSOL Multiphysics: Version 3.3*: Comsol, 2006.

[18] A. Yariv, "Critical coupling and its control in optical waveguide-ring resonator systems," *IEEE Photonics Technology Letters,* vol. 14, pp. 483-485, 2002.

[19] A. Yariv and P. Yeh, *Photonics: optical electronics in modern communications (the oxford series in electrical and computer engineering)*: Oxford University Press, Inc., 2006.

[20] D. G. Rabus, *Integrated ring resonators*: Springer, 2007.

[21] V. Braginsky, *et al.*, "Quality-factor and nonlinear properties of optical whispering-gallery modes," *Physics Letters A,* vol. 137, pp. 393-397, 1989.

[22] E. Takaoka and K. Kato, "Thermo-optic dispersion formula for AgGaS 2," *Applied optics,* vol. 38, pp. 4577-4580, 1999.

A Single Phase SEPIC AC-DC Converter with Improved Power Factor and Input Current THD

Mohammad Rubaiyat Tanvir Hossain,[1,*] Amina Hasan Abedin,[2,$] and M. A. Choudhury[3,‡]

[1]Department of Electrical and Electronic Engineering, Chittagong University of Engineering and Technology
Chittagong-4349, Bangladesh

[2]Department of Electrical and Electronic Engineering, BRAC University
66 Mohakhali, Dhaka-1212, Bangladesh

[3]Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology
Dhaka-1205, Bangladesh

Email: *mrthossain@cuet.ac.bd, $prity@bracu.ac.bd, ‡mac@eee.buet.ac.bd

Abstract—**A new topology of single-phase SEPIC (Single Ended Primary Inductor Converter) AC-DC converter with low input current THD and high input power factor is proposed. Instead of using a single phase rectifier followed by a boost DC-DC converter, or using a bridge rectifier with two diodes and two switches, the rectifier's input is chopped at high frequency during positive and negative cycles by a single bi-directional switch to get step-up/step-down AC-DC conversion. Total Harmonic Distortion (THD) of the input current of the proposed circuit with small input current filter is low. Input power factor and the efficiency of conversion of the circuit are high. No additional control is required to achieve the advantages.**

Index Terms—ac-dc converter, sepic converter, power factor correction, total harmonic distortion, efficiency.

I. Introduction

Single-phase full wave rectifiers in bridge configuration suffer from problems of non-sinusoidal input current and low input power factor. Various methods have been proposed to solve these problems. One solution is to use input filter. Filter inductor and capacitor required in such solution are large. It improves THD but power factor remains low. To overcome the problems, switch mode regulators have been proposed. Conventionally a single-phase AC-DC switch mode converter is made of a bridge rectifier followed by a DC-DC converter [1]-[6]. The most common of these is a single-phase rectifier followed by a Boost DC-DC converter. But Buck, Buck-Boost, Ĉuk and SEPIC converters may be employed for the same purpose with different input/output voltage gain relationships [6]-[10]. In DC-DC converter regulated rectifiers the bridge rectifier converts AC to uncontrolled DC and the DC-DC converter provides controlled and regulated DC output. The output DC-DC converter acts as high frequency output current/voltage chopper which is reflected at the input as high frequency chopped AC current [1]-[2]. High frequency chopped input current is then filtered by a small filter to obtain near sinusoidal input current with high input power factor [1]-[2]. This is possible only when the output is full wave pulsed DC. If the output is desired to be DC with low ripple, large output capacitor filter is necessary. This large capacitor filter draws pulsed current, and the pulsed AC current is reflected at the input side. Additional control is then required to maintain high frequency chopped sinusoidal shape of the input current.

The DC-DC converters use unidirectional switch for the switching purpose. In the boost-regulated AC-DC conversion, the boost rectifier must be operated in critical mode [11]-[13] where, the power switch should be turned ON at the instant of boost diode's current reaches zero. This needs variable switching frequency operation of the DC-DC converter as the load or the input voltage change. Another approach for boost-regulated rectifier involves controlling to a constant level the average current of the boost diode [11]-[13]. In order to keep the average current constant through the boost diode, the duty cycle must be modulated over the line cycle. This method requires current sensor in the scheme. Bridge-less configurations [14], [15] and two-diode, two-switch rectifiers [1] are also reported in literatures for AC-DC conversion having the above advantages and disadvantages of boost-regulated rectifiers. The reported bridgeless single phase AC-DC converters use more than one unidirectional switches or one bidirectional switch composed of two unidirectional switches anti-parallels with two diodes [16]-[18].

In this paper, a new single-phase AC-DC converter topology based on SEPIC (Single Ended Primary Inductor Converter) conversion is reported. SEPIC converter has the advantages of having non-inverted output (the output voltage is of the same polarity as the input voltage) that can operate in either buck or boost mode, Moreover, it offers isolation between its input and output (provided by a capacitor in series), inherent inrush current limitation during startup and overload conditions, lower input current ripple, and less electromagnetic interference (EMI) associated with the DCM topology [18]-[20]. The proposed converter chops input AC current by a single bidirectional switch. The circuit differs from the mentioned DC-DC converter regulated rectifiers or bridgeless rectifiers in two senses. The first is the use of a single bidirectional switch composed of a unidirectional switch and proper positioning of diodes. The second is the input current switching. The second feature ensures high constant frequency switching of input AC almost in phase with input AC voltage without any additional control scheme ensuring little dependence on the input voltage and the output load. This would result shaping of input current to near sinusoidal by use of small filter and will ensure good input power factor without having much impacts of output filter, input voltage and change of load.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

II. Proposed SEPIC Regulated Single Phase AC-DC Converter

Fig. 1 shows the proposed circuit for AC-DC conversion using SEPIC topology. The circuit consists of five diodes (D1 to D5), three inductors (L1_p, L1_n and L2), two capacitors (C1 and C2), a switch (S) and a resistive load (RL). A filter is also connected at the input side formed with an inductor (Lif) and Capacitor (Cif). Since the supply voltage is AC, conversion on both positive and negative cycle of the input voltage waveform is required. The conversion is achieved by the power switch 'S' and proper positioning of diodes. An extra inductor L1_n is used in the circuit which provides current path on the negative half cycle of the source voltage. The operation of the circuit is given in the section below.

III. Principle of Operation

Input AC chopping at high frequency provides switched AC current that requires small filter to make the current waveform nearly sinusoidal and in phase with input voltage. As a result, the input current THD reduces and the power factor improves. High frequency conversion for both positive and negative cycles of the input signal achieves this goal. In the proposed topology AC source is used as the input, so bi-directional switching is necessary.

The circuit has four states according to ON-OFF condition of the switch in the positive half cycle and negative half cycle of the supply. These four states are shown in Fig. 2, along with the direction of current flow. The circuit performs SEPIC conversion of the positive half cycle of the supply by state 1 and state 2, while it performs conversion of the negative half cycle of the supply by state 3 and state 4 respectively. The principle of operation of the proposed circuit is described below.

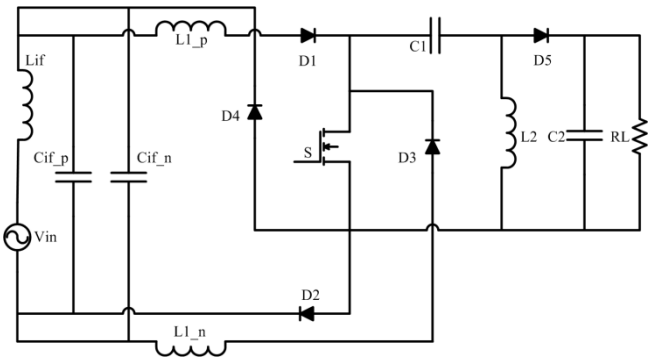


Fig. 1 Proposed single phase SEPIC AC-DC converter

State 1. During the positive cycle of the input signal when the switch S is turned on, as shown in Fig. 2(a), the energy from the source Vin, is stored in the inductor L1_p, and the capacitor C1 transfer its energy to the inductor L2. During this stage the diode D5 is kept blocked and the capacitor C2 supplies energy to the load.

State 2. In the positive cycle of the input signal when the switch S is turned off and the diode D5 is turned ON, the inductor L1_p charges the capacitor C1 and also provides current to the load, as shown in Fig. 2(b). The inductor L2 is also connected to the load during this time.

State 3. Similar to State 1 operation, during the negative cycle of the input signal when the switch S is turned on, inductor L1_n is charged from the source, and the inductor L2 is charged from capacitor C1. The capacitor C2 supplies energy to the load.

State 4. In the negative cycle of the input signal when the switch S is turned off and the diode D5 is turned on, the inductor L1_n and L2 transfer their storage energy to the load by charging C2 like State 2 operation in the positive cycle.

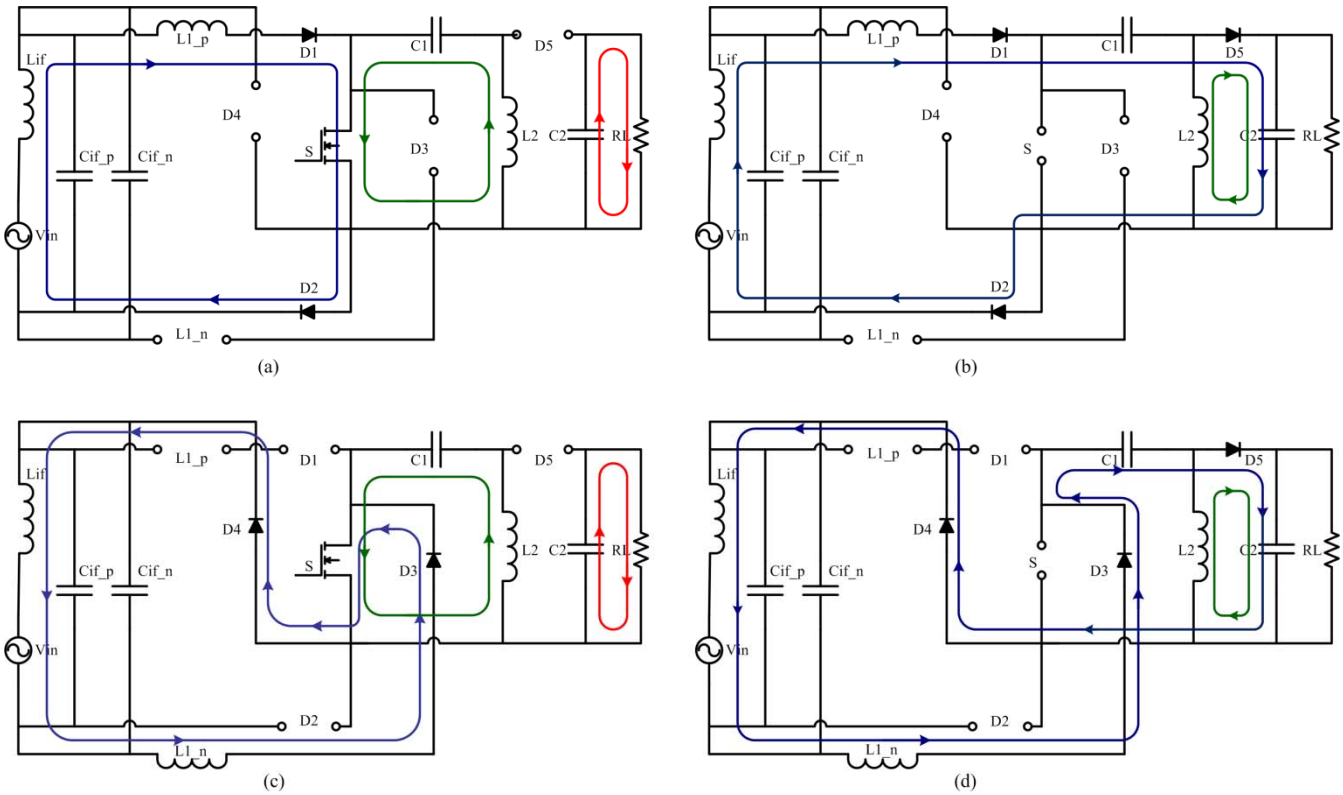


Fig. 2 Four states of operation of proposed single phase AC-DC converter, (a) State 1, circuit when switch S is ON during positive half cycle of the AC supply, (b) State 2, circuit when switch S is OFF during positive half cycle of the AC supply, (c) State 3, circuit when switch S is ON during negative half cycle of the AC supply, (d) State 4, circuit when switch S is OFF during negative half cycle of the AC supply.

IV. Simulation and Results

The simulation of the proposed circuit is performed with Psim professional version 9.0.3.400. The source is an AC source with input frequency of 50 Hz having amplitude of 300V and an IGBT is considered as the switching device. The switching frequency is chosen as 6 kHz. The inductor and capacitor values for the circuit in Fig. 1 are L1_p=L1_n= 2mH, L2= 2mH, C1= 1μF, Co= 20μF. The inductor Lif= 5mH and capacitors Cif_p= Cif_n= 1μF have been used to work as input filter both in positive and negative cycle of the supply voltage. A resistor of RL= 100Ω has been used as load in the circuit. The proposed circuit has been compared with conventional single phase AC-DC diode bridge SEPIC rectifier topology using the same circuit parameter values. Simulation has been performed and typical results are presented in the next subsection.

A. Results from Simulation

Fig. 3 shows typical simulation results of the waveforms for output and input voltages and input current for the proposed single phase AC-DC rectifier using SEPIC topology. Frequency spectrum of the input current is illustrated in Fig. 4 which shows the fundamental component and small magnitude of current harmonics near 6 kHz. It is evident that the proposed circuit can reduce the harmonic distortion.

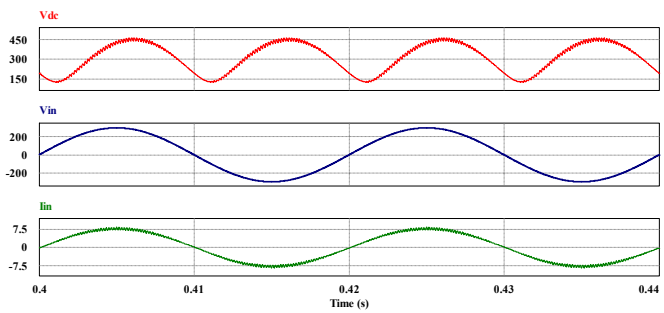


Fig. 3 Typical output voltage (Vdc), input voltage (v_{in}) and input current (i_{in}) wave shapes of the proposed scheme at 50% duty cycle

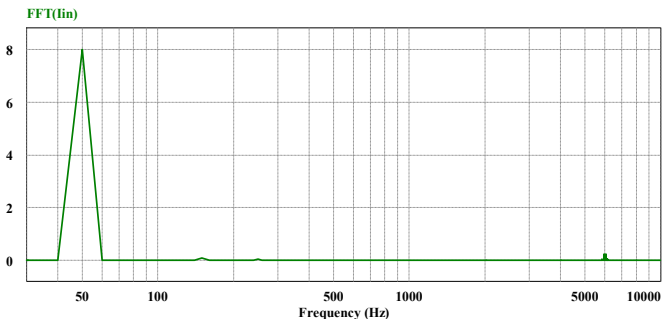


Fig. 4 Frequency Spectrum of input current (i_{in}) at 50% duty cycle

B. Performance Under Duty Cycle Variation

The performances of the proposed scheme is compared with conventional single phase AC-DC SEPIC converter of Fig. 5 in which a power switch exists between the bridge rectifier and load. In this conventional circuit high frequency chopping is performed on the rectified output prior to going at the load. Performances of both proposed and conventional circuits are compared for same circuit parameter values. Moreover, the switching frequency is also maintained the same (6 kHz). The results of these comparisons are presented in Fig. 6 to 8.

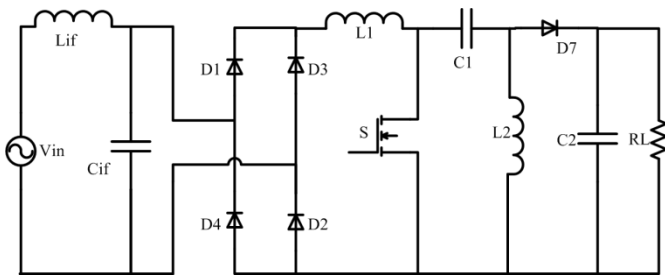


Fig. 5 Conventional single phase diode bridge SEPIC AC-DC converter.

The comparison of total harmonic distortion (THD) of input current for the conventional and proposed method under various duty cycles is presented in the bar chart as shown in Fig. 6. The result depicted in the chart shows that input current THD (%) of the proposed rectifier circuit is much better than that of conventional one throughout the range of duty cycle. THD (%) of the conventional circuit may be reduced by employing extra control scheme, whereas, the proposed scheme has low THD (%) without such control schemes for the load under consideration.

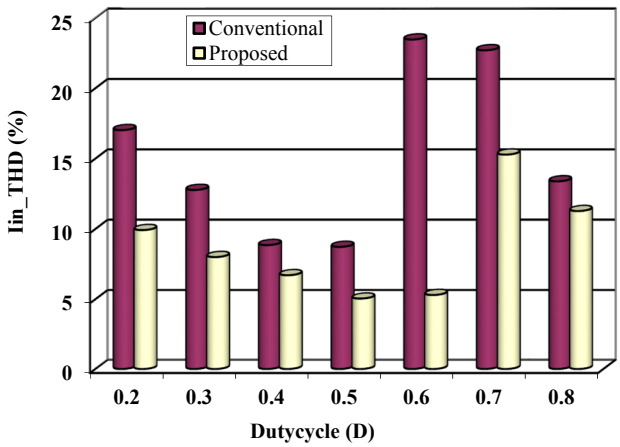


Fig. 6 Comparison of THD (%) of input current between conventional and proposed scheme

For input power factor, Fig. 7 shows that the proposed scheme gives high input power factor all over the range of duty cycles. The power factor of the proposed circuit remains almost constant near unity whereas the same of the conventional circuit degrades at higher duty cycles (D > 0.5). However, within 50 % of the duty cycle the performance of the conventional scheme is compatible with the proposed scheme.

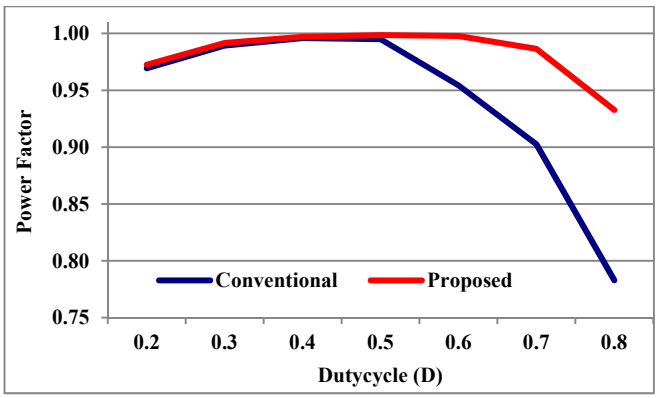


Fig. 7 Comparison of input power factor between conventional and proposed scheme

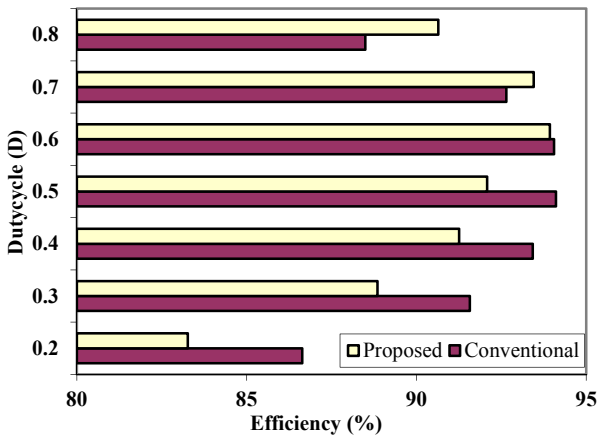


Fig. 8 Comparison of efficiency (%) between conventional and proposed scheme

The performance of the circuits in terms of efficiency (%) is presented by the horizontal bar diagram as illustrated in Fig. 8. For duty cycle in the range from 20% to 50% the conventional circuit gives slightly better efficiency than the proposed circuit. Nevertheless the difference is reduced with increase in duty cycle and after D>0.5 the efficiency of the proposed circuit is superior.

C. Performance Under Load Variation

Both conventional and proposed single phase AC-DC Sepic converter circuits are subjected to load variation at constant switching frequency and duty cycle. The load resistance is varied from 20Ω to 100Ω and the performance is monitored in terms of power factor, THD (%), efficiency (%) and voltage gain. Table I shows the outcome of this analysis at the switching frequency set to 6 kHz and duty cycle maintained at 50%. The proposed and the conventional circuits maintain high power factor and efficiency with variable load values. Performance of the proposed circuit is consistently better. The THD (%), of the proposed circuit does not vary much under load variation while in the conventional circuit THD (%) increases with increase in load.

TABLE I

PERFORMANCE COMPARISION UNDER LOAD VARIATION

Load (Ω)	Proposed Sepic				Conventional Sepic			
	Power factor	THD (%)	η (%)	Voltage gain	Power factor	THD (%)	η (%)	Voltage gain
20	1.00	4.59	94.00	0.67	0.99	4.98	93.92	0.55
40	1.00	4.89	94.46	0.94	0.99	9.84	94.88	0.69
60	1.00	5.06	94.26	1.14	0.99	13.56	94.52	0.77
80	1.00	5.15	94.06	1.32	1.00	9.87	94.15	0.87
100	1.00	5.01	93.80	1.48	1.00	6.94	93.83	0.96

V. CONCLUSIONS

A new scheme for single phase AC-DC conversion based on SEPIC topology is proposed. The circuit chops the input current at the AC side in contrast to the conventional scheme which chops the rectified output. Comparison with conventional circuit has been made. The proposed scheme demonstrates better performance having lower input current THD, higher power factor and high efficiency. Moreover, it does not require any additional control schemes in order to reduce the harmonics. Based on the results obtained from simulation, the proposed scheme can be considered for AC-DC conversions with low distortion and high input power factor and efficiency.

REFERENCES

[1] B. Singh, B. N. Singh, A. Chandra, K. Al-Haddad, A. Pandey, and D. P. Kothari, "A review of single-phase improved power quality AC-DC converters," *IEEE Transactions on Industrial Electronics*, vol. 50, no. 5, pp. 962 – 981, Oct. 2003.

[2] O. Garcia, J. A. Cobos, R. Prieto, P. Alou, and J. Uceda, "Single phase power factor correction: a survey," *IEEE Transactions on Power Electronics*, vol. 18, no. 3, pp. 749-755, May 2003.

[3] H. Wei and I. Batarseh, "Comparison of basic converter topologies for power factor correction," in *Proc. Southeastcon'98*, 1998, pp. 348-353.

[4] J. Salmon, "Techniques for minimizing the input current distortion of current-controlled single-phase boost rectifiers," *IEEE Transactions on Power Electronics*, pp. 509–520, 1993.

[5] R. W. Erickson and M. Madigan, "Design of a simple high- power-factor rectifier based on the flyback converter," in *Proc. IEEE Applied Power Electronics Conference (APEC)*, 1990, pp. 792–801.

[6] D. Tollik and A. Pietkiewicz, "Comparative analysis of 1-phase active power factor correction topologies," in *Proc. Int. Telecommunication Energy Conference*, Washington DC, USA, Oct. 1992, pp. 517-523.

[7] O. L´opez, L. Vicu˜na, M. Castilla, J. Matas, and M. L´opez, "Sliding-mode-control design of a high-power-factor buck–boost rectifier," *IEEE Transactions on Industrial Electronics*, vol. 46, no. 3, pp. 604 – 612, Jun. 1999.

[8] M. Brkovic and S. Cuk, "Input current shaper using Cuk converter," in *Proc. 14th International Telecommunications Energy Conference*, INTELEC '92, Oct. 1992, pp. 532- 539.

[9] P. F. Melo, R. Gules, E. F. R. Romaneli, and R. C. Annunziato, "A modified SEPIC converter for high-power-factor rectifier and universal input voltage applications," *IEEE Transactions on Power Electronics*, vol. 25, no. 2, pp. 310-321, Feb. 2010.

[10] K. Singh, A. N. Tiwari, K. P. Singh, "Performance analysis of modified sepic converter with low input voltage," *International Journal of Electronics & Communication Technology (IJECT)*, vol. 3, no. 1, pp. 21-25, Jan.-Mar. 2012.

[11] D. Maksimovic, J. Yungtaek, and R. W. Erickson, "Nonlinear-carrier control for high power factor boost rectifiers," in *proc. Applied Power Electronics Conference and Exposition*, Dallas, Texas, USA, Mar. 1995, vol. 2, pp. 635-641.

[12] T. Tanitteerapan and E.Thanpo, "Negative slope ramp carrier control for high power factor boost converters in ccm operation," *World Academy of Science, Engineering and Technology*, pp. 1084-1089, 2009.

[13] M. Kazerani, P. D. Ziogas and G. Joos, "A novel active current wave shaping technique for solid-state input power factor conditioners," *IEEE Transactions on Industrial Electronics*, vol. 38, no.1, pp. 72-78, Feb. 1991.

[14] L. Huber, Y. Jang, and M. Jovanovic, "Performance evaluation of bridgeless pfc boost rectifiers," *IEEE Transactions on Power Electronics*, vol. 23, no. 3, pp. 1381–1390, May 2008.

[15] W. Y. Choi, J. M. Kwon, E. H. Kim, J. J. Lee, and B. H. Kwon, "Bridgeless boost rectifier with low conduction losses and reduced diode reverse-recovery problems," *IEEE Transactions on Industrial Electronics*, vol. 54, no. 2, pp. 769–780, Apr. 2007.

[16] A. J. Sabzali, E. H. Ismail, M. A. Al-Saffar, and A. A. Fardoun, "New bridgeless dcm sepic and cuk pfc rectifiers with low conduction and switching losses," *IEEE Transactions on Industry Applications*, vol. 47, no. 2, pp. 873-881, Mar./Apr. 2011.

[17] M. R. Sahid, A. H. M. Yatim, and T. Taufik, "A new ac-dc converter using bridgeless sepic," in *Proc. 36th Annual Conference on IEEE Industrial Electronics Society (IECON 2010)*, Nov. 2010, pp. 286 – 290.

[18] E. H. Ismail, "Bridgeless sepic rectifier with unity power factor and reduced conduction losses," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 4, pp. 1147-1157, Apr. 2009.

[19] V. Naithani, A. N. Tiwari, and S. Dobhal, "Simulation of sepic converter fed leds," *International Journal of Engineering Science and Technology (IJEST)*, ISSN: 0975-5462, vol. 4, no.03, pp. 1015-1020, Mar. 2012.

[20] S. Singh, and B. Singh, "Single-phase sepic based pfc converter for pmbldcm drive in air-conditioning system," *Asian Power Electronics Journal*, vol. 4, no.1, pp. 16-21, Apr. 2010.

Modified H-Bridge Multilevel Inverter for Photovoltaic Micro-grid Systems

Md Multan Biswas,[1,*] Shuvankar Podder,[2] and Md. Ziaur Rahman Khan[2]

[1]Department of Electrical Engineering, University of South Carolina
Columbia, SC 29208, USA

[2]Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology
Dhaka 1000, Bangladesh

[*]mbiswas@email.sc.edu

Abstract—**A modified multilevel inverter (MLI) structure has been presented for Photovoltaic (PV) fed 2.3 kV micro-grid applications. To feed the cascaded multilevel micro-grid connected inverter (CM-MGCI) voltage, a PV array is designed considering the environmental effects of Bangladesh. We have estimated the performance of the classical and a modified cascaded H-bridge (CHB) MLI with level-shifted carrier sinusoidal pulse width modulation (LSC-SPWM) technique in MATLAB/Simulink environment. The voltage and total harmonic distortion (THD) profile for the proposed eleven-level modified H-bridge MLI topology has been compared with the conventional CHB topology. The harmonic profile has also been compared with other modified MLI topologies. Compared to the conventional CHB and other recent modified MLI topologies, the proposed H-bridge CM-MGCI shows better THD profile.**

Index Terms—**Cascaded multilevel inverter; harmonics distortion; level-shifted pulse width modulation; micro-grid; photovoltaic.**

I. Introduction

At the present time, the application of micro-grid connected PV systems is broadly preferable compared to standalone PV systems [1-3]. With the reduction of installation cost and other financial supports, the capacities of PV power stations are continuously mounting. In this consequence, more researches are being carried out on PV-fed micro-grid connected inverters. With the advancement of new power electronics devices, digital controllers, and sensors, the role of voltage source inverters is also envisaged and acknowledged in frontiers such as greater penetration of PV based power generation and futuristic smart power grids in Bangladesh [3], [4].

Due to the higher harmonics distortion in conventional bipolar inverters, the multilevel inverters (MLI) topologies are extensively used in renewable energy system applications in recent times [5-7]. Different types of MLI configurations have been proposed in the literatures [8-12]. Among the three conventional MLIs, the diode clamped MLI and capacitor clamped MLI uses higher number of power switches than the cascaded H-bridge (CHB) MLI. Thus, cascade H-bridge inverter topology is more popular for micro-grid connected PV applications. However, due to a considerable large number of power switches and related switching and conduction losses, the efficiency of the CHB MLIs are still low and costs are still high.

The domain of cascaded multilevel micro-grid connected inverter (CM-MGCI) and reduced switch configuration in MLI are gaining its attention in terms of diminishing expenses and operational optimization. From such a motivation, some researchers are addressing their efforts in outlining new MLI configurations or in modifying the existing ones, aiming at enhancing the power quality available at the inverter output. Some asymmetrical approaches of reduced number of power switch for MLIs are presented in their literatures. However, for medium voltage PV fed grid application, asymmetrical approach is not appropriate due to high voltage stress and higher THD [5], [7]. Also, for same level of output voltage, the power losses in semiconductor components are more than twice for asymmetrical topology than symmetrical one [6], [8]. In this work, the design and analysis of the performance of a symmetrical CM-MGCI is presented. A comparison of power switch requirement and harmonic distortion with classical CHB and other modified MLIs topologies is also provided to explain the advantages of the proposed circuit.

II. PV-Fed CM-MGCI System

For PV based medium voltage micro-grid systems CHB MLI is prominently used. The functional block diagram of the PV-fed MLI system is illustrated in Fig. 1. The inverter input voltage is maintained at constant value by using maximum power point tracking (MPPT) controller with a DC-DC converter. Despite of environmental variations, the nearly fixed output voltage from DC-DC converter is fed to the inverter to connect with ac micro-grid system.

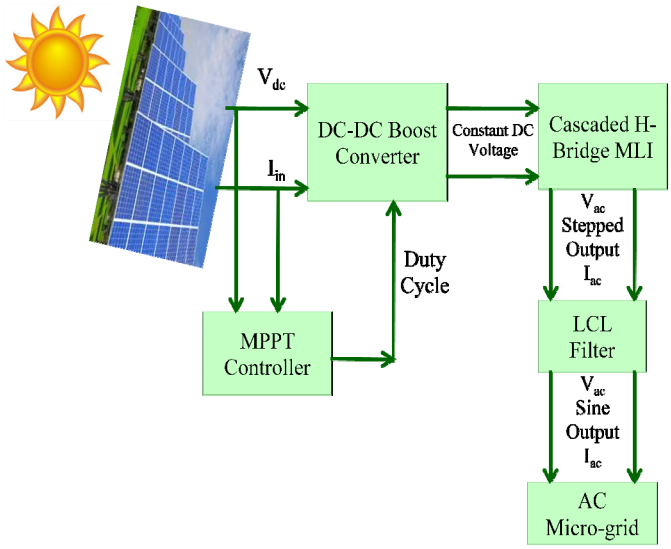


Fig. 1. Photovoltaic system with CM-MGCI.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

III. Calculation of PV Array

In this work Sunpower E20/435 solar panel has been used for modeling and simulation of 600 kW three phase 2.3 kV micro-grid connected (MGC) PV system. The principal goal of this work is to simulate the performance of this 600 kW three phase PV fed MGC eleven-level conventional and modified H-bridge inverters. Accordingly, the single phase power of the ac micro-grid becomes 200 kW with line to line voltage of 2.3 kV and frequency of 50 Hz. Then the peak voltage for single phase can be calculated as follows:

$$v_{pm} = \frac{V_{ll}}{\sqrt{3}}\sqrt{2} = \frac{2.3}{\sqrt{3}}\sqrt{2} = 1.878 \text{ kV} \qquad (1)$$

To simulate the design and MGC eleven-level MLIs performance an array of 5 strings is considered in each phase and the strings are connected in series. For high efficiency operation of higher order MLIs, to stipulate a grid phase voltage of v_p the inverter dc link voltage should be limited as

$$\frac{(m-1)v_{pm}}{m-2} \le V_{dc} \le \frac{(m-1)v_{pm}}{m-5/2} \qquad m>5 \qquad (2)$$

where m is the inverter output voltage level. Since the peak value of phase voltage of the micro-grid is 1.878 kV and 5 strings will be simulated, so each string voltage should have a value in between 417.3 V to 441.88 V.

In Bangladesh the ambient temperature varies generally from 10 ^{0}C to 36 ^{0}C over the year. Since, the maximum power voltage significantly decreases with the increase of environment temperature so solar standard test conditions (STC) string voltage should have sufficient magnitude to stipulate required output grid voltage. Again, the PV module output power also diminishes with decrease in solar irradiation and increase in cell operating temperature. Thus the total PV array must have considerably large power than the station's rated output power. To fulfill this requirement each string is subdivided into 6 modules characterized by a rated optimal voltage of 72.9 V connected in series and 17 modules with a rated optimal current of 5.95A and connected in parallel. Thus, the total output voltage of single PV string is 437.4 V and the output current is 101.15A in each phase. And, the maximum optimal output dc voltage of 5 PV strings arrays is 2.19 kV.

IV. Performance of Classical H-Bridge

The output voltage of the PV string at maximum power point (MPP) is around 438 V. The peak amplitude of single phase voltage for desired 2.3 kV micro-grid is estimated as 1.878 kV. The PV based micro-grid configuration of desired CHB MLI is shown in Fig. 2. The output voltage and harmonic profile for this MGC traditional CHB MLI will be evaluated and analyzed with level-shifted carrier sinusoidal pulse width modulation (LSC-SPWM) technique.

A. Level Shifted Carrier Sinusoidal PWM (LSC-SPWM)

The LSC-SPWM technique is universally applied for grid interfacing in CHB configurations. To control the frequency and harmonics of the output voltage of MLIs the LSC-SPWM method has been applied to the power switches, in which a reference sinusoidal wave of fundamental frequency is compared to high frequency triangular carrier waves. The amplitude levels of the carrier signals are varied. To apply the PWM two modulation index are considered. Amplitude modulation index is the ratio of the amplitude of the reference sine wave to the amplitude of the triangular carrier waves

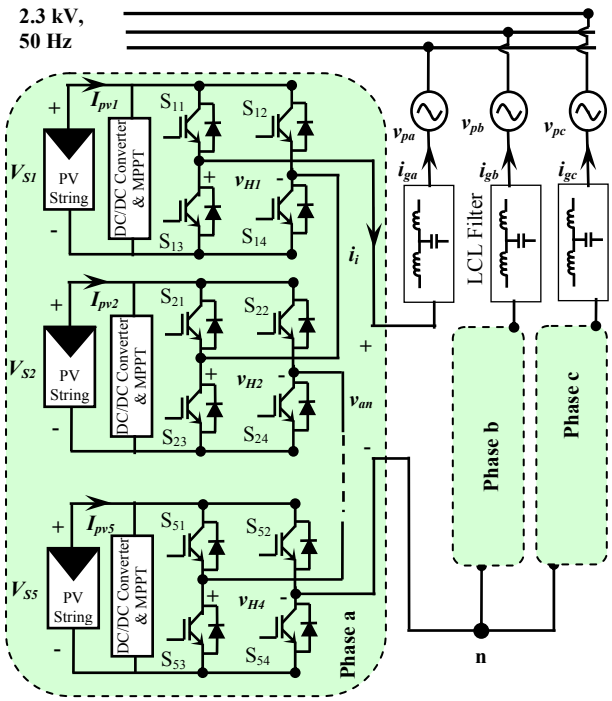


Fig. 2. CHB multilevel inverter topology with PV strings.

is expressed as $m_a=A_m/A_c$. And, the frequency modulation index, is defined as the ratio of the frequency of carrier wave to the frequency of the modulating wave as $m_f=f_c/f$. With the five separate PV strings, 10 triangular carriers are required for LSC-SPWM method which is shown in Fig. 3.

B. Simulation Results of Conventional CHB

The performance of 11-level conventional CHB MLI has been evaluated with R-L load in MATLAB/Simulink environment using simulation parameters calculated in section III. The load active power used for the simulation is 100 kW per phase. Each of the carriers has a frequency of 2 kHz, so the switching frequency of the inverter is 10 kHz. Applying LSC-SPWM, the 11-level output voltage and corresponding harmonic profile have been illustrated in Fig. 4 with an amplitude modulation index as 0.85. The unfiltered output voltage consists of THD of 13.03%.

V. Proposed Modified H-Bridge CM-MGCI

In this work, we have proposed a modified H-bridge for symmetrical approach and the performances of the proposed 11-level twelve-switch topology have been simulated and

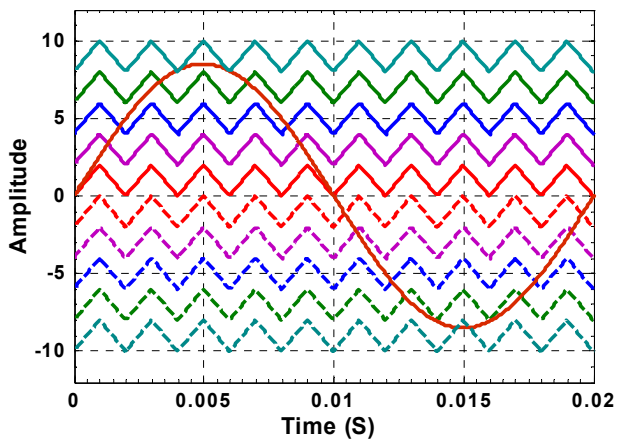
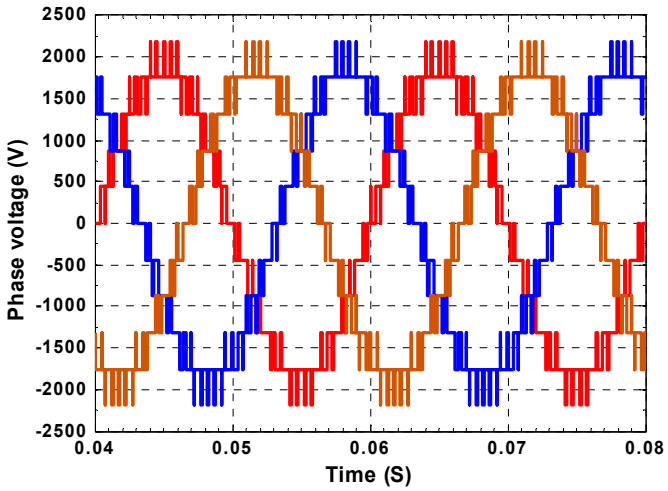
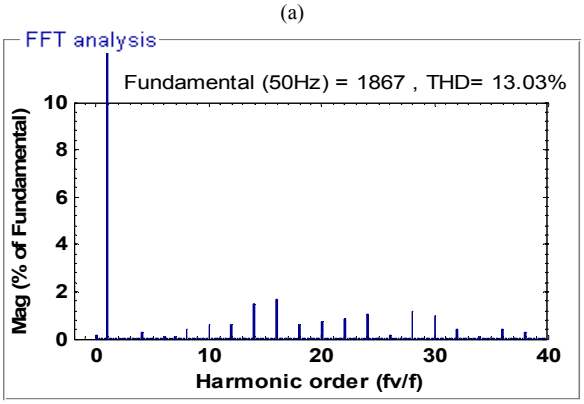


Fig. 3. LSC-SPWM ten triangular carrier for 11-L (5 Module String) CHB inverter with m_a = 0.85.



(a)



(b)

Fig. 4. Inverter response for phases; (a) voltage waveform, (b) voltage spectrum for $m_a = 0.85$ and $m_f = 40$.

investigated with LSC-SPWM in MATLAB/Simulink environment. The topology of desired modified H-bridge MLI for three phase PV fed micro-grid system is demonstrated in Fig. 5. The basic specifications for the proposed micro-grid connected H-bridge inverter structure are identical as traditional CHB topology.

A. Operation of Reduced Switched CM-MGCI

The suggested reduced switch CM-MGCI structure consists of a series connected H-bridge blocks, which decide the level of output waveform, in other words it is the level generation circuit. Fig. 6 demonstrates the conducting switches at different operating voltages. In the Fig. 6, the green thick line indicates that the corresponding IGBT is in on states. The series combination among the five symmetrical PV strings as designed in section III is used to produce eleven DC levels at the inverter output. For example, $5V_{dc}$ level is obtained by turning on S_{11}, S_{22}, S_{31}, S_{42}, S_{51}, and S_{62} switches as shown in Fig. 6(a) with green line indicating the turned on switches. Again, $-2V_{dc}$ level is obtained by turning on S_{11}, S_{21}, S_{32}, S_{41}, S_{51}, and S_{61} switches as shown in Fig. 6(b) with green thick line indicating the turned on switches. With various combination of power switches conducting situation, one of the possible eleven level of $5V_{dc}$, $4V_{dc}$, $3V_{dc}$, $2V_{dc}$, V_{dc}, 0, $-V_{dc}$, $-2V_{dc}$, $-3V_{dc}$, $-4V_{dc}$, $-5V_d$ can be achieved with this modified reduced switch H-bridge inverter.

B. Comparison of Modified MLI Topologies

The proposed modified H-bridge MLI shown in Fig. 5 is compared with the classical symmetrical CHB and recent cascaded symmetrical MLI topologies proposed in [10]-[12].

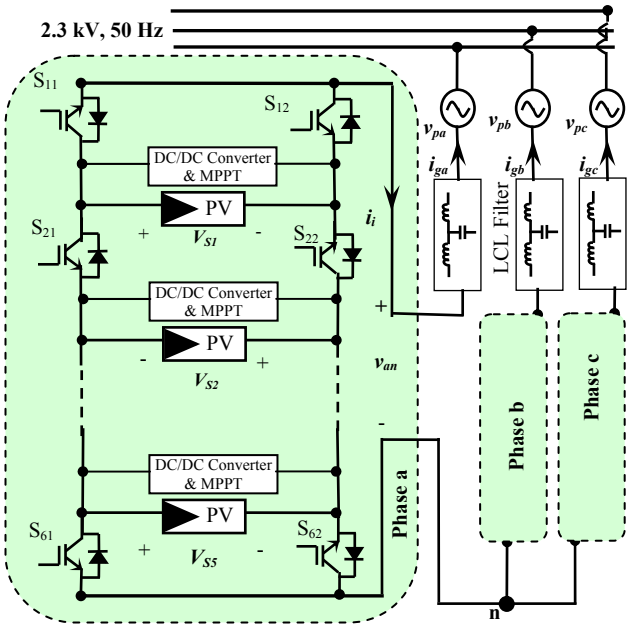


Fig. 5. Proposed reduced switch CM-MGCI.

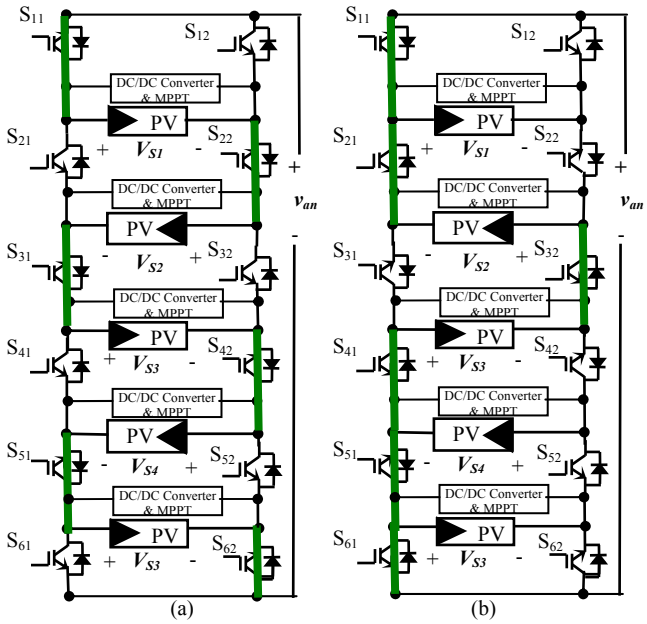


Fig.6. IGBT switching states corresponding to different output voltage levels; (a) $+5V_s$, (b)-$2V_s$.

The comparison in illustrated in Fig. 7. It can be seen from the Fig. 7, that the proposed topology needs fewer switches

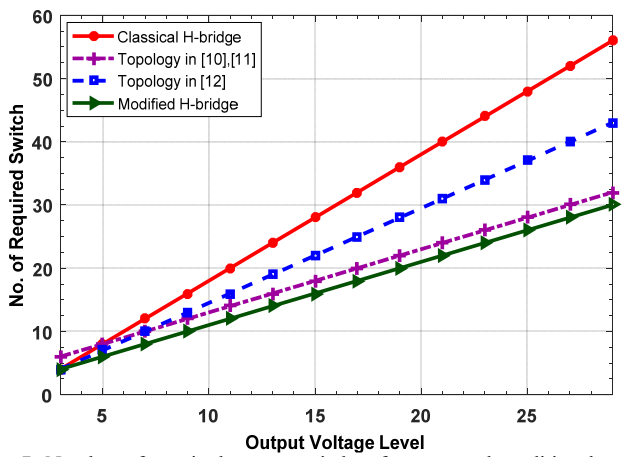


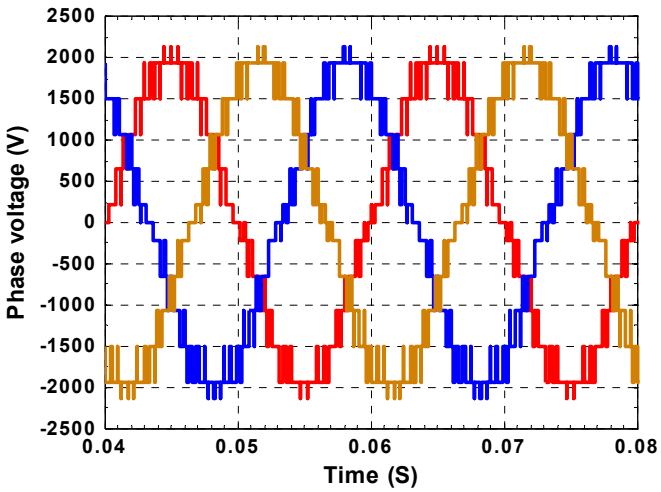
Fig. 7. Number of required power switches for proposed, traditional, and proposed cascaded symmetric multilevel inverter presented in [10]-[12].

for realizing same output voltage levels. This sufficiency reduces the installation area and the number of the gate driver circuits. Therefore, the cost of the proposed topology will be lower than the traditional topology for realizing the same output voltage levels.

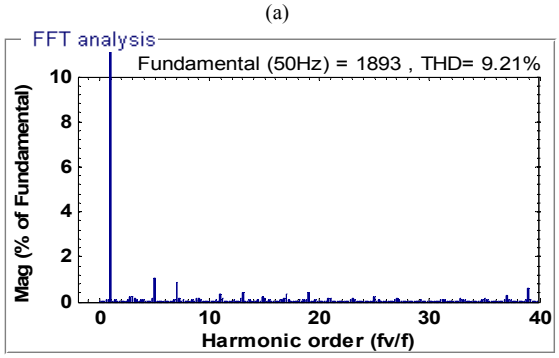
VI. Circuit Response of Modified H-bridge CM-MGCI

The proposed modified H-bridge CM-MGCI is developed in MATLAB/Simulink to verify the configuration response and to analyze the performance. Since, there are six different half H-bridges, so 12 level-shifted carriers are required for this MLI configuration. The output voltage and corresponding harmonic distortion profile of the modified reduced switch H-bridge MLI is illustrated in Fig. 8. With m_a as 0.85 and m_f as 40, the THD of the proposed topology is 9.21% which is 3.82% less than that of the value of traditional CHB MLI.

The THD value of the proposed design is also compared with the other reported designs in Table I. From the table, it is clearly depicted that our proposed modified H-bridge has better THD profile than those recent related works presented in [9]-[12]. In addition to this, our topology uses less power switches than the topologies in [9]-[12]. With topology presented in [10], the 13-level output voltage has a THD of 10.70%. Since, with the decrease in voltage level, the THD is increased, so with this topology the 11-level voltage must gives a THD of higher than 10.70%. Again, the configuration in [11] has a 9-level output voltage which consists of a THD of 17.18% for modulation index m_a as 0.85. It can be clearly observed that, for 11-level configuration of this topology, the THD will not be less than 13%. Thus our modified H-bridge has less THD than the recent reduced switch MLI topologies.



(a)



(b)

Fig. 8. Modified inverter response for phases; (a) voltage waveform, (b) voltage spectrum for m_a = 0.85 and m_f =40.

TABLE I
THD Profile in Recently Proposed MLI Models.

Topology Proposed	Number of Power Switches for 11-L	Unfiltered THD (%)
Classical CHB	20	13.03
Circuit in [9]	20	11.35
Circuit in [10]	14	>10.70
Circuit in [11]	14	>13
Circuit in [12]	16	11.64
Modified H-bridge	12	9.21

VII. Conclusions

A modified design for the multilevel H bridge inverter has been proposed to improve the performance of system. The effectiveness and validity of the proposed version is demonstrated with MATLAB/Simulink environment. The proposed inverter has less THD, requires less number of power devices, and thus has low power loss. Due to these features, proposed topology is superior over the classical structures. Thus, this system can be a good alternative to the inverters used in PV power based micro-grid applications.

References

[1] E. Villanueva, P. Correa, J. Rodriguez, and M. Pacas, "Control of a single-phase cascaded H-bridge multilevel inverter for grid-connected photovoltaic systems," *IEEE Trans. Ind. Electron.*, vol. 56, no. 11, pp. 4399-4406, Nov. 2009.

[2] P. Vijayarajan, A. Shunmugalatha, and H. H. Sait, "Development of modified carrier based PWM scheme for single phase H-bridge inverter fed isolated wind–PV systems," *Solar Energy*, vol. 126, pp. 208-219, Mar. 2016.

[3] L. Wang, D. Zhang, Y. Wang, B. Wu, and A. H. Athab, "Power and voltage balance control of a novel three-phase solid-state transformer using multilevel cascaded H-bridge inverters for Microgrid applications," *IEEE Trans. Power Electron.*, vol. 31, no. 4, pp. 3289 - 3301, Apr. 2016.

[4] S. R. Akhter and M. M. Biswas, "Roadmap of smart grid for Bangladesh based on functions and technological challenges," *Electric Power Components Systems*, vol. 44, no. 8, pp. 864-872, Apr. 2016.

[5] M. Chithra, and S. G. B Dasan, "Analysis of cascaded H bridge multilevel inverters with photovoltaic arrays," in *Proc. IEEE Int. Conf. Emerging Trends Electrical Comp. Tech.*, Tamil Nadu, India, Mar. 2011, pp. 442-447.

[6] A. Mokhberdoran and A. Ajami, "Symmetric and asymmetric design and implementation of new cascaded multilevel inverter topology," *IEEE Trans. Power Electron.*, vol. 29, no. 12, pp. 6712-6724, Dec. 2014.

[7] M. B. Latran, and A. Teke, "Investigation of multilevel multifunctional grid connected inverter topologies and control strategies used in photovoltaic systems," *Renewable and Sustainable Energy Reviews*, vol. 42, pp. 361-376, Feb. 2015.

[8] K. K. Gupta, A. Ranjan, et al., "Multilevel inverter topologies with reduced device count: A review," *IEEE Trans. Power Electron.*, vol. 31, no. 1, pp. 135 - 151, Jan. 2016.

[9] K. V. K. Varma, K. Sirisha, G. Satyanarayana, and K. L. Ganesh, "Optimal PWM strategy for 11-level series connected multilevel converter using Hybrid PV/FC/BESS source," in *Proc. IEEE Int. Conf. Circuit, Power Computing Tech.*, India, Mar. 2014, pp. 686 - 691.

[10] M. T. Khosroshahi, "Crisscross cascade multilevel inverter with reduction in number of components," *IET Power Electron.*, vol. 7, no. 12, pp. 2914 - 2924, Dec. 2014.

[11] R. Sridhar, S. Dhar, and S. S. Dash, "Performance analysis of a standalone PV system with reduced switch cascaded multilevel inverter," *Int. J. Power Energy Conv*, vol. 6, no. 2, pp. 107-127, 2015.

[12] S. H. Hosseini, K. Varesi, et al., "An attempt to improve output voltage quality of developed multi-level inverter topology by increasing the number of levels," in *Proc. IEEE Int. Conf. Electric. Electron. Engr.*, Bursa, Turkey, Nov. 2015, pp. 665-669.

Saturation Throughput Using Different Backoff Algorithms In IEEE 802.15.4

Md. Mohibur Rahaman[1*], Kazi Ashrafuzzaman[2] , Mohammad Sanaullah Chowdhury[1], Osiur Rahman[1]

[1]Department of Computer Science & Engineering, University of Chittagong
Chittagong-4330, Bangladesh

[2]Department of Electrical Engineering & Computer Science, University of Calgary
Canada

[*]mohib.csecu@yahoo.com

***Abstract*—The IEEE 802.15.4 standard for the Low Rate Wireless Personal Area Networks (LR-WPANs)is popular and widely used for many areas of applications. The standard uses slotted CSMA/CA protocol in it's contention access period (CAP) in the beacon enabled mode. The protocol adopted a Binary Exponential Backoff (BEB)algorithm. In this paper, we investigate the saturation throughput of this standard using the existing BEB algorithm and compare it with three other backoff schemes - Exponential Increase Exponential Decrease (EIED), Exponential Increase Linear Decrease (EILD) and Exponential Increase Multiplicative Decrease (EIMD) algorithms. From the simulation results, it is found that EIED, EILD and EIMD perform better than the BEB for the higher loads. The EIED produces the highest throughput among all of these schemes.**

Keywords—IEEE 802.15.4; backoff algorithms; contention access period; CSMA/CA.

I. INTRODUCTION

The IEEE 802.15.4 is a standard for Low Rate Wireless Personal Area Networks (LR-WPANs) [1]. This isa simple, low-cost communication network standard that allows wireless connectivity in applications with limited power. The principal objectives are low cost wireless communication with moderate data rates, reliable data transfer, ease of installation, smaller-range operation and a reasonable battery life. It maintains a simple and flexible protocol.LR-WPANs support those simple devices that consume minimal or low power.

The major application area of this technology is the implementation of Wireless Sensor Networks (WSNs), which are systems with miniature devices deployed in an area and monitoring some physical phenomena, reporting the information achieved through wireless links upon reception of queries or detection of events to one or more sinks.

IEEE 802.15.4 Medium Access Control(MAC) allows two types of channel access mechanisms: non-beacon and beacon enabled. The first mode employs an unslotted Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA).On the other hand, the beacon enabled mode uses superframes separated by periodic beacons. The superframe consists of Contention Access Period (CAP) which employs non-persistent slotted CSMA/CA in a distinct manner.

The slotted CSMA/CA protocol in the CAP employs a channel access mechanism and Binary Exponential Backoff (BEB) scheme. The saturation throughput results for the CAP using this protocol of the standard can be found in the reference[2].The BEB aims to exploit the medium while still react rapidly to the signs of network congestion. This is achieved by doubling the channel arbitration time(backoff period) every time a node experiences a failure to deliver a packet. The channel arbitration time decreases rapidly to its minimum value after a successful transmission. The cutback in the backoff period can lead to performance degradation if there are a large number of competing nodes, as it causes more collisions after any successful transmission. As a consequence, there are two schemes that suggest a slower reduction in a backoff period after a successful transmission. The first scheme is called Exponential Increase Exponential Decrease (EIED) [3], doubles the backoff period after a collision and halves it after a successful transmission. The second algorithm doubles the backoff period after collision, but linearly decreases the backoff period after a successful transmission is known as Exponential Increase Linear Decrease (EILD). Also there exists a third scheme known as Exponential Increase Multiplicative Decrease (EIMD) [4]. However, the IEEE 802.15.4 MAC employs BEB as the backoff scheme.

In this paper, we investigate and compare the performances of the slotted CSMA/CA protocol in the IEEE 802.15.4 MAC using the different backoff mechanisms - EIED, EILD and EIMD along with the BEB.

The rest of the paper is organized as follows. IEEE 802.15.4 MAC Protocol is explained in section II. The different backoff algorithms are described in section III. The results are discussed in the next section. Finally, we conclude the paper in section V.

II. THE IEEE 802.15.4 MAC PROTOCOL

In this section, we briefly discuss the MAC protocol in the IEEE 802.15.4 standard. More details of the protocol can be found in [1] and [5].

The IEEE 802.15.4 MAC specifies two modes: non-beacon-enabled mode and a beacon-enabled mode. Non-beacon mode uses a non-slotted CSMA with collision avoidance (CSMA/CA) for channel access. If the channel is idle, the transmission of a frame will begin immediately, otherwise the node will backoff and try to access the channel in a future slot. This mechanism has been widely studied in the literature [6]. In the beacon-enabled mode, a coordinator

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

transmits a beacon periodically to form the superframe as shown in Fig. 1.The superframe can have an active and an inactive portion and it consists of a beacon, Contention Access Period (CAP), Contention Free Period (CFP), and an optional inactive portion. The CAP and CFP together form the active portion of the superframe. All communication among the nodes should take place during the active portion. In the CFP, the coordinator controls the contention free channel access by assigning Guaranteed Time Slots (GTS) to those nodes with their GTS requests granted. The assignment of the GTS tothose nodes is determined by the scheduling scheme decided by the coordinator.

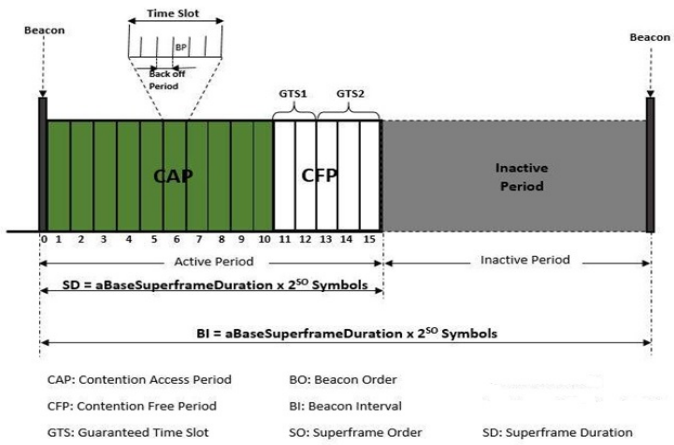


Fig. 1. IEEE 802.15.4 superframe and Backoff period in CAP

In the CAP, a slotted CSMA/CA with binary exponential backoff protocol is used. Each frame need to maintain three variables before it is transmitted successfully. They are - the number of random backoffstages experienced (NB), the current backoff exponent(BE), and the contention window (CW). NB is the number of times the node requires to delay while attempting the current transmission. It is initialized to 0 before every new transmission. CW defines the number of slot periods that need to be clear of activity before the start of transmission. It is set to 2 before each transmission attempt and reset to 2 each time the channel is assessed to be busy. BE is related to how many slot periods a device must wait before attempting to assess the channel.

The slotted CSMA/CA of IEEE 802.15.4 works as follows and depicted in Fig. 2.

- Step 1: NB, CW and BE are initialized.
- Step 2: The node delays for a random number of complete slot periods drawn from a uniform distribution over $[0, 2^{\mathrm{BE}} - 1]$.
- Step 3: Perform a CCA. Then the MAC sub layer proceeds provided that the remaining steps (frame transmission and acknowledgement) can be completed before the end of the CAP. If it cannot proceed, it must wait until the start of the CAP in the next superframe.
- Step 4: If the channel is busy, both NB and BE are incremented by one, ensuring that BE is not more than aMaxBE, and CW is reset to 2. If the value of NB is less than or equal to macMaxCSMABackoffs, the CSMA/CA must return to step 2, else the CSMA-CA end with a Channel-Access-Failure status.
- Step 5: If the channel is idle, the CW is decremented by one. If CW is not equal to 0, it must go to step 3, else start transmission on the boundary of the next slot period.

The standard specifies the following default parameter values - $amacMinBE$=3, $aMaxBE$=5and $NB_{max} = 5$. If the node succeeds in accessing the channel during the backoff procedure, it will reset the parameters NB، BE and CW to the default values for initial transmission of the next frame.

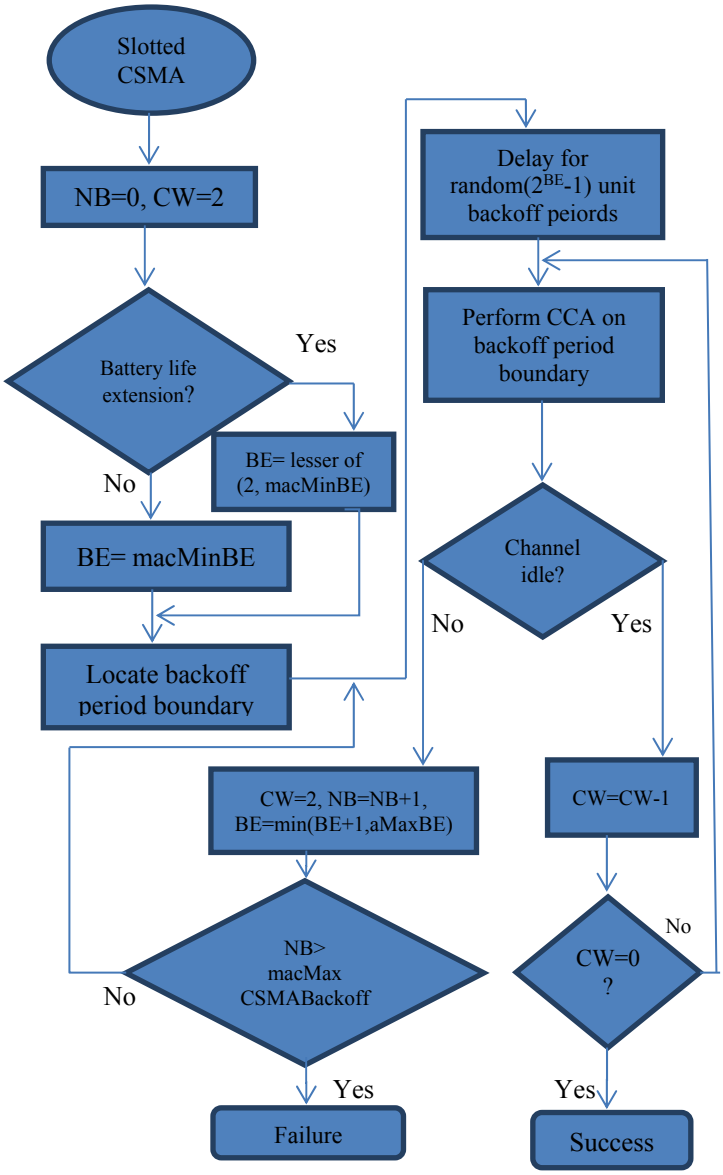


Fig. 2. Slotted CSMA/CA mechanism for IEEE 802.15.4

III. DIFFERENT BACKOFF ALGORITHMS

In previous section, we have described about the backoff algorithm for the IEEE 802.15.4 which is known as BEB. In this section, we briefly discuss three more backoff algorithms namely EIED, EIMD and EILD. Then we adopt these algorithms in the slotted CSMA/CA of IEEE 802.15.4.

A. Exponential Increase Exponential Decrease Algorithm

Step1: Initialization- minimum backoff exponent, maximum backoff exponent, maximum number of backoffstages.
*Step2:*At the First Transmission Attempt
- Set: Backoff Period Set to the Minimum.

Step3:For Each Collision or Unsuccessful Transmission
- EIED doubles the Backoff Period
- It retries until maximum number of backoff occurs

Step4:For Each Successful Transmission
- EIED halves the Backoff Period

*Step5:*End

B. Exponential Increase Multiplicative Decrease Algorithm

Step 1: Initialization-minimum backoff exponent, maximum backoff exponent, maximum number of backoffstages.
Step 2: At the First Transmission Attempt
- Set: Backoff Period Set to the Minimum.

Step 3: For Each Collision or Unsuccessful Transmission
- EIMD doubles the Backoff Period
- It retries until maximum number of backoff occurs

Step 4: For Each Successful Transmission
- EIMD decrease the Backoff Period by a Factor(here 1.5).

Step 5: End

C. Exponential Increase Linear Decrease Algorithm

Step1: Initialization-minimum backoff exponent, maximum backoff exponent, maximum number of backoffstages.

*Step 2:*At the First Transmission Attempt
- Set: Backoff Period Set to the Minimum.

*Step3:*For Each Collision or Unsuccessful Transmission
- EILD doubles the Backoff Period
- It retries until maximum number of backoff occurs

*Step4:*For Each Successful Transmission
- EILD decrease the Backoff Period Linearly, that is current Backoff period=Backoff period-1.

Step5: End

IV. SIMULATION RESULTS

In this section, we present simulation results and compare the saturation throughputs by applying the four backoff algorithms in slotted CSMA/CA of IEEE 802.15.4.

The simulation results presented in this section are obtained using the same simulator used in [2] for the BEB.We extend the simulator adding other backoff algorithms. It is a discrete event simulator using C++ programming language that closely follows the MAC layer procedure of the IEEE 802.15.4.
We have used two scenario – the first one is the using the default values in the standard and other one is using different values as shown in the Table 1.

TABLE I. SIMULATION PARAMETERS

Parameters	Default Scenario	Proposed Scenario
amacMinBE	3	4
aMaxBE	5	9
NB_{max}	5	5

The Fig. 6 shows the normalized throughput for all the four backoff algorithms using the default scenario. The EIED algorithm outperforms the other three. As the load (number of nodes) increases, the relative performance of the EIED increases. However, other three schemes – BEB, EILD and EIMD show almost similar performance for the default parameters. This implies, EIED would be better choice among all in this scenario.

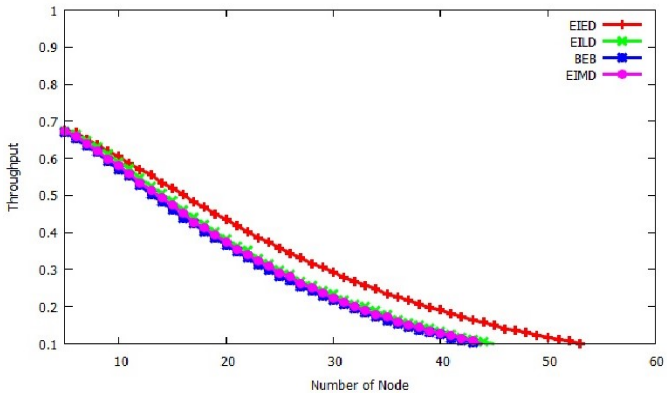


Fig. 6. Normalized Saturation throughput (*amacMinBE* = 4, *aMaxBE* = 9 and NB_{max} = 5)

On the other hand, the Fig. 7 shows the performance results for another scenario as the values are given at the Table 1. The throughputs of BEB are better up to a certain number of nodes (around 25). After that, the performance is decreased. This concludes that for higher load, BEB would not be good choice. However, the EIED closely performs same as BEB up to the first 25 nodes and takes lead afterwards. It outperforms other backoff algorithms for higher number of nodes. The EILD and EIMD also perform better than the BEB for higher amounts of nodes. It happens because of taking a random number from a small distribution range yield a higher collision probability.In BEB, after a successful transmission it decreases the backoff window into the minimum. Hence, the distribution range is small and collision probability is higher. But in the cases of EIED, EILD and EIMD, it decrease the backoff window into halves, linearly, and by a factor(here 1.5) respectively. As a result, a larger distribution range produced and collision probability become lower. So, the throughputs of

EIED, EILD, and EIMD algorithm are better than that of BEB algorithm after certain amount of nodes.

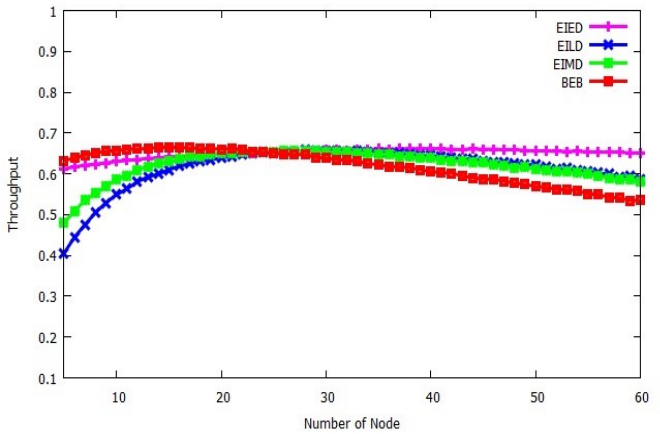


Fig. 7. Normalized Saturation throughput ($amacMinBE$ = 4, $aMaxBE$ = 9 and NB_{max} = 5)

TABLE II. THROUGHPUT COMPARISON(USING DEFAULT PARAMETER)

Node Ranges	Normalized Throughput (Average)			
	BEB	Throughput increase in % with respect to BEB		
		EIED	EIMD	EILD
5-14	5.801938	5%	1%	3%
15-24	3.769715	18%	2%	5%
25-34	2.251157	33%	3%	6%
35-44	1.277758	52%	4%	8%
45-60	0.951412	79%	4%	11%

TABLE III. THROUGHPUT COMPARISON(USING PROPOSED PARAMETER)

Node Ranges	Normalized Throughput (Average)			
	BEB	Throughput increase/decrease in % with respect to BEB		
		EIED	EIMD	EILD
5-14	6.526819	-4%	-13%	-20%
15-24	6.603314	-2%	-2%	-4%
25-34	6.395326	3%	2%	2%
35-44	6.076873	9%	5%	6%
45-60	8.976842	17%	8%	9%

Table II shows the increased/decreased percentage of throughput for BEB, EIED, EILD, EIMD using the default parameter ($amacMinBE$ = 3, $aMaxBE$ = 5 and NB_{max} = 5).EIED performs significantly better than BEB, EIMD, EILD. Moreover, EIMD and EILD show better performance than BEB.

Similarly, we compare the changed throughput of all the proposed algorithms in percentage with that of BEB using the proposed parameter ($amacMinBE$ = 4, $aMaxBE$ = 9 and NB_{max} = 5), which is shown in Table III. It is evident from the table that BEB performs well only in the case of lower number of nodes[5-24]. But as the number of node increases [25-60], EIED performs much better than any other algorithm. EIMD and EILD also perform better than BEB for higher number of node[25-60].

V. CONCLUSION

We have compared the throughputs of existing BEB algorithm with three other backoff algorithms - EIED, EILD, EIMD for the slotted CSMA/CA in IEEE 802.15.4 MAC. Simulation results show that EIED algorithms give better throughputs over BEB. For the higher load, the EIED, EILD, EIMD perform better than that of BEB algorithm. Throughput is considered as one of the most important aspects for analyzing the performance. The other features are the delay and the energy. This work can be extended to investigate the effects on these features by applying these backoff schemes in IEEE 802.15.4.

REFERENCES

[1] LAN/MAN Standards Committee. "Part 15.4: wireless medium access control (MAC) and physical layer (PHY) specifications for low-rate wireless personal area networks (LR-WPANs)." *IEEE Computer Society* (2003).

[2] Kazi Ashrafuzzaman and SUP KWAK KYUNG. "On the performance analysis of the contention access period of IEEE 802.15. 4 MAC." *IEEE communications letters* 15.9 (2011): 986-988.

[3] Song, Nai-Oak, et al. "Enhancement of IEEE 802.11 distributed coordination function with exponential increase exponential decrease backoff algorithm."*Vehicular Technology Conference, 2003. VTC 2003-Spring. The 57th IEEE Semiannual*. Vol. 4. IEEE, 2003.

[4] Vukovic, Ivan N., and Natt Smavatkul. "Saturation throughput analysis of different backoff algorithms in IEEE802. 11." *Personal, Indoor and Mobile Radio Communications, 2004. PIMRC 2004. 15th IEEE International Symposium on*. Vol. 3. IEEE, 2004.

[5] Zheng, Jianliang, and Myung J. Lee. "Will IEEE 802.15. 4 make ubiquitous networking a reality?: a discussion on a potential low power, low bit rate standard." *IEEE Communications magazine* 42.6 (2004): 140-146.

[6] Kleinrock, Leonard, and Fouad Tobagi. "Packet switching in radio channels: Part I-carrier sense multiple-access modes and their throughput-delay characteristics." *IEEE transactions on Communications* 23.12 (1975): 1400-1416.

A Phase-Locked Loop Based Technique for Accurate Estimation of Grid Voltage Fundamental Parameters

S. M. Kamruzzaman*, Ausmita Sarker, Rahul Ghosh and Md. Shamim Reza**
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology, Dhaka 1205, Bangladesh
sm.kamruzzaman@gmail.com* and shamim_reza@eee.buet.ac.bd**

Abstract: **Phase–locked loop based on an anticonjugate decomposition and a cascaded delayed signal cancellation (ACDSC-PLL) strategy is a recognized technique for estimating the single-phase grid voltage fundamental parameters such as amplitude, frequency and phase angle. However, it struggles with high order harmonics and increases complexity in real time implementation when the number of delayed signal cancellation (DSC) operators is increased. To overcome these problems, this paper proposes a technique where the conventional ACDSC-PLL is combined with a simple selective filtering stage. The addition of selective filtering makes the proposed technique capable of removing high order harmonics with a dynamic delay slightly greater than one fundamental cycle. When compared to the conventional one, the proposed technique is relatively simple as it needs less number of DSC operators. Simulation results are presented to verify the performance of the proposed technique.**

***Index Terms*—Anticonjugate decomposition, Cascaded delayed signal cancellation, and phase-locked loop.**

I. Introduction

With flourish of industrial plants, heating equipment, and photovoltaic cells connected to the grid, nonlinear electrical loads are growing. In addition to that, the grid connected power converter [1] and electric vehicles [2] also fall victim to nonlinearity, which consequently distorts the grid voltage waveform. It is prerequisite for smooth power transmission and high quality power system that the voltage distortions (voltage sag, swell, flicker, frequency drifts, harmonics [3] etc.) are rapidly minimized and efficiently restored to the standard values. Therefore, an efficient technique is required for accurate estimation of the grid voltage parameters.

There are various estimation techniques reported in the technical literature [4][5]. The frequency domain based estimation techniques experience inefficient computation, slow transient and many more issues [4][6]. These complications can be avoided by using time domain techniques [5][7]-[11]. Among the time domain techniques, phase-locked loop (PLL) based on a cascaded delayed signal cancellation (CDSC) integrated with an anticonjugate decomposition (ACD) strategy is highly accepted for its ability to reject low order harmonics from the single-phase grid voltage waveform and estimate the voltage amplitude, frequency and phase with better accuracy and faster dynamic response [9]-[11]. Despite its ability to estimate the grid voltage parameters to an appreciable level of accuracy, the ACD-CDSC based PLL (ACDSC-PLL) system cannot reject high order harmonics effectively when a small number delayed signal cancellation (DSC) operator is cascaded. On the other hand, the computational complexity increases when a large number of DSC operators are used for rejecting high order harmonics present in the input signal.

The objective of this paper is to propose an enhanced ACDSC-PLL for estimating the single-phase grid voltage fundamental parameters under distorted grid conditions. In the proposed technique, an additional selective filter is integrated with the conventional ACDSC-PLL. The addition of selective filter helps to use a small number of DSC operators to be cascaded, thus reduces the system complexity and computational burden. Moreover, the technique can successfully reject the high order harmonics resulting in a better estimation of grid voltage parameters.

II. Grid Voltage Parameter Estimation

The proposed ACDSC-PLL for estimating the single-phase grid voltage fundamental amplitude, frequency and phase angle is shown in Fig. 1, where the details of frequency adaptive $H_{17,33}$ and $CDSC_{4,8,16}$ blocks are shown in Fig. 2(a) and Fig. 2(b), respectively.

A. Orthogonal Voltage Waveform Generation

The proposed technique, as shown in Fig. 1, is inspired from [11], where the ideas of ACD and CDSC (ACDSC) are reported. The ACD refers to the decomposition of a single-phase voltage into two space vectors: a positive sequence and a negative sequence one. The amplitudes of both of them are half of the single-phase voltage amplitude and both of them are symmetrical along the imaginary axis, hence the name anticonjugate. The positive sequence component is in phase with the input voltage. In [11], it is proven that the real component of their vectorial sum is zero and the imaginary component equals the original input voltage. In other words, we can decompose the vectorial sum back into the anticonjugate pair and use them to estimate the amplitude and

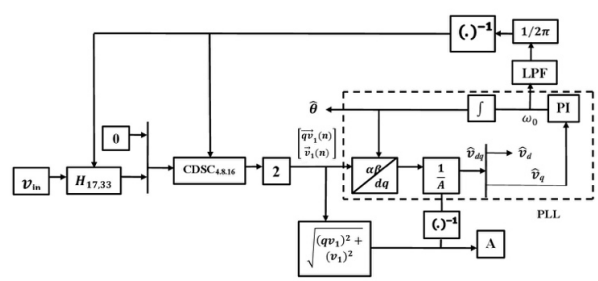
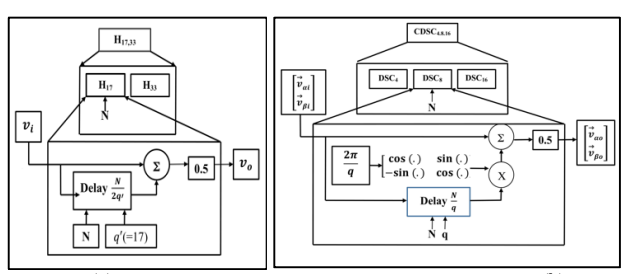


Fig. 1. Proposed ACDSC-PLL technique.



(a) (b)

Fig. 2. (a) $H_{17,33}$ filter structure. (b) $CDSC_{4,8,16}$ structure.

phase angle of the single phase voltage signal using PLL [9]-[11]. However, most often the signal received by the controller is distorted by various harmonics and noises which, if not managed properly, introduce considerable error while estimation. Therefore, to reject the negative effects of harmonics, CDSC is integrated with the ACD [11].

The CDSC is an advanced and one of the most effective selective filtering techniques used to extract the grid voltage fundamental component from the input stationary $\alpha\beta$-frame signals [7]-[11]. It consists of multiple DSC operators like DSC_p, DSC_q, DSC_r and so on, cascaded in series forming the operator $CDSC_{p,q,r,\ldots}$ where each DSC operator, say: DSC_q takes the signal $\vec{v}_{\alpha\beta i}$ as input and conducts the following operation [7][8],

$$\begin{bmatrix}\vec{v}_{\alpha o}(n)\\ \vec{v}_{\beta o}(n)\end{bmatrix} = \frac{1}{2}\left\{\begin{bmatrix}\vec{v}_{\alpha i}(n)\\ \vec{v}_{\beta i}(n)\end{bmatrix} + \begin{bmatrix}\cos\left(\frac{2\pi}{q}\right) & \sin\left(\frac{2\pi}{q}\right)\\ -\sin\left(\frac{2\pi}{q}\right) & \cos\left(\frac{2\pi}{q}\right)\end{bmatrix}\begin{bmatrix}\vec{v}_{\alpha i}\left(n-\frac{N}{q}\right)\\ \vec{v}_{\beta i}\left(n-\frac{N}{q}\right)\end{bmatrix}\right\} \quad (1)$$

where, $\vec{v}_\alpha(n)$ and $\vec{v}_\beta(n)$ are the quadrature and in-phase components respectively and the subscripts i and o refer to the input and output variables, respectively. $2\pi/q$ is the rotational angle and N/q is the delay time where, N is the number of samples per fundamental period (T) and q is the delay factor associated with the particular DSC_q operator. The implementation of CDSC for single-phase voltage system can be represented by the block diagram shown in Fig. 2(b) [7][8], where, the details of $CDSC_{p,q,r}$ and the DSC_q operators are shown. Here, the value of N is updated at every sample using the estimated fundamental angular frequency ω_o[7][8]. It is important to note that, to perform a delay operation in a discrete environment the delay amount N/q has to be an integer value [7]. Therefore, if the calculated N/q value is not integer i.e.$M < N/q < M+1$, where M is a positive integer, then the following interpolation operation, in equation (2) is performed to estimate the value of $\vec{v}_{\alpha\beta i}(n - N/q)$ [7].

$$\vec{v}_{\alpha\beta i}(n - {}^{N}/_{q}) = \left[1 - ({}^{N}/_{q} - M)\right]\vec{v}_{\alpha\beta i}(n-M) + ({}^{N}/_{q} - M)\vec{v}_{\alpha\beta i}(n-M-1) \quad (2)$$

The transfer function for the DSC_q block is defined as: [9]

$$DSC_q(s) = \frac{1}{2}\left(1 + e^{-\frac{T}{q}s}.e^{-j\frac{2\pi}{q}}\right) \quad (3)$$

which has the following magnitude and phase response: [9]

$$DSC_q(j\omega) = \left|\cos\left(\frac{h\omega_o T - 2\pi}{2q}\right)\right| \measuredangle -\left(\frac{h\omega_o T - 2\pi}{2q}\right) \quad (4)$$

Here, h is the harmonic order. The magnitude response plots for the DSC_4, DSC_8, DSC_{16} and $CDSC_{4,8,16}$ are shown in Fig. 3(a) to Fig. 3(d), respectively. It can be seen from Fig. 3(a) that, for $q = 4$ the magnitude response for the harmonics of order $h = 3, 7, 11, \ldots\ldots$ becomes zero. In other word these harmonics get eliminated which also comes with the cost of a dynamic delay of $T/4$ introduced in the estimation time [9]. For $CDSC_{4,8,16}$ the overall dynamic delay is: ($T/4+T/8+T/16$) ≈0.48T. However, the $CDSC_{4,8,16}$ operator introduce 0° phase delay at the tuning fundamental frequency [9][11].

B. Proposed Filtering Procedure

Despite the advantages, there is a major drawback of $CDSC_{4,8,16}$. As can be seen from Fig. 3(d), this system cannot eliminate the harmonics of order h = 17 and 33. Although we can eliminate these by introducing another two DSC blocks i.e. DSC_{32} and DSC_{64} along with the existing one creating $CDSC_{4,8,16,32,64}$ operator, however, it introduces much complexity to the system, since we must create individual

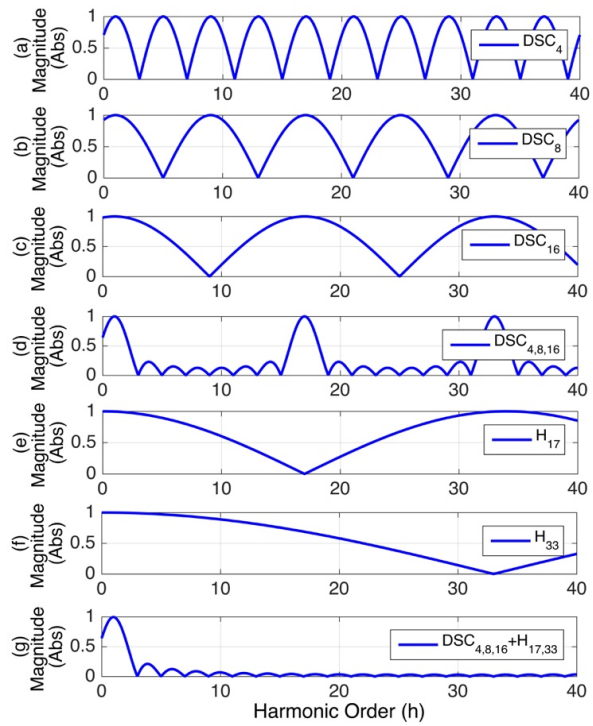


Fig. 3. Magnitude responses of (a) DSC_4, (b) DSC_8, (c) DSC_{16}, (d) $CDSC_{4,8,16}$ (e) H_{17} (f) H_{33} (g) $CDSC_{4,8,16}$ + $H_{17,33}$.

delay stages for each orthogonal input signals $\vec{v}_\alpha(n)$ and $\vec{v}_\beta(n)$ for each DSC block while implementing this operator in practice. In our proposed method, we add a pre-filtering stage prior to the ACD-$CDSC_{4,8,16}$ system which consists of two frequency selective filters H_{17} and H_{33} each having the transfer function expressed by the following equation [13]:

$$H_{h'}(s) = \frac{1}{2}\left(1 + e^{-\frac{T}{2h'}s}\right) \quad (5)$$

where h' is the order of harmonic that is to be eliminated. Fig. 2(a) exhibits the structure of H_{17} and H_{33}. These two filters eliminate the 17th and 33rd harmonic from the input voltage waveform before it is fed to the ACDSC system. The magnitude responses of H_{17}, H_{33} and $CDSC_{4,8,16}$+$H_{17,33}$ are shown in Fig. 3(e) to Fig. 3(g), respectively. The frequency adaptive H_{17} and H_{33} filters introduce a fixed delay at the fundamental frequency in steady-state condition. This delay can be computed from the phase response of $H_{17,33}$ and can be subtracted from the phase angle estimated by the PLL. Using the filters H_{17} and H_{33} instead of two additional DSC operators give us the advantage of less system complexity over DSC, since each pre-filtering stage consists of only one delay stage instead of two like a DSC operator.

C. Fundamental Voltage Amplitude, Frequency and Phase Estimation

The fundamental voltage amplitude can be estimated from the orthogonal voltage waveform and is given by [7][8]

$$A_1 = \sqrt{(q\tilde{v}_1)^2 + \tilde{v}_1{}^2} \quad (6)$$

A conventional PLL is used to estimate the fundamental frequency and phase angle. Any PLL system consists of a phase detector (PD), a loop filter (LF) and a voltage control oscillator (VCO). In this paper, the PLL uses park transformation as the PD [10]. This eventually shifts the input

TABLE I
HARMONICS BASED ON THE EUROPEAN STANDARD EN 50160

Harmonics Order	3	5	7	9	11	13	15	17
Relative Voltage (%)	5	6	5	1.5	3.5	3	0.5	2

Harmonics Order	19	21	23	25	27	29	31	33
Relative Voltage (%)	1.5	0.5	1.5	1.5	0.5	1	1	1.5

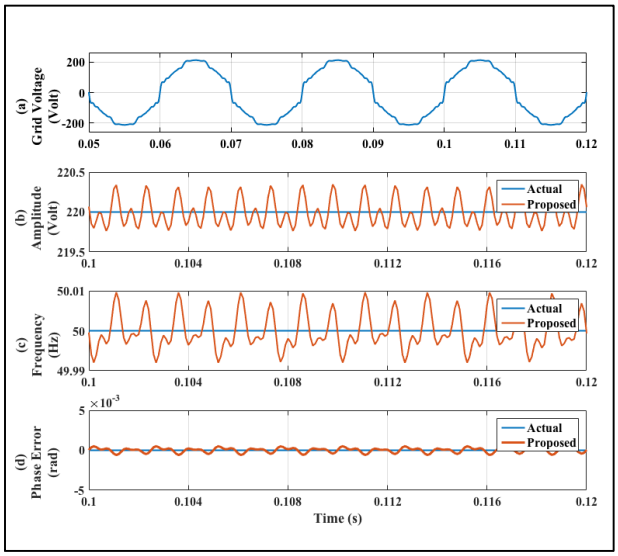


Fig. 4. Case I: Steady State with Harmonics (a) Grid voltage waveform (b) Fundamental voltage amplitude (c) Fundamental frequency (d) Phase Error.

signal from $\alpha\beta$ stationary reference frame to dq synchronous reference frame [10]. The LP is a basic Proportional Integral (PI) controller with a transfer function of [10]:

$$F(s) = k_p + \frac{k_i}{s} \tag{7}$$

where k_p and k_i are the proportional and integral gains, respectively. The VCO is a simple integration operator, turning the angular frequency to the value of an angle [10].

III. SIMULATION RESULTS

Simulation results in MATLAB/SIMULINK are presented in this section. The distortions and variations are emulated in a traditional 220V, 50Hz system and are based on the European standard EN50160 [12]. The following cases are studied to evaluate the performance of the proposed technique, where the harmonic contents are shown in Table I.

i. Steady-State with Harmonics
ii. Frequency Step and Harmonics
iii. Frequency Sweep and Harmonics
iv. Voltage Sag and Harmonics
v. Voltage Swell and Harmonics
vi. Voltage Flicker and Harmonics

CASE I: Steady-State with Harmonics

Fig. 4 deals with the steady-state of a distorted grid voltage waveform. The harmonics introduced are only odd harmonics, as the even harmonics are comparatively negligible. The relative harmonic voltages are chosen according to the European standard EN50160 [12] and given in Table I. It can be seen from Fig. 4 that the estimated parameters contain very little error at steady-state under harmonics.

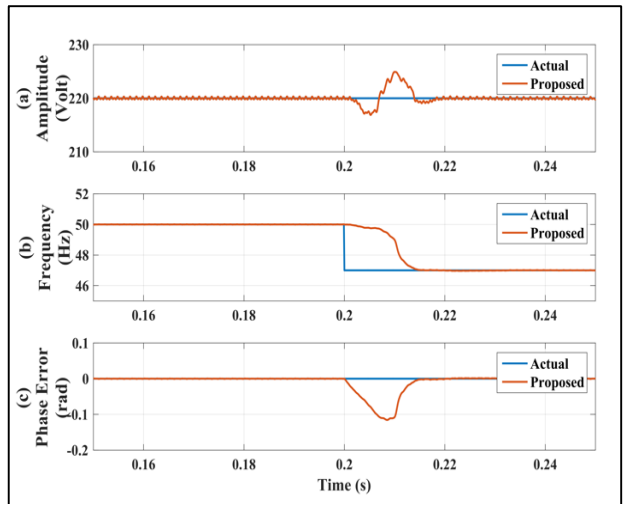


Fig. 5. Case II: Frequency Step and Harmonics (a) Fundamental voltage amplitude (b) Fundamental frequency (c) Phase error.

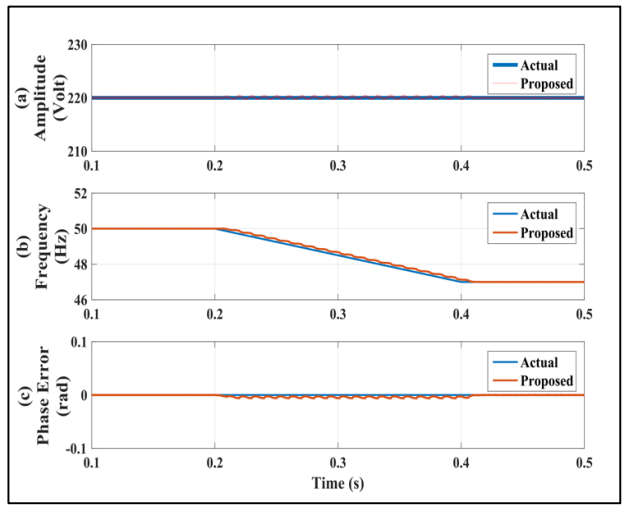


Fig. 6: Case III: Frequency Sweep and Harmonics (a) Fundamental voltage amplitude (b) Fundamental frequency (c) Phase error.

Case II: Frequency Step and Harmonics

According to the standard EN 50160, the grid frequency should be within +2Hz to -3Hz from the nominal value (50 Hz) for 100% of time. Therefore, the worst-case scenario is most likely to occur at frequency 47 Hz or 52 Hz. Fig. 5 deals with the circumstance when the grid frequency suddenly changes from 50 Hz to 47 Hz with the presence of harmonics. As it can be noticed, the technique takes around 0.015s (0.75 fundamental cycles) to track the fundamental voltage parameters.

Case III: Frequency Sweep and Harmonics

A frequency sweep of -10 Hz/s and 35.5% THD are considered for this case study. It can be noticed from Fig. 6 that the technique can track the frequency sweep accurately with a phase error less than 0.008 radian (0.5°).

CASE IV: Voltage Sag and Harmonics

The performance of the proposed technique under 50% of voltage sag and harmonics is shown in Fig. 7. It is visible that, it takes around 0.01s (0.5 fundamental cycles) to estimate the voltage sag. On the other hand, it requires around 0.025s (1.25 fundamental cycles) to settle the frequency and phase angle.

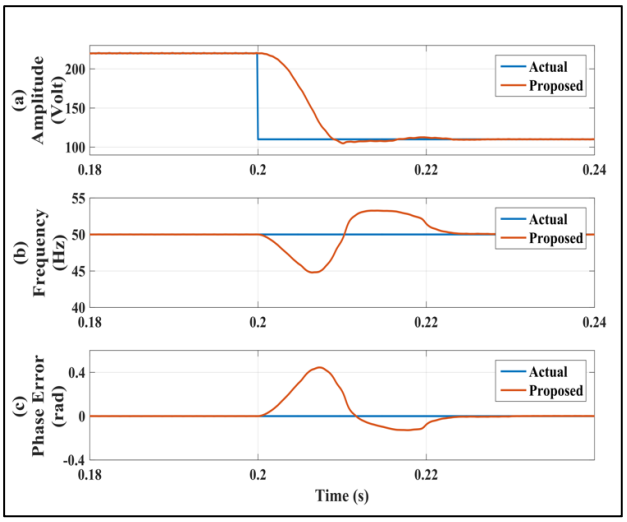


Fig. 7. Case IV: Voltage Sag and Harmonics (a) Fundamental voltage amplitude (b) Fundamental frequency (c) Phase error.

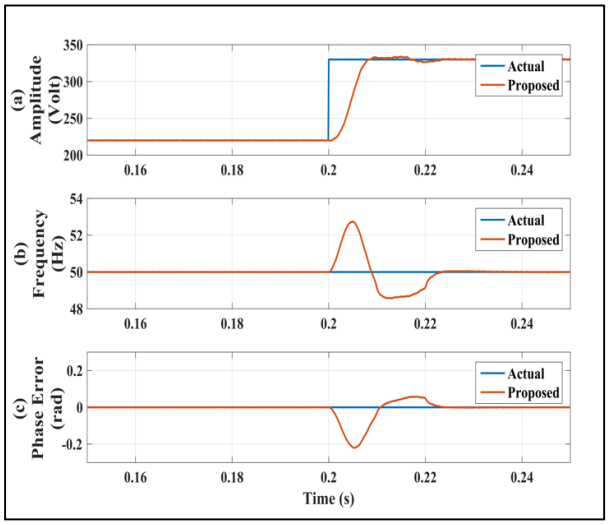


Fig. 8. Case V: Voltage Swell and Harmonics (a) Fundamental voltage amplitude (b) Fundamental frequency (c) Phase error.

CASE V: Voltage Swell and Harmonics

The performance of the proposed technique under 50% voltage swell is shown in Fig. 8. In this case, the proposed technique takes around 0.01s (0.5 fundamental cycles) to estimate the voltage swell and around 0.025s (1.25 fundamental cycles) to track the frequency and phase.

CASE VI: Voltage Flicker and Harmonics

The frequency and amplitude of the voltage flicker are 25 Hz and ±0.1 p.u. respectively. The simulation results are presented in Fig. 8. Here, it is evident that the proposed technique can track the amplitude, frequency and phase with very few ripples. The ripples are ±0.020 Hz and ±0.0008 rad (±0.05°) for the estimated frequency and phase respectively.

IV. Conclusion

An enhanced ACDSC-PLL has been documented in this paper to provide an accurate estimation of single-phase grid voltage fundamental parameters under distorted grid conditions. In the proposed technique, a selective filtering stage has been combined with the conventional ACDSC-PLL. The addition of selective filter helps to use a small number of DSC operators to be cascaded, thus reduces the system

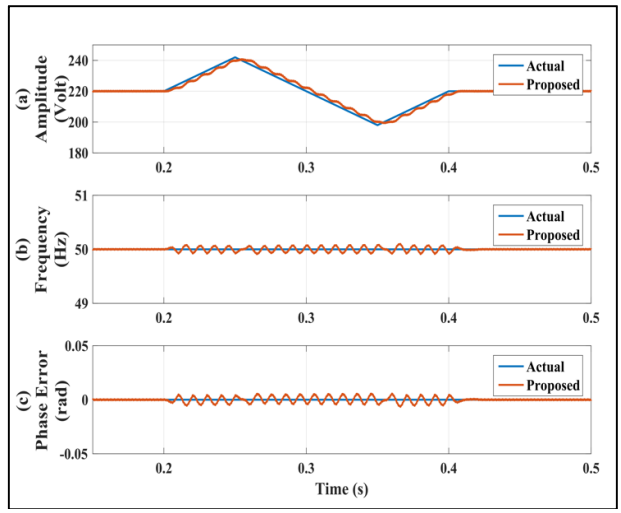


Fig. 9. Case VI: Voltage Flicker and Harmonics (a) Fundamental voltage amplitude (b) Fundamental frequency (c) Phase error.

complexity and computational burden, which makes it more efficient and more suitable for real-time implementation when compared to the conventional one. Moreover, the proposed technique can successfully reject the high order harmonics resulting in a better estimation of grid voltage fundamental parameters.

V. References

[1] Seung-Ho Song, Shin-il Kang, Nyeon-kun Hahm "Implementation and control of grid connected AC-DC-AC power converter for variable speed wind energy conversion system", Div. of Electron. & Inf. Eng., Chonbuk Nat. Univ., Jeon-Ju, South Korea.

[2] Saeid Haghbin, Sonja Lundmark, Mats Alakula, Ola Carlson"Grid-Connected Integrated Battery Chargers in Vehicle Applications: Review and New Solution" *IEEE Trans. Ind Elect.*, vol. 60, no.2, pp. 459-473, Feb 2013.

[3] S. Gonzalez, R. Garcia-Retegui, and M. Benedetti, "Harmonic computation technique suitable for active power filters," *IEEE Trans. Ind. Electron.*, vol. 54, no. 5, pp. 2791–2796, Oct. 2007.

[4] B. McGrath, D. Holmes, and J. Galloway, "Power converter line synchronization using a discrete Fourier transform (DFT) based on a variable sample rate," *IEEE Trans. Power Electron.*, vol. 20, no. 4, pp. 877–884, July 2005.

[5] X. Yuan, W. Merk, H. Stemmler, and J. Allmeling, "Stationary-frame generalized integrators for current control of active power filters with zero steady-state error for current harmonics of concern under unbalanced and distorted operating conditions," *IEEE Trans. Ind. Appl.*, vol. 38, no. 2, pp. 523–532,Mar./Apr.2002.

[6] A. Timbus, R. Teodorescu, F. Blaabjerg, and M. Liserre, "Synchroniza- tion methods for three phase distributed power generation systems. An overview and evaluation," in *Proc. IEEE 36th Power Electron. Spec. Conf., 2005*, Jun., pp. 2474–2481.

[7] Md. Shamim Reza, Mihai Ciobotaru, "A Frequency Adaptive Technique for Accurate Estimation of Single-Phase Grid Voltage Fundamental Parameters." *Industrial Electronics Society, IECON 2013 - 39th Annual Conference of the IEEE.*

[8] Md. Shamim Reza, Mihai Ciobotaru, Vassilios Georgios Agelidis, "Grid synchronization technique without using trigonometric functions for accurate estimation of fundamental voltage parameters," *IET Generation, Transmission & Distribution (Volume:9 , Issue: 12).*

[9] Yi Fei Wang and Yun Wei Li, "Three-Phase Cascaded Delayed Signal Cancellation PLL for Fast Selective Harmonic Detection," *IEEE Trans. Ind. Elect.*, vol. 60, no. 4, pp. 1452-1463, Apr. 2013.

[10] Yi Fei Wang and Yun Wei Li, "Grid synchronization PLL based on cascaded delayed signal cancellation," *IEEE Trans. Power Elect., vol. 26, no. 7, pp. 1987-1997, July 2011.*

[11] Yi Fei Wang and Yun Wei Li, "A grid fundamental and harmonic component detection method for single-phase systems," *IEEE Trans. Power Elect., vol. 28, no. 5, pp. 2204-2213, May 2013.*

[12] Henryk Markiewicz & Antoni Klajn, "Standard EN 50160 - Voltage Characteristics in Public Distribution Systems," *July 2004.*

[13] P. Roncero-Sanchez, X. del Toro Garcia, A. P. Torres and V. Feliu, "Robust frequency estimation method for distorted and imbalanced three-phase systems using discrete filters," *IEEE Trans. Power Electron.*, vol. 26, no.4, pp. 1089-1001, 2011.

Transport characteristics of Double Gate N-Channel Junctionless Transistor

Tanmoy Kumar Paul * and Quazi D. M. Khosru

Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology,
Dhaka, Bangladesh
*tanmoyeee09@gmail.com

Abstract—**In this work, Electrostatic and transport characteristics of symmetric double gate n-channel junctionless transistor have been investigated through numerical simulation. A 1-D self consistent analysis of coupled Schrodinger-Poisson equation has been performed. Channel mobile charge density has been obtained and studied for variations in device dimensions and biasing voltages. A simplified ballistic transport assumption has been applied to show the variations of drain current with device parameters. Finally transconductance has been extracted with respect to biasing gate voltage.**

Index Terms—**Junctionless, Partial Depletion Mode, Self Consistent Solution, Ballistic Transport, Transconductance.**

I. Introduction

The invention of MOSFET consisting of two semiconductor junctions is one of the revolutionary incidents in the field of electronics. The rapidly decreasing device size resulted in losing proper gate control of MOSFET due to short channel effects. Invention of multi-gate structure paved the way for better control against the effects [1]. However, high gradient in doping concentration is a serious challenge in fabrication of such devices with one or more junctions. The first junctionless Transistor was proposed by Colinge et al. at 2009 which was a Tri-Gate N channel junctionless transistor [2]. The device was fabricated by Colinge et al. in 2010 which showed the experimental I-V relationship [3]. The on-off current ratio and subthreshold slope of junctionless transistor is better than traditional MOSFETs which has triggered the large scale study of the device. In such a transistor channel doping concentration is very high so that the metal-semiconductor junction acts as ohmic contact with very negligible resistance. Moreover, channel thickness is very low in order to achieve the off state by obtaining fully depleted channel. Here, surface conduction of current is replaced by bulk conduction [4]. Threshold voltage of the transistor depends on doping concentration, bulk silicon thickness as well as oxide gate thickness. A mathematical model of threshold voltage of a tri-gate junctionless transistor was first given by Trevisoli by solving two dimensional Poisson's equation [5]. The symmetry in a double gate transistor is the main working force to bring much light to analyzing its characteristics. A conduction mechanism model in all three regions of operation- subthreshold, linear and saturation regions is available in [6] which has been obtained by extending the previous models of [7]-[9]. Current flow in the transistor almost follows ballistic transport mechanism due to reduced phonon scattering above threshold voltage [10]. In this paper, we have performed self consistent numerical analysis on a double gate junctionless transistor to show various device properties.

II. Device Structure

Double gate junctionless transistor channel consists of heavily doped semiconductor. In our simulation we have used n type silicon as the semiconductor. There is no pn junction in the channel region along the direction of current flow unlike a conventional MOSFET. The two edges of channel region are termed as source and drain. The terminal where comparatively higher voltage is applied is named as drain and the other is named as source similar to NMOS. The channel is covered from two opposite directions with the gate structures which are coupled together. The gates are placed along an axis perpendicular to current flow. We have used p+ polysilicon as gate material although metal gate structure is also feasible. An insulator layer, in this case, SiO_2 separates the gate structure from the channel. Fig. 1 shows the cross section of the device along a direction perpendicular to transport direction.

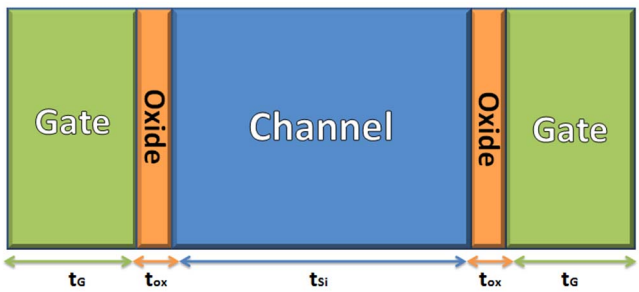


Fig. 1 Cross section of double gate junctionless transistor in the plane perpendicular to current flow direction.

III. Device Working Principle

Operating principle of the device is best understood from its energy band diagram. When contact is formed, Fermi levels align and a depletion region is formed in the channel region. The channel region may be fully depleted or partially depleted as depending on the device parameters. As positive voltage is applied in gate, the band bending declines and depletion width tends to decrease. If the device was initially fully depleted, then it becomes partially depleted at some value of gate voltage. Therefore, current can flow through the non-depleted centre of the channel. Fig. 2 shows the band diagram of the device in partial depletion mode. If voltage is increased further, at some condition, there is no depletion region and the device reaches flat band condition. When applied voltage crosses flat band voltage, charges accumulate in the surface of the channel. Surface current is then governed by this accumulation charge.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

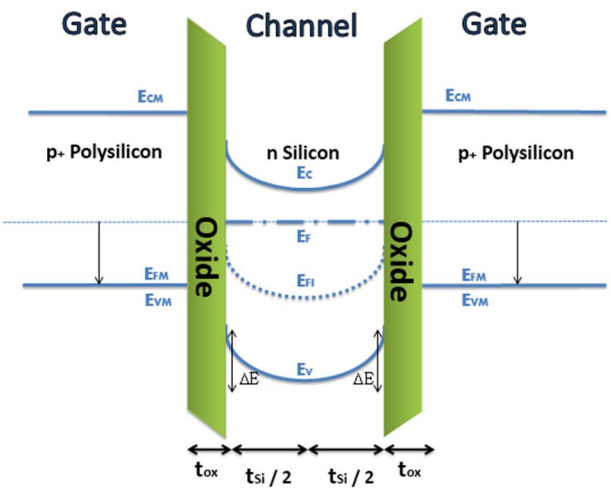


Fig. 2 Energy band diagram along the cross section in partial depletion mode.

IV. Device Simulation

COMSOL assisted MATLAB simulation has been used to perform self consistent analysis through coupled schrodinger-poisson equation solving. For long channel approximation we can assume uniform charge density along transport direction (with small drain-source voltage) whereas charge density varies along the direction perpendicular to transport direction. So we have taken the cross section along the gate-oxide-channel direction and performed 1D simulation. First, for a particular gate voltage as input, an initial potential drop profile has been assumed depending on the band diagrams of the materials used for the device. Solving the schrodinger equation, we obtain wave function and eigen energy of each sub-band of conduction band. For electrons confined in an ideal 2d plane, the total electron concentration profile in conduction band can be obtained by considering contributions from all sub-bands using the following formulae [11]:

$$n(x) = \sum_j n_j(x) = N_C \sum_j |\psi_j(x)|^2 \ln\left(1 + e^{\left(\frac{E_F - E_j}{k_B T}\right)}\right) (m^{-3}) \quad (1)$$

$$Here, N_C = \frac{m_{ds}^*}{\pi\hbar^2} k_B T \quad (2)$$

Similarly, hole concentration profile, $p(x)$ is determined for valance band. The net carrier concentration profile depends on the following conditions:

$$\rho(x) = \begin{cases} p(x) - n(x) + N_D - N_A & ; \text{if x is in depletion region} \\ p(x) - n(x) & ; otherwise \end{cases} \quad (3)$$

The depletion layer width is calculated as [12]:

$$x_d(V_G) = \frac{\epsilon_{si}}{C_{ox}}\left[\sqrt{1 - \frac{2C_{ox}^2}{\epsilon_{si} q N_{si}}(V_G - V_{FB})} - 1\right] \quad (4)$$

Now, ρ(x) is used to determine the potential drop in the device by solving the poisson equation and then potential profile is updated using a mixing coefficient for numerical convergence. The process is repeated until the solution reaches the pre-assumed convergence limit. The total charge concentration per unit area is obtained by integrating the final ρ(x) along the simulation direction.

A current-voltage characteristic is obtained using simplified transport model. The source contact is kept at ground potential, while drain contact Fermi level is lowered by applied drain bias. Drain Current per unit length of the device is obtained as [13]:

$$I_D = \sum_i \frac{q}{h^2}\sqrt{\frac{m_{ti}^* k_B T}{2\pi^3}} \int_{-\infty}^{+\infty} \left(\mathfrak{F}_{-1/2}(\mu_S - \varepsilon) - \mathfrak{F}_{-1/2}(\mu_D - \varepsilon)\right) \times T_i(\varepsilon) d\varepsilon \quad (5)$$

Transconductance is obtained using the following formulae:

$$g_m = \frac{\delta I_D}{\delta V_g} \quad (6)$$

V. Results

A. Normalised Mobile charge density

Total charge concentration per unit area, Q_{total}, is obtained from self consistent simulation. Normalised mobile charge density is then obtained from the following equations:

$$Q_{mobile} = Q_{total} - q \times N_d \times t_{si} \quad (7)$$

$$Q_{mobile}^{norm} = \frac{-Q_{mobile}}{q \times N_d \times t_{si}} \quad (8)$$

Fig. 3 shows the variation of normalised mobile charge density with gate voltage as well as thickness of gate oxide (t_{ox}). For 2nm t_{ox}, 10 nm Si thickness and $1\times 10^{19}\ /cm^3$ doping concentration, 0.3V is the voltage where the device starts to open a non-depleted channel. For, increasing t_{ox}, the reduced effect of gate voltage causes smaller depletion width. So the device starts to show mobile charge at a lower gate voltage. In this case, for 7nm t_{ox}, device is normally on and we have to apply negative voltage to turn it off. So the device shows higher mobile charge density in the partially depleted condition for larger t_{ox}. In accumulation, larger t_{ox} means less accumulation charge, so mobile charge density is also lower. The result is consistent with the result obtained from [14] which is shown in Fig. 4.

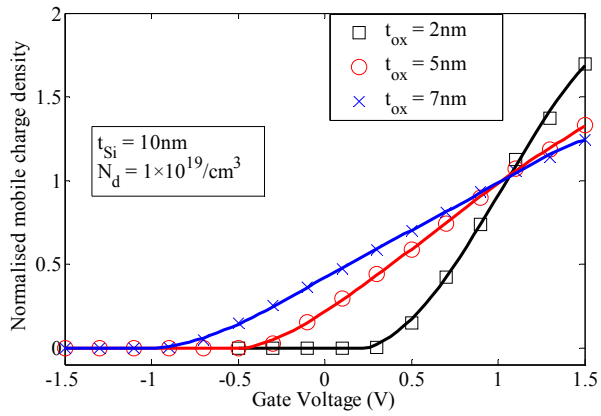


Fig. 3 Normalised mobile charge density variation with oxide thickness.

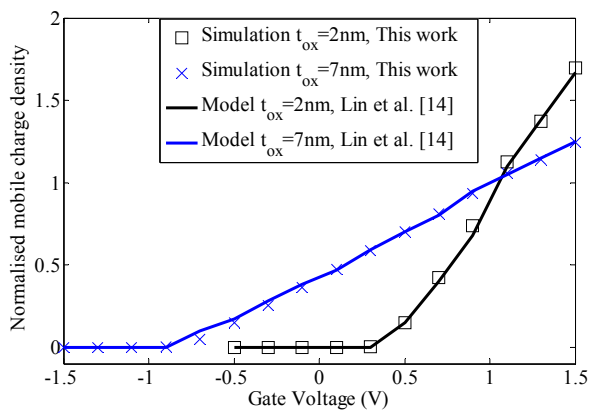


Fig. 4 Normalised mobile charge density compared with mathematical model in Lin et al. [14]

Non-depleted channel thickness changes as the silicon layer thickness (t_{Si}) when t_{ox} is kept constant. So, mobile charge appears in the channel at low gate voltage for a higher t_{Si}. Accumulation charge is independent of t_{Si} which causes lower mobile charge density due to increased t_{Si}. Depletion layer width shows inverse relation with doping concentration (N_d), so mobile charge appears early at higher doped semiconductor. Accumulation charge is more for higher doped semiconductor. When it is normalised, it becomes less for larger N_d. Fig. 5 and Fig. 6 show the effect of t_{Si} and N_d on normalised mobile charge density respectively.

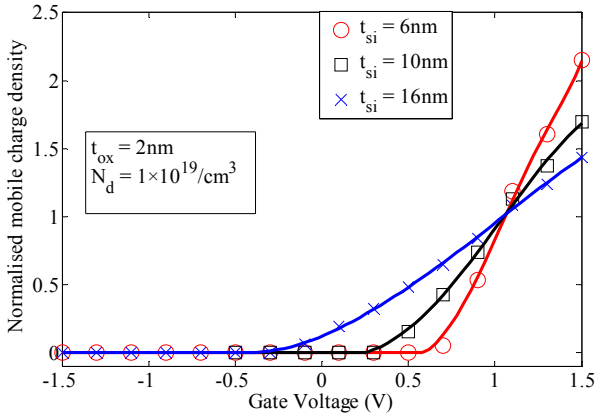


Fig. 5 Normalised mobile charge density variation with Si layer thickness.

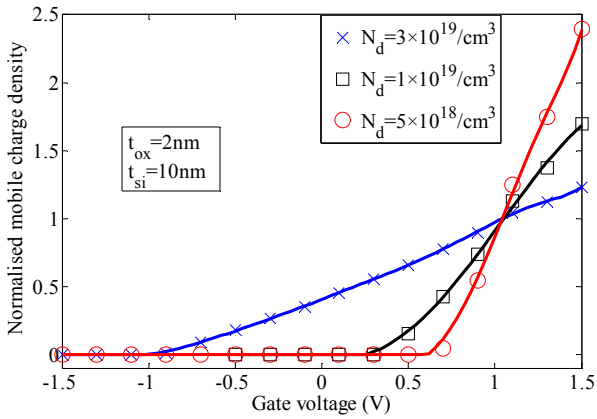


Fig. 6 Normalised mobile charge density variation with doping concentration.

B. Potential

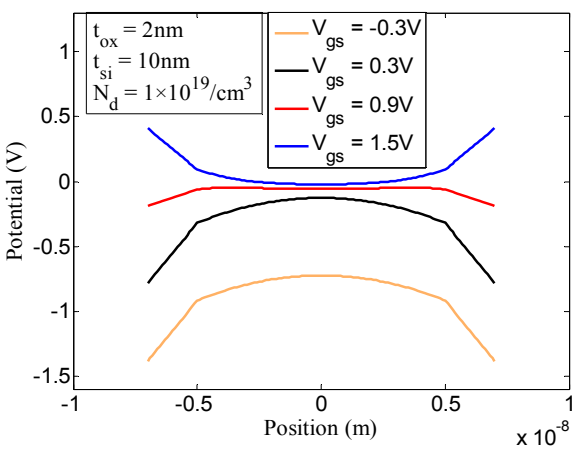


Fig. 7 Potential drop for different gate voltages

For zero gate voltage, potential drop across the oxide and semiconductor region is approximately equal to polysilicon semiconductor work function difference. This value encounters the applied gate voltage which is easily observable from Fig. 7. In the non depleted semiconductor region, potential drop is zero. So for an applied gate voltage, during which device is on, part of the semiconductor potential drop lies along zero value. As the gate voltage is increased potential drop also increases, so the curve shifts upward.

C. Drain Current

Current in the device depends on the gate voltage as shown in Fig. 8 which is similar to the dependence of mobile charge. As the gate voltage crosses the threshold value, current increases due to increase in non-depleted channel width. The transfer characteristics curve is shown for a particular drain voltage of 1V.

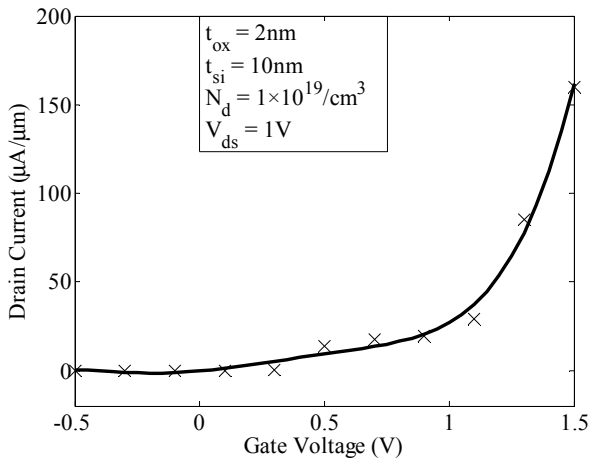


Fig. 8 Drain current dependence on gate voltage

Fig. 9 shows that when the gate voltage is above threshold voltage, increase in drain to source voltage causes increase in drain to source current until saturation is reached. Saturation occurs when drain voltage reaches a value such that the channel near drain region becomes fully depleted. As the gate voltage is increased, channel is widened, so larger drain voltage is required to close the channel near drain region. That's why saturation voltage is higher for higher gate voltage.

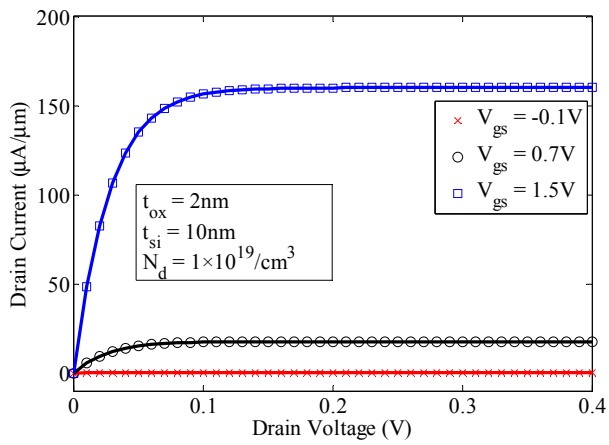


Fig. 9 Drain current variation with drain voltage for different gate voltages.

Threshold voltage is lower for larger oxide thickness because of the comparatively thin depletion region. Current in partially depleted mode is larger for larger oxide thickness while current in accumulation mode is less for larger oxide

thickness. These phenomena occur in the same logic as mobile charge. Fig. 10 and Fig. 11 show the effect of oxide thickness in partial depletion mode and accumulation mode respectively.

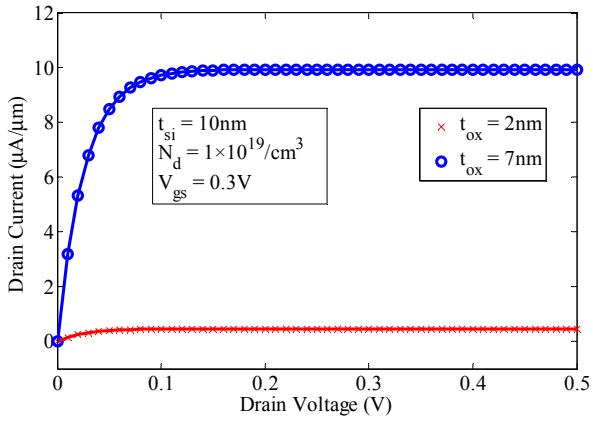


Fig. 10 Drain current variation with oxide thickness in partial depletion mode.

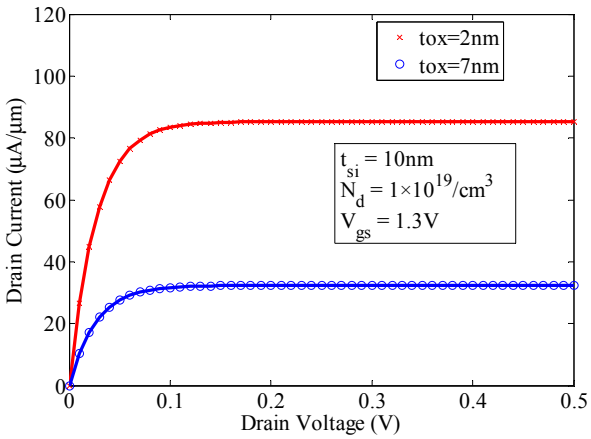


Fig. 11 Drain current variation with oxide thickness in accumulation mode.

D. Transconductance

Transconductance of junctionless transistor is zero up to threshold voltage. When gate voltage crosses threshold value, transconductance starts to increase. The transconductance curve in Fig. 12 has been extracted from the result obtained it Fig. 8.

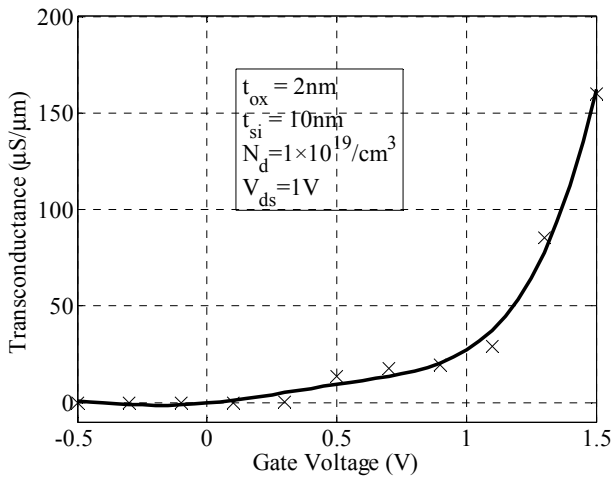


Fig. 12 Dependence of transconductance on gate voltage

VI. Conclusions

Self consistent simulation has been performed for n-type Si as semiconductor and SiO_2 as insulator oxide. The method is equally applicable for both p-type and n-type material which may be of any semiconductor and oxide material. Results obtained in the analysis are consistent with the mathematical models that have been presented in previous works on junctionless transistor. So, the method acts as a useful tool for investigating junctionless transistor and its application in modern electronic devices.

References

[1] J. Colinge, X. Baie, V. Bayot, E. Grivei, "A Silicon-on-Insulator Quantum Wire", *Solid-State Electronics* 1996; 39:49–51, 22 March 1995.

[2] Lee C., A. Afzalian, N. Akhavan, Yan R., I. Ferain, J. Colinge, "Junctionless Multigate Field-Effect Transistor", *Applied Physics Letters* 94, 053511, 6 February 2009.

[3] J. Colinge, Lee C., A. Afzalian, N. Akhavan, Yan R., I. Ferain, P. Razavi, B. O'Neill, A. Blake, M. White, A. Kelleher, B. McCarthy, R. Murphy, "Nanowire Transistors Without Junctions", *Nature Nanotechnology*, 21 February 2010.

[4] J. Colinge , A. Kranti, Yan R., Lee C., I. Ferain, Yu R., N. Akhavan, P. Razavi, "Junctionless Nanowire Transistor (JNT): Properties and Design Guidelines", *Solid-State Electronics*, 27 July 2010, pp 33–37.

[5] R. Trevisoli, R. Doria, M. Souza, M. Pavanello, "Threshold Voltage in Junctionless Nanowire Transistors", *Semiconductor Science and Technology* 26(2011) 105009 (8pp), 6 September 2011.

[6] J. Duarte, Choi S., Choi Y., "A Full-Range Drain Current Model for Double-Gate Junctionless Transistors", *IEEE Transactions on Electron Devices*, Vol. 58, No. 12, December 2011.

[7] J. P. Duarte, S.-J. Choi, D.-I. Moon and Y.-K. Choi, "Simple analytical bulk current model for long-channel double-gate junctionless transistors," *IEEE Electron Device Lett.*, vol. 32, no. 6, pp. 704–706, Jun. 2011.

[8] Y. Taur, X. Liang, W. Wang, and H. Lu, "A continuous, analytic drain-current model for DG MOSFETs," *IEEE Electron Device Lett.*, vol. 25, no. 2, pp. 107–109, Feb. 2004.

[9] J.-M. Sallese, N. Chevillon, C. Lallement, B. Iniguez, and F. Pregaldiny, "Charge-based modeling of junctionless double-gate field-effect transistors," *IEEE Trans. Electron Devices*, vol. 58, no. 8, pp. 2628–2637, Aug. 2011.

[10] N. Akhavan, I. Ferain, P. Razavi, YuR., J. Colinge, "Improvement of carrier ballisticity in junctionless nanowire transistors", *Applied Physics Letters* 98, 103510, 10 March 2011

[11] Peter Blood, *Quantum Confined Laser Devices: Optical Gain and Recombination in Semiconductors*, 1st Ed, Oxford University Press, 2015.

[12] J. Colinge, "Conduction mechanisms in thin-film accumulation-mode SOI p-channel MOSFETs", *IEEE Transaction Electron Devices*, vol. 37, no. 3,pp. 718–723, March 1990.

[13] Z. Ren, "Nanoscale MOSFETs: Physics, Simulation and Design", Ph.D. thesis, Purdue University, October 2001.

[14] Xinnan Lin, "Analytical Current Model for Long-Channel Junctionless Double-Gate MOSFETs", *IEEE Transaction Electron Devices*, vol. 63, no. 3, March 2016.

Vulnerability Analysis of Bangladesh Power System Network

Fahim Al Hasnain*, Atique Mahmud and A. H. Chowdhury

Department of EEE, Bangladesh University of Engineering and Technology
Dhaka-1000, Bangladesh

*himeldohar@gmail.com

***Abstract*—An electric power system network is always under stress. Because of its overall interconnectivity it becomes very much necessary to protect the power system if there is any major failure. Otherwise, it will cause the whole system to collapse. Vulnerability analysis of the power system network indicates the major vulnerable nodes that should be given special protective measures for reliable power transmission and operation. Centrality approach is an effective way to determine the most sensitive edges and nodes of a power system network. This mathematical based criterion has been applied with proper algorithm for Bangladesh power system network to evaluate the most vulnerable parts of the network.**

***Index Terms*— Nodes, Edges, Centrality, Edge Between-ness Centrality, Closeness Centrality, Floyd Warshall algorithm, Shortest Path, Weighted Matrix.**

I. Introduction

The electric power grid is one of the most critical infrastructures. The inter-connectivity of the power grid enables long-distance transmission for more efficient system operation; however, it also allows the propagation of disturbances in the network. The past instances of North American blackout, the Italian blackout and the blackout in India has cascading effect on other engineered system that depend on electric power grid. Therefore, it is very important to determine the vulnerable nodes or edges in a power system network.

Studies on the power grid system structures and vulnerability analysis have attracted many research efforts in the past years [1-6]. From a topological viewpoint, various measures of the importance of a network element (edge or node), i.e. of the relevance of its location in the network with respect to a given network performance, can be introduced. Metrics such as degree, closeness and between-ness are well-known and commonly used [7,8]. These utilize either local or shortest path information. Among them closeness and between-ness are cases which utilize shortest paths to estimate the importance of network components. The closeness is the reciprocal of farness and calculated as the reciprocal of sum of the distances from the given node to all the other nodes. The between-ness centrality is the number of shortest paths going through a node or a link.

In [9] Rajashigh et al. developed a formula to compute the between-ness centrality for a regular grid network. Torres and Anders discussed different methods of graph theory, mainly the topology centralities, for ranking the relative importance of substations in a power grid and illustrated the procedure on a synthetic 5-node test system [10]. In [11] Gorton et al. proposed a new method for contingency selection based on the concept of graph edge between-ness centrality which can be used for the contingency analysis of large scale power grids. Zio and Piccinelli in their work on vulnerability improved the model of power flow distribution in the power grid network [12]. That is, the flows are not concentrated only along the shortest paths; instead, they are randomly distributed on all the paths between nodes, as random walks. Also power flow between-ness centrality [5], multi path consideration and flow based approach have also been done to determine vulnerability.

In this paper, the closeness centrality and edge between-ness centrality measurement considering the flows concentrated only along the shortest paths has been used to find the most vulnerable nodes and edges in the Bangladesh power system network (BPSN).

II. Vulnerability Analysis Techniques

A power grid is typically composed of a lot of nodes and edges. Vulnerability of these nodes and edges are to be determined for reliability assessment. A node means a point in the system where two or more circuit elements meet and for a node to be unique its voltage is different than the other node. So a total line may be considered as a node if voltage is same everywhere in that line.

On the other hand, an edge is a connection between two nodes. Current flows from node to node through edges and generators, loads, capacitors etc. electrical elements are connected in various nodes. Impact of various nodes and edges is not equally the same because their interconnectivity varies and contribution to the overall grid in case of load, generation varies as well. To determine each of their impact on the overall system centrality is a powerful tool. Centrality refers to the extent to which a single node impacts on the overall network. In graph theory and network analysis the indicators of centrality ultimately show the significance of a node or edge in the big picture.

There are many types of centrality such as degree centrality (Topological approach), closeness centrality, edge between-ness centrality etc. We have used closeness centrality and edge between-ness centrality with Floyd-Warshall algorithm in our calculation to find out the vulnerability of Bangladesh Power System Network.

A. Closeness Centrality

Closeness centrality measures the closeness of different nodes in a connected system. Closeness is the opposite of farness, which is the summation of distances of every other connected node from a particular node. If $d(x,y)$ is the distance between nodes x and y, then closeness centrality is defined by (1).

$$C(x) = \frac{1}{\sum_{y=1}^{y=y} d(x,y)} \quad (1)$$

The node with high closeness centrality is thus less far from other nodes and has high vulnerability.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

B. Edge Between-ness Centrality

Edge between-ness centrality points out an edge's centrality in a network by indicating the number of shortest paths from all vertices to all others that pass through that particular edge. The more the centrality is, the more influence it has on the whole network. It is calculated by determining and using the shortest path from weighted matrix.

Weighted matrix denotes the minimum weight between all the nodes of a network. For power system network the per-unit value of resistances and reactances of all the edges are taken as weight between all the nodes. To calculate weighted matrix of a network, an adjacency matrix is taken as input, which is an initial weighted matrix defined by (2),

$$D(i,j) = 0 \quad \text{if, } i = j$$
$$= \infty \quad \text{if there is no edge between } i \text{ and } j$$

Otherwise,

$$D(i,j) = \text{"weight of edge"} \tag{2}$$

The D matrix is then updated for all the intermediate nodes. The completely updated D matrix contains the shortest weights between all nodes. When updating D matrix, every element of the matrix is updated for all intermediate node k.

The Floyd Warshall algorithm [8] determines the shortest path between all pairs of vertices. The *path* matrix stores the last vertex visited on the path from i to j. So $path(i,j) = k$ means that in the shortest path from vertex i to vertex j, the last vertex on that path before j is k. Based on this definition, the path matrix is initialized by (3).

$$path(i,j) = i, \text{ if } i \neq j \text{ and edge exists from } i \text{ to } j$$
$$= \text{NIL otherwise} \tag{3}$$

At each step of Floyd's, it is checked whether using vertex k improves an estimate of the distance vertices i and j. The *path* matrix is updated using the following algorithm.

$$\text{if,} \quad D(i,k)+D(k,j) < D(i,j)$$
$$\text{then,} \quad path(i,j) = path(k,j).$$

The shortest path matrix and the weighted matrix is then used to find edge between-ness centrality. For this method, three input matrices are taken, shortest path number matrix denoting the number of shortest paths from one node to every other node, the shortest path matrix and the weighted matrix. The next step is to find the head nodes and leaf nodes from shortest paths to calculate edge between-ness centrality values. To apply the edge between-ness analysis algorithm, the pathway directions are reversed such that the paths begin at the leaf nodes and end at the head node. The edge between-ness centrality of the edges is then calculated as described by Girvan and Newman [13]. The final results show how many times each edge occurs in between all the shortest paths from one node to other nodes.

III. Bangladesh Power System Network

The techniques described in section II have been used to find the most vulnerable nodes and edges of the BPSN. The transmission system of the BPSN mainly consists of transmission lines having voltages of 132 kV, 230 kV. A few kilometers of 400 kV lines are also there. In this work only the 132 kV and 230 kV lines have been considered. To determine edge between-ness centrality, the shortest paths from only the source nodes (nodes with generators) have been considered.

Table I presents the results for closeness centrality for BPSN. The calculated values have been normalised that means that the closeness centrality score of every node are divided by (n-1), n being the total number of nodes in BPSN. Closeness centrality is a topological measurement as it only counts if one node has connection to other nodes. It does not consider other parameters such as resistance or reactance or distance between two nodes. From table I, Ghorashal has a closeness centrality score of 0.2633 whereas Ashuganj has a score of 0.2567. It means that Ghorashal is more connected to other nodes than Ashuganj. So greater value of closeness centrality means a more centrally located node as well as more vulnerable node.

TABLE I
Closeness Centrality For BPSN

Node name	Closeness centrality (Normalised)
Ghorashal	0.2633
Ashugonj	0.2567
Comilla North	0.2455
Isurdi	0.2451
Baghabari	0.2394
Rampura	0.2255
Meghnaghat	0.2214
Tongi	0.2214
Narsingdi	0.2206
Bhulta	0.2206

Table II presents the edge between-ness centrality of BPSN. It shows top 10 edges which are more centrality located and number of occurrence of these edges in all the shortest paths from the source nodes is higher. If power flows from the source node to another node, it is highly probable that power will flow through these edges. Failure of any of these edges will have a greater effect on the power system operation. From table II, Isurdi-Natore edge has an edge between-ness centrality score of 96. This result represents that among all the shortest paths from the generator nodes to all the other nodes, Isurdi-Natore edge happens to occur 96 times. Similarly, Ghorashal-Ishurdi edge has a less score of 83 that means that it occurs 83 times. Higher edge between-ness centrality score means more vulnerability. Ishurdi-Natore edge is more vulnerable than Ghorashal-Ishurdi edge in BPSN.

TABLE II
Edge Between-ness Centrality of BPSN

Edge	Edge between-ness centrality
Ishurdi-Natore	96
Ghorashal-Ishurdi	83
Ashuganj-Ghorashal	82
Raozan-Hathazari	77
Natore-Bogra Old	72
Noapara-Jessore	70
Jessore-Jhenaidah	68
Bheramara-Ishurdi	66
Jhenaidah -Kushtia	62
Hathazari-Comilla North	62

From the two tables we can see that there are some similar names coming out of the two different approach such as Ishurdi, Ghorashal, Ashuganj, Comilla North etc. as the most vulnerable nodes in BPSN. Although closeness centrality is totally a topological measurement whereas the edge between-ness centrality approach uses weighted matrix and shortest paths to calculate vulnerability, they produce quite similar result.

Figure 1 shows the most vulnerable edges in the BPSN. In the whole network these paths can be called the backbone of BPSN as they have more edge between-ness centrality values according to our research. If any of these edges collapse, the whole network will be affected severely.

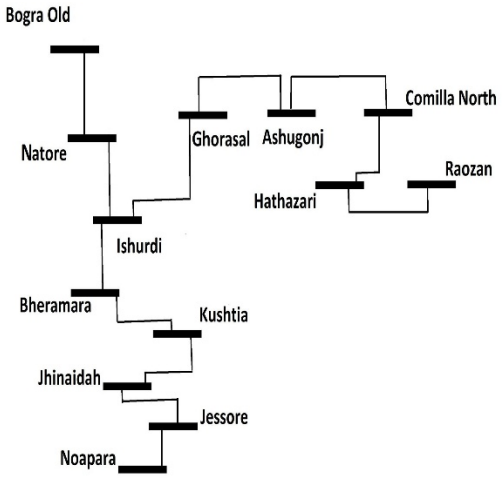


Fig. 1 The most vulnerable edges in BPSN.

IV. Conclusions

This paper evaluates the vulnerable nodes and edges in the Bangladesh power system network using two different types of centrality measurements. More protective measures should be taken in these nodes or edges to prevent sudden system collapse or blackout.

References

[1] Z. Wang, A. Scaglione and R. J. Thomas, "Electrical centrality measures for electric power grid vulnerability analysis," *49th IEEE conference on decision and control (CDC)*, pp. 5792-5797, 2010.

[2] P. Hines, S. Blumsack, "A centrality measures for electrical networks," *Hawaii international conference on system sciences, proceedings of the 41st annual*, pp. 185, 2008.

[3] R. Kong, C. Han, T. Guo and W. Pei, "An enegy-based centrality for electrical networks," *Electrical and Power Engineering*, vol. 5, pp. 597-602, 2013.

[4] F. Cadini, E. Zio and C. A. Petrescu, "Using centrality measures to rank the importance of the components of a Complex Network Infrastructures," *Critical Information Infrastructure Security, Third International Workshop, CRITIS 2008*, Rome, Italy, October 13-15, 2008.

[5] E. Zio, R. Piccinelli, M. Delfanti, V. Olivieri and M. Pozzi, "Performance analysis of a power transmission system under uncertain load conditions and network configurations," HAL Id: hal-00721062,26 Jul 2012. [Online]. Available: https://hal.archives-ouvertes.fr/hal-00721062

[6] A. S. Bhave, M. L. Crow, E. K. Cetinkaya, "Robustness of power grid topologies against centrality-based attacks," Missouri University of Science and Technology, Rolla, MO 65409, USA. [Online]. Available: http://web.mst.edu/~mstconets

[7] C. Mascolo lecture on centrality measures. [Online]. Available: https://www.cl.cam.ac.uk/teaching/1314/L109/stna-lecture3.pdf

[8] Network Centrality based on materials by Lada Adamic. [Online]. Available: http://cs.brynmawr.edu/Courses/cs380/spring2013/section02/slides/05_Centrality.pdf

[9] I. Rajasingh, B. Rajan, and F. I. D., "Betweeness-centrality of grid networks," in International Conference on Computer Technology and Development, Kota Kinabalu, Malaysia, Nov. 13–15 2009.

[10] A. Torres and G. Anders, "Spectral graph theory and network dependability," in Fourth International Conference on Dependability of Computer Systems, Brunow, Jun 30 – Jul 2 2009, pp. 356 – 363.

[11] I. Gorton, Z. Huang, Y. Chen, B. Kalahar, S. Jin, D. ChavarraMiranda, D. Baxter, and J. Feo, "A high-performance hybrid computing approach to massive contingency analysis in the power grid," in International Conference on e-Science and Grid Computing, Oxford, UK, Dec.09 – 11 2009, pp. 277–283.

[12] E. Zio and R. Piccinelli, "Randomized flow model and centrality measure for electrical power transmission network analysis," Reliability Engineering and System Safety, vol. 95, no. 4, pp. 379 – 385, Apr. 2010.

[13] M. Girvan, M. E. J. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 12, pp. 7821–7826, 2002.

Feature Selection Techniques for Increasing Reliability of Fault Diagnosis of Bearings

Md. Rashedul Islam[1,*], M M Manjurul Islam[2], Jong–Myon Kim[3]

[1]Department of Computer Science and Engineering
University of Asia Pacific, Dhaka, Bangladesh
[2, 3]School of Electrical, Electronics, and Computer Engineering,
University of Ulsan, Ulsan, South Korea
[*,1]rashed.cse@gmail.com, [2]m.m.manjurul@gmail.com, [3]jmkim07@ulsan.ac.kr

***Abstract*— Features selection (FS) techniques have an apparent need in many complex engineering applications especially the bearing fault diagnosis of low-speed industrial motor. The main goal of an FS algorithm is to select the most discriminant features subset from a high-dimension features vector that increases the model performance by reducing the redundant and irrelevant fault features. This paper proposes an efficient fault diagnosis model of bearing by incorporating the optimal feature selection approach for increasing the reliability of fault diagnosis of bearing. Also, this paper investigates the feature selection approaches including sequential forward selection (SFS), sequential floating forward selection (SFFS), and genetic algorithm (GA) for identifying the most discriminant subset. The effectiveness of this discriminant features subset is verified with a low-speed bearing fault diagnosis application for identifying bearing failures. The experimental shows up-to-mark diagnosis performance using GA based optimal feature selection method.**

Keywords—Acoustic emssion(AE); signal processing; feature selection; machine learning; Fault detection and diagnosis

I. INTRODUCTION

Bearings are the most crucial component of low-speed machinery because of its support to provide stationary rotational speeds with heavy load in industry [1]. The continuous heavy load and environmental contamination are subject to bearing defects [2], [3]. The undetected bearing defects eventually lead to unexpected machine failure and great economic losses in industry [4], [5].

Over the last several decades, model-based fault detection and diagnosis (FDD) techniques for rolling element bearings (REBs) have been widely employed for enhancing the reliability and availability of industrial systems subject to faults [6]. Due to the complexity of modern industrial process, data-driven approach has recently gained lots of attention instead of model-driven in research community because of its simplicity and effectiveness [7]. The essence of data-driven fault diagnosis is consisting of three basic steps: (1) data acquisition to extract meaningful features, (2) feature analysis to select effective fault features and (3) fault classification.

Thus, to realize highly reliable data driven FDD methods, it is necessary to deploy different signal processing techniques to extract intrinsic information about bearing defects from the signal of defective bearings [8], [9]. This paper, therefore, uses a heterogeneous feature model, which comprises statistical features from the time domain, frequency domain, and envelope spectrum of an AE signal to extract the as much as fault information.

In practices, a high-dimension feature vector may contains either redundant or irrelevant features, which can be one of the primary reasons for degraded diagnosis [7]. To address this issue, discriminant feature analysis technique, which is can be either feature selection (FS) or dimensionality reduction (DR) of the feature vector, is required to find the most discriminative fault features in the given high-dimensional feature vector [9]-[11].

The optimal feature selection process is a combination of two sub processes namely, i) feature subsets generation and ii) evaluate the feature subset based on some evaluation criteria. Based on the variation of techniques, the FS algorithms do trade off between the effectiveness and computational cost [8]. In this study, we investigate most widely used selection algorithms, i.e., sequential forward selection (SFS) [7], sequential forward floating selection (SFFS) [3], and genetic algorithm (GA) [8]. On the other hand, the feature selection process is categorized in wrapper and filter approach [3]. In this study the filter approach is employed for evaluating the feature selection methods for reliable bearing fault diagnosis.

Finally, the effectiveness of selected features via SFS, SFFS and GA is further verified with a fault diagnosis application by estimating classification accuracy using k-nearest-neighbour (k-NN) classifier.

The rest of this paper is structured as follows. Section 2 presents the experimental setup and AE signal acquisition. Section 3 describes the proposed fault diagnosis methodology, including heterogeneous feature extraction models, and optimal feature selection schemes. Section 4 discusses experimental results, and Section 5 presents concluding remarks

II. EXPERIMENTAL DATA ACCUSITION

Last several decades, the current signal, voltage signal, temperature measurement, and vibration signal are used in FDD [12]. However, Acoustic emission (AE) signal is appropriate for early stage fault detection and diagnosis by capturing low-energy signal [1], [13]. Take into consideration

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

the effectiveness, this study employs AE sensor for acquiring signal from bearing.

To verify the effectiveness of feature selection model, we use the experimental setup depicted in Fig. 1. In this experimental setup, cylindrical rolling elements bearings are used with both drive and non-drive end shafts. For collecting the bearing signal, an AE sensor is placed on the top of the bearing house of NDE shaft. AE signals are collected form sensor and stored digital signal in computer with 250 kHz sampling frequency.

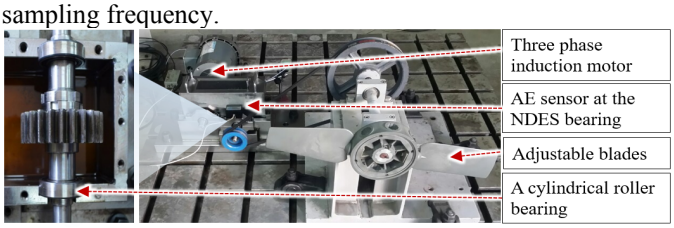


Fig. 1. (a) Experimental setup for collecting bearing fault signal and (b) data sampling and storing system

In this study, AE signals are collected from normal healthy bearing (NB) and cracked bearing (i.e., bearing crack of inner raceway (BCI), bearing crack on outer raceway (BCO), bearing crack on roller element (BCR)) under various shaft rotational speed (i.e., 300rpm, 350rpm, 400rpm, 450rpm, and 500rpm) conditions. Fig. 2 shows pictures the cracked bearings.

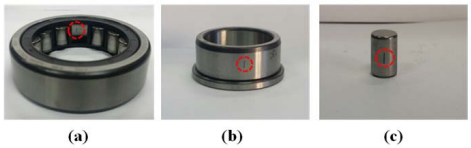


Fig. 2. Different cracked bearings

TABLE I. SUMMARY OF DATASETS

	Datasets				
	Dataset 1	*Dataset 2*	*Dataset 3*	*Dataset 4*	*Dataset 5*
RPMs	300	350	400	450	500
Crack sizes	Length: 3mm, Width 0.35mm, and Depth: 0.3mm				

Finally, we construct the experimental datasets of collected AE signals under different shaft rotational speed. Each dataset contains four types AE signal, each type contains 90 signals, and each signal is 5 sec length. Table I represents the summary of datasets.

III. COMPREHENSIVE FAULT DIAGNOSIS MODEL

Fig. 3 presents the flow diagram of comprehensive fault diagnosis model. This model is mainly two folds process (i.e., analysis process and evaluation process). The analysis process extracts the high dimensional hybrid features from known AE signals of normal and cracked bearing. In practices, the high-dimension feature vector may contain redundant or irrelevant information. So that, the optimal feature selection is applied in analysis process to select the optimal features for a machine condition. In evaluation process, only optimal features are extracted from unknown signal and *k*-NN classification algorithm is used to identify faults of unknown signal.

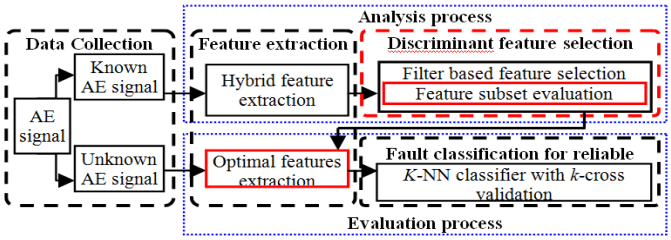


Fig. 3. Comprehensive feature selection mdoel for bearing fault diagnosis

A. An Hybrid Feature Model

To find the effective fault information, hybrid feature extraction form different signal domain is very much effective [8], [9]. To do this, this study extracts statistical feature from time and frequency domain of signals. Table II and Table III present list of extracting statistical features (parameters) from time domain and frequency domain of signals respectively.

TABLE II. EXTRACTED TIME-DOMAIN STATISTICAL FEATURES FROM AE SIGNAL

Root mean square (RMS)	Square root of the amplitude (SRA)
Kurtosis value (KV)	Skewness value (SV)
Peak-to-peak value (PPV)	Crest factor (CF)
Impulse factor (IF)	Margin factor (MF)
Shape factor (SF)	Kurtosis factor (KF)

TABLE III. EXTRACTED FREQUENCY-DOMAIN STATISTICAL FEATURES FROM AE SIGNAL

Frequency center (FC)	RMS frequency (RMSF)	Root variance frequency (RVF)

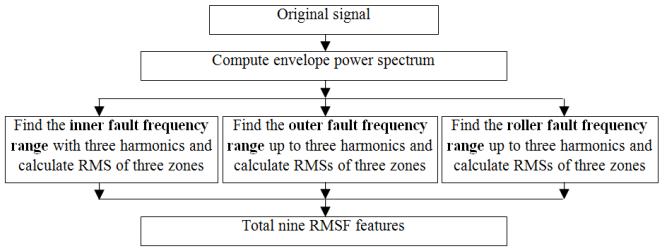


Fig. 4. RMSFs features extracting process from envelope spectrum.

Generally, the more intrinsic bearing fault information must be found around bearing fault frequency of signal. To extract that information, this study extracts RMSFs features from fault frequency ranges from envelope spectrum of signal. The determination of ranges is formulated by equation (1). Fig. 4 presents the RMSFs features extracting process from envelope spectrum.

$$\begin{aligned} N_{OFreqR,i} &= (1-RV_{order})\cdot(RPFO_i) \sim (1+RV_{order})\cdot(RPFO_i) \\ N_{IFreqR,i} &= (1-RV_{order})\cdot(RPFI_i - 2.F_{shaft}) \sim (1+RV_{order})\cdot(RPFI_i + 2.F_{shaft}) \\ N_{RFreqR,i} &= (1-RV_{order})\cdot(2\times RSF_i - 2.FTF) \sim (1+RV_{order})\cdot(2\times RSF_i + 2.FTF) \end{aligned} \quad (1)$$

where *RPFO* is roller pass on outer fault, *RPFI* is roller pass on inner fault, $2\times RSF$ is roller fault frequency, *FTF* is fundamental train frequency, F_{shaft} is shaft rotational speed, and RV_{order} if random variation of frequencies.

Finally, this study extracts 22 features consisting of 10 time-domain, 3 frequency-domain statistical features, and 9 envelop RMSFs features.

B. Feature Selection

The Fig. 5 depicts the methods and detail steps of optimal feature selection process.

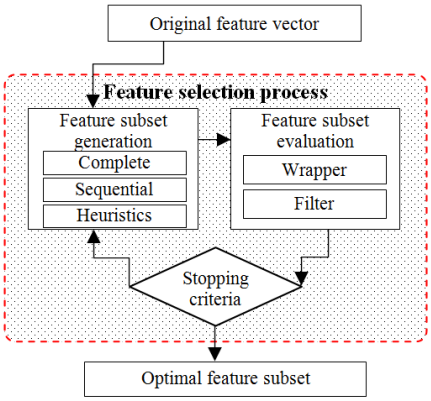


Fig. 5. Methodes and steps of optimal feature selection process.

An optimal feature selection method is the process of selecting an optimal subset of features which represents better distribution of samples in classification process. Based on the feature subset evaluation function, the feature selection processes are classified into two groups, i.e., wrapper selection approach and filter selection approach.

The wrapper based approaches evaluate the feature subsets based on performance measurement of a specific classification algorithm [14]. However, it is more computational costly. Alternatively, the filter approach evaluates the feature subsets based on some intrinsic information of distribution of samples in feature space [7]. Now a day the filter approach became more popular for its effectiveness. Thus, this paper considers the filter based selection approach.

The feature subset searching process of those feature selection techniques can be carried out using brute-force (i.e. complete search), sequential search, and heuristic search. The complete search ensures a high-quality subset, but it is computationally expensive.

The sequential search technique, especially in SFS the search process starts with empty subset and incrementally features are added to the subset and evaluate the goodness of that subset. Alternatively, the SFFS is an improvement of SFS with reverse feature dependency check. This process repeats until the all feature added or no better feature found [9]. The randomized approach, especially heuristic search (i.e., GA, hill climbing, branch and bound and particle swarm) used a large number of generations to find optimal subset form randomly generated subset called chromosome [11]. The above mentioned feature selection methods are good to select optimal features. However, there are some performance variations. Therefore, to validate the efficacy of feature selection process, this paper employed sequential forward selection (SFS), sequential floating forward selection (SFFS), and genetic algorithm (GA).

In feature selection approaches, there is a crucial task to assess the feature subsets on the basis of some evaluation function or rules from classifier. Most of the previous studies regrading feature evaluation methods have been proposed based on the classification accuracy or average intra-class compactness and inter-class separability based on Euclidean distance. One of the major shortcomings of this method is that this evaluation technique cannot consider proper feature distribution in the high-dimensional space. To overcome these apparent limitations of average distance based evaluation method Islam et al. recently proposed an efficient feature subset evaluation algorithm, maximum class separability (MCS), where the inter class separability and intra class compactness are calculated in effective way and the final objective value is calculated as the ratio of inter-class separability and intra-class compactness as Equation (2). In this study, the evaluation function proposed in [8] is adopted for evaluating the feature subsets in SFS, SFFS, and GA.

$$objective_value = \frac{between_class_distance}{within_class_compactness} \tag{2}$$

C. Identifing Fault using k-NN classifer

As depicted in Fig. 3, the ultimate goal of this study is to increase the fault classification performance. Once the most discriminant features are selected based on our feature selection algorithms that are further utilized to identify fault of unknown signal. In the diagnosis process, this study considers the k-NN classifier [15], an inexpensive and most widely used classifier, to validate our optimal feature sets in terms of the classification accuracy.

IV. Experimental Results

To validate the effectiveness, each dataset contains four types of signals including normal bearing signal and three cracked bearing signals are formed each dataset. Further, each dataset, consists of 90 sample signals, is equally divided into two parts. One part is used in analysis process and remaining part is used as unknown signal in evaluation process. In analysis process, the hybrid feature vectors are extracted for each signal and the optimal feature selection process is used for selecting optimal features for each conditional datasets. As we mentioned in previous section, the SFS, SFFS, and GA are used in feature selection process for comparing the performance of feature selection algorithms.

TABLE IV. SELECTED OPTIMUM FEATURES FOR DIFFERENT DATASETS VIA VARIOUS FEATURE SELECTION ALGORITHMS

Datasets	SFS	SFFS	GA
Dataset-1	$\{f_2, f_3, f_9, f_{13}, f_{21}\}$	$\{f_2, f_3, f_{13}\}$	$\{f_2, f_{13}\}$
Dataset-2	$\{f_2, f_6, f_9, f_{11}, f_{13}\}$	$\{f_2, f_3, f_9, f_{13}\}$	$\{f_2, f_3, f_9\}$
Dataset-3	$\{f_2, f_3, f_9\}$	$\{f_2, f_{13}\}$	$\{f_1, f_2\}$
Dataset-4	$\{f_2, f_6, f_{11}, f_{14}, f_{21}\}$	$\{f_2, f_9, f_{13}, f_{14}\}$	$\{f_2, f_9, f_{11}\}$
Dataset-5	$\{f_2, f_3, f_9, f_{13}\}$	$\{f_2, f_9, f_{13}\}$	$\{f_2, f_9\}$

For each dataset the feature selection algorithms select the optimal feature subset based on feature subset searching and MCS feature evaluation algorithm. Table IV presents the most discriminant features for different datasets using different feature selection algorithm that utilized in this study.

According to this result, the GA selects less number of effective features. In the evaluation process of proposed fault diagnosis model, the optimal features are used in *k*-NN classification algorithm (i.e., *k* set to 3 in *k*-NN in this study) for identifying faults of unknown samples. In classification process the *k*-cross validation (*k*-cv) is used for increasing the reliability of experimental result. In *k*-cv the evaluation dataset is randomly divided into *k* folds (here *k*=2). One fold is used as training and another for testing. The whole process is executed in $N_{iteration}$ time (here $N_{iteration}$=10). Finally, the generalized classification accuracy is calculated as the average of all estimated classification accuracies. Table V presents the classification performances of different feature selection algorithms using different datasets.

TABLE V. DETAILS SYSTEM PERFORMANCE OF DIFFERENT FEATURE SELECTION APPROACHES USING CONDITIONAL DATASETS

# Dataset	Feature selection algorithms	Average classification accuracy per class				ACA(%)
		BCI	**BCR**	**BCO**	**NB**	
Dataset-1	SFS	92.56	94.15	92.56	96.45	**93.93**
	SFFS	93.54	94.75	95.12	97.41	**95.21**
	GA	96.94	96.30	97.32	97.50	**97.02**
Dataset-2	SFS	93.10	94.25	95.68	97.87	**95.23**
	SFFS	94.52	96.26	94.58	97.45	**95.70**
	GA	98.17	97.72	98.60	98.89	**98.35**
Dataset-3	SFS	92.12	93.68	93.56	97.19	**94.14**
	SFFS	95.38	96.12	98.40	97.53	**96.86**
	GA	98.60	98.89	100.00	100.00	**99.37**
Dataset-4	SFS	94.10	94.89	95.68	97.19	**95.47**
	SFFS	96.38	97.89	98.61	100.00	**98.22**
	GA	100.00	99.16	99.80	100.00	**99.74**
Dataset-5	SFS	95.06	98.00	96.94	94.00	**96.00**
	SFFS	96.49	99.65	99.72	100.00	**98.97**
	GA	100.00	100.00	100.00	100.00	**100.00**

The experimental result indicates that the classification performances are varied depending on effectiveness of the feature selection algorithms. According to the experimental results, GA shows the better performance in most of the cases than that of SFS and SFFS.

V. CONCLUSION

This paper presents a highly reliable fault diagnosis model with filter based feature selection method for improving the reliability of bearing in industrial motor. In the proposed model, hybrid features are extracted from AE bearing signal and select optimal features for classifying faults in the diagnosis process. In this paper the SFS, SFFS, and GA based feature selection models are introduced for selecting optimal features. After embedding the feature selection in fault diagnosis process, the diagnosis performance is increased to 100%. However, SFS, SFFS, and GA show their highest accuracy 96.00%, 98.97%, and 100% for dataset-5.

ACKNOWLEDGEMENTS

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning(KETEP) and the Ministry of Trade, Industry & Energy(MOTIE) of the Republic of Korea (No. 20162220100050), in part by The Leading Human Resource Training Program of Regional Neo industry through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT and future Planning(NRF-2016H1D5A1910564), in part by Business for Cooperative R&D between Industry, Academy, and Research Institute funded Korea Small and Medium Business Administration in 2016 (Grants No. C0395147, Grants S2381631).), in part by the development of a basic fusion technology in electric power industry(Ministry of Trade, Industry & Energy, 201301010170D), and in part by a study on the "Leaders INdustry-university Cooperation" Project, supported by the Ministry of Education(MOE).

REFERENCES

[1] A. Widodo, E. Y. Kim, J.-D. Son, B.-S. Yang, A. C.C. Tan, D.-S. Gu, B.-K. Choi, and J. Mathew, “Fault Diagnosis of Low Speed Bearing Based on Relevance Vector Machine and Support Vector Machine,” Expert Systems with Applications, vol. 36, no. 3, part 2, pp. 7252–7261, 2009.

[2] M. Kang, J. Kim, J.-M. Kim, A. C. C. Tan, E. Y. Kim, and B.-K. Choi, “Reliable Fault Diagnosis for Low-Speed Bearings Using Individually Trained Support Vector Machines with Kernel Discriminative Feature Analysis”, IEEE Trans. on Power Elec., vol. 30, no. 5, pp. 2786–2797, 2015.

[3] M. Kang, R. Islam, J. Kim, J.-M. Kim, and M. Pecht, “A Hybrid Feature Selection Scheme for Reducing Diagnostic Performance Deterioration Caused by Outliers in Data-Driven Diagnostics”, IEEE Trans. Ind. Elec., vol. 63, no. 5, pp. 3299–3310, 2016.

[4] M. Zhao, X. Jin, Z. Zhang, B. Li, “Fault Diagnosis of Rolling Element Bearings via Discriminative Subspace Learning: Visualization and Classification,” Expert Systems with Applications, vol. 41, no. 7, pp. 3391–3401, 2014.

[5] J. Uddin, R. Islam, J. Kim, “Texture Feature Extraction Techniques for Fault Diagnosis of Induction Motors,” Journal of Convergence, vol. 5, no. 2, pp. 15–20, 2014.

[6] J. Seshadrinath, B. Singh, and B. K. Panigrahi, “Vibration Analysis Based Interturn Fault Diagnosis in Induction Motors,” IEEE Trans. Ind. Informat., vol. 10, no. 1, pp. 340–250, 2014.

[7] R. Islam, S. A. Khan, and J. Kim, “Discriminant Feature Distribution Analysis-based Hybrid Feature Selection for Online Bearing Fault Diagnosis in Induction Motors,” Journal of Sensors, vol. 2016, pp. 1–16, 2016.

[8] R. Islam, S. A. Khan, and J.-M. Kim, “Maximum Class Separability-based Discriminant Feature Selection Using A GAfor Reliable Fault Diagnosis of Induction Motors,” Lecture Notes in Artificial Intelligence, vol. 9227, pp. 526–537, 2015.

[9] T. W. Rauber, F. A. Boldt, and F. M. Varejao, “Heterogeneous Feature Models and Feature Selection Applied to Bearing Fault Diagnosis,” IEEE Trans. Ind. Elec., vol. 62, no. 1, pp. 637–646, 2015.

[10] Z. Xia, S. Xia, L. Wan, and S. Cai, “Spectral Regression Based Fault Feature Extraction for Bearing Accelerometer Sensor Signals,” Sensors, vol. 12, pp. 13694–13719, 2012.

[11] M. Mahrooghy, H. Y. Nicolas,“On the Use of the Genetic Algorithm Filter-Based Feature Selection Technique for Satellite Precipitation Estimation,” IEEE Geoscience And Remote Sensing Letters, vol. 9, no. 5, 2012.

[12] M. Kang , J. Kim , B.-K. Choi , J.-M. Kim, “Envelope Analysis with a Genetic Algorithm-Based Adaptive Filter Bank for Bearing Fault Detection,” JASA-EL, vol. 138, no. 1, pp. 65 – 70, 2015.

[13] A. Widodo, B.-S. Yang, E. Y. Kim, A. C. C. Tan, and J. Mathew, “Fault Diagnosis of Low Speed Bearing Based on Acoustic Emission Signal and Multi-Class Relevance Vector Machine,” Nondestruct. Test. Eval., vol. 24, no. 4, pp. 313–328, 2009.

[14] B. Li, P. –L. Zhang, H. Tian, S. –S. Mi, D. –S. Liu, and G. –Q. Ren, “A New Feature Extraction and Selection Scheme for Hybrid Fault Diagnosis of Gearbox,”Expert Syst. Appl., vol. 38, pp. 10000–10009, 2011.

[15] S. A.Dudani,“The Distance-Weighted k-Nearest-Neighbor Rule, IEEE Transactions on Systems,” Man and Cybernetics, vol. SMC-6, no. 4, pp. 325,327, 1976.

Optical Characterization of Spincoated Multiwall Carbon Nanotube Films on Silicon Substrates

Muhammad Athar Uddin[1*, 2], T. Soga[3] and Sharif M Mominuzzaman[1]

[1]Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology, Dhaka-1000, Bangladesh

[2]Department of Electrical and Electronic Engineering, International Islamic University Chittagong, Chittagong-4203, Bangladesh

[3]Department of Environmental Technology and Urbane Planning, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya, 466-8555, Japan

*mau_iiuc@yahoo.com

Abstract—**Multiwall Carbon Nanotube (MWCNT) films were spincoated on Silicon (Si) substrate. The effect of 0.1% and 0.2% solutions of MWCNT incorporation with methanol were investigated. The spincoated MWCNT films were characterized by SEM and UV-VIS-IR spectroscopy. In SEM micrographs, changes were observed in the spincoated films compared to pure Si substrates. Existence of carbon nanowire was also observed in SEM micrograph. From the UV-VIS-IR measurement it was observed that the absorption of light has been increased by the increase of MWCNT. So the optoelectronic properties of the electrodeposited films are tunable by spincoating using MWCNT on Si substrates.**

***Keywords*—Thin Film, Spincoating, Multiwall Carbon Nanotube.**

I. Introduction

The unique and tunable properties of carbon based nanomaterials like bucky ball (C_{60} or fullerene), Carbon nanotubes (CNTs) enable new technologies for optoelectronic devices. Carbon nanotubes [1] are gaining prominence as potential components of future nanoelectronic and nano-mechanical systems [2]. The CNTs exceptional electronic [3] and mechanical [4-7] properties can be very sensitive functions of the wrapping indices (n, m). As a new member of the carbon family fullerene (C_{60}), single-walled carbon nanotubes (SWCNT) and multi-walled carbon nanotubes (MWCNT) have exhibited great potentials in myriad applications including electronics, composite materials, fuel cells, sensors, optical devices and biomedicine [8]. Mechanical, electrical and electrochemical properties of the carbon nanotubes are well established and have immediate applications. The tunability of properties is potentially useful in optics and photonics. In particular, light-emitting diodes (LEDs) [9]- [10] and photo-detectors [11] based on a single nanotube have been produced in the lab. The control of carriers in carbon nanotubes is of critical importance prior to an industrial application in transistors [12].

Interest in deposition of carbon thin film has been motivated by properties of this material and the demand of modern technologies. The precursors and method of deposition of carbon films are the two dominating factors that dictate the optoelectronic properties of the film. Therefore, researchers on finding alternative precursor materials and simple method of deposition have been getting priority all the time. In connection with this research, MWCNT has been used as an alternative precursor material.

Hao wang *et al.* [13] deposited DLC films by using methanol solution as precursor material. Teresa *et al.* deposited [14] CNTs on Si substrates by CVD. Tomoya *et al.* electrodeposited [15] carbon nanotubes on Zinc Oxide. Dong-Jin Yun *et al.* [16] deposited MWCNT by spin coater. Yi Jia *et al.* [17] fabricated silicon-carbon nanotube heterojunction solar cells by acid doping. Jiangtao Di *et al.* [18] achieved 10.5% power conversion efficiency (PEC) from schottky solar cell composed of aligned carbon nanotubes on silicon substrates. Steve *et al.* [19] attempted for fabricating logic gates using fullerene with single-walled carbon nanotubes. Taketo *et al.* [20] investigated intersection of carbon nanotubes on the carrier accumulation under an external electric field. Mominuzzaman *et al.* [21] reported semiconducting carbon films by PLA from camphor ($C_{16}H_{16}O$). So it can be concluded that semiconducting thin films of carbon nanomaterials can be deposited on Silicon (Si) by very simple electroplating or spincoating technology to tune and improve their optoelectronic properties.

In connection to this few electrodeposion works have been reported [22]-[26] where carbonic thin films were deposited on silicon (Si), cupper (Cu) and aluminium (Al) substrates by electrolysis of methanol.

In the present article considering all these continuations from electrodeposition in methanol, deposition of CNT on Si substrate by CVD, electroplating of CNTs on ZnO and deposition of MWCNT by spincoating an attempt has been made for spincoating of MWCNT on Si substrate. Varying concentration of MWCNT materials were mixed to prepare carbon nanotube-methanol solution. As per our knowledge this type of work has been reported first time.

II. Experimental Details

The spincoating has been conducted by the vacuum spin coater VTC 100. Prior to deposition, the substrates were cleaned in electronic grade acetone (CH_3COCH_3) and methanol (CH_3OH) for 5 minutes at 55°C. MWCNT with varying amount (0.1% and 0.2%) was mixed in methanol solvent to prepare the electrolytes. The substrate with the electrolytes was rotated at 500 rpm for 60 seconds and at 3000 rpm for 60 seconds.

III. Results and Discussions

The spincoated MWCNT (0.1% and 0.2%) films were characterized by SEM spectroscopy. Figure 1, 2 and 3 are the

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

SEM micrographs of pure silicon, silicon substrate with 0.1% MWCNT film and silicon substrate with 0.2% MWCNT film respectively. Changes are observed between the pure Si substrate and spincoated MWCNT films. CNT nanowires are also observed in MWCNT thin films.

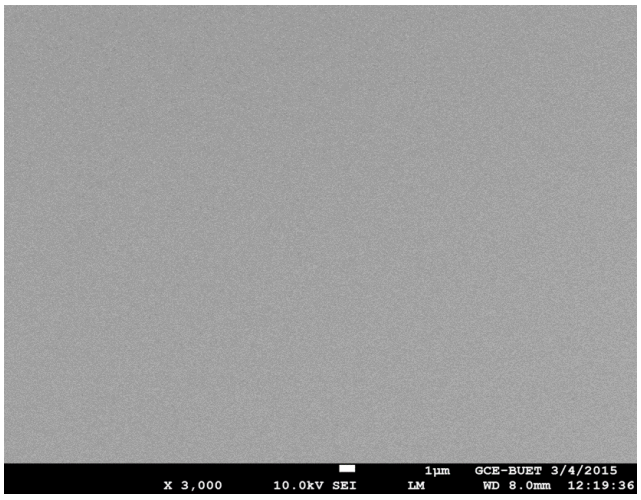


Fig. 1 SEM micrograph of pure silicon substrate

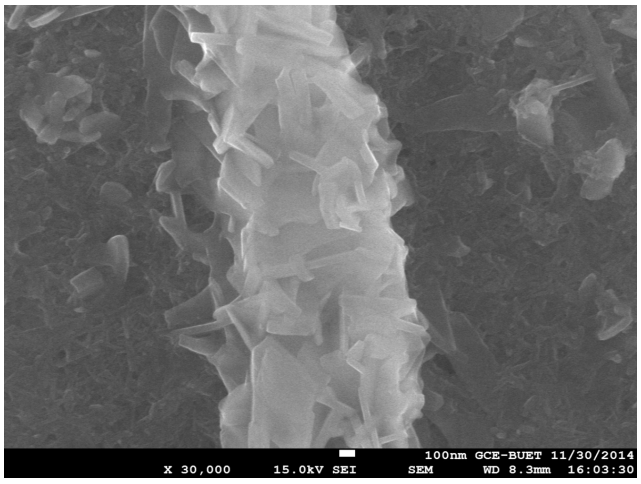


Fig. 2 SEM micrograph of 0.1% MWCNT spincoated Si substrate.

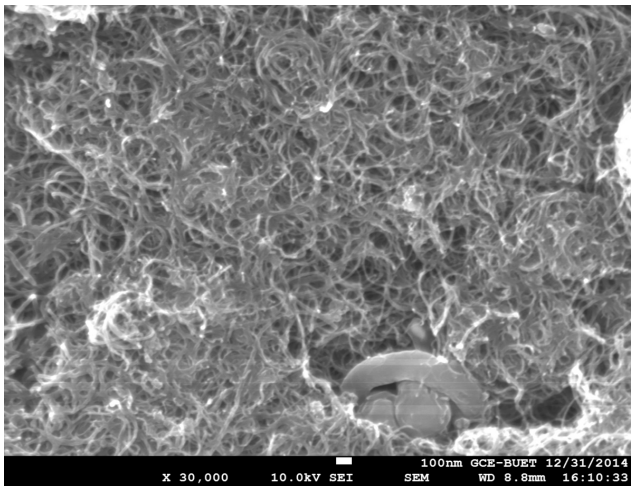


Fig. 3 SEM micrograph of 0.1% MWCNT spincoated Si substrate.

The MWCNT spincoated thin films are also characterized by UV-VIS-IR spectroscopy. Figure 4 shows the comparative analysis of the spectrum of % transmission of photon as a function of photon wavelength. Figure 5 shows the similar comparative analysis for absorption of photon.

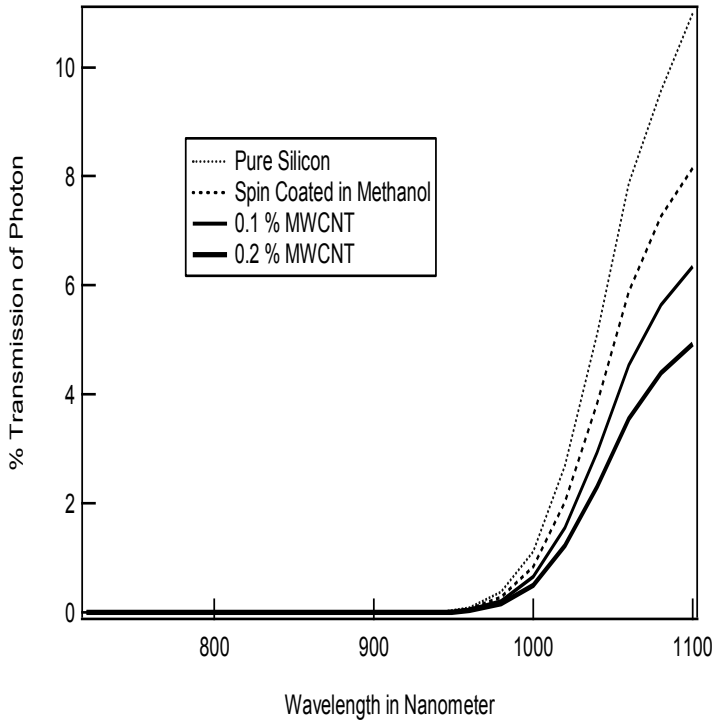


Fig. 4 % Transmission as a Function of Photon Wavelength for Spin Coated Films Using MWCNT

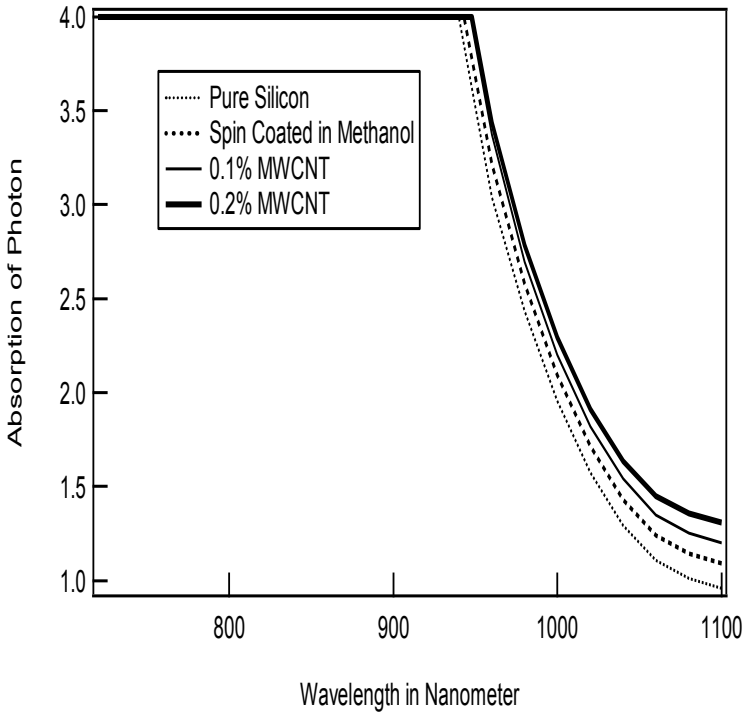


Fig. 5 Absorption Spectrum as a Function of Photon Wavelength for Spin Coated Films Using MWCNT

Transmission of photon is higher for pure Si substrates. Transmission decreases for the films spincoated using only methanol solution. Transmission decreases again for spincoated films using 0.1% MWCNT. After increasing the content of MWCNT from 0.1% to 0.2% the transmission also

decreases. So a successive decreasing transmission pattern is observed for successive incorporation of MWCNT in methanol solution. A complete reverse picture is found in case of absorption spectrum, which shows successive increasing of absorption by successive increasing of MWCNT in methanol solution. It reveals that, optoelectronic properties of carbon thin films are tuneable by MWCNT. The optical Bandgap is decreasing by increasing of photon absorption. So optical Bandgap is decreasing by increasing of MWCNT through spincoating on Si substrates. It was reported earlier [26] during formation of MWCNT thin films by electrodeposition, where carbon nano materials have almost similar influence or effect regarding tuning the optoelectronic property. So in both electrodeposition and spincoating, increasing of MWCNT increases the photon absorption and decreases the optical Bandgap.

IV. Conclusions

MWCNT films were spincoated on Si substrates. SEM micrographs show the existence of carbonic film for spincoating by methanol solution compared with pure Si substrate. CNT nanowires are observed by MWCNT incorporation through spincoating on Si substrate. UV-VIS-IR spectroscopy of spincoated MWCNT film shows tunability of optical transmission and absorption photon. Increase of MWCNT incorporation in spincoated thin films increases photon absorption and obviously decreases the optical Bandgap, which often will show increase of conductivity. The effect of electrodeposition is almost similar to that of spincoating regarding MWCNT incorporation in methanol solution to deposit thin films on Si substrates.

V. Acknowledgements

The authors would like to thank Dr. Abdul Gafur for providing UV-VIS-IR spectroscopy at BCSIR and also thankful to the Dept. of GCE, BUET for providing SEM facility.

References

[1] Iijima, S., "Helical microtubules of graphitic carbon", Nature (London), Vol-354, pp. 56-58, November 1991.

[2] Baughman, R. H., Zakhidov, A.A., de Heer, W.A., "Carbon Nanotubes - the Route Toward Applications", Science, Vol-297, No. 5582, pp. 787-792, August 2002.

[3] Hamada, N., Sawada, S. and Oshiyama, A., "New one-dimensional conductors: Graphitic Microtubules", Phys. Rev. Lett., Vol-68, No. 10, pp. 1579-1581, March 1992.

[4] Samsonidze, G. G. and Yakobson, B.I., "Kinetic Theory of Symmetry-Dependent Strength in Carbon Nanotubes", Phys. Rev. Lett., Vol-88, No. 6, pp. 065501-4, February 2002.

[5] Buongiorno Nardelli, M., Yakobson, B. I. and Bernholc, J., "Mechanism of strain release in carbon nanotubes", Phys. Rev. B, Vol-57, No. 8, pp. R4277-R2280, February 1998.

[6] Walters, D. A., et al., "Elastic strain of freely suspended single-wall carbon nanotube ropes", Appl. Phys. Lett., Vol.-74, No. 25, pp. 3803-3805, June 1999.

[7] Lammert, P. E., Zhang, P. and Crespi, V. H., "Gapping by Squashing: Metal-Insulator and Insulator-Metal Transitions in Collapsed Carbon Nanotubes", Phys. Rev. Lett., Vol-84, No. 11, pp. 2453-2456, March 2000.

[8] Banerjee, Hemraj-Benny, S., T., Wong, S. S., "Covalent Surface Chemistry of Single-Walled Carbon Nanotubes", Adv. Mater., Vol-17, No. 1, pp. 17-29, January 2005.

[9] J. A. Misewich et al., "Electrically Induced Optical Emission from a Carbon Nanotube FET," *Science* vol. 300, issue 5620, pp. 783-786, 2003.

[10] J. Chen et al., "Bright Infrared Emission from Electrically Induced Excitons in Carbon Nanotubes," *Science*, vol. 310, issue 5751, pp. 1171-1174, 2005.

[11] M. Freitag et al., "Photoconductivity of Single Carbon Nanotubes," *Nano Letters*, vol. 3, issue 8, pp. 1067–1071, 2003.

[12] Woo Jong Yu, Young Hee Lee, "Strategy for Carrier Control in Carbon Nanotube Transistors," *ChemSusChem,* vol. 4, no. 7, pp. 890-904, 2011.

[13] Hao wang, Ming-Rong Shen, Zhao-Yuan Ning, and Chao Ye., "Deposition of diamond-like carbon films by electrolysis of methanol solution," *Appl. Phys. Lett.,* vol. 69, pp. 1074-1076, 1996.

[14] Teresa de les Arcos, Peter Oelhafen and Daniel Mathys, "The importance of catalyst oxidation for the growth of carbon nanotubes on Si substrates", *J. of Carbon*, vol. 47, issue 8, pp. 77-82, 2009.

[15] Tomoya Kuwajima, Yusuke Nakanishi, "Fabrication of Carbon Nanotube/Zinc Oxide Composite Films by Electrodeposition", *Jpn. J. Appl. Phys.*, vol. 50, pp. 085504-085506, 2011.

[16] Dong-Jin Yun, KiPyo Hong, Se hyun Kim, Won-Min Yun, Jae-young Jang, Woo-Sung Kwon, Chan-Eon Park, Shi-Woo Rhee, "Multiwall Carbon Nanotube and Poly(3,4-ethylenedioxythiophene): Polystyrene Sulfonate (PEDOT:PSS) Composite Films for Transistor and Inverter Devices" *Applied Materials and Interfaces.*, Vol-3, No. 1, pp. 43-49, 2011.

[17] Yi Jia, Anyuan Cao, Xi Bai, Zhen Li, Luhui Zhang, Ning Guo, Jinquan Wei, Kunlin Wang, Hongwei Zhu, Dehai Wu and P. M. Ajayan, "Achieving High Efficiency Silicon-Carbon Nanotube Heterojunction Solar Cells by Acid Doping" *Nano Letr.* XXXX, XXX, 000–000, 2011.

[18] Jiangtao Di, Zhenzhong Yong, Xinhe Zheng, Baoquan Sun and Qingwen Li, "Aligned Carbon Nanotubes for High-Efficiency Schottky Solar Cells" *small 2013*, 9, No. 8, 1367–1372, March 2013.

[19] Steve Park, Ji Hyun Nam, Ja Hoon Koo, Ting Lei, and Zhenan Bao, " Enhancement of ambipolar characteristics in single-walled carbon nanotubes using C60 and fabrication of logic gates," *Applied Physics Letters* 106, 103501 (2015).

[20] Taketo Kochi and Susumu Okada, "Effect of an intersection of carbon nanotubes on the carrier accumulation under an external electric field," *The Japan Society of Applied Physics*, vol. 9, no. 8, 2016.

[21] S. M. Mominuzzaman, T. Soga, T. Jimbo, and M. Umeno, "Camphoric carbon soot: a new target for deposition of diamond-like carbon films by pulsed laser ablation," *Thin Solid Films*, vol. 376, pp. 1-4, 2000.

[22] Md. Mahmud Hasan, Muhammad Athar Uddin, S. M. Mominuzzaman, "An Experimental Approach of DLC Film Deposition on Metal Substrates,"*5th International Conference on Electrical and Computer Engineering ICECE 2008,* pp.803-806,20-22 December 2008, Dhaka, Bangladesh.

[23] Muhammad Athar Uddin, S.M. Mominuzzaman, "Semiconducting Carbon Thin Film Deposition on Silicon by Electroplating,"*5th International Conference on Electrical and Computer Engineering ICECE 2008*, pp. 846-850, 20-22 December 2008, Dhaka, Bangladesh.

[24] Muhammad Athar Uddin, Md. Shamimul Haque Choudhury and Sharif M Mominuzzaman, "Characterization of Carbon Thin Films Electrodeposited on Aluminum, Copper and Silicon Substrates,"*6th International Conference on Electrical and Computer Engineering ICECE 2010*, pp. 206-209, 18-20December 2010, Dhaka, Bangladesh.

[25] Muhammad Athar Uddin, M. Shamimul Haque Choudhury and Sharif M Mominuzzaman, "Optical properties of electrodeposited carbon thin films using camphor,"*Bangladesh Journal of Physics*, vol. 9 & 10, pp. 9-19, 2011.

[26] Mohammad Wahidur Rahman, Muhammad Athar Uddin, Ahmedullah Aziz, Tanvir Mustofa, Zahid Hasan Mahmood, T. Soga and Sharif M Mominuzzaman, " Characterization of Electrodeposited Multiwall Carbon Nanotube Films on Silicon Substrates,"*8th International Conference on Electrical and Computer Engineering ICECE 2014*, pp. 373-376, 20-22December 2014, Dhaka, Bangladesh.

Gate dielectric dependent performance of GNR MOSFET: A tight binding study

Md. Shamim Sarker*, Al-Mohtashim Sabbik+, Muhammad Mainul Islam $, Md. Nur Kutubul Alam&, Md. Rafiqul Islam#
Department of Electrical and Electronic Engineering,
Khulna University of Engineering & Technology,
Khulna-9203, Bangladesh
E-mail: *sarkershamimbd@gmail.com, + almothashim@gmail.com, $ mainul0803004@gmail.com, &alamjhilam@gmail.com
mri@eee.kuet.ac.bd

***Abstract*—This paper presents the gate oxide dielectric strength and its thickness-dependent performance of a graphene nanoribbon MOSFET (GNRMOSFET). Here we have studied the transfer characteristics, on/off current (I_{ON}/I_{OFF}) ratio, subthreshold slope and drain induced barrier lowering (DIBL) of the device using Non Equilibrium Greens Function (NEGF) formalism in tight binding frameworks. The results are obtained by solving the NEGF and Poisson's equation self-consistently in NanoTCAD ViDES environment and found to have strong dependence on the oxide thickness as well as its dielectric strength.**

Keywords- GNR MOSFET; subthreshold slop; DIBL; Oxide thickness; Tight Binding

I. INTRODUCTION

Silicon oriented devices are very much popular in semiconductor industries due to their extraordinary matured process technology and low cost. However, following the Moor's law when device size goes well below 100 nm, this technology has reached to its physical limit [1] .This limit imposes different adverse effects like short channel effects, parasitic capacitance, and leakage currents etc. , and these effects become more severe when gate length is scaled continuously[2]-[3].To reduce these effects and at the same time to keep the rhythm with the Moore's law, researchers are searching for alternate materials to replace the existing silicon technology. Graphene nanoribbon (GNR) is being considered as a potential candidate as channel material that would allow further device scaling due to their Nano scale size and unique electronic properties e.g. large carrier velocity, near ballistic transport and reduced parasitic capacitance, low effective mass and compatible with planner processing[4]-[7]. Owing to the large carrier velocity and reduced parasitic capacitance, GNR MOSFETs are expected to have several times faster intrinsic speed than that of Si MOSFET. Keeping these in mind GNR based MOSFET with top-gate control has recently been reported experimentally [8], that prove the superiority of the GNR. However at the ultra-scaled gate length, thickness of the gate oxide needs to be scaled significantly in order to get the required value of gate capacitance that is mandatory for achieving proper control over the channel according to Dennard's scaling rule[9]. But thickness scaling causes some severe size induced problem such as gate leakage, threatening the reliability of the device. One way of getting required gate capacitance without reducing the oxide thickness is to use oxides/insulators having higher dielectric constant[10].

In this article we report the performance of GNRMOSFET taking into account of various high-k oxide materials and their thicknesses. The GNR FET under study has an n-type doped source and drain reservoirs. Note that, the GNR does not have a fixed energy gap in its band structure; rather it shows an oscillating behavior which is roughly a function of the ribbon width [11]. In order to take this fact into consideration, the current transport is calculated in the tight binding approach. The current voltage (I-V) characteristics are obtained self-consistently solving the NEGF and Poisson's equation in NanoTCAD ViDES environment. Other performances of the device like I_{ON}/I_{OFF} ratio and subthreshold slope and drain induced barrier lowering are studied from the transfer characteristics in conjunction with band diagram.

This article is organized as follows: device structure and simulation methods are explained under section II, while section III demonstrate the details results and discussion, and section IV contains the summery.

II. SIMULATION DETAILS

A. Device Structure

The double gate MOSFET structure used in our simulation is shown in "Fig .1". As channel, we have used GNR with chirality (6,0) having a length of 10 nm. GNR in this chirality has a band gap of 0.61 eV. As source and drain we used the doped extension of the channel with n type dopant. The doping density of the source and drain reservoirs is taken $f = 3\times10^{-3}$. On the top and bottom of channel we used gate insulator that is extended over the source, and drain as well.

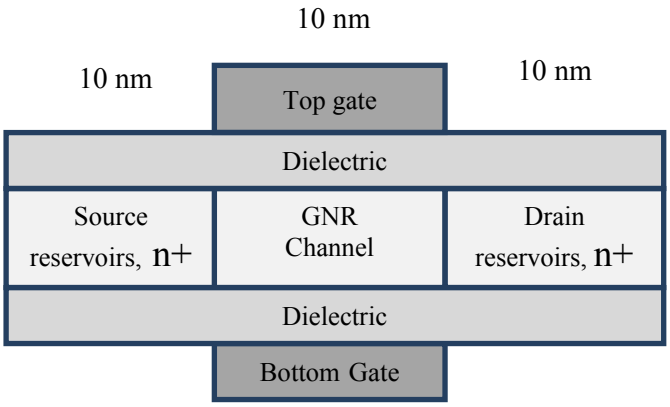


Figure 1. Proposed GNR MOSFET structure.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

B. Simulation approach

The GNR MOSFET characteristics are determined by solving Schrodinger equation using NEGF formalism, which are solved self consistently with 3-D Poison's equation. The 3D Poison's equation is given by [12]-[13]

$$\nabla[\varepsilon(\vec{r})\nabla\phi(\vec{r})] = -q[p(\vec{r}) - n(\vec{r}) + N_D^+(\vec{r}) - N_A^+(\vec{r}) + \rho_{fix}] \quad (1)$$

where $\phi(\vec{r})$ is the electrostatic potential, $\varepsilon(\vec{r})$ is the dielectric constant, N_D^+ and N_A^+ are the concentration of ionized donors and acceptors respectively and ρ_{fix} is the fixed charge. Remaining terms are the electron and hole concentration, $n(\vec{r})$ and $p(\vec{r})$, respectively. These are evaluated by solving the Schrodinger equation using open boundary condition considering the tight binding Hamiltonian with an atomistic (p_z) basis,

The greens function can be expressed as [12]-[13]

$$G(E) = [EI - H - \Sigma_S - \Sigma_D]^{-1} \quad (2)$$

where 'E' is the energy, 'I' the identity matrix, 'H' is the Hamiltonian of the GNR written in tight binding approach and Σ_S and Σ_D are the self-energies of source and drain contacts, respectively. Here transport is assumed completely ballistic and charge is considered as point charge.

Assuming that the chemical potential of the reservoirs are aligned at equilibrium with the Fermi level of the GNR, and given that there is no confined states. The electron and hole concentrations in the channel are

$$n(\vec{r}) = 2\int_{E_i}^{+\infty} dE[|\psi_S(E,\tilde{r})|^2 f(E-E_{FS}) + |\psi_D(E,\tilde{r})|^2 f(E-E_{FD})] \quad (3)$$

and

$$p(\vec{r}) = 2\int_{-\infty}^{Ei} dE\{[|\psi_S(E,\tilde{r})|^2 [1-f(E-E_{FS})] + |\psi_D(E,\tilde{r})|^2 [1-f(E-E_{FD})]\} \quad (4)$$

Here '$\vec{r}$' is the coordinate of carbon site, 'f' is the Fermi-Dirac occupation factor, and $|\psi_S|^2$ $(or\,|\psi_D|^2)$ is the probability that states injected by the source (or drain) reach the carbon site $(\vec{r})$ and $E_{FS}(orE_{FD})$ is the Fermi level of the source (or drain)

The current is computed by

$$I = \frac{2q}{h}\int_{-\infty}^{+\infty} dET(E)[f(E-E_{FS}) - f(E-E_{FD})] \quad (5)$$

Where 'q' is the electron charge, 'h' is the planks constant and 'T(E)' is the transmission coefficient that is given as-

$$T = -T_r[(\sum_S - \sum_s^+)G(\sum_D - \sum_D^+)G^+] \quad (6)$$

Here T_r is the trace operator. The simulation is carried out at temperature 300K

III. RESULT AND DISCUSSION

"Fig.2" shows the transfer characteristics of the proposed GNR MOSFET for the SiO_2 gate insulator when its thickness is varied from 2 nm to 4.5 nm with the step of 0.5 nm. The results are evaluated varying the gate bias V_{gs} from 0 to 0.5 V while keeping the drain bias V_{ds} fixed at 0.3 V. The curves is plotted in logarithmic and as well as linear scale to understand the difference between I_{ON} and I_{OFF} clearly. It is observed that when the thickness of the gate oxide is 2nm, the current at drain terminal is found to be highest. And with the increase in oxide thickness, the drain current gradually decreases almost linearly. Each transfer curve has two state: ON state and OFF state. The ratio of currents between ON and OFF state is termed as on/off current (I_{ON}/I_{OFF}) ratio, which has a falling characteristics with the oxide thickness, shown in "Fig.3(a)". These characteristics can easily be explained from their band diagram shown in "Fig.4(a)" and "Fig.4(b)". From respective OFF state band diagram shown in "Fig.4(a)", it is seen that the oxide with thickness of 2 nm offer higher energy barrier for electron than 4.5 nm. As a result the channel become more resistive for electron incase of 2 nm oxide thickness than the 4.5 nm thickness, that leads to have less OFF state currents for 2 nm thickness than 4 nm thickness. On the other hand, from the ON state band diagram shown in "Fig.4(b)", we observe that the 2 nm oxide thickness offer lower energy barrier for electron conduction than 4.5 nm. So higher ON current is obtained for 2 nm. This analysis is also true for other oxide thickness. The similar tendency in energy band diagram is also found for other oxide materials having different dielectric strength. This implies that whatever may be the oxide material, smaller oxide thickness provides higher ON current and lower OFF current. As the oxide thickness is increased, ON current starts to decrease while the OFF current starts to increase. This can also explain the falling trends of I_{ON}/I_{OFF} ratio with increase in oxide thickness. Here the band diagram strongly supports these behaviors of transfer characteristics and I_{ON}/I_{OFF} ratio. The results presented here demonstrate that increasing the oxide thickness results in poorer control of the gate over the channel

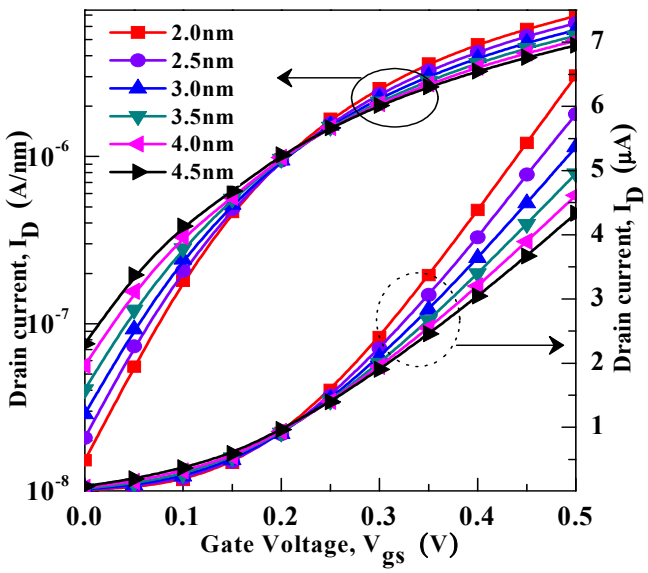


Figure 2. I_D-V_{gs} characteristics of the proposed GNR MOSFET as a function different oxide thicknesses. SiO_2 is used as gate oxide in this case.

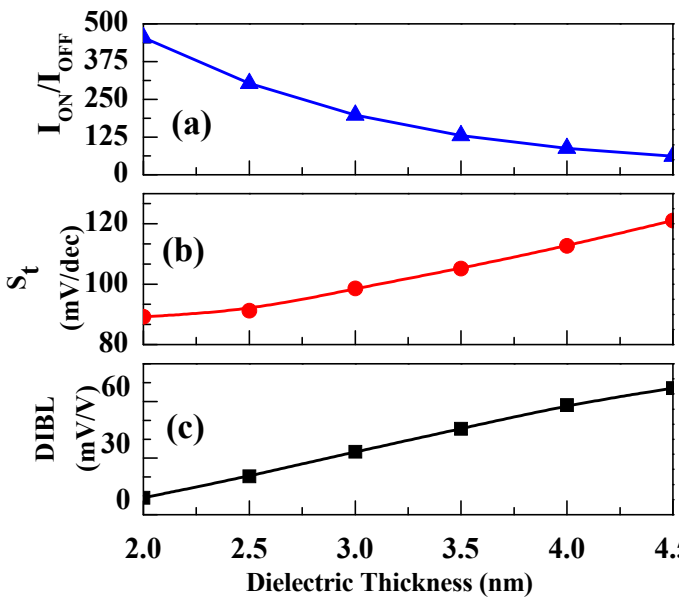


Figure 3. Oxide thickness dependent (a)I_{ON}/I_{OFF} ratios, (b) Subthreshold Slop and (c) DIBL of GNR MOSFET. SiO_2 is used as gate oxide in this case.

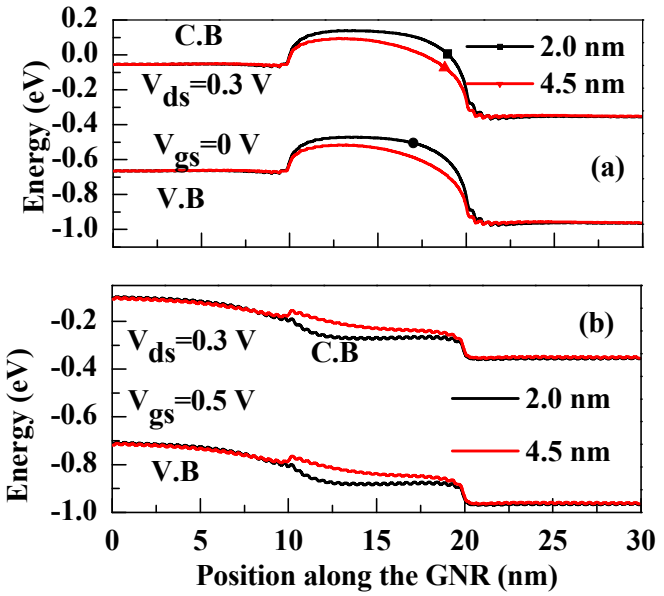


Figure 4. (a) OFF state and (b) ON state band diagramm for oxide thickness of 2.0 nm and 4.5 nm.

From the the transfer characteristics we also found that the sub-threshold slop of the GNR MOSFET follows a proportional relation with oxide thickness as shown in "Fig.3(b)". This means, as the oxide thickness is increased, more voltage will be required to obtain a certain amount of ON current. Here minimum subthreshold slope is found 72 mV/dec for 2 nm oxide thickness, and it can be reduced further by reducing the oxide thickness.

Another short channel effect known as drain induced barrier lowering (DIBL) is also studied here. And it is seen that the value of DIBL increases gradually with increase in oxide thickness shown in "Fig.3(c)". This can be explained by the results shown in "Fig.4(a)", where it is seen pictorially that for a particular drain voltage, the OFF state current faces stronger barrier potential for lower oxide thickness than the higher oxide thickness. This is due to the gradually losing the control of gate over the channel with increase in oxide thickness.

To understand the oxide strength and thickness dependent performance of the device a similar study is also carried out for different oxide materials. The dielectric materials used in our study are listed here along with their dielectric constants, k : SiO_2 (k=3.9), Si_3N_4 (k=7), Al_2O_3 (k=9), $HfSiO_4$ (k=21), HfO_2 (k=25), La_2O_3 (k=30).

Dielectric strength dependent transfer characteristics of GNR MOSFET are shown in "Fig.5" where thickness of the oxide is kept fixed at 2 nm. Here it is seen that the ON current is lowest for smaller dielectric strength and gradually increased with increase in dielectric strength although the variation in OFF current is negligible. As a result the I_{ON}/I_{OFF} ratio increases gradually with higher k shown in "Fig.6(a)". These results can be explained from the respective band diagram shown in "Fig.7(a)" and "Fig.7(b)". From "Fig.7(a)" it is seen that the magnitude of barrier height for both low-K and high-K gate dielectric are almost equal. As result the variation in OFF state current is almost negligible. While in ON state "Fig.7(b)" shows that high-K gate oxide paves the lower energy path for electron. As a result ON state current for high-K gate oxide become higher than the low-K gate oxide. These results clearly explain the rising trends of I_{ON}/I_{OFF} ratio with increase in dielectric strength. On the other hand the sub-threshold slope of the device decreases with increase in dielectric strength which is shown in "Fig.6(b)". DIBL of the GNRMOSFET follows a falling trends shown in "Fig.6(c)", which indicates the better gate control. All of these results prove that the control of the gate over the channel gets better with increasing dielectric constant, and these observation is consistent with other groups as well [14]-[15].

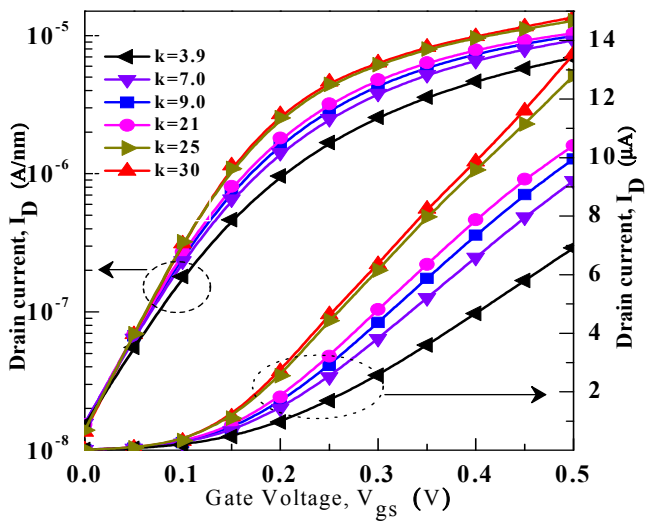


Figure 5. I_D-V_{gs} characteristics of the proposed GNR MOSFET as a function different gate oxide materials having thickness of 2 nm.

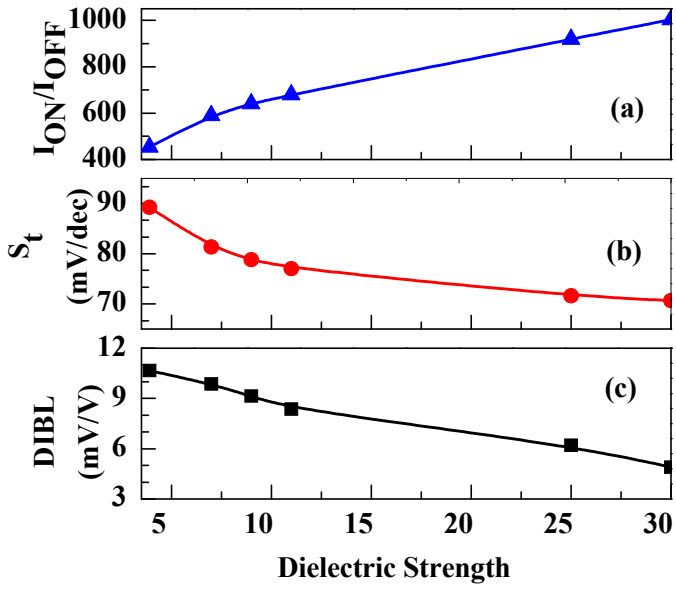


Figure 6. Dielectric strength dependent (b) on/off ratios, (b) Subthreshold Slop and (c) DIBL of GNR MOSFET for the oxide thickness 2 nm.

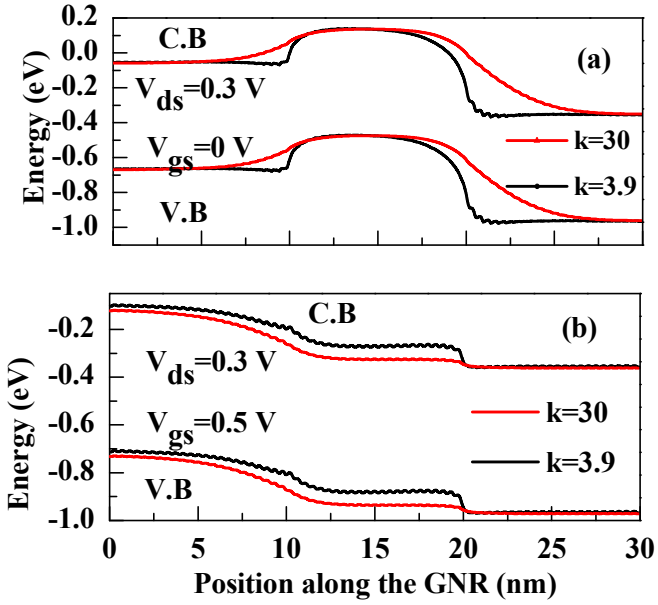


Figure 7. (a) OFF state (b) ON state band diagramm for SiO_2 (k=3.9) and La_2O_3 (k=30)

IV. CONCLUSION

Oxide thickness and its dielectric strength dependent performance is studied for a GNR MOSFET. It is seen that the performance of the GNR MOSFET deteriorates with increasing the oxide thickness. The results obtained in the present study demonstrate that the drain current and I_{ON}/I_{OFF} ratio decrease while the subthreshold slope increases. At the same time short channel effect, which is reflected by DIBL also increases with increase in oxide thickness. On the other hand, for a particular oxide thickness, the overall performance of the GNR MOSFET improves with increasing the dielectric strength of the gate oxide material.

REFERENCES

[1] F. Assad, Z. Ren, D. Vasileska, S. Datta, and M. Lundstrom, "On theperformance limits for Si MOSFETs: A theoretical study IEEE Trans.Electron Devices, vol. 47, no. 1, pp. 232–240, Jan. 2000.

[2] Md. Nur Kutubul Alam, Muhammad Shaffatul Islam, Md. Golam Kibria, and and Md. Raf qul Islam, "Anomalous Staircase CV Characteristicsof InGaSb-on-Insulator FET," IEEE TRANSACTIONS ON ELECTRON DEVICES, vol. 61, no. 11, pp. 3910-3913, November 2014.

[3] G. K. Celler and S. Cristoloveanu, "Frontiers of silicon-on-insulator,"J. Appl. Phys., vol. 93, no. 9, pp. 4955–4978, May 2003

[4] Q. Zhang et al., "Graphene Nanoribbon Tunnel Transistors",IEEE Elec. Dev. Lett., vol. 29, no. 12, pp. 1344-1346, 2008.

[5] P. Zhao et al., "Computational Study of Tunneling Transistor Based on Graphene Nanoribbon",Nano Lett., vol. 9, no. 2,pp. 684-688, 2009

[6] G. Fiori and G. Iannaccone, "Simulation of Graphene Nanoribbon Field-Effect Transistors," IEEE Electron Device Letters, vol. 28, no. 8, pp. 760-762, Aug. 2007.

[7] Y. Ouyang, P. Campbell, and J. Guo, "Analysis of ballistic monolayer and bilayer graphene field-effect transistors," Applied Physics Letters, vol. 92, no. 6, p. 063120, 2008.

[8] M.S Ballon. (2008, May 28). "Carbon nanoribbons hold out possibility of smaller, speedier computer chips". Stanford Report [Online].

[9] ITRS Roadmap, 2009, available online at http://www.itrs.net

[10] Ghader Darbandy1, François Lime, Antonio Cerdeira, Magali Estrada, Ivan Garduño and Benjamin Iñiguez"Study of potential high-k dielectric for UTB SOI MOSFETs using analytical modeling of the gate tunneling leakage," Semiconductor Science and Technology, Vol. 26, pp. 1-8, Number 2011

[11] Zhihong Chen, Yu-Ming Lin, Michael J. Rooks, Phaedon Avouris "Graphene Nano-Ribbon Electronics"Physica E: Low-dimensional Systems and Nanostructures, Vol. 40/2, pp 228-232 (2007)

[12] Y. Yoon, G. Fiori, S. Hong, J. Guo and G. Iannaccone, "Performance Comparison of Graphene Nanoribbon FETs With Schottky Contacts and Doped Reservoirs", IEEE Transaction on Electron Devices, Vol. 55, pp. 2314-2323, 2008.

[13] Gianluca Fiori, Giuseppe Iannaccone, and Gerhard Klimeck, "A three dimensional simulation study of the performance of Carbon nanotube Field Effect transistor with doped reservoir and realistic geometry," IEEE TRANSACTIONS ON ELECTRON DEVICES, vol. 53, no. 8, pp. 1782-1788, August 2006.

[14] Sheikh Ziauddin Ahmed, Mashiyat Sumaiya Shawkat, Md. Iramul Hoque Chowdhury, Sharif Mohammad Mominuzzaman, "Current-Voltage Characteristics of Ballistic SchottkyBarrier GNRFET and CNTFET: Effect of RelativeDielectric Constant," Proceedings of the 10th IEEE International Conference onNano/Micro Engineered and Molecular Systems (IEEE-NEMS 2015)Xi'an, China, April 7-11, 2015, pp. 384-387.

[15] Sheikh Ziauddin Ahmed, Mashiyat Sumaiya Shawkat, Md. Iramul Hoque Chowdhury, Sharif Mohammad Mominuzzaman, "Current-Voltage Characteristics of Ballistic Schottky Barrier GNRFET and CNTFET: Effect of Gate Oxide Thickness," Proceedings of the 10th IEEE International Conference on Nano/Micro Engineered and Molecular Systems (IEEE-NEMS 2015)Xi'an, China, April 7-11, 2015, pp. 388-390.

Nonlinear Backstepping Controller Design for a Three-Phase Grid-Connected Photovoltaic System Using DPC Approach

M. F. Pervej, T. K. Roy*, F. K. Tumpa, and M. I. Sarkar

Dept. of Electronics & Telecommunication Engineering
Rajshahi University of Engineering & Technology, Rajshahi-6204, Bangladesh
Email: roy.kanti03@gmail.com

Abstract— **A new direct power control (DPC) approach for a voltage source inverter (VSI) based three-phase grid-connected solar photovoltaic (PV) system is proposed in this paper. The proposed DPC is designed using the nonlinear backstepping control scheme based on the power dynamics rather than the conventional *dp* current dynamics. As a result, the inner current control loop is eradicated in this paper which simplifies the system modeling and controller design and enhances the transient performance. Conversely, the outer voltage loop is used to generate the active power reference from the closed loop control of the dc-link voltage dynamics. The proposed controller is designed recursively based on the Lyapunov function theory and overall stability of the whole system is confirmed with the formulation of control Lyapunov functions (CLFs) at every design stage of the controller. Finally, the performance of the designed controller is tested on a three-VSI based three-phase grid-connected PV system under different operating conditions. The performance is also compared with an existing controller and it is clear that designed controller can better performance than the existing controller in terms of the total harmonic distortion (THD) of the injected current and quality of power.**

***Index Terms*—Active power, backstepping controller, direct power control, control Lyapunov function, power quality, reactive power, total harmonic distortion**

I. INTRODUCTION

Due to the global concerns about the climate change (e,g., greenhouse gas emissions) and diminishing the fossil fuel reserves, renewable energy sources (RESs) such as solar PV, wind power, fuel cell are exponentially increasing their contributions to the utility grid. However, among them, the solar PV system is one of the most promising RESs due to its distinctive advantages as mentioned in [1]. But, the integration of solar PV system with the existing grid might be posed several challenges such as grid stability, power quality, power mismatch etc. [2]. Thus, to transfer high-quality power with low THD and effectively to maintain the dynamic stability of the system, a high-performance converter is essential which can operate more efficiently and effectively. In this sense, the use of VSI based grid-connected PV system is increased enormously due to several advantages such as high quality sinusoidal current with controllable power factor and steady output power control with bi-directional power flow [3], [4].

The most commonly used control methods for a VSI based grid-connected PV system are the voltage oriented control and DPC. However, DPC approach is a widely used control strategy over the last few years because it can regulate the instantaneous active and reactive power without extra current loop [5]. In general, look up table (LUT) based DPC is widely used due to its simplicity. In this technique, the switching signals are selected based on the instantaneous errors between the desired and actual values of the active and reactive power [6]. Nevertheless, the variable switching frequency is the main limitations of this technique. Due to the variable switching frequency, an unexpected harmonic spectrum is generated and then designs of a filter for grid-connected PV system it quite difficult. To overcome the aforementioned problem of LUT-based DPC, a predictive based DPC approach is proposed in [7], [8]. However, although it can achieve satisfactory control performance, it is very sensitive to the variation of system parameters as well as external disturbances, and consequently, makes it challenging for the future research work [9].

Due to the robustness against the system parameters variation, and external disturbance a sliding mode controller (SMC) based DPC is proposed in [4], [10], [11]. But, though the SMC can provide fast dynamic response with acceptable performance, high-frequency oscillation always exists in the control input due to the chattering phenomenon. Moreover, the overall stability of the system is confirmed to the sliding surface which is pretty difficult for wide variations of operating points. Motivated from the aforementioned literature review a nonlinear backstepping control scheme is proposed in this paper as discussed in [12]-[16] for various application.

In most cases, the grid-connected solar PV systems operate at the maximum power point to extract the maximum extracted power from the PV arrays under any atmospheric conditions. Various maximum power point technique (MPPT) algorithms have been proposed in the literature [17], [18]. The perturb and observe (P&O) is one of the most popular methods due to its simple control structure which is commonly used in solar PV systems to extract the maximum power where the operating point of MPPT moves toward the maximum power point by periodically increasing or decreasing the output voltage of PV arrays [17]. However, the P & O method has some limitations such as slower response times and oscillations in the output power of PV arrays. To overcome these problems, an incremental conductance (IC) method is used in [18] for its conciseness and higher tracking efficiency which is also used in this paper to extract maximum power from PV arrays.

The aim of this paper is to control the active and reactive power injection into the grid from a VSI based three-phase solar PV system using the DPC approach. The proposed controller is designed based on the active and reactive power dynamics rather than the conventional *dq* current components using the nonlinear backstepping control technique. The

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

proposed control scheme also ensures the improved power quality which is measured in terms of total harmonic distortion (THD). The effectiveness of the proposed scheme is analyzed on a three-phase grid-connected solar PV system and compared with an existing nonlinear control approach.

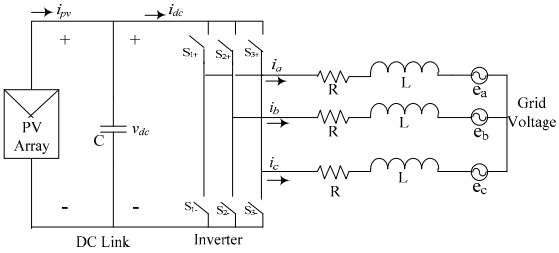


Fig.1. Three-phase grid-connected PV system

II. Dynamic Modeling of The Systems

The typical configuration of a voltage source inverter (VSI) based three-phase grid-connected PV system is shown in Fig. 1. As, this paper is focused on the VSI, the dynamics of the VSI based three-phase grid-connected PV system in the *dq*-reference frame can be written as follows:

$$\frac{dI}{dt} = \frac{1}{L}(V_i - RI - V_g) \quad (1)$$

where R and L represent the resistance and inductance of the line filter, respectively, $V_g = [V_{gd} \quad V_{gq}]^T, V_i = [V_{id} \quad V_{iq}]^T$ are the grid voltage and inverter output voltage, respectively, and $I = [I_d \quad I_q]^T$ is the grid current. In the rotating *dq*-reference frame, the active and reactive power can be described by the following equations:

$$P_g = \frac{3}{2}\left(V_{gd}I_d + V_{gq}I_q\right) \quad (2)$$

$$Q_g = \frac{3}{2}\left(V_{gq}I_d - V_{gd}I_q\right) \quad (3)$$

In this paper, the proposed model is developed based on the active and reactive dynamics rather than the conventional *dq* current components. Thus, the active and reactive power derivative with respect to time can be written as follows:

$$\frac{dP_g}{dt} = \frac{3}{2}\frac{dV_{gd}}{dt}I_d + \frac{3}{2}\frac{dI_d}{dt}V_{gd} + \frac{3}{2}\frac{dV_{gq}}{dt}I_q + \frac{3}{2}\frac{dI_q}{dt}V_{gq} \quad (4)$$

$$\frac{dQ_g}{dt} = \frac{3}{2}\frac{dV_{gq}}{dt}I_d + \frac{3}{2}\frac{dI_d}{dt}V_{gq} - \frac{3}{2}\frac{dV_{gd}}{dt}I_q - \frac{3}{2}\frac{dI_q}{dt}V_{gd} \quad (5)$$

For a balanced system, the grid voltage can be described by the following equations:

$$V_{gd} = V\sin\omega t \quad (6)$$

$$V_{gq} = -V\cos\omega t \quad (7)$$

As the line voltage variation is also required in equations (4) and (5), the grid voltage derivative with respect to time can be written as follows:

$$\frac{dV_{gd}}{dt} = -\omega V_{gq} \quad (8)$$

$$\frac{dV_{gq}}{dt} = \omega V_{gd} \quad (9)$$

Now by combining equations (1), (4), (5), (8), and (9) yields

$$\frac{dP_g}{dt} = -\frac{R}{L}P_g - \omega Q_g + \frac{3}{2L}[V_{gd}V_{dc}S_d + V_{gd}V_{gq} - (V_{gd}^2 + V_{gq}^2)] \quad (10)$$

$$\frac{dQ_g}{dt} = -\frac{R}{L}Q_g + \omega P_g + \frac{3}{2L}[V_{gd}V_{gq} - V_{gd}V_{dc}S_q] \quad (11)$$

where $V_{id} = V_{dc}S_d$, and $V_{iq} = V_{dc}S_q$. On the other hand, at the PV side, the dynamic of dc-link voltage can be obtained by applying the KCL at the node where the capacitor is connected and this can be written as:

$$\frac{dV_{dc}}{dt} = \frac{1}{CV_{dc}}(P_{pv} - P_{dc}) \quad (12)$$

where P_{dc} is the dc-side power. During a steady-state operation of the system, power exchange at the dc side will be equal to the power at the *ac* side if it is assuming that the losses due to switching actions of the inverter are negligible. Hence, the following relation can be obtained based on the active power balance between the *dc* side and *ac* side of the converter.

$$P_{dc} = P_g \quad (13)$$

Based on equation (13), equation (12) can be written as:

$$\frac{dV_{dc}}{dt} = \frac{1}{CV_{dc}}(P_{pv} - P_g) \quad (14)$$

Now from equations (10), (11), and (14), new dynamic model of the three-phase grid-connected PV system in the *dq*-reference frame can be written as:

$$\frac{dV_{dc}}{dt} = \frac{1}{CV_{dc}}\left(P_{pv} - P_g\right) \quad (15)$$

$$\frac{dP_g}{dt} = -\frac{R}{L}P_g - \omega Q_g + \frac{3}{2L}[V_{gd}V_{dc}S_d + V_{gd}V_{gq} - (V_{gd}^2 + V_{gq}^2)] \quad (16)$$

$$\frac{dQ_g}{dt} = -\frac{R}{L}Q_g + \omega P_g + \frac{3}{2L}[V_{gd}V_{gq} - V_{gd}V_{dc}S_q] \quad (17)$$

Since the *d*-axis of the synchronously rotating *dq* frame is aligned with the grid voltage vector, *d*- and *q*-axis components of the grid voltage can be obtained as follows:

$$V_{gd} = E_d \text{ and } V_{gq} = 0$$

As a consequence, equations (15)-(17) can be further simplified as follows:

$$\dot{x}_1 = \frac{1}{Cx_1}(P_{pv} - x_2) \quad (18)$$

$$\dot{x}_2 = -\frac{R}{L}x_2 - \omega x_3 + \frac{3}{2L}(E_d x_1 S_d - E_d^2) \quad (19)$$

$$\dot{x}_3 = -\frac{R}{L}x_3 + \omega x_2 - \frac{3}{2L}E_d x_1 S_q \quad (20)$$

where $x_1 = V_{dc}$, $x_2 = P_g$, and $x_3 = Q_g$ are the new state variables. Equations (18)-(20) represent the complete dynamical model of a three-phase grid-connected solar PV system and the proposed controller is designed based on this model in the following section.

III. Proposed Controller Design

The main purpose of this section is to find out the control laws for controlling active and reactive power injection into the grid. The designed procedure of the proposed controller is elaborately discussed in the following.

Step 1: According to the design purpose, the dc-link bus voltage tracking error can be defined as:

$$e_1 = x_1 - x_{1d} \quad (21)$$

where x_{1d} is the dc-link reference voltage. By considering equation (18), the dynamics of e_1 can be written as:

$$\dot{e}_1 = \frac{1}{Cx_1}(P_{pv} - x_2) \quad (22)$$

where x_2 is a stabilizing control input corresponding to the dynamics of dc-link voltage tracking error. Now, the active power tracking error can be defined by the following error equation:

$$e_2 = x_2 - \alpha \quad (23)$$

In term of equation (23), equation (22) can be written as:

$$\dot{e}_1 = \frac{1}{Cx_1}(P_{pv} - e_2 - \alpha) \quad (24)$$

Now, in order to achieve $x_1 \to x_{1d}$ (zero tracking error), the CLF can be chosen as follows:

$$W_1 = \frac{1}{2}e_1^2 \quad (25)$$

Taking the derivative of W_1 with respect to the time, it can be written as:

$$\dot{W}_1 = e_1\dot{e}_1 \quad (26)$$

Inserting the value of $\dot{e}_1$ from equation (24) into equation (26), it can be written as:

$$\dot{W}_1 = e_1\frac{1}{Cx_1}(P_{pv} - e_2 - \alpha) \quad (27)$$

At this point, the stabilizing function α needs to be selected in such a way that $\dot{W}_1$ would be negative semi-definite, i.e., $\dot{W}_1 \leq 0$. Now, to achieve fast error convergence and better tracking precision, the stabilizing function α can be chosen as follows:

$$\alpha = Cx_1(P_{pv} + k_1 e_1) \quad (28)$$

where $k_1 > 0$ is a positive constant which is used to tune the output responses. After selecting this, equation (27) can be written as:

$$\dot{W}_1 = -k_1 e_1^2 - \frac{e_1 e_2}{Cx_1} \quad (29)$$

From equation (29), it is clear that the first term on the right side is negative definite. Finally, the second term on the right side of equation (29) will be zero if $e_2$is zero. The derivation of final control law along with the stability and robustness analysis of the whole system is shown in the following step.

Step 2: Since α is not the actual control input, one can seek only the convergence of the error between α and x_2 to zero. To this end, the derivative of e_2 along with equation (19) can be written as:

$$\dot{e}_2 = -\frac{R}{L}x_2 - \omega x_3 + \frac{3}{2L}(E_d x_1 S_d - E_d^2) - \dot{\alpha} \quad (30)$$

Now, lets define the reactive power tracking error as:

$$e_3 = x_3 - x_{3d} \quad (31)$$

whose dynamics can be written as:

$$\dot{e}_3 = -\frac{R}{L}x_3 + \omega x_2 - \frac{3}{2L}E_d x_1 S_q - \dot{x}_{3d} \quad (32)$$

In backstepping control scheme, a Lyapunov function is usually used for driving the conditions on the control law that will drive the state trajectory to the equilibrium point. In this case, the following CLF is considered:

$$W_2 = W_1 + \frac{1}{2}e_2^2 + \frac{1}{2}e_3^2 \quad (33)$$

After differentiating W_2 with respect to time, it can be written as:

$$\dot{W}_2 = \dot{W}_1 + e_2\dot{e}_2 + e_3\dot{e}_3 \quad (34)$$

Substituting the values of $\dot{W}_1$ from equation (29), $\dot{e}_2$ from equation (30), and $\dot{e}_3$ from equation (32) into equation (34) yields

$$\dot{W}_2 = -k_1 e_1^2 + e_2\{-\frac{e_1}{Cx_1} - \frac{R}{L}x_2 - \omega x_3 + \frac{3}{2L}(E_d x_1 S_d - E_d^2) - \dot{\alpha}\} + e_3(-\frac{R}{L}x_3 + \omega x_2 - \frac{3}{2L}E_d x_1 S_q - \dot{x}_{3d}) \quad (35)$$

To ensure the stability of the whole system, the derivative of W_2 should be negative semi-definite, i.e., $\dot{W}_2 \leq 0$. At this point, the actual control inputs S_d and S_q can be selected as follows:

$$S_d = \frac{2L}{3E_d x_1}(\omega x_3 + \frac{e_1}{Cx_1} + \frac{R}{L}x_2 + \frac{3}{2L}E_d^2 - k_2 e_2) \quad (36)$$

$$S_q = \frac{2L}{3E_d x_1}\left(\omega x_2 - \frac{R}{L}x_3 - \dot{x}_{3d} + k_3 e_3\right) \quad (37)$$

Then, equation (35) reduces to

$$\dot{W}_2 = -k_1 e_1^2 - k_2 e_2^2 - k_3 e_3^2 \leq 0 \quad (38)$$

From equation (38), it is clear that the $\dot{W}_2$ is negative semi-definite. Therefore, it can be concluded that the whole VSI based grid-connected PV system will be stable with the designed controller. The performance of the designed controller is evaluated in the following section.

IV. Controller Performance Evaluation

In this section, the performance of the designed controller is evaluated on a VSI based three-phase grid connected PV system as shown in Fig.1. To do that, an SPR 305-WHT based 100 kW PV module is considered in this paper. Table I represents the corresponding array electrical characteristics of module SPR 305-WHT. The value of the dc-link capacitor is 470 µF and reference dc-link voltage is 400V which is generated from the MPPT. The VSI is connected to the grid through an L filter whose inductance is 20 mH and 100-kVA 260V/20kV step-up three-phase transformer is used as a coupling transformer which is connected to a utility grid through a 20-kV distribution feeder along with 120-kV equivalent transmission system. To tune the output responses of the system, the parameters of the designed backstepping controller are chosen as follows: $k_1 = 5$, $k_2 = 8$, and $k_3 = 10$. The performance of the designed nonlinear backstepping controller is also compared to that of an existing sliding mode controller as proposed in [11]. However, due to the space constraint, the performance of the controller is evaluated by considering only the following condition:

- Controller performance at 0.8 power factor under changing atmospheric conditions.

TABLE 1 Electrical Characteristics for one module of SPR 305-WHT

Parameter	Symbol	Value
Maximum power	P_m	100 kW
Open circuit voltage	V_{oc}	64.2 V
Short circuit current	I_{sc}	5.96 A
Maximum power voltage	V_{max}	54.7 V
Maximum power current	I_{max}	5.58 A
Number of series connected module per string	N_{ser}	5
Number of parallel strings	N_{par}	66

In a practical implementation, the atmospheric conditions (solar irradiation and temperature) are not constant and consequently the output power of the solar PV array is not constant. Under this condition, to show the effectiveness of the designed controller, the simulation is carried out with the change of solar irradiation.

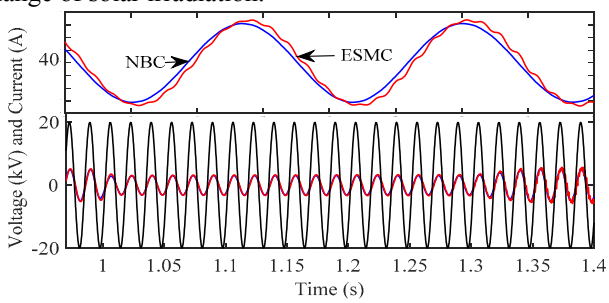
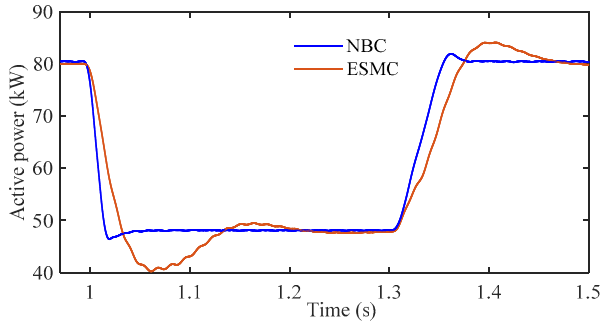


Fig.2. Grid voltage and current of phase A at power factor 0.8 with the change of solar irradiation

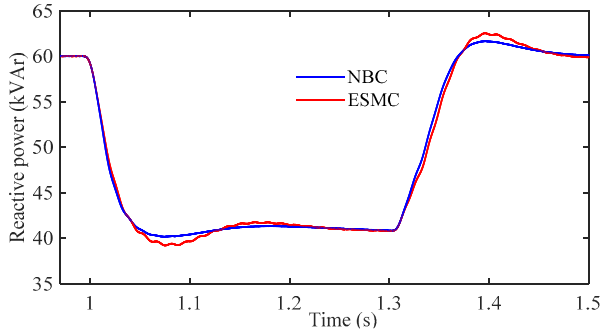
In this case, it is assumed that initially the solar PV array operating at standard condition i.e., solar irradiation of 1000 W/m^2 and temperature of 25^{0}C up to 1 s and then at t=1 s, the solar irradiance level step downs from 1000 W/m^2 to 500 W/m^2, maintains this level up to 1.3 s, and again steps up from 500 W/m^2 to 1000 W/m^2 at t=1.31 s. The corresponding grid voltage and current responses are shown in Fig. 2. From Fig. 2, it can be seen that the designed controller can inject a high-quality current into the grid as compared to the existing controller. The injected active and reactive power into the grid is shown in Fig. 3. From Fig. 3, it can be observed that their amplitudes change whenever the solar radiation is varied. Yet, active and reactive power responses are more stable with the designed controller as these responses quickly settle down to their steady-state values as compared to the existing controller.

When the system works under solar irradiation variations, the quality of the injected power will be deteriorated which can be described with the THD as shown in Fig. 4. From Fig. 4, it can be seen that THD of the injected grid current with the designed controller is 3.20% whereas; it is 3.60% with the

existing controller. From the above case studies, it is obvious that the designed controller can enhance the dynamic stability and power quality of a three-phase grid-connected PV system by sharing both active and reactive power into the grid in a better way than the existing controller under various atmospheric and operating conditions.

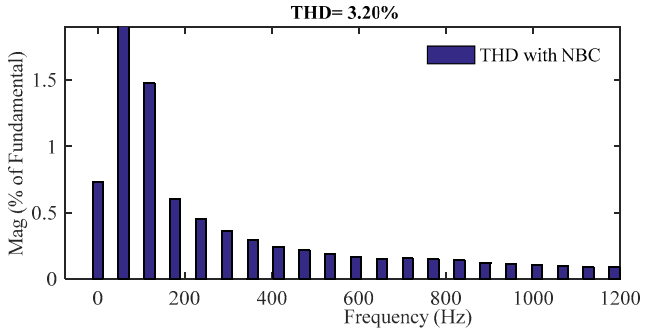


(a) Active power

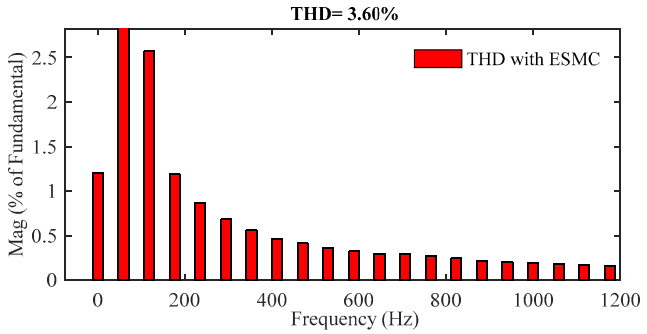


(b) Reactive power

Fig. 3. Active and reactive power injected into the grid at power factor 0.8 with the change of solar irradiation



(a) THD with designed controller



(b) THD with existing controller

Fig. 4. THD in grid current at 0.8 power factor with the change of solar irradiation

V. Conclusion

In this paper, a new recursive method to control the active and reactive power injection into the grid from a three-phase grid-connected PV system using the DPC approach is proposed. The proposed controller is designed considering all parameters of the system as unknown and these unknown parameters are estimated through the adaptation laws. From the simulation results, it is clear that the designed nonlinear backstepping controller provides a very satisfactory performance in terms of maintaining steady-state operation under various operating conditions as compared to the existing controller. The simulation results also indicate that the designed controller improves the power quality of the system by reducing the THD of the injected current into the grid. Future work will be devoted to the design of an adaptive backstepping controller considering the parametric uncertainties of the system.

References

[1] A. E. Fadili, F. Giri, and A. E. Magri, "Reference voltage optimizer for maximum power point tracking in triphase grid-connected photovoltaic systems," *Int. Journal of Electrical Power & Energy Systems,* vol. 60, pp. 293-301, 2014.

[2] B. Yang, W. Li, Y. Zhao and X. He, "Design and analysis of a grid-connected photovoltaic power system," *IEEE Trans. on Power Electronics*, vol. 25, no. 4, pp. 992-1000, April 2010.

[3] P. Cheng, and H. Nian, "Direct power control of voltage source inverter in a virtual synchronous reference frame during frequency variation and network unbalance," *IET Power Elect.,* vol. 9, no. 3, pp.502-511, 2016.

[4] J. Hu, L. Shang, Y. He, and Z. Zhu, "Direct active and reactive power regulation of grid-connected dc/ac converters using sliding mode control approach," *IEEE Trans. on Power Electronics*, vol. 26, no. 1, 2011.

[5] G. H. Lee. Y. Gui, and C. C. Chung, "Direct power control for three-phase grid-connected inverter via port-controlled Hamiltonian method," *41st Annual Conf. on Ind. Electronics Society,* pp. 2312-2317, 2015.

[6] T. Noguchi, H. Tomiki, S. Kondo, and I. Takahashi, "Direct power control of PWM converter without power-source voltage sensors," *IEEE Trans. on Industry Applications*, vol.34, no.3, pp.473-479, 1998.

[7] S. Larrinaga, M. Vidal, E. Oyarbide, and J. Apraiz, "Predictive control strategy for dc/ac converters based on direct power control," *IEEE Trans. on Ind. Electronics,* vol. 54, no. 3, pp. 1261-1271, 2007.

[8] S. Aurtenechea, M. A. Rodriguez, E. Oyarbide, and J. R. Torrealday, "Predictive control strategy for DC/AC converters based on direct power control," *IEEE Trans. on Ind. Electron.*, vol. 54, no. 3, pp. 1261–1271, 2007.

[9] P. Cortes, J. Rodriguez, P. Antoiewicz, and M. Kazmierkowski, "Direct power control of an AFE using predictive control," *IEEE Trans. on Power Electronics*, vol. 23, no. 5, pp. 2516-2523, September 2008.

[10] S. Huseinbegovic, and B. P.-Drazenovic, "A sliding mode based direct power control of three-phase grid-connected multilevel inverter," *13th Int. Conf. on OEEE,* pp. 790-797, 2012.

[11] J. Hu, and B. Hu, "Direct active and reactive power regulation of grid connected voltage source converters using sliding mode control approach," *IEEE Int. Symposium on Ind. Elect.*, pp. 3877-3882, 2010.

[12] T. K. Roy, "Position control of a small helicopter using robust backstepping," in *Proc. of the 7th ICECE*, pp. 787-790, 2012.

[13] T. K. Roy, and M. Garrat, "Altitude control of an unmanned autonomous helicopter via robust backstepping controller under horizontal wind gusts," in *Proc. of 7th ICECE*, pp. 771-774, 2012.

[14] T. K. Roy, M. Garrat, H. R. Pota, and H. Teimoori, "Hover flight control of a small helicopter using robust backstepping and PID," in *Proc. of the 10th WCICA*, 6-8 July 2012.

[15] T. K. Roy, M. Garrat, H. R. Pota, and H. Teimoori, "Robust backstepping control for longitudinal and lateral dynamics of small scale helicopter," *Journal of University of Science and Technology of China*, vol. 42, no. 7, July 2012.

[16] T. K. Roy, "Robust controller for vertical-longitudinal-lateral dynamics control of small helicopter," *Universal Journal of Control and Automation*, vol.2, no.1, pp. 14-24, 2014.

[17] J. Ahmed and Z. Salam, "An improved perturb and observe (P&O) maximum power point tracking (MPPT) algorithm for higher efficiency," *Applied Energy*, vol. 150, pp. 97–108, 2015.

[18] B. Liu, S. Duan, F. Liu, and P. Xu, "Analysis and improvement of maximum power point tracking algorithm based on incremental conductance method for photovoltaic array," in *Proc. IEEE Int. Conf. Power Electron. Drive Syst.*, pp. 637–641, 2007.

Bandwidth Enhancement of a Compact Planar Antenna for Wireless Capsule Endoscopy

Farhadur Arifin,[1,*] Pran Kanai Saha,[2]
[1]Dept. of Electrical and Electronic Engineering,
American International University Bangladesh, Dhaka, Bangladesh
[2]Dept. of Electrical and Electronic Engineering,
Bangladesh University of Engineering and Technology, Dhaka, Bangladesh
* arifin@aiub.edu

***Abstract*—Antennas are important components of Wireless Capsule Endoscope (WCE) that has been explored due to its non-invasive nature compare to the traditional endoscopy. The main challenges of these antennas include data rate of the telemetry system, miniaturizing the capsule, propagating efficiency of the antenna in the vicinity of human body. In this paper, a miniature Ultra-wideband (UWB) antenna is proposed and evaluated for WCE applications. The proposed antenna consists of slotted circular radiating patch having a sawtooth partial ground plane with 50 Ohm microstrip feed line. The substrate of the proposed antenna is chosen as Roger R03010. The bandwidth enhancement is achieved by proper selection of dimensions and positions of slots on the radiating patch and introducing a sawtooth partial ground plane. The in-body performances of the antenna are investigated inside a homogeneous human muscle layer. The antenna is able to achieve an impedance bandwidth of 1.84 GHz from 4.6554 GHz to 6.4953 GHz for return loss less than -10 dB with Omni-directional pattern. Maximum power of 44.2051 mW can be set as input to the proposed antenna in order to comply with the IEEE C95.1-2005 safety standards.**

***Index Terms*—Wireless Capsule Endoscopy, UWB technology, Antenna**

I. Introduction

With the tremendous advancement of communications and VLSI technology , Wireless Capsule Endoscope (WCE) shows a great promise for inspecting the human gastrointestinal (GI) tract for diseases in a non-invasive way [1]. The WCE is an ingestible capsule that can able to capture images during its travel through the GI tract including the stomach, large bowel or colon and part of the small bowel after being swallowed. The capsule is composed of several key parts such as image sensors, lighting, control unit, wireless communication unit, power source, and mechanical actuator.

For real-time data transfer from and to the capsule, antennas with wide-bandwidth capabilities are required. Also miniaturization of the antenna is a key issue to save the precious space of capsules. In addition, the antenna is to design with omnidirectional radiation pattern for reliable transmission of data regardless of the orientation and location of the capsule or receiver [2].

Recently many research works have been proposed for the development of compact and high performance antenna for WCE in different frequency bands. The spiral and cylindrical helical types antennas proposed in [3], [4] and the conformal meandered type antennas reported in [5], [6], [7] are developed in different frequency bands such as in Medical Implant Communication Services (MICS) band, Industrial, Scientific, and Medical (ISM) band and in Wireless Medical Telemetry Service (WMTS) band. However all the antennas show low data rates for real-time image transmitting. This is due to the narrow band of operation of the antennas.

Due to FCC's authorization of unlicensed frequency band of 3.1 GHz to 10.6 GHz, UWB technology become one of the promising technologies for the short range wireless biometry. A UWB antenna in the frequency range between 3.1 GHz and 4.5 GHz is proposed for the WCE [8]. However the antenna has complex geometry and the feeding technique which might hinder the commercial development of the antenna.

In this paper, a compact planar slotted patched antenna is designed and optimized to operate inside a human body at UWB frequency band. To enhance the bandwidth of the antenna, multiple slots with different shapes such as ring, square and rectangle are etched on the patch. Furthermore, the top edge of the partial ground plane is reshaped to form a sawtooth shape. The bandwidth obtained of our proposed antenna is 1.84 GHz inside a simplified human body model. Compare to other reported WCE antennas, it exhibits higher bandwidth with a miniature size. The finite integration technique (FIT) in Computer Simulation Technology's (CST's) Microwave Studio is used in the simulation.

The paper is organized as follows. In Section II the paper presents the simplified human body model with frequency dependent characteristics. The proposed antenna is introduced in section III. The performances of the antenna are evaluated and comparison with the other reported antennas is discussed in section IV. Concluding remarks with future directions are given in section V.

II. Human body model

For the WCE antenna, it is important to measure the antenna performance in a lossy and dispersive human body environment. However analysing inside a real tissue environment such as a full human GI tract is hard and time consuming. Therefore the simpler model like a homogeneous human tissue model helps better understanding the interactions between the antenna and the human body and acquires meaningful results of the antenna parameters [9].

In this paper, a single layer muscle tissue model has been chosen because it is simple to model, and exhibits much faster computational times compared to more complex human body models. Also it provides an approximate result for any proof of concept. A numerical human body model of the CST MWS is exploited to evaluate the antenna parameters. Fig. 1 shows the frequency dependency of relative permittivity and conductivity

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

of the human muscle tissue [10]. For example, this muscle tissue has relative permittivity, $\varepsilon_r = 48.883$, conductivity, $\sigma = 4.609$ S/m, and loss tangent, $tan\delta = 0.30815$ at 5.5 GHz. The antenna performances including return loss, radiation pattern, and, SAR in the vicinity of the human body are evaluated based on this frequency dependent human body model.

In Fig. 2, a rectangular box of the human muscle tissue is considered as the human body model for the performance analysis of the antenna. The dimensions of the muscle tissue are $40mm \times 40mm \times 90mm$. The antenna is placed in the center of the muscle tissue box.

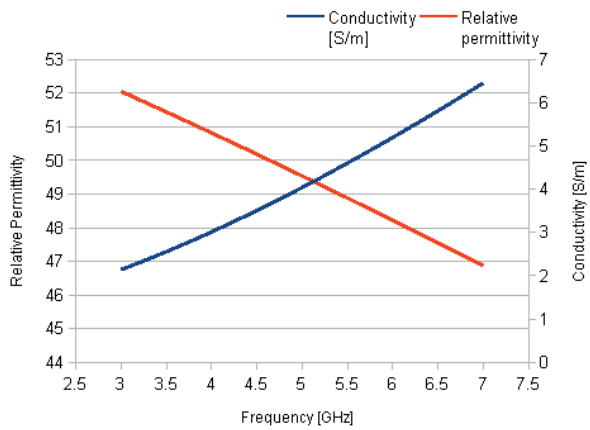


Fig. 1. Frequency response of relative permittivity and conductivity inside the human muscle tissue.

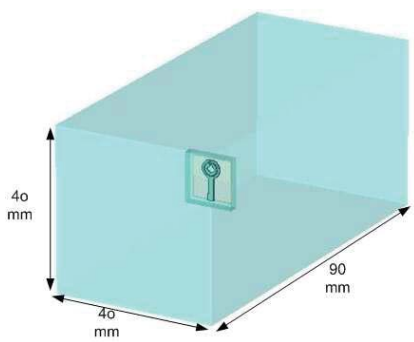


Fig. 2. A single-layer homogeneous muscle box with the antenna inserted in the box.

III. Antenna Design

The main motivation of the design of the antenna is to improve the bandwidth of the WCE antenna. The miniaturization of the antenna will be the another key issue to be addressed. The material Roger R03010 is chosen as the substrate of the antenna due to its high relative permittivity, $\epsilon_r = 10.2$ and loss tangent, $tan\delta = 0.0023$. The antenna is covered with a bio-compatible material, teflon having $tan\delta = 0.001$, and relative permittivity $\varepsilon_r = 2.25$ to preserve patient safety.

The structure of the proposed antenna is shown in Fig. 3 and in Table I, sizes of different antenna components are shown. The antenna is fabricated on a $10 \times 10\ mm^2$ substrate with thickness of 0.60 mm. The circular patch of the antenna is fed by a microstrip feed line of width 1.7 mm. The designed antenna has a circular disc of radius 2.3 mm. The radius of the patch, a , is calculated using the equation 1 [11].

$$a = \frac{1.8412v}{2\pi f_r \sqrt{\varepsilon_r}} \tag{1}$$

where v is the speed of light in the medium, f_r is the resonance frequency, and ε_r is the relative permittivity of the substrate.

In order to enhance the bandwidth as well as to optimize the design, a square shape slot is introduce at the centre and a rectangular slot which is merged with the square slot is at the top of the circular patch. A split ring slot is surrounded outside the square slot and the rectangular slot. The overall slotted area of circular patch is shown in Fig. 3 (left). Optimized values for slotted area of the patch are also provided in Table I. The study [12] shows that due to the presence of slots, the Q-factor of the patch antenna decreases as patch inductance increases. As the Q-factor is inversely proportional to the bandwidth, it results the bandwidth enhancement of the antenna. At the same time, these slots enable the antenna to operate in lower frequency range for a given size due to the fact that the resonance frequency is inversely proportional to the inductance of the patch.

For further improvement of the bandwidth, the top edge of the partial ground plane is redesigned to a symmetrical sawtooth shape by cutting of the edge of plane as shown in Fig. 3 (right) . The study shows that this technique change the distance between the ground plane and lower part of planar antenna and tune the capacitive coupling between them. It causes the bandwidth enhancement [13].

Finally, the volume of the optimized antenna is set to only 110 mm^3($10mm \times 10mm \times 1.1mm$). The dimensions are within the limit of the capsule size. All the parameters except the radius of the patch of the optimized antenna are obtained using CST Microwave Studio.

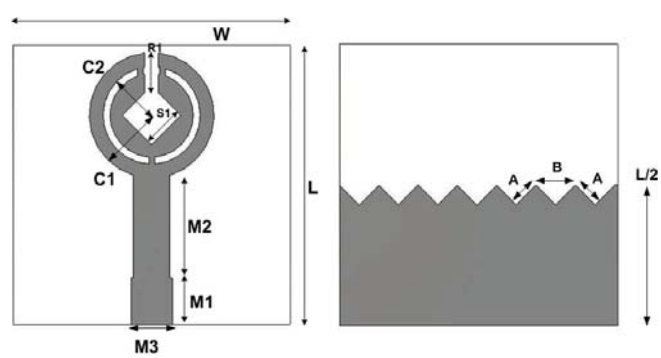


Fig. 3. Top view (left) and bottom view (right) of the proposed antenna.

IV. Performance of the Proposed UWB antenna

A. S-parameter analysis of the antenna

The frequency response of reflection co-efficient($|S_{11}|$) of the proposed antenna inside human body shown in Fig. 4. It indicates that the antenna resonates at 5.3841 GHz with a reflection coefficient magnitude ($|S_{11}|$) of -20.4 dB and a wide -10 dB-bandwidth of 1.84 GHz ranging from 4.65 GHz to 6.49 GHz. The antenna has a fractional bandwidth of 34.45%. Therefore the antenna can be considered as an UWB antenna.

TABLE I
GEOMETRICAL DIMENSIONS OF THE OPTIMIZED ANTENNA

Parameters	Value (mm)
Substrate	Length, L = 10; Width, W = 10; Thickness = 0.6.
Ground	Length, L = 5; Width , W/2= 10; Sawtooth, arm A = 1; Sawtooth, Base, B = 1.414; Thickness = 0.25.
Circular Patch	Circle of radius , C2 = 2.3; Thickness = 0.25.
Feedline	Length, M1 = 1.7; Length, M2 = 3.5; Width, M3 = 1.7; Thickness = 0.25.
Square slot	Each side of the square, S1 = 1.0.
Rectangular slot	Length, R1 = 15; Width, R2 = 5.
Slotted rings	Inner radius, C2 = 1.7; Outer radius = 1.9; Distance between a ring and the circle = 0.4.

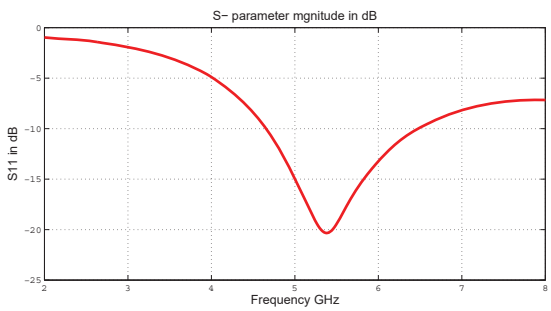


Fig. 4. Reflection co-efficient, S11, frequency response of the proposed optimized antenna inside the human muscle.

B. Radiation pattern analysis

The three dimensional far field radiation pattern are shown in Fig. 5 for the resonant frequency of 5.3841 GHz. The directivity of the antenna is found 3.11 dBi. The radiation efficiency and the total efficiency of the antenna in free space are shown -0.2221 dB (95.01 %) and -0.2802 dB (93.75 %) respectively. However, due to lossy muscle tissues surrounded by the antenna, the radiation efficiency and the total efficiency of the antenna reduce to -41.13 dB and -41.17 dB respectively. The antenna shows omnidirectional radiation patterns in the human muscle model both in E-plane and H-plane as shown in Fig. 6. This is a prime requirement for ingestible antennas as the direction of the transmitted signal should not depend on the orientation and location of the transmitter.

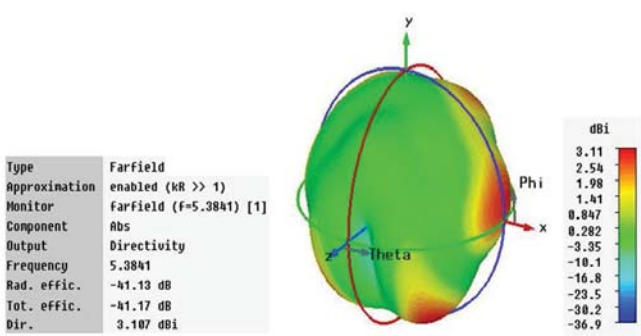


Fig. 5. Three dimensional far-field radiation pattern inside the human muscle at 5.3841 GHz.

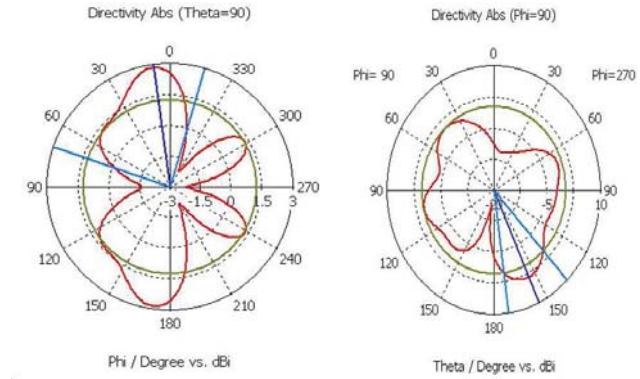


Fig. 6. Far-field radiation pattern in H-plane (left) and E-plane (right) inside the human muscle at 5.3841 GHz.

C. Electromagnetic Characteristics

The magnetic field distribution and the electric field distribution are shown in Fig. 7. It is observed that the maximum electric field equals to 6993 V/m and the maximum magnetic field equals to 33.9 A/m.

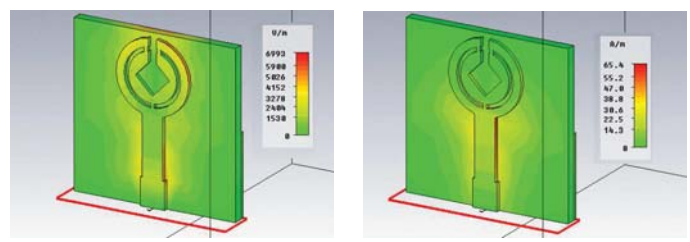


Fig. 7. Simulated electric field distribution(left) and magnetic field distribution(right) at frequency of 5.3841 GHz inside the human muscle.

D. SAR analysis

Specific Absorption Ratio (SAR) is a parameter for electromagnetic energy absorption by biological tissue mass when exposed to radiating elements. IEEE C95.1-1999 standard limits 1g-avg SAR to 1.6 W/Kg and the IEEE C95.1-2005 standard limits 10g-avg SAR to 2 W/Kg. The Fig. 8 shows the maximum SAR distribution of the proposed antenna embedded into the human muscle at 5.3841 GHz. For the antenna, maximum SAR is found to 45.2436 W/Kg averaged over 10g of tissue for input power of 1 W. To comply with IEEE safety standards, the maximum allowed net input power for the proposed antenna is limited to 44.2051 mW.

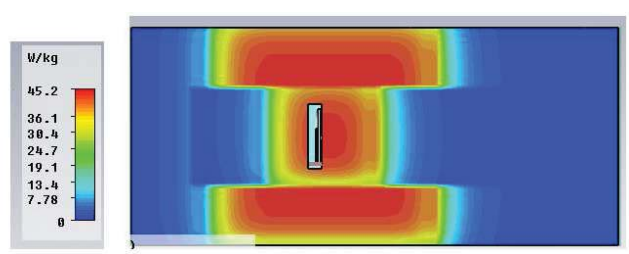


Fig. 8. 10-g average SAR distribution calculated in the simplified human body model at frequency of 5.3841 GHz (input power = 1 W).

E. Comparison of selected antennas

The performance of the proposed antenna is compared with research works mentioned in Section I. The performance parameters include type, size and the frequency of operation. In table II. , the proposed antenna shows the highest bandwidth improvement compare to all the antennas. Also the size of the antenna are smaller than all the antennas except one mentioned in [5]. However, comparing with [5], the proposed antenna is simple in design. Moreover, our proposed antenna shows a significant bandwidth enhancement.

TABLE II
COMPARISON OF SELECTED ANTENNAS

Reference	**Antenna shape**	**Bandwidth**	**Volume**(mm^3)
[3]	helical	80 MHZ (2.4GHz -2.48 GHz)	362
[4]	spiral	75 MHZ (460 MHz-535 MHz)	346
[5]	conformal	120 MHz (390 MHz-510 MHz)	57.35
[6]	conformal	420 MHZ (330 MHz-750 MHz)	436
[7]	conformal	785 MHz (215 MHz-1 GHz)	950
[8]	conformal	1.7 GHz (3.1 GHz-4.8 GHz)	134
this work	patch	1.84 GHz (4.6554 GHz to 6.4953 GHz)	110

V. CONCLUSIONS

In this paper, an UWB ingestible capsule endoscope antenna is proposed. The antenna is small in size compare to the reported works. The directivity of the antenna is 3.107 dBi which can be considered as an omnidirectional antenna. The proposed antenna has wide -10 dB reflection coefficient bandwidth of 1.84 GHz ranging from around 4.6554 GHz to 6.4953 GHz having resonate at 5.34 GHz with a reflection coefficient magnitude, ($|S_{11}|$)of -20.4 dB. A significant bandwidth improvement has been achieved by the antenna. Therefore, it is a good candidate for the antenna in the WCE . To manufacture a prototype antenna and to test and to verify the antenna will be considered as the future work of the research.

REFERENCES

[1] M. R. Basar, F. Malek, K. M. Juni, M. S. Idris, and M. I. M. Saleh, "Ingestible wireless capsule technology: A review of development and future indication," *International Journal of Antennas and Propagation*, vol. 2012, 2012.

[2] G. Ciuti, A. Menciassi, and P. Dario, "Capsule endoscopy: from current achievements to open challenges," *Biomedical Engineering, IEEE Reviews in*, vol. 4, pp. 59–72, 2011.

[3] C. Liu, Y.-X. Guo, and S. Xiao, "Circularly polarized helical antenna for ism-band ingestible capsule endoscope systems," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 12, pp. 6027–6039, 2014.

[4] S. H. Lee, J. Lee, Y. J. Yoon, S. Park, C. Cheon, K. Kim, and S. Nam, "A wideband spiral antenna for ingestible capsule endoscope systems: experimental results in a human phantom and a pig," *Biomedical Engineering, IEEE Transactions on*, vol. 58, no. 6, pp. 1734–1741, 2011.

[5] X. Cheng, J. Wu, R. Blank, D. E. Senior, and Y.-K. Yoon, "An omnidirectional wrappable compact patch antenna for wireless endoscope applications," *IEEE Antennas and Wireless Propagation Letters*, vol. 11, pp. 1667–1670, 2012.

[6] E. G. Lim, J. C. Wang, Z. Wang, T. Tillo, and K. L. Man, "The uhf band in-body antennas for wireless capsule endoscopy," *Engineering Letters*, vol. 21, no. 2, pp. 72–80, 2013.

[7] R. Alrawashdeh, Y. Huang, P. Cao, and E. Lim, "A new small conformal antenna for capsule endoscopy," in *Antennas and Propagation (EuCAP), 2013 7th European Conference on*, pp. 220–223, IEEE, 2013.

[8] Q. Wang, K. Wolf, and D. Plettemeier, "An uwb capsule endoscope antenna design for biomedical communications," in *Applied Sciences in Biomedical and Communication Technologies (ISABEL), 2010 3rd International Symposium on*, pp. 1–6, IEEE, 2010.

[9] A. Kiourti and K. S. Nikita, "Miniature scalp-implantable antennas for telemetry in the mics and ism bands: Design, safety considerations and link budget analysis," *Antennas and Propagation, IEEE Transactions on*, vol. 60, no. 8, pp. 3568–3575, 2012.

[10] S. Gabriel, R. Lau, and C. Gabriel, "The dielectric properties of biological tissues: Ii. measurements in the frequency range 10 hz to 20 ghz," *Physics in medicine and biology*, vol. 41, no. 11, p. 2251, 1996.

[11] C. A. Balanis, *Antenna theory: analysis and design*, vol. 1. John Wiley & Sons, 2005.

[12] A. Munir, G. Petrus, and H. Nusantara, "Multiple slots technique for bandwidth enhancement of microstrip rectangular patch antenna," in *QiR (Quality in Research), 2013 International Conference on*, pp. 150–154, IEEE, 2013.

[13] M. T. Islam, R. Azim, and A. T. Mobashsher, "Triple band-notched planar uwb antenna using parasitic strips," *Progress In Electromagnetics Research*, vol. 129, pp. 161–179, 2012.

Human and Carried Baggage Detection & Classification based on RSD-HOG in Video Frame

Tahmina Khanam,[1] Kaushik Deb[2,*]

[1]Institute of ICT, CUET
Chittagong, Bangladesh
[2]Department of CSE, CUET
Chittagong, Bangladesh
[*]debkaushik99@gmail.com

Abstract: **In the twenty-first century, crimes are destructively increased in public spaces. Besides, oblivious human movement sometimes dug a death trap for them. These issues are gradually dragging the attention of the computer vision researchers' to automatic video surveillance and warning system (AVSWS). As a fresh branch of AVSWS, human and carried baggage detection and classification has a wide area of applications that are, monitoring suspicious movement of human in public places, warn human in human restricted areas like in atomic power station, detect carriage of illegal materials like weapons, gold into baggage in airport and detect baggage in baggage restricted super shop. However, in this paper an accurate detection & classification framework of human and carried baggage is proposed. Initially, background subtraction is performed to speed up the system and use HSI model to cope up with different illumination condition. Then, rotational signal descriptor (RSD-HOG) is extracted which make the detection accurate. Finally, dynamic human body parameter setting enables the system to detect & classify single or multiple baggage even if some portions of human are occluded or disappear from window. The experimental results discover the system has satisfactory accuracy and faster comparative to others.**

***Index Terms*—Carried baggage, HSI model, RSD-HOG, Dynamic human body parameter;**

I. Introduction

In automatic video surveillance and warning system (AVSWS), real-time and post event analysis is done by manipulating digital video data with standarized descriptors. In this paper, as a fresh branch of active AVSWS research, a contribution is deposited to automatic human and carried baggage detection and classification field. This field becomes challenging under uneven illumination condition, ocllusion and absense of human parts. From some era ago, this active research field is dragging the attention of the researchers and as an outcome, now this field has been reached to a successful state-of-the-art. Among those, one successful system of detecting and classifying backpack is developed in [1]. They used human part based model with HOG descriptor. It classifies baggage depending on the head's height. However, its leaving scopes are, it has the chance of misclassification of baggage when head and leg is occluded and it cannot detect multiple baggage carried by same person. In [2], Local Binary Pattern (LBP) with its variants is used in recognition and tracking of human along with their activity. A new dimension in feature descriptor is added in [3] by adding Rotational Signal Descriptor (RSD) in this field. In [4] proposed baggage detection and classification scheme by estimating moving direction of human carrying baggage and the alignment of a temporal human like template with the best matched view specific exemplars. HOG scheme in [7] gets matured when it is combined with Scale Invariant Feature Transform (SIFT) approach in [8]. However, in [9], Speeded–UP Robust Features (SURF) is found faster than SIFT in [8].

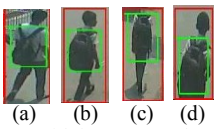


Fig. 1 Analysis of human with baggage: a) human with large height, b) human with small height, c) head occluded, and d) leg absent.

The key contribution of this research are, rotational descriptor make the detection accurate. Besides, in classification of baggage dynamic human body parameter setting approach enables the system to cope up with different sized person & deal with the problem when some portion of human disappear from window as shown in Fig. 1. In addition to this, this approach enables the system to detect and classify multiple baggage carried by same person. The remaining portion of the paper is organized as follows: the whole system structure of human with baggage detection and classification is described in section II. Then experimental results are shown in section III. Finally, concluding the system in section IV.

II. Proposed Framework

Human and carried baggage detection and classification becomes diffficult under uneven illumination condition, ocllusion and absense of hunan parts. In this paper, first taken up the help of HSI model to cope up with uneven illumination condtion. Then adopting the descriptor used in [3] in detection purpose since it is a most recent efficient descriptor as described. In classification of baggage, we gather knowledge about human body parameter from [1] and add dynamic parameter setting approach with it. The overall proposed system structure is shown in Fig. 2. The framework is divided into 5 primary steps, (a) background subtraction and filtering foreground, (b) training SVM (c) human detection with RSD-HOG feature and SVM, (d) detection of human part and set dynamic parameter and (e) detection and classification of baggage.

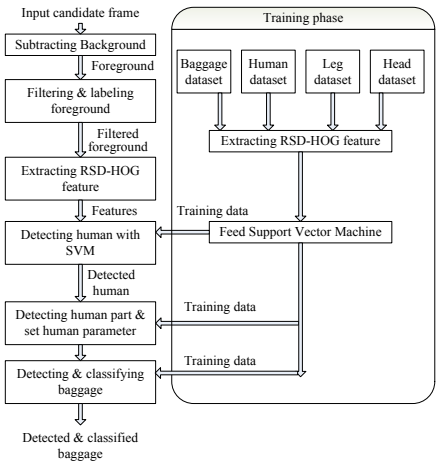


Fig. 2 Proposed human and carried baggage detection and classification framework.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

A. Background Subtraction & Filtering Foreground

In background subtraction, a challenging issue is different illumination condition of the input scene. Also, extraction of feature with sliding window introduces time complexity. To overcome these problems, in this paper a different use of HSI model is introduced. Initially, a memory of background is built up with different illumination condition. Then convert them into HSI model. Among all frames, keep only those in memory which satisfies eq. (1).

$$\mu_{R(I)} > \mu_{R(H|S)} \tag{1}$$

Where, $R(I), R(H|S) \rightarrow$ ratio between channels (I, H, S) of consecutive memory frames after applying threshold. Then for an input frame, find out the best match of it with a memory. However, best match found where memory generates the minimum value of $\mu_{R(I)}$. Here, $R(I)$ is the ratio of I between input frame and memory frame. After selecting memory, for each incoming consecutive frames subtract background using the procedure shown in Fig. 3(a). Here, threshold $T1$ & $T2$ are set using the eq. (2).

$$\left.\begin{aligned} T1 &= \mu\big(R(I|H|S)\big), where\ R(I|H|S) < 1 \\ T2 &= \mu\big(R(I|H|S)\big), where\ R(I|H|S) > 1 \end{aligned}\right\} \tag{2}$$

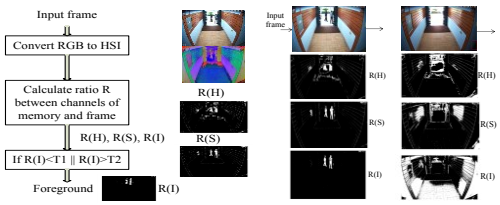
1) Sensing Illumination Change of Input Frames: If any input frame satisfies eq. (1) (where $R(I|H|S)$ is the ratio of I, H, S respectively between input frame and memory frame), then update memory frame with best matched memory for current frame. Fig. 3(b) demonstrates the scenario when input frame has no illumination change and $\mu_{R(I)} < \mu_{R(H|S)}$ where Fig. 3(c) demonstrates when illumination is changed and $\mu_{R(I)} > \mu_{R(H|S)}$. After that, with the motivation of speeding up the detection system by reducing false ROI, false foreground regions are removed which are not human like. In this purpose, one human shape is analyzed here after some casual filtering approach i.e. human head is circular and leg is rectangular. Fig. 4 describes the overall process where each labeled object is divided into upper $H/4$ portion and lower $H/4$ portion, where, H is the height of the object. Here, circularity, C of an object is calculated using eq. (3) and rectangularity, Rec is weighted using bounding box.

$$C = \frac{4A\pi}{P^2} \tag{3}$$

Where, $A \rightarrow$area of foreground object, $P \rightarrow$perimeter of the object. Fig. 5 discussing some processing examples.

2) Removal of Light Shadow: It is observed that when shadow is light, I channel change is high in shadow region and changes of H & S channels are approximate to zero. So, in this stage extracted ROI is further filtered for shadow removal with the eq. (4) where, $f_{p(x,y)}$ is the foreground pixels found after filtering process. Fig. 6 describes some processing examples of this stage.

$$f_{p(x,y)} = 0; \qquad if\ R(H|I) \approx 1 \tag{4}$$



(a) (b) (c)

Fig. 3 Proposed Background subtraction process: a) proposed module, b) input frame in same illumination with memory, and c) illumination of input frame is changed.

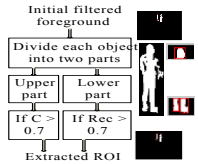
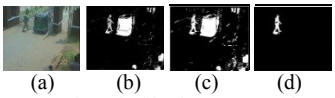


Fig. 4 Proposed foreground filtering approach



(a) (b) (c) (d)

Fig. 5 Processing example of filtering approach: a) input frame, b) background subtracted foreground image, c) after dilation operation, and d) after filtering.

(a) (b) (c)

Fig. 6 Processing example of shadow removal process: a) input frame, b) ROI with shadow, and c) ROI without shadow.

B. Training Phase

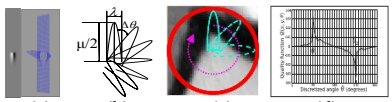
Every classifier performance depends on accurate training data. Since each object of interest varies according to its orientation and appearances, training of classifier with probable orientation and appearances of object could provide better result. In this paper, SVM classifier is used since it is an efficient one and trained with four types of interest object that are shown in Fig. 2. Here, in training period each object of interest is trained up with its 4 z-axis orientation shown in section III. This training approach provides a detection system with more accuracy.

C. Human Detection with RSD-HOG Feature and SVM

Baggage detection & classification depends on human parameter setting. So, human detection is most important part to direct the system in correct way. Since, Gaussian anisotropic filtering overcomes the blurring problems introduced in isotropic Gaussian filters by preserving the features along a particular direction, it is used here. Fig. 7(a) shows an anisotropic half Gaussian derivative filter. As shown in Fig. 7 at pixel (*x, y*), a derivative kernel is applied to obtain a derivative information $\varphi(x, y, \theta)$ in a function of orientation $\theta \epsilon [0; 360°]$. The derivative kernel is defined in eq. (5).

$$\varphi(x, y, \theta) = I * C.H(-y).x.e^{-\left(\frac{x^2}{2\lambda^2}+\frac{y^2}{2\mu^2}\right)} \tag{5}$$

Where $I \rightarrow$ image, $\theta \rightarrow$ orientation of the filter $[0; 360°]$, $C \rightarrow$ normalization coefficient, $(x, y) \rightarrow$ pixel coordinates, and (μ, λ) the standard deviation of the anisotropic Gaussian kernel called height and width of the filter. Only the casual part of this filter along the Y axis is used. This is obtained by cutting the kernel in the middle, in an operation that corresponds to the Heaviside function H. Thus rotational signal descriptor (RSD) is obtained by rotating the filter around a point in steps of 5° here. The step is selected 5° because increasing the rotation step results in loss of information. We compute the gradient $||\nabla I||$ and the two angles for the pixel at location (xp,yp) by considering the global extrema of the function $Q(xp, yp, \theta)$. Two of these global extrema are combined to maximize $||\nabla I||$ shown in eq. (6).



(a) (b) (c) (d)

Fig. 7 Description of filters: a) half anisotropic Gaussian derivative filter, b) a thin rotating Gaussian derivative half-filter, c) half Gaussian kernel applied to a keypoint (*xp, yp*), and d) extrema of a function $Q(xp, yp, \theta)$ where $\mu = 10, \lambda = 1$ and $\Delta\theta = 10°$.

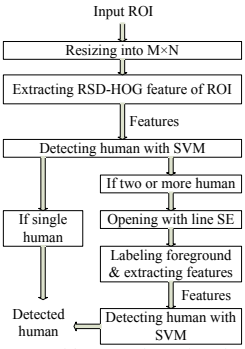


Fig. 8 Proposed human detection framework

Fig. 9 Processing examples of human detection: a) single human detected, b) multiple human detected, c) extracted multiple human region, d) result after erosion, e) result after dilation, and f) separated single human from multiple.

The resulting RSD-HOG is generated using two angles maxima θ_1, and minima, θ_2. In this purpose, we bin the two angles separately using eq. (7) and eq. (8). However, fusing these two descriptors as in eq. (9) resulting in a robust description. Fig. 8 elaborates the human detection scheme where an approach is introduced to separate slight overlapping of multiple human by opening operation and at the same time Fig. 9 shows some processing examples.

$$\begin{cases} ||\nabla I|| = \max\limits_{\theta\epsilon[0,360°]} Q(xp,yp,\theta) - \min\limits_{\theta\epsilon[0,360°]} Q(xp,yp,\theta) \\ \theta_1 = \arg\max\limits_{\theta\epsilon[0,360°]} Q(xp,yp,\theta) \\ \theta_2 = \arg\min\limits_{\theta\epsilon[0,360°]} Q(xp,yp,\theta) \end{cases} \quad (6)$$

$$RSD-HOG-\theta_1 = \{\theta_1 bin1, \theta_1 bin2, \ldots . \theta_1 bin\ n\} \quad (7)$$

$$RSD-HOG-\theta_2 = \{\theta_2 bin1, \theta_2 bin2, \ldots . \theta_2 bin\ n\} \quad (8)$$

$$RSD-HOG = \{\theta_1 bin1, . \theta_1 binN, \theta_2 bin2, . \theta_2 bin\ n\} \quad (9)$$

D. Human Part Detection & Dynamic Parameter Setting:

Some basic human parameters that used here are, bend line B which is crossing in the middle of the human horizontally and vertical line C which is crossing in the middle of the human vertically shown in Fig. 10(a). However, an effort is contributed here in setting of bend line B dynamically which can accurately set bend line B even if head or leg is absent as shown in Fig. 10(c & d). In this purpose, like Fig. 11, initially detected human region is further processed to detect head and leg region of a human, where H →height of bounding box and +ve denotes positive result of SVM and –ve is negative result. Then the bend line & vertical line of the human are set dynamically using eq. (10) Where, T is the top position of the bounding box and W is the width of the bounding box.

$$\begin{cases} B = T + \frac{7H}{10}; & if\ \ head, leg \rightarrow +ve \\ B = T + H; & if\ head \rightarrow +ve, leg \rightarrow -ve \\ B = T + \frac{H}{2}; & if\ head \rightarrow -ve, leg \rightarrow +ve \\ & C = \frac{W}{2} \end{cases} \quad (10)$$

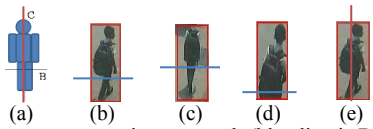


Fig. 10 Dynamic parameter setting approach (blue line is B & yellow line is C): a) human body parameter, b) set B when all parts are present, c) set B in absense of head, d) set B in absense of leg, and e) set vertical line, C.

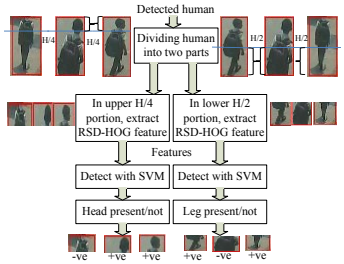


Fig. 11 Proposed human body part detection framework.

Fig. 12 Classification of baggage into two category: a) category 1: bagpack reside up on bend line (green color box), and b) category 2: handbag reside below the bend line (blue color box).

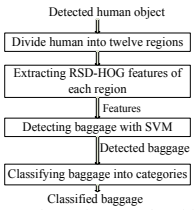


Fig. 13 Proposed baggage detection & classification framework.

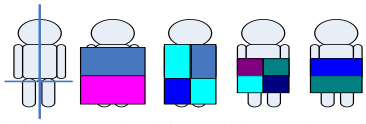


Fig. 14 Dividation approach of detected human into tweleve region.

E. Detection and Classification of Baggage:

Generally, baggage which are carried in hand usually reside below the bend line and which are carried in shoulder reside on the upper portion of bend line. However, with the consideration of these facts, in this work baggages are classified into two categories like in Fig. 12: 1) bagpack and 2) handbag. Before classification of baggage the system first divides the region based on bend line and vertical line into twelve regions. Fig. 14 demonstrates the region dividation scheme where Fig. 13 eleborates the overall process of baggage detection and classification.

III. EXPERIMENTAL RESULT

In this paper, initially each ROI is resized into 256×256. Extra enlarging of ROI could cause the loss of information, on the other side too small size arise difficulties in shape representation. After dividing into 16×16 blocks we get 256 blocks. Then each block is bin separately into 8 bins with a span of 45°. The block size is selected as 16×16, because 5° rotation of the filter around a point can sense the variation in this region only. Thus, resulting feature space size is 8×256=2048 bin. The proposed methadology has been implemented in MATLAB environment with corei5 processor and 4GB ram.The system is trained up with about 500 human, 423 category1 baggage, 221 category2 baggage, 198 head and 153 leg images. A bunch of training data is given in Fig. 15(a). The training datasets are feeded into SVM in their 4 z-axis orientation as shown in Fig. 15(b).

Fig. 15 Training data of four types object: a) training data and b) four orientation of training data.

The system module is applied on own dataset as well as on some standard datasets including 524 iLids video frames and 121 CAVIAR video frames. Table I summarizes the result of testing data under normal situation, different illumination, occlusion of human parts and the situation when multiple baggage are carried by the same person.

TABLE I
EXPERIMENTAL RESULTS

Dataset	Testing		Detected	Classified
Human	Normal	1920	1770	--
	Illumination changed	361	335	--
Category 1 baggage	Normal	861	750	706
	Illumination changed	223	195	183
	occluded	116	101	95
Category 2 baggage	Normal	992	864	814
	Illumination changed	237	207	195
	occluded	51	45	42
Category 1&2 baggage	Multiple baggage	141	123	116

The detection ability of the system under illumination change is visualized in Fig. 16(a). In addition to this, Fig. 16 (b & c) visualized the detection and classification success for category 1 & category 2 baggage under occlusion of human parts which is the shortcomings of [1].The detection and classification accuracy when multiple baggages (including category 1 and 2) are carried by same person is shown in Fig. 16(d) which is another limitation of [1]. Fig. 17 shows some random experimental results. The average detection accuracy of the proposed system for 4902 frame samples of 20 videos in different environment is found about 87.01% and classification accuracy is found about 81.06%. These results are compared with [1] in terms of detection and classification accuracy and time complexity in Table II and found the proposed system has superior performance than [1].

TABLE II
PERFORMANCE EVALUATION

	Detection accuracy	Classification accuracy	Average computational time in second (s)
[1]	77.02%	59.40%	--
Proposed method	87.01%	81.60%	0.98002

(a) (b) (c) (d)

Fig. 16 Analysis of detection & classification success under difficult scenerios: a) illumination change, b) leg absent and head occluded for category 1 baggage, c) occlusion of head and leg for category 2 baggage, and d) multiple baggage detection and classification carried by same person.

(a)

(b) (c)

Fig. 17 Experimental results of detected and classified human and baggage on: a) own dataset, b) iLids dataset, and c) CAVIAR dataset.

The system is also evaluated for False Negative (FN) rate and it is found 86.83%. The better detection accuracy is found due to rotational descriptor and SVM training approach. In addition to this, better classification accuracy is found because of dynamic human body parameter setting. The system has been proved more faster in terms of time complexity because we performed background subtraction with illumination changing challenge instead of using sliding window.

IV. CONCLUSION

In this paper, initially, background subtraction has performed to speed up the system and used HSI model to cope up with different illumination condition. Then, rotational signal descriptor (RSD-HOG) is extracted and feed to SVM and finally, dynamic human body parameter has set. The proposed system has a satisfactory classification rate with multiple baggage detection and classification carried by same person. This efficiency is found because of dynamic human body parameter setting. In detection purpose, RSD-HOG is an efficient standard which introduces better rated detection system. However, some of its' future scopes are, when mutiple human overlapped too much tightly it fails to detect baggage, addition of big noise with interest object misguide the system in wrong direction, dark shadow can not be removed which introduced big noise. Removing these limitaions can improve the system performance in terms of dectection and classification accuracy.

REFERENCES

[1] Wahyono and Jo K.H., *Carried Baggage Detection and Classification Using Part-Based Model*, ser. Lecture Notes in Artificial Intelligence, Springer International Publishing Switzerland, 2015, vol. 9227.

[2]] Nigam S., Khare A., "Analysis and Understanding of Human Behavior in Video Sequences", PhD thesis, University of Allahabad, India, 2014.

[3] Venkatrayappa D., Montesinos P., Diep D. and Magnier B., *RSD-HoG: A New Image Descriptor*, ser. Lecture Notes on Computer Science, Springer International Publishing Switzerland, 2015, vol. 9127.

[4] Tzanidou G., Edirisinghe E. A., "Automatic baggage Detection and Classification", *in Proc. IEEE International Conference on Intelligent Systems Design and Applications*, 2011, p. 825-830.

[5] Nigam S., Deb K., Khare A., "Moment Invariants based Object Recognition for Different Pose and Appearances in Real Scenes", *in Proc. IEEE International Conference on Informatics, Electronics & Vision*, 2013, p. 1-5.

[6] Zhao C., Liu C., Lai Z., "Multi-scale gist feature manifold for building recognition", *Neurocomputing*, vol. 74, pp. 2929–2940, 2011.

[7] Dalal N., Triggs B., "Histograms of Oriented Gradients for Human Detection", *IEEE Computer Society*, vol. 1, pp. 886-893, 2005.

[8] Lowe D.G., "Distinctive Image Features from Scale-Invariant Keypoints", *International Journal of Computer Vision,* vol. 60, pp. 91–110, 2004.

[9] Abeles P., *Speeding Up SURF*, ser. Lecture Notes on Computer Science, vol. 8034, Springer-Verlag Berlin Heidelberg, 2013.

FLC Based Optimization of AGC Controller Parameters for an Interconnected Power System

Asif Ferdoush,* M. G. Rabbani

Department of Electrical & Electronic Engineering, Eastern University
Dhaka 1205, Bangladesh
*asifferdoush@gmail.com

Abstract—**The aim of Automatic Generation Control (AGC) in any power system is to restore the frequency and tie line power in a smooth way to its nominal value in the shortest possible time if any load disturbance occurs at any area of the system. A two area interconnected power system is investigated in this paper under several load conditions, taken into account the various system nonlinearities to ensure the robustness of the proposed control strategy. A dc tie line is integrated in parallel with the ac tie line to stabilize the frequency of oscillation and based on it a comprehensive mathematical model is developed. Initially the gains and frequency bias parameters of the controller are tuned on the basis of performance index. Online Fuzzy Logic Controller (FLC) is then employed to generate the controller parameters of each area, in accordance to the rapid changes of small load perturbation (SLP). It may be apprehended from the simulation results that the proposed FLC controller is more efficient to minimize the effect of SLP.**

Index Terms—**Automatic Generation Control (AGC), Fuzzy Logic Controller (FLC), Performance Index, Governor Dead band (DB), GRC.**

I. Introduction

Interconnected power system is effectual under the matching of total generation with the total load demand and the associated system losses. Owing to the continuous growth of electrical power system in size and complexity with increasing interconnections, the problem of tie line power and frequency oscillations due to unpredictable load changes, has become increasingly serious [1]. Moreover, wind and solar energy plays significant role in interconnected power system and the energy delivered from them are uncertain and comparatively low. The increased number of PV penetration in the system also affects the tie line power damping oscillations, frequency changes and harmonic currents [2]. Therefore, Automatic Generation Control (AGC) is incorporated in each area to marrow the real power generation with the load demand (with losses). There are two variables of interest, namely, frequency and tie-line power exchanges. Their variations are weighted together by a linear combination to a single variable called the area control error (ACE) [3]. This ACE is minimized by the integration of proper controller parameters of AGC. Since different regional grids may belong to different utilities, they may not cooperate with each other when frequency deviates. In certain circumstances, this will probably deteriorate either performance or economy of AGC regulation. To circumvent this problem, it is important to mathematically characterize the interaction among AGC regulations in different control areas, in turn, reveal how their non-cooperative or cooperative behaviours influence the frequency regulation [4]. Many investigations have been reported over the past few decades with simplified approach excluding nonlinearities, such as governor dead band, practical limits on the rate of increase of turbine output, delays associated with stored energy of steam in the boilers are often neglected [5]. These nonlinearities have been considered in the model of this study to envisage a realistic environment. Now a day, it is obtained from literature review, that the controller parameters of AGC are tuned using different intelligent algorithm, where the optimized controller parameters remain fixed over a wide range of operating conditions. For a certain changes in load, the controller parameters such as, integral control gains and frequency bias parameters also should be changed to minimize the ACE. In this study the controller parameters are optimized on the basis of performance index for a wide range of load perturbation. Then the fuzzy logic controller generates the appropriate value of gains in response to the load changes.

II. Modelling of Proposed System

The dynamic model of AGC for a two area interconnected power system [1] is presented in this section. Each area of the power system consists of speed governing system, turbine, Boiler effects and generator. The scheme of interconnection for the case study has been shown in Fig. 1. The two areas are interconnected through an AC-DC tie line. Each generating system consists of nonlinearities, such as GRC, governor dead band and time delays. The output power of an area is given by [6],

$$P_{Gi} = P_{THi} \tag{1}$$

For small load perturbation in area i can be written as,

$$\Delta P_{Gi} = \Delta P_{THi} \tag{2}$$

Where, i is the area index, ΔP_{Gi} is the deviation in thermal turbine output and ΔP_{THi} is the Deviation in intermediate state of reheat turbine.

The tie line power flow from area 1 to area 2 for a small load perturbation is given by,

$$\Delta P_{tieAC} = \frac{2\pi T_{12}}{s}(\Delta F_1 - \Delta F_2) \tag{3}$$

Where, ΔP_{tieAC} is the tie line power deviation of AC, $2\pi T_{12}$ is the tie line power coefficient of area 1 and 2; ΔF_1 and ΔF_2 are the frequency deviation of area 1 and 2 respectively.

For small changes in DC tie line flow, ΔP_{tieDC} is given by,

$$\Delta P_{tieDC} = \frac{K_{dc}}{1 + sT_{dc}}(\Delta F_1) \tag{4}$$

Where, ΔP_{tieDC} is the tie line power deviation of DC, K_{dc} is the gain constant of rectifier and converter unit and T_{dc} is the time constant of rectifier and converter unit.

$$\therefore \Delta P_{tie12} = \Delta P_{tieAC} + \Delta P_{tieDC} \tag{5}$$

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

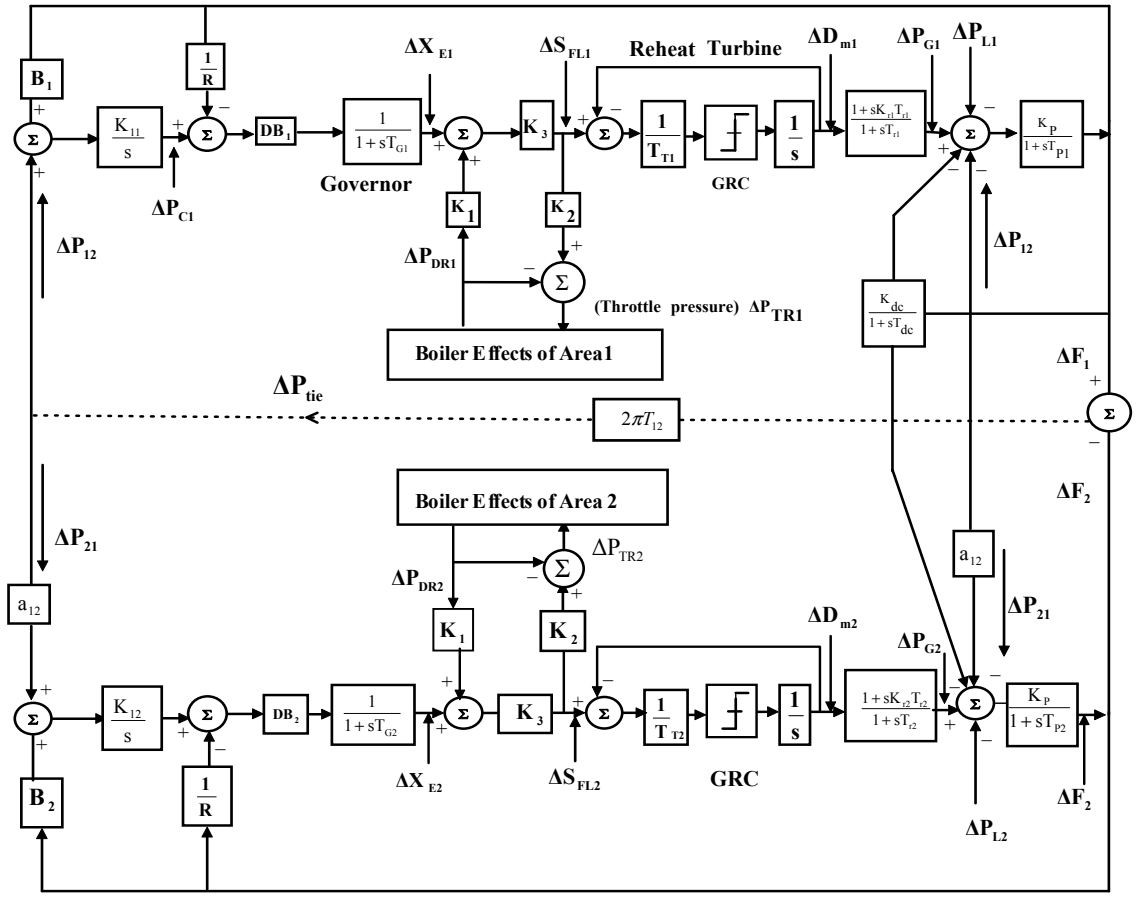


Fig. 1 Linear model of two area interconnected power system.

Where, ΔP_{tie12}=tie line power deviation between area 1 & 2.

The nonlinearity due to the governor dead band is modelled using the describing function and the approximate solution for representing backlash is given by (6) [7, 8],

$$DB_x = N_1 X + \frac{N_2}{w_0}\dot{X} \tag{6}$$

Where, DB is the dead band. For this study, N_1 is taken as 0.8 and N_2 is taken as -0.2 [7].

In this investigation, the effects of the boiler dynamics at each area of the power system is also incorporated as detailed configuration shown in Fig. 2 to generate steam under pressure. Changes in the steam flow and deviations in pressure are sensed and the corresponding action is initiated by the turbine control valves and boiler control [9].

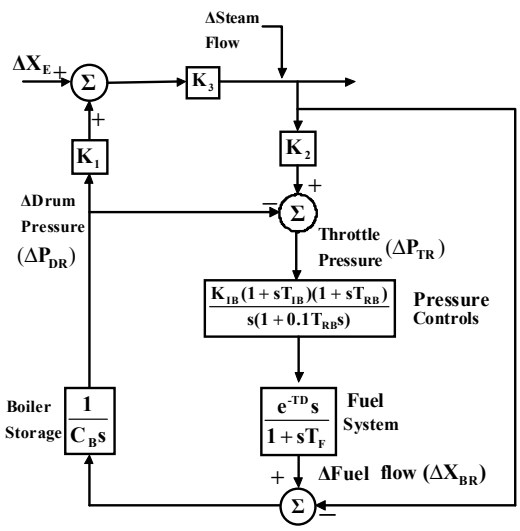


Fig. 2 The boiler system.

In practical steam turbine systems, due to thermodynamic and mechanical constraints, there is a limit to the rate at which its output power *(dP_g/dt)* can be changed. This limit is referred as generator rate constraint (GRC) [7]. In order to project physical constraints, a generation rate limitation of 0.1 p.u. per minute (i.e. 0.0017 p.u. MW per second) for thermal areas and 4.5% per second for hydro unit are deliberated [10].

III. Control Strategy and Parameter Optimization

The tie line power and frequency deviation of each area of the interconnected power system following a SLP must be zero. Each area tries to regulate its area control error (ACE) to zero and can be expressed by the following equations as in [6],

$$ACE_1 = B_1 \Delta F_1 + (\Delta P_{tieAC} + \Delta P_{tieDC}) \tag{7}$$

$$ACE_2 = B_2 \Delta F_2 + a_{12}(\Delta P_{tieAC} + \Delta P_{tieDC}) \tag{8}$$

Where area ratio a_{12} is given by,

$$a_{12} = -\frac{P_{r1}}{P_{r2}}$$

Where P_{r1} and P_{r1} are the power rating of area 1 & 2, B_1 and B_2 are the frequency bias parameter of each area and ΔF_1 & ΔF_2 are the frequency deviation of the area 1 & 2 respectively.

The speed changer commands will thus be of the form,

$$\Delta P_{c1} = -K_{I1}\int(\Delta P_{tie1} + B_1 \Delta F_1)dt \tag{10}$$

$$\Delta P_{c2} = -K_{I2}\int(\Delta P_{tie2} + B_2 \Delta F_2)dt \tag{11}$$

Where, K_{I1} and K_{I2} are the integral control gains of area 1 & 2 respectively. The studied area has equal area capacity, 1000 MW. These four parameters B_1, B_2, K_{I1} and K_{I2} are optimized on the basis of performance index, which is given in (12),

$$PI = \int_0^{40} (\Delta P_{tie}{}^2 + w_1 \Delta F_1{}^2 + w_2 \Delta F_2{}^2) dt \qquad (12)$$

On the above equation value of the weight factor w_1 and w_2 both are chosen as 0.25.

In the optimization procedure of the controller parameters the range of integral gain is taken from 0.005 to 0.2 and the range of frequency bias parameter is taken from 0.1 to 0.6. In both of the area the SLP is introduced in the range from -5% to +5%. But maximum 5% of SLP in total is accepted of both areas. If the arithmetic sum of SLP of both areas is more than 5%, then it is not taken into account. Therefore, 91 number of SLP combinations are formed for both areas as ΔP_{L1} and ΔP_{L2}. For each combination of SLP the corresponding four controlling parameters B_1, B_2, K_{I1} and K_{I2} are optimized on the basis of PI on the procedure described in [1]. Fig. 3 shows 91 numbers of optimized parameter sets against each of the combination of SLP.

IV. Controller Parameters Generation by FLC

The AGC controller parameters are optimized on the basis of performance index which is described in the previous section. In the power system, load is varying continuously which is uncertain and not constant. Nevertheless, the controller parameters should be switched according to the position of SLP. Therefore, a varying range of SLP is deliberated in the tuning procedure. Fuzzy logic controller will be able to generate each of the four of the optimized parameters for an unknown combination of SLP in the range from -5% to +5% disturbances. Moreover, known value of small load change ΔP_{L1} and ΔP_{L2} are taken as the input of FLC and any of the controller parameters is the output of FLC. For each of the input or output there are 11 membership function taken as, NVB, NB, NM, NS, NVS, Z, PVS, PS, PM, PB and PVB. The experience for the rule base is considered from the previous section. Mamdani Implication technique is used and centroid defuzzification procedure is considered to get the output from FLC. Fig. 4 shows surface plot for different optimized controller parameters obtained from fuzzy logic controller. In this studied model, four fuzzy logic controllers are connected and SLP of both areas are sensed as FLC input and instantly four optimized outputs of B_1, B_2, K_{I1} and K_{I2} are generated.

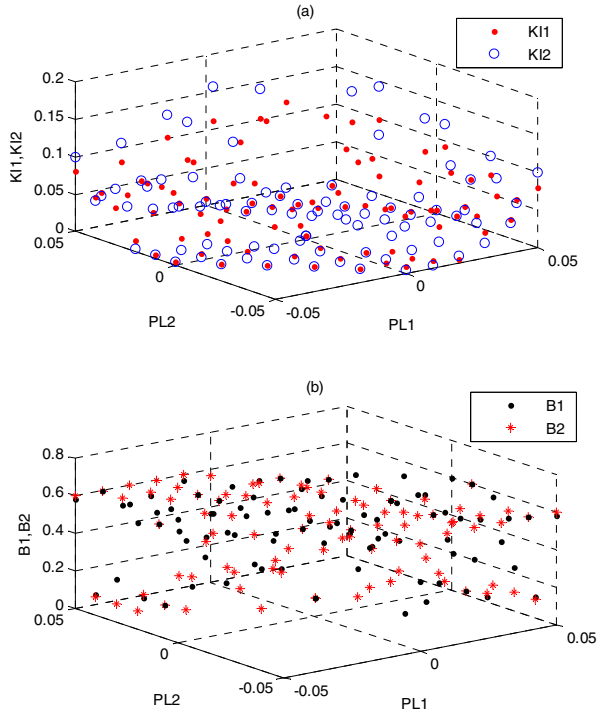


Fig. 3 Optimized controller parameters for -5% to +5% SLP on both area a) integral control gain parameters, b) frequency bias parameters.

V. Transient Response and Discussion

The transient response of frequency deviation of both areas and tie line power deviation of the interconnected power system is deliberated in this paper with governor dead band effects as depicted in Fig. 5. Three test cases are observed. At first in column a, responses are shown for +1.0% step load change at area 1, for 40 seconds. The optimized gain

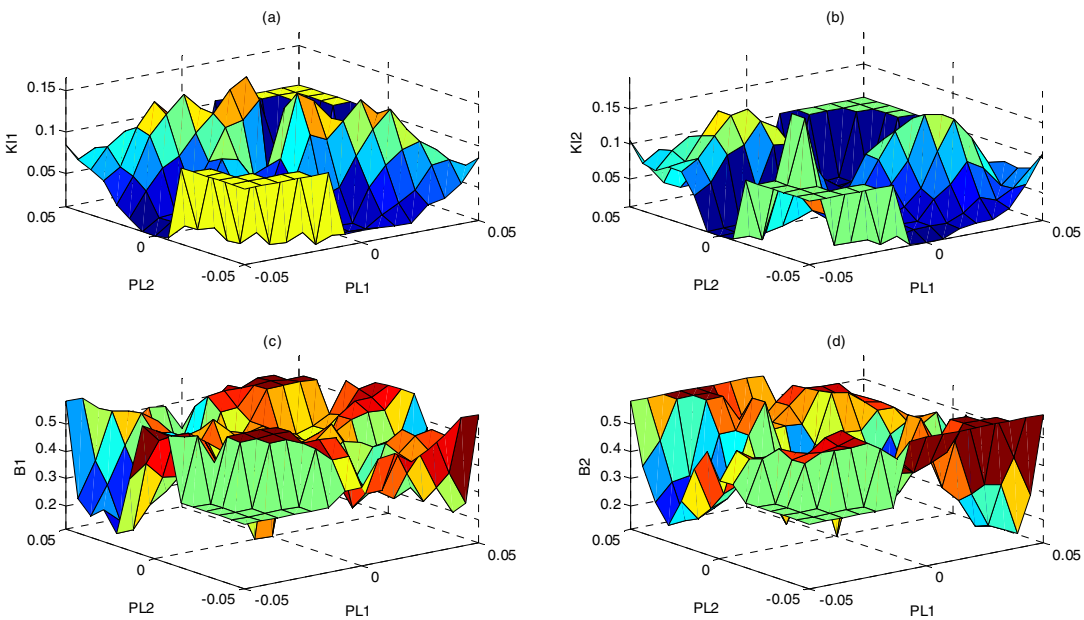


Fig. 4 FLC generated optimized controller parameters, a) & b) integral control gain parameter KI_1 & KI_2, c) & d) frequency bias parameters B_1 & B_2.

parameters are found after tuning on the basis of performance index, as B_1 =0.6, B_2 =0.39, K_{I1} =0.065 and K_{I2} =0.065. The parameters from FLC output found for +1.0% SLP at area 1 are B_1 =0.585, B_2 =0.4, K_{I1} =0.065 and K_{I2} =0.065 which is very similar to the parameters found from the tuning through *PI* basis. So the figure shows almost same responses due to this SLP for the system with and without FLC. Secondly, +1.8% SLP in area 1 and +1.2% SLP at area 2 are introduced in the system and responses are shown by column b, in the same figure. But the optimized parameters are not provided by *PI* based tuning procedure, but found from FLC output. Therefore, parameters found are B_1 =0.388, B_2 =0.388, K_{I1} =0.0402 and K_{I2} =0.0244. Finally, +1.0% SLP is introduced both at area 1&2 and responses are shown by column c and optimized parameters found by FLC are B_1 =0.585, B_2 =0.585, K_{I1} =0.045 and K_{I2} =0.025. The parameters found for +1.0% SLP at area 1 through *PI* basis is taken as reference parameters for all load disturbances. This is sometimes done by many researchers, to tune the control parameters for only one sample SLP and is used to the system. It is demonstrated in the figure that, a single set of control parameters does not give a proper solution for all SLP and is highly unaccepted. Therefore, FLC generated variable control parameters in response to the SLP gives the exact solution.

VI. Conclusions

A comprehensive mathematical model for AGC of two area interconnected power system with AC-DC tie line is considered for this investigation. Nonlinearities such as, governor dead band, generator rate constraints are also considered to obtain a realistic environment. Initially the controller parameters are optimized for each individual SLP combination on the basis of performance index. Then fuzzy logic controller is designed to generate the optimized control parameters in accordance to the load disturbance. Simulation results are performed under the MATLAB environment.

References

[1] Hasan Masrur, Asif Ferdoush, and M. G. Rabbani, "Automatic generation control of two area power system with optimized gain parameters," in the 2nd *International conference on Electrical Engineering and Information & Communication Technology*, Dhaka, 21-23 May, 2015.

[2] L.V.S. Kumar, G.V.N. Kumar, and S. Madichetty, "Pattern search algorithm based automatic online parameter estimation for AGC with effects of wind power," *International Journal of Electrical Power & Energy Systems*, vol. 84, pp. 135-142, Jan. 2017.

[3] Ibraheem, P. Kumar, and D. P. Kothari, "Recent philosophies of automatic generation control strategies in power systems," *IEEE Trans. On Power System,* vol. 20, pp. 346–357, Feb. 2005.

[4] Z. Wang, F. Liu, L. Chen, and S. Mei, "Distributed economic automatic generation control: a game theoretic perspective," in the *IEEE PES General Meeting,* Denver, 26-30 Jul, 2015.

[5] S. C. Tripathy, and K. P. Juengst, "Sampled data automatic generation control with superconducting magnetic energy storage in power systems," *IEEE Trans. on Energy Conversion*, vol. 12, No. 2, pp. 187-192, Jun. 1997.

[6] N.J.V. Kumar, and M.M.T. Ansari, "A new design of dual mode type-II fuzzy logic load frequency controller for interconnected power systems with parallel AC-DC tie-lines and capacitor energy storage unit," *International Journal of Electrical Power & Energy Systems*, vol. 82, pp. 579-598, Nov. 2016.

[7] J. B. X. Devotta, M. G. Rabbani, and S. Elangovan, "Effect of SMES unit on AGC dynamics," in the *International Conference on Energy Management and Power Delivery*, Singapore, 3-5 Mar. 1998.

[8] S. C. Tripathy, G. S. Hope, and O. P. Malik, "Optimization of load frequency Parameters for power system with reheat steam turbines and governor deadband nonlinearity," in *IEE proceeding*, vol. 129, pp. 10-16, 1982.

[9] R. K. Sahu, T. S. Gorripotu, and S. Panda, "Automatic generation control of multi-area power systems with diverse energy sources using teaching learning based optimization algorithm," *International Journal of Engineering Science and Tech.,* vol. 19, pp. 113-134, Mar. 2016.

[10] E. Yesil, A. Demiroren, and Y. Erkin, "Automatic generation control with fuzzy logic controllers in the power system including three areas," in the 2nd *International Conference on Electrical and Electronic Engineering*, Bursa, Turkey, 7-11 Nov. 2001.

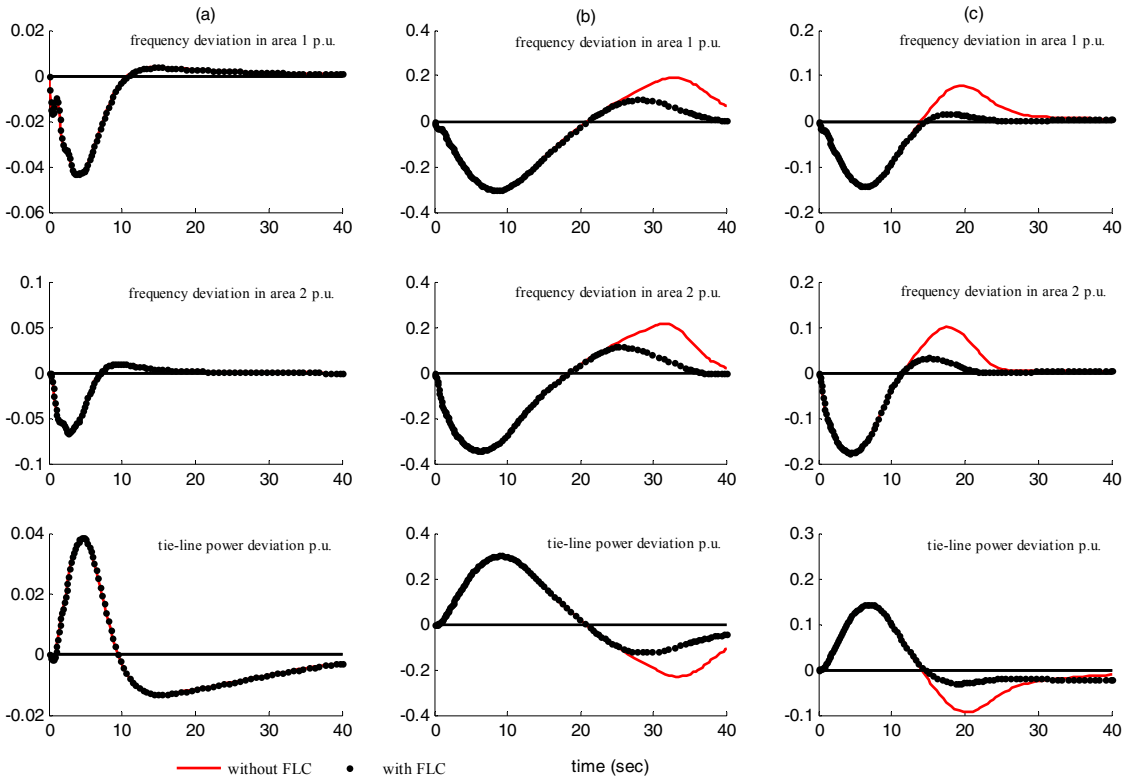


Fig. 5 System transient responses (frequency deviation at area 1&2 and tie line power deviation).

Reactive Power Reserve Management to Prevent Voltage Collapse in Bangladesh Power System

A. T. M. Mustafizur Rahman,* A. H. Chowdhury

Department of EEE, Bangladesh University of Engineering and Technology
Dhaka-1000, Bangladesh
*mustafizur.rahman@wartsila.com

Abstract—**Voltage stability has become major concern in power system planning and operation, and deals with system adequacy and security. Amount of reactive power reserves at generating stations is a measure of the degree of voltage stability. In Bangladesh power system generating units are mostly running at their highest capacity, no active or reactive power reserve is being kept for supporting system contingencies. This makes the system vulnerable to voltage stability and collapse. This paper studies the Bangladesh power system network characteristics from voltage stability perspective and proposes a reactive power reserve management scheme for operating as well as upcoming power plants. The proposed scheme has been tested in an operating power plant and shows effective results.**

Index Terms— **Voltage stability, reactive power reserve, generator capability limits, AVR.**

I. Introduction

Voltage stability of a power system is concerned with the ability of the system to maintain acceptable voltages at all buses in the system under normal conditions and after being subjected to a disturbance. A system enters a state of voltage instability when a disturbance, increase in load demand or change in system condition causes a progressive and uncontrollable decline in voltage. The main factor causing instability is the inability of the power system to meet the demand for reactive power [1].

In the modern competitive electric market power systems are more heavily loaded than ever before because of the growing demands, maximum economic benefits and efficiency of usage of transmission capacity. In this perspective many power systems are operated more often longer and closer to voltage stability resulting higher probability of voltage instability or collapse [2-3]. The evidence of these circumstances is widespread blackouts in the recent decades. Insufficient voltage and reactive power support was an origin or a factor in the major power outages worldwide [4-5] including Bangladesh [6]. Adequate voltage stability margin needs to be obtained through appropriate scheduling of the reactive power resources and ensuring sufficient reactive power reserve (RPR). Both reactive power generation and its reserve should be considered simultaneously in the procurement and the scheduling of the reactive power resources [7]. The RPR can be taken into account from the load or the generator point of view which is called LRPR and GRPR respectively [8-19]. This paper particularly focuses on GRPR and taking the relevance of Bangladesh Power System.

Bangladesh power system is evolving and expanding very fast. Its generation capacity has increased three folds in the last eight years. Transmission capacity also increased with interconnection to India. In 2030 the target growth of generation is estimated about 30 GW. Power generating units are being installed having different sizes, capacities and different types of control dynamics. Considering these facts, Bangladesh Power System Network (BPSN) requires in depth study on system stability, protection coordination and centralized generation and load dispatch control scheme.

This paper proposes a new generation reactive power reserve management scheme. An extra outer loop has been added to the AVR control blocks, where generators' reactive power are being controlled by system voltage. If system voltage drops, more reactive power will be injected. When system voltage remains in nominal level, reactive power generation is adjusted and kept the margin available for future demand of contingency. The proposed scheme has been tested in an operating power plant and shows satisfactory performance.

II. Bangladesh Power System

A. Simplified Model of Bangladesh Power System Network

The transmission network of BPSN has been simplified for study and analysis purposes. The simplified network consists of 12 buses and kept its basic connectivity scheme. The buses have been chosen so that major segments or areas of the network are covered. Each bus represents a certain area, generation and loads of that area are considered lumped and assigned to that bus only. Fig. 1 shows the connectivity of these buses.

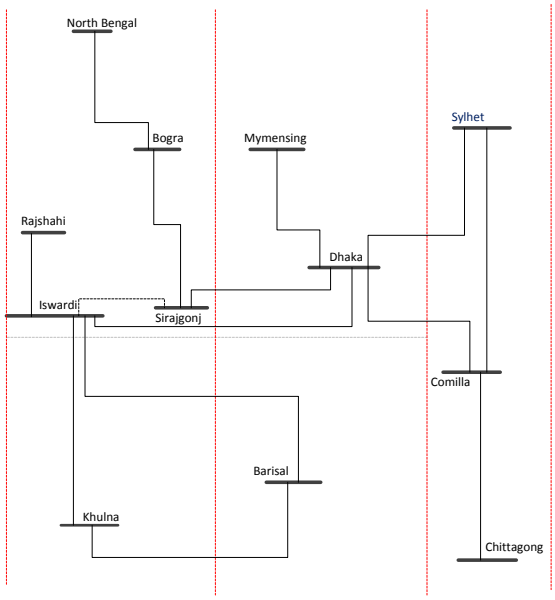


Fig.1 Simplified model of BPSN with 12 buses

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

B. *Voltage profile of different buses*

Voltage and load profile of BPSN have been studied from operation data of PGCB and BPDB. Data for the whole year of 2015 has been studied. It is observed that in the 230 kV level voltage varied from 210 kV to 235 kV. Dhaka and Chittagong areas are highly stressed, North Bengal and Khulna region have lower voltage profile over the year. The 132 kV system is highly stressed and has less regulation. North Bengal, Khulna, Barisal, Mymensing, Comilla and Dhaka areas are mostly affected. In April 2015, maximum and minimum both values occurred to their peak, it means the voltage control of the network is not well regulated. Average minimum voltage of all buses is about 115kV which is 12.8% less than nominal.

C. *Voltage Control Simulation of BPSN*

To understand the nature of voltage profile of BPSN, a simulation tool has been developed based on Newton Raphson load flow algorithm. Bus voltages at different load and operating conditions have been determined using the tool. With incremental change of load, $\frac{\Delta v}{\Delta Q}$ can be identified for a bus. The effect of generator reactive power reserve has also been investigated. These investigation showed that, if generators' operating mode are triggered so that they respond to system voltage variation and utilize their remaining reactive power reserve greater voltage stability can be ensured.

Different scenarios have been tested:

- How the bus voltage changes when PV buses utilize their full reserve capacity and turn into PQ Bus
- How the bus voltage being affected by utilizing reactive power reserve of the PQ buses.

Two simulation results have been presented in Figs. 2 and 3. In Fig. 2, it can be seen that when PV bus has utilized its full capacity, the rate of fall of voltage at Khulna Bus is very sharp with change in load demand. In Fig. 3, it is seen that when local generator reactive power reserve is being utilized, voltage profile improves significantly.

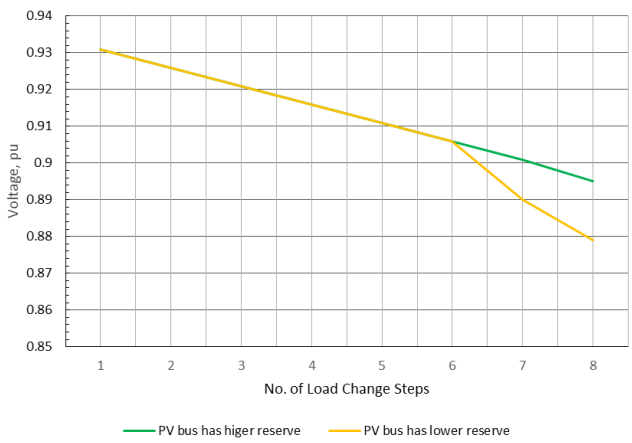


Fig.2: Voltage at Khulna bus when PV bus has utilized its full reactive power reserve.

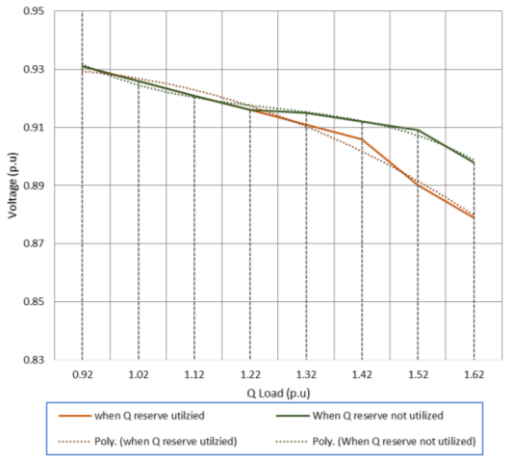


Fig.3: Voltage profile comparison of Khulna bus when local generator reactive power reserve has been utilized.

Form these test and simulation results it has been clearly understood that generators should have sufficient reactive power reserve and will respond to system voltage deviation for better voltage stability. Fixed active and reactive power generation scheme (PQ mode) is not suitable for the purpose. Generators should run in fixed active power mode but reactive power should be system voltage depended. In Bangladesh most of the power plants are running in fixed active power with obligatory fixed power factor. Thus the owners or operators run their power plants at the highest rated point keeping no reserve for any contingency.

III. Proposed Reactive Power Management Scheme

Generator terminal voltage and reactive power are controlled by excitation current which is regulated by automatic voltage controller (AVR). AVR always monitors generator terminal voltage and machine current for maintaining voltage and reactive power to the desired level. This is standard AVR control loop. In the proposed control scheme, an additional control loop is included to this standard control model. The proposed control function produces a biasing signal (U_{bias}) by monitoring and comparing switchyard voltage magnitude at HT side of power transformer. If HT side voltage drops below a threshold level, the control function generates a positive bias signal which eventually being added to the reference value of AVR and directs AVR for producing more reactive power. The additional reactive power injection helps to improve the system voltage. While system voltage stays within an acceptable limit, there will be zero bias signal and no additional reactive power is delivered. In Figure-4 this control function is illustrated.

Additional to the bias signal there should be associated control functions to be performed like selection of MV, HV bus and generators selection of operating mode, enabling disabling of the function, evaluation of generator reserve capacities etc. Also, there should be safety function integrated to this control scheme, so that generators are not being operated exceeding their capability limit. This control scheme is termed as Generator Reactive Power Management (GRPM) scheme.

The control algorithm is presented in Fig. 5. GRPM algorithm can be implemented in most of the power plant having automatic voltage regulator and SCADA or PLC system. A Matlab Simulink model has been designed with this algorithm and simple SIMSCAPE model the dynamic

behaviour has been simulated. Figs. 6 and 7 show the basic Simulink blocks. When reactive power generation is controlled with reference to the system voltage, generator capability and reserve margin is evaluated in real time based on active power generation, derating factors, temperature rise etc. This capability limit dynamically sets the upper limit of the saturation block in Fig. 6. Generators which are being operated with full capacity will not respond to the U_{bias} signal.

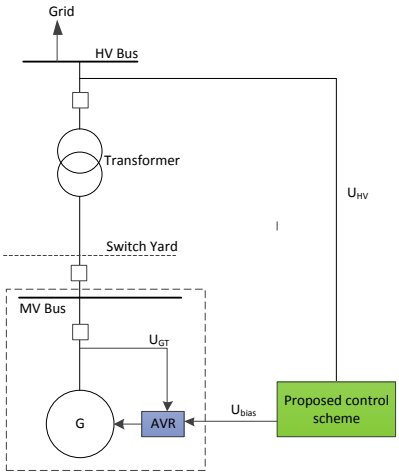


Fig. 4: Proposed control function in simple block diagram

In the SIMSCAPE model a generator has been considered in parallel with a grid system where variable load switching causes voltage variations. Generator AVR is fed by bias signal generated by the GRPM control block to respond the voltage variations. In the simulation result, it is seen that when GRPM scheme is active, system voltage gets more stability. When GRPM is disabled, system voltage drops at the point of switching and takes more time to stabilize. Simulation results are presented in Figs. 8 and 9.

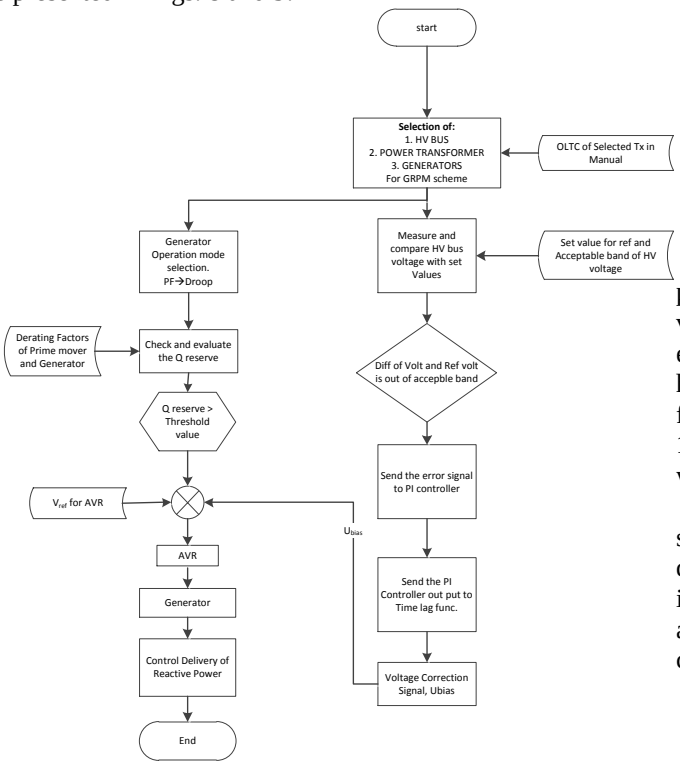


Fig. 5: Flow diagram of control algorithm of GRRM

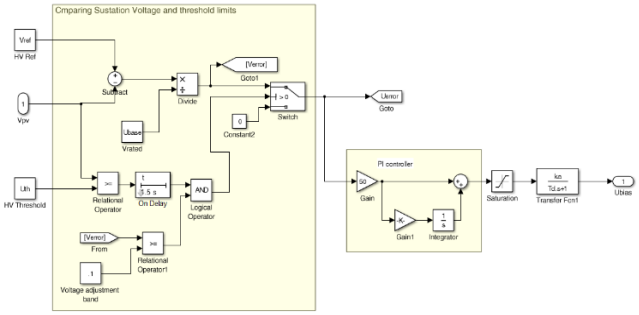


Fig. 6: Simulink Model for producing voltage bias signal, U_{bias}

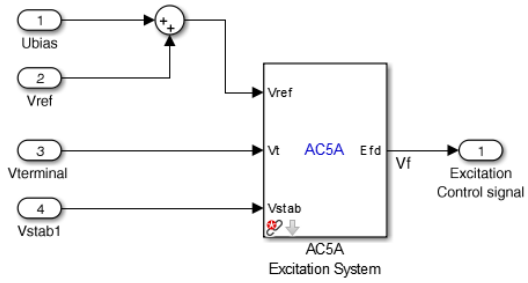


Fig. 7: AVR control block with voltage bias signal integration

In Fig. 8, when GRPM has been disabled it is sheen that system voltage drops more when a load has been switched on at 5 sec. But in the case when GRPM ais active, voltage drop is less (Fig. 9).

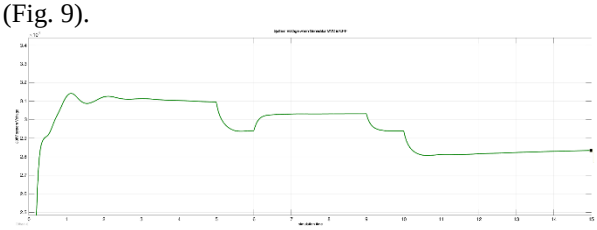


Fig. 8: Simulation result without GRPM

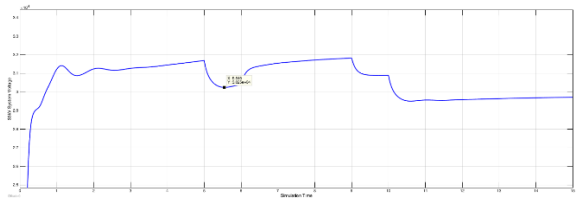


Fig. 9: Simulation result with GRPM active.

This control algorithm has been tested in an operational power plant, which connected to 33kV grid system with a quite variation of reactive power demand due to nearby induction and electric furnaces. The test power plant has four generators each having 8.7 MW capacities and a rated voltage of 11 kV. The four generators are grouped into two at two different buses, bus-1 and bus-2. During this test, transformer on-load tap changer was disabled by keeping it in ‘manual’ mode.

The test results are presented in Fig. 10 that shows satisfactory performance. The system voltage gets corrected quite fast by reactive power injection and when system voltage is within acceptable limits, generators reduces VAr injection and keep the reserve ready for any contingency or fast load change.

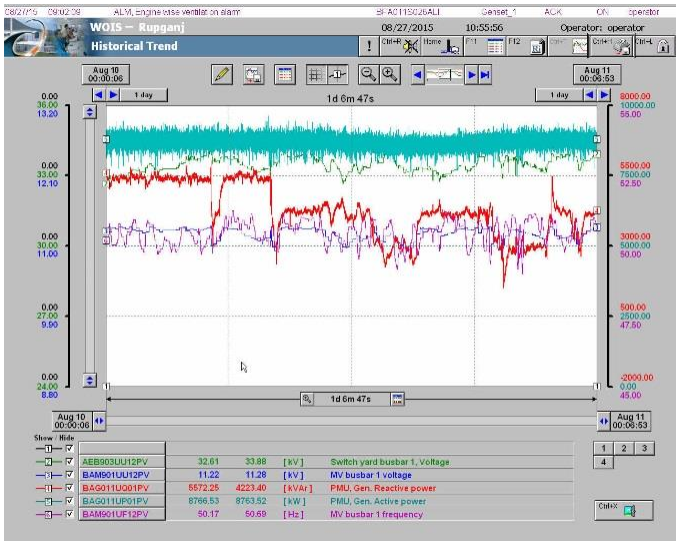


Fig. 10: 24 hours test result for bus-1 with GRPM enabled.

Similar test result has been achieved for bus-2. In the real life, with this test it has been proved that generators can be run in GRPM mode and better voltage control can be achieved. Another feature of this test is that, when network voltage is healthy and within limit, generator reactive power will be less thus over stressing can be avoided.

IV. Conclusions

In this paper, the voltage profiles of Bangladesh Power system have been analyzed. A simplified network model has been used to find out voltage stiffness at different key buses. A new reactive power control scheme has been proposed. The proposed scheme has been implemented an operating power plant and shows very good performance. These scheme can be implemented in different power plants located remotely to give better voltage support to the grid.

References

[1] P. Kundur, Power System Stabiltiy and Control, New York: McGraw-Hill, 1994

[2] H. H. Goh, Q. S. Chua, S. W. Lee, B. C. Kok, K. C. Goh, K. T. K. Teo, " Comparative Study of line Voltage Stability Indices for Voltage Collapse Forecasting in Power Transmission System," *IEEE, Proceeding of the Conference, Rio de Janeiro, Brazil*, 11-12 Feb 2015, Vol. 13, Issue.2, Part-II

[3] C.Y. Lee, S.H. Tsai, Y.K. Wu, "A new approach to the assessment of steady-state voltage stability margins using the P–Q–V curve," *ELSEVIER, Electrical Power and Energy Systems*, pp. 1091-1098, 32 (2010)

[4] O. A. Mousavi, R. cherkaoui, "Literature survey on fundamental issues of voltage and reactive power control," October 2011 *[Online]. Available at: http://mars.etz.ch/en/research-and-publications/publications.html*

[5] Y.V. Makarov, V.I. Reshetov, A. Stroev and I. Voropai, "Blackout prevention in the United States, Europe, and Russia," *IEEE, Proceedings*, vol. 93, pp. 1942-1955, Nov.2005.

[6] Q. Ahsan, A.H. Chowdhury, S.S. Ahmed, I.H. Bhuyan, A. Haque and H. Rahman, "Technique to Develop Auto Load Shedding and Islanding Scheme to Prevent Power System Blackout," *IEEE Transaction on Power Systems*, Vol. 27. No. 1, pp. 198-205, February 2012

[7] "Planning Against Voltage Collapse," CIGRE Task Force 38-01-03, 1987

[8] O. A. Mousavi, M. Bozorg, R. Cherkaoui ,"Preventive Reactive Power Management for Improving Voltage Stability Margin," *Elsevier Electric Power Systems Research*, pp. 36– 46, 96 (2013)

[9] Y. L. Chen, " Weak bus oriented reactive power planning for system security," *IEEE Proc. Generation Transmission Distribution*,Vol. 143, No. 6, pp.541-545 Nov. 1996

[10] P.A. Lof, T. Smed, G. Andersson, D. J. Hill, " Fast Calculation of a Voltage Stability Index," *IEEE, Transactions on Power System*, Vol. 7, No. 1, 1992

[11] P. Kessel, H. Glavitsch, " Estimating the Voltage Stability of a Power System," *IEEE, Transactions on Power Delivery*, Vol. PWRD-1, No. 3, July 1986

[12] V. Ajjarapu, C. Christy, "The Continuation Power Flow: A Tool for Steady Stage Voltage Stability Analysis," *IEEE, Transaction on Power Systems*, Vol. 7, No. 1, February 1992

[13] M. K. M. Rabby, A. H. Chowdhury, M. A. Azamand, M. A. Towfiq, "Bifurcation analysis to identify voltage collapse in Bangladesh Power System,", *IEEE, International Conference on ICIEV*, Dhaka, pp. 1-5, 2013

[14] C.Y. Lee, S.H. Tsai, Y.K. Wu, "A new approach to the assessment of steady-state voltage stability margins using the P–Q–V curve," *ELSEVIER, Electrical Power and Energy Systems*, pp. 1091-1098, 32 (2010)

[15] A. Capasso, E. Marinai, " Influence of Generator Capability Curves on System Voltage and Reactive Power Control Studies," *IEEE, Transactions on Power Apparatus and Systems*, Vol. PAS-97, No. 4, July/Aug 1978

[16] "Investigation of Various Generator Reactive Power Reserve (GRPR) Definitions for Online Voltage Stability/Security Assessment," *Supported by Electric Power Research Centre (EPRC)*, Ames, IA 50010, USA. IEEE 2008

[17] S.Yao, Z. Jian , L. Baiqing, W. Qi, L. Weifang, L. Liping, L. Mingsong, "Study on the mechanism of dynamic reactive power reserves in respect of voltage stability," *IEEE, International Conference on Power system Technology (POWERCON)*,2014, Chengdu, 20-22 Oct. 2014, pp. 1026-1032

[18] X.Fan, G. Wang, L. Fu, H. Yang, F. Ji, "Control strategy of voltage stability based on generator dynamic reactive power releasing," *IEEE, Int. Conference on Advanced Power System Automation and Protection (APAP)*,2011. Pp.460-465, Vol. 1, 2011

[19] Wagar, N., "Voltage Control System, Master Voltage Control," Wartsila's Control Documents for Plant Automation and Control, Doc ID: DBAB709304, 27 May 2011

FCSEA: A Floodless Carrier-grade Scalable Ethernet Architecture

A S M Asadujjaman,[1,*] Shafika Showkat Moni,[2] and Mohammad Shah Alam [1]

[1]IICT, Bangladesh University of Engineering and Technology
[2]CSE, Rajshahi University of Engineering and Technology
[*]eng.asm.asadujjaman@ieee.org

Abstract—**This paper proposes Floodless Carrier-grade Scalable Ethernet Architecture (FCSEA) as a completely floodless enhancement of EtherProxy to meet scalability and protection switching requirements of Ethernet in carrier networks. FCSEA uses a priority based cache retention principle, exploits attempt from end node/host to verify connectivity during commissioning and intercepts dynamic configuration (DHCP) messages to ensure an entirely floodless environment for end hosts. This work enhances the ability of EtherProxy to suppress broadcast traffic and adds protection switching support enabling it to function properly in a ring or mesh protected network. FCSEA retains EtherProxy's advantages of autoconfiguration and backward compatibility with existing hardware. Simulation result and analysis with real data is presented to verify effectiveness of FCSEA and to show its suitability over recent SDN based solutions.**

Index Terms—**Ethernet; scalability; service availability; ring protection, autoconfiguration, ARP.**

I. Introduction

Ethernet is the most widely used and dominant technology for Local Area Network (LAN). The fast switching speed, cost effectiveness and simplicity of Ethernet makes it also desirable in other networks areas such as carrier networks. For instance, Ethernet switching is preferred to routing for high performance of LTE X2 interface [9]. However, a simple Ethernet network is not scalable to serve a large number of nodes as required in telecommunication carrier networks. This is because of broadcast traffic caused by protocols like ARP and DHCP running in an Ethernet network for service and resource discovery.

Currently, protection technologies in carrier Ethernet networks such as Ethernet Ring Protection Switching (ERPS), employ VLANs to keep broadcast within acceptable limit. While this approach requires complex VLAN planning and configuration for every switch, it serves the purpose as a replacement of point to point TDM technologies (e.g. SDH). But, in next generation networks, such as LTE, where nodes require connectivity with their peers, VLAN based segmentation will prohibit nodes under one VLAN to communicate with their neighbours in a different VLAN. FCSEA can operate in carrier networks to replace complex VLAN based scalability solutions while providing unrestricted peer to peer connectivity.

This work evolves from EtherProxy [1], a device for large enterprise networks that uses caching to suppress broadcast traffic. The distinct contributions of this paper are as below,

- A proactive learning and prioritized cache retention policy to prevent initial flooding of entire network.
- Carrier grade protection switching support.
- Simulation and analysis of EtherProxy, SDN based ARP proxy and current work to evaluate broadcast suppression capability and query response time.

The remainder of this paper is organized as follows. A review of works in pursuit of achieving a scalable layer-2 network and their drawback is presented in section II. We present the design of FCSEA in section III. Our work is evaluated in section IV and we conclude this paper in section V.

II. Related Work

To achieve a floodless layer-2 network, SEATTLE [2] uses a link-state routing to maintain switch level topology. Although SEATTLE provides end host backward compatibility, it is not backward compatible with existing Ethernet switches. PortLand [3] uses a fabric manager as a central controller to suppress ARP broadcast in a multi rooted tree topology. The main drawback of this approach is topology dependency and scalability problem of a single controller. The work in [4] proposes a Distributed Registration based Address Resolution Protocol (DRARP) and an End user enabled Mac-in-MAC (EMiM). This approach requires modification of existing hardware in end hosts. NetLord [5] uses a push based ARP proxy model but broadcast still occurs on new virtual machine (VM) boot. Moreover its memory requirement is proportional to number of servers and therefore may lead to scalability issue as number of servers grow. Torii-HLMAC [6] attempts to improve PortLand as a distributed alternative but it is also dependent on a specific topology, fat tree. Moreover Torii-HLMAC uses a new MAC address, Hierarchical Local MAC (HLMAC), and tree based forwarding scheme which makes it incompatible with traditional Ethernet switches. FSDM [7] and SEASDN [8] propose SDN based ARP proxy. The limitation of SDN based approach is that, it suffers from slower failure recovery compared to protection mechanism available for existing Ethernet technology such as ITU-T G.8032. Moreover, these approaches are subject to scalability problem of a single controller and communication with the central controller, located across backbone network, involves a high latency.

III. The Design of FCSEA

The FCSEA partitions a network into segments to protect a segment of the network from broadcast generated by other segments. Each FCSEA segment consists of a subset of hosts and one or more FCSEA enabled devices. This subset of hosts are said to be served by the FCSEA for this segment. Fig. 1 shows an example topology where hosts H1.1 and H1.2 are served by FCSEA device U1 to form a segment

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

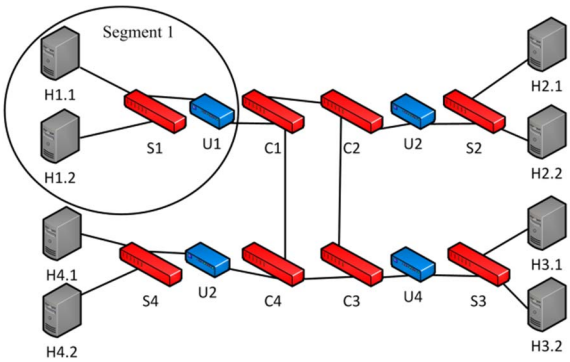


Fig. 1 An example topology to illustrate the design of FCSEA.

(Segment 1). Effective positioning of the protocol requires:

- ability to intercept all traffic towards and from a subset of hosts,
- multiple devices in case of a ring or mesh topology for a subset of hosts.

A. Floodless ARP proxy

EtherProxy uses an on demand pull method to learn MAC addresses. This results in broadcast for the first time request of a target IP. Moreover, if the target host IP is not live in the network, broadcast will continue to happen. The FCSEA protocol proactively learns the logical (IP) address to hardware address (MAC) mapping before a request is made for the first time. Thus FCSEA can serve ARP requests without ever require flooding the hosts for their reply. If a target IP for an ARP request coming from outside the segment is not known, FCSEA ARP module simply drops the request.

Proactive learning of hardware addresses is done by intercepting DHCP address assignments, inspired by [7] and [8]. If DHCP is not used for IP address assignment then it is assumed that the host is commissioned manually. This is usually the case in carrier networks where a new node such as LTE eNodeB is commissioned manually. After commissioning, a PAT (Pre-Acceptance Test) is performed where to check the connectivity a ping test is performed. FCSEA uses the ARP request generated by this ping test to learn the new host.

B. Priority Cache Retention and Learning

FCSEA needs the logical-address-to-hardware-address mapping to reply to queries, to forward broadcast message to unicast destination and to be able to drop a request for an unknown ARP target. However, as the number of nodes in the network increases, some of the learnt mapping may need to be overwritten to keep the memory requirement within feasible limit.

This is achieved by prioritizing the mappings based on three criteria. Firstly, mappings are classified into two groups – internal and external. Internal mapping is - mapping learnt by ARP requests arriving on internal port. To clarify, in Fig. 1, the port of U1 toward S1 is the internal port of U1. Internal mappings are assigned highest priority and never overwritten. External mappings are those learnt from ARP requests and replies arriving on the external port (in Fig. 1, the port of U1 towards C1). Secondly, external ARP requests are not learnt if the cache is already full. Thirdly, LRU replacement policy is implemented similar but unlike EtherProxy, only on external response mappings (External RESP).

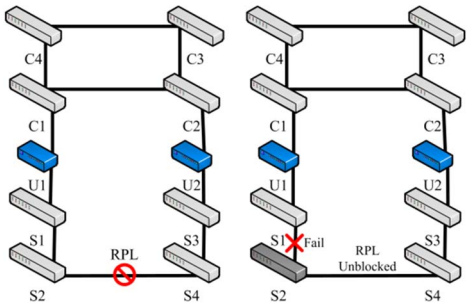


Fig. 2 ERPS ring topology in normal (left) and failure (right) scenario.

C. Protection switching consideration

In order to ensure high service availability, topology with redundant links is deployed in Ethernet carrier networks. A protection switching protocol then ensures a loop free logical topology. ERPS, as discussed in section I, is a protection switching mechanism where multiple rings interconnect to span the entire network. One ring in ERPS is considered as the main ring while rest of the rings are called sub-ring. A sub-ring connects the main ring or other sub-rings at two interconnection nodes. In Fig. 2 C1-C2-C3-C4 is the main ring and C1-S1-S2-S4-S3-C2 is a sub-ring. Every switch in Fig. 2 is connected to one host, not shown in the figure for simplicity. The host connected to S1 is called H1 and so on. C1 and C2 are interconnection nodes. In the terminology of FCSEA, we can consider each sub-ring as a segment and position FCSEA devices (U1 and U2 in Fig. 2) at interconnection nodes to separate a sub-ring from the main ring or to separate one sub-ring from another. Now to ensure a loop free topology ERPS will block one link in the segment (S2-S4) and thus create two partitions in the segment. If a link fails in the segment, the blocked link will be unblocked thus the partitions will change. The segment is now divided at the failed link instead of the blocked link causing some hosts (H2) to migrate from one partition to the other partition. We assume that U1 and U2 is replaced by EtherProxies E1 and E2 respectively. EtherProxy (E2), at this condition, replies to ARP request by an idle host (H4) querying a migrant host (H2), if corresponding entry is in a valid status. This is because, the EtherProxy (E2) learns about the migrant host (H2) from the external interface before link failure. This causes switches to think that the migrant host is towards the EtherProxy. Future traffic from H4 will be forwarded towards E2 by S2 leading to packet loss. In FCSEA, we reflectively issue an ARP request for the original target on the internal interface and only reply to the original request if the reflective request is not answered within a timeout period (Reflective timer < 4ms). If there is a reply to the reflective ARP request then FCSEA drops the original request.

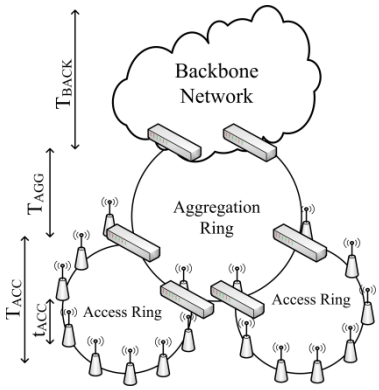


Fig. 3 Telecommunication carrier network architecture.

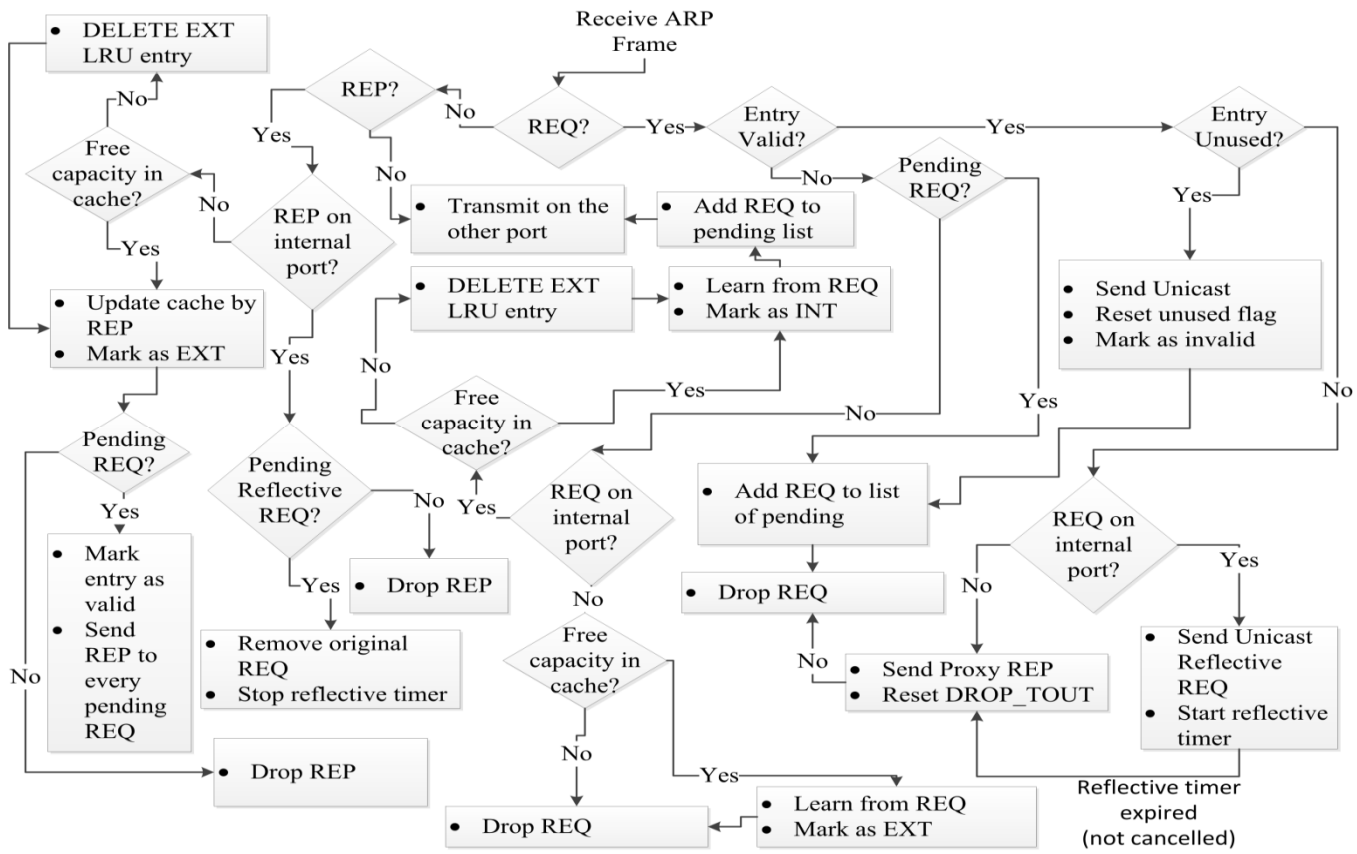


Fig. 4 FCSEA ARP request and response processing logic (includes floodless ARP proxy, priority based cache retention and protection switching logic).

D. Response time

As shown in Fig. 3, let round-trip delay of each access ring, aggregation ring and backbone network link be t_{ACC}, t_{AGG}, and t_{BACK} respectively. Thus total delay to cross access ring, $T_{ACC} = \sum t_{ACC}$. Similarly, delay in aggregation ring, $T_{AGG} = \sum t_{AGG}$ and delay in backbone network, $T_{BACK} = \sum t_{BACK}$. Two neighbours in a network may be in the same access ring or in different access rings. In the latter case aggregation network must be travelled. Therefore response time of FCSEA can be written as,

$$T_{FCSEA} = T_{ACC} + T_{AGG} + T_{ACC} = \sum T_{ACC} + T_{AGG}$$
$$\text{Or, } T_{FCSEA} = \sum\sum t_{ACC} + \sum t_{AGG} \quad (1)$$

For SDN based proxy FSDM, delay of backbone network is included to reach central controller. Therefore response time of FSDM can be written as,

$$T_{FSDM} = T_{ACC} + T_{AGG} + T_{ACC} + \sum T_{BACK}$$
$$\text{Or, } T_{FSDM} = \sum\sum t_{ACC} + \sum t_{AGG+} \sum\sum t_{BACK} \quad (2)$$

IV. EVALUATION

In this section we evaluate the effectiveness of the FCSEA protocol using event based simulation.

A. Cache hit rate

In this experiment (Fig. 5 and Fig. 6) we evaluate cache hit rate over time to see effectiveness of the system to suppress broadcast traffic. We consider a successful hit if no host outside the segment from where ARP request is generated, receive the request. This way a hit is equivalent to a suppressed broadcast. We have performed two experiments with request generation rate of 60 requests per minute.

Results shows that, as the number of hosts increase, hit rate of EtherProxy is decreased whereas FCSEA always have 100% cache hit which suggest a floodless network with respect to hosts.

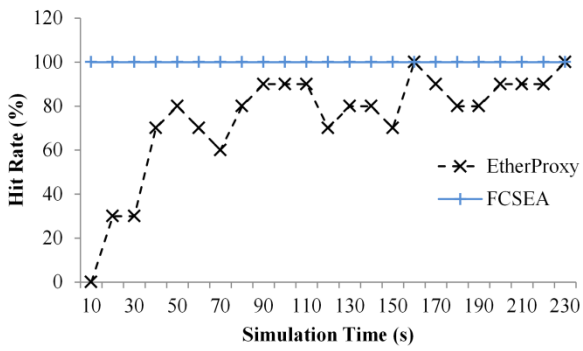


Fig. 5 Cache hit rate comparison with 130 hosts, in this experiment 30 hosts in the segment generating ARP request for 100 hosts in other segments.

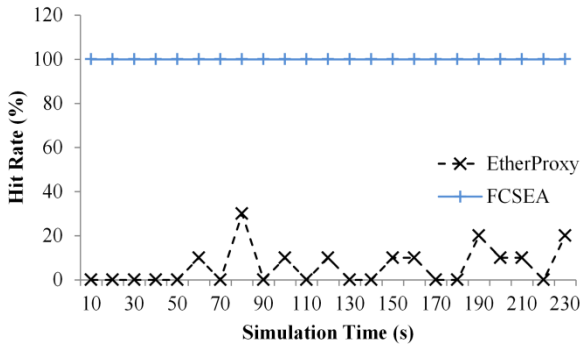


Fig. 6 Cache hit rate comparison with 50.5K hosts, in this experiment 500 hosts in the segment generating ARP request for 50K hosts in other segments.

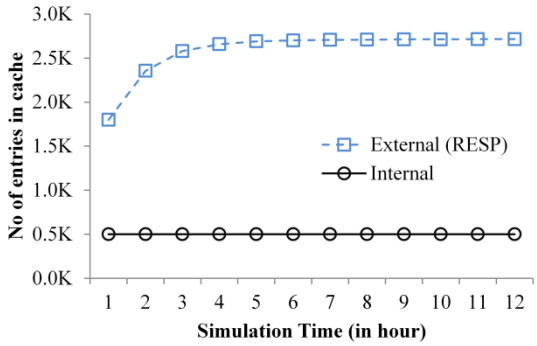


Fig. 7 Cache size evaluation with 50.5K hosts. In this experiment cache size is observed over duration of 12 hour with 500 hosts in the segment sending ARP request for 50K hosts in other segments in the network.

TABLE I
ROUND-TRIP DELAYS IN TELECOM BACKBONE NETWORK

Region	1	2	3	4	5	6	7	8
Delay (ms)	6.2	9.7	14.3	6.6	6.4	3	9.8	7.2

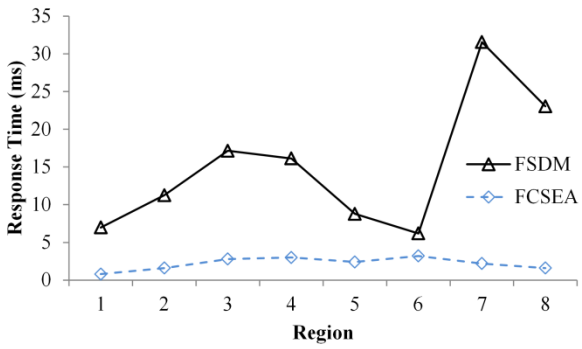


Fig. 8 Response time of FCSEA to reply to ARP requests compared to SDN based proxy.

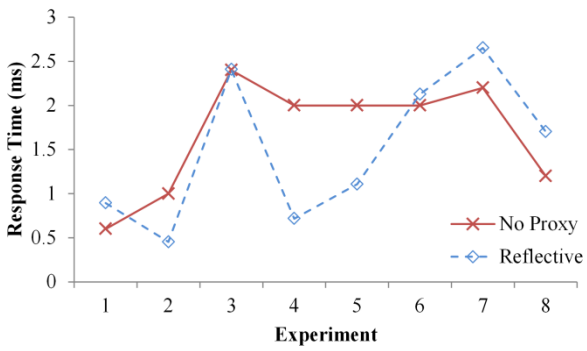


Fig. 9 Performance evaluation of Reflective ARP mechanism.

B. Cache size

In this experiment (Fig. 7), we observe characteristics of the mappings of different types: Internal, External (RESP) as defined in section III (B) to confirm the priority cache retention and learning principle's limited memory requirement.

C. Response Time

This experiment (Fig. 8) measures the response time, that is, the delay between generation of an ARP request by a host and receiving the corresponding ARP reply. Delays shown in analysis in section III (D) are individually measured for eight regions of a telecommunication carrier network. Backbone network delays found for different regions of the carrier network are listed in Table 1. These delays are used as channel delays in the simulation. As we can see from Fig. 8, FCSEA always have a lower delay compared to SDN based proxy FSDM (Floodless Service Discovery Mechanism). A sudden rise of delay in FSDM is noticed in region 7. This is because more backbone network links were involved during the experiment in this region.

D. Reflective ARP Delay

In this experiment (Fig. 9) we assess the delay caused by reflective ARP mechanism of FCSEA. Eight different experiments were performed by generating ARP requests between eight different randomly selected pairs of source and target hosts. Query response time using FCSEA and without using FCSEA is measured for same pair of hosts. Results show that reflective ARP doesn't introduce any significant delay overhead on FCSEA. As we can see in Fig. 9, query response time of FCSEA (Reflective ARP) is higher than response time of no proxy in some experiments and vice versa. This is because when FCSEA device is closer to ARP source than target, reflective ARP require less time.

V. CONCLUSIONS

In this paper, we propose FCSEA, a protocol to scale Ethernet to large carrier networks. Our work is based on EtherProxy, a protocol to scale Ethernet by suppressing broadcast traffic. With the aim to scale Ethernet for carrier network topologies, we improve EtherProxy in several ways. Firstly, we introduce a floodless proxy mechanism to overcome the initial-flooding limitation of EtherProxy. Secondly, we also employ a priority based caching mechanism to achieve a scalable memory requirement. Finally, we add support for redundancy protected networks for high service availability with a reflective ARP request mechanism. FCSEA is a simple protocol with limited memory requirement. Therefore it can be implemented as a software up-gradation on existing Ethernet hardware. It can also be implemented as a standalone device. Simulation results show that our proposed protocol can effectively ensure a floodless environment for end nodes. We also show that FCSEA outperforms SDN based broadcast suppression protocols due to its fast query response time.

REFERENCES

[1] K. Elmeleegy and A. L. Cox, "EtherProxy: Scaling Ethernet By Suppressing Broadcast Traffic," In IEEE INFOCOM, May 2009.
[2] C. Kim, M. Caesar, and J. Rexford, "Floodless in SEATTLE: A Scalable Ethernet Architecture for Large Enterprises," in ACM SIGCOMM, Aug. 2008.
[3] R. Niranjan Mysore et al., "Portland: a scalable fault-tolerant layer 2 data center network fabric," in ACM SIGCOMM, 2009.
[4] X. Sun and Z. Wang, "An efficient and Scalable Metro-Ethernet Architecture," in Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 3, no. 4, Dec., 2008.
[5] J. Mudigonda et al., "NetLord: a scalable multi-tenant network architecture for virtualized datacenters," ser. ACM SIGCOMM, 2011.
[6] E. Rojas and G. Ibáñez, "Torii-HLMAC: A Distributed, Fault-tolerant, Zero Configuration Fat Tree Data Center Architecture with Multiple Tree-based Addressing and Forwarding," Globecom 2012.
[7] W. Jian, L. Jiang, H. Tao and L. Yunjie, "A Novel Floodless Service Discovery Mechanism Designed for Software-Defined Networking", in IEEE Communications, China, vol.11, No.2, pp.12-25, Feb., 2014.
[8] N. Jehan and A. M. Haneef, "Scalable Ethernet Architecture using SDN by Suppressing Broadcast Traffic," 5th Intl. Conf. on Advances in Computing and Comm., 2015.
[9] "Backhauling X2", Cambridge Broadband Networks, White Paper 2011.

A Linearly Polarized (LP) Reader Antenna for LP and Orientation Insensitive (OI) Chipless RFID tags

Md. Aminul Islam,[1,*] and Nemai Karmakar[2]

[1]Electrical, Electronic and Communication Engineering (EECE),
Military Institute of Science and Technology (MIST), Mirpur, Dhaka 1216, Bangladesh
[2]Electrical and Computer Systems Engineering (ECSE), Monash University, Clayton, VIC 3800, Australia
[*]aminul9274@gmail.com

***Abstract*—A single element linearly polarized (LP) aperture coupled microstrip patch antenna (ACMPA) in microwave frequency band is presented in this paper. The paper starts with the requirement for developing a low profile, low cost, lightweight, wideband LP antenna suitable for commercial chipless RFID tag readers. Then, a thorough review is done on existing low profile wideband antenna designs to find out a suitable match with all required reader antenna specifications. From the literature analysis, ACMPA design technique is found as the best match for developing the reader antenna, which is followed next. First, antenna specifications are listed for developing a single element LP ACMPA. Then, a single element is designed and optimized to achieve the targeted values, which is fabricated and assembled next. After that, this ACMPA is measured in a vector network analyzer (VNA), which fulfills the set targets. Next, tag reading performance of the antenna in three different measurement setups is evaluated through measurement of 'O' slot loaded orientation insensitive (OI) tags. These tag measurement results have confirmed the usability of the developed antenna in commercial tag reader.**

***Index Terms*—Aperture-coupled antennas, radio-frequency identification (RFID), RFID tags.**

I. INTRODUCTION

Radio Frequency Identification (RFID) system is a wireless data capturing technology that uses radio frequency (RF) waves for extracting the encoded data from remotely placed tags. This system consists of two main elements, the RFID tag, where data are encoded, and the RFID reader, which consists of reader antennas, a digital section, and an RF section, and is used for extracting the encoded data from the tags. RFID tags have the potential to replace barcodes, but they are still not used in low-cost applications because of their higher prices compared to that of barcodes. The cost of the widely-used passive RFID tags depends mainly on the silicon *chips* used in them [1]. Therefore, research has been focused on developing *chipless* printable RFID *tags* and reliable *readers* for detecting those tags.

A number of printable chipless RFID tags have been reported using simple or hybrid of time, frequency, phase domain and image-based encoding techniques. Among the time domain based tags, it is not feasible to encode sufficient bits using printable microstrip delay lines [2] within a credit card sized area. Therefore, time domain-based tags cannot be used for high data density tag manufacturing. Phase domain- and image-based tags are still in the experimental phase. In contrast, frequency domain-based tags have higher data density than time-domain based tags. Hence, researchers around the world are now concentrating more on the frequency domain and the number of proposed tags using the UWB microwave frequency band is increasing significantly. Linearly polarized (LP) compact printable tags are proposed in [3] and [4], where the tag placement polarity has to be the same as the polarity of the LP reader antennas to be read by the reader. Orientation insensitive (OI) tags are proposed later in [5] and [6], which can be read in any orientation with the same polarized LP antennas of the reader. Besides, to improve the isolation between the transmitted and received signal, a cross polarized (CP) tag is presented in [7], where the re-transmitted signal from the tag is received in orthogonal polarity in the reader, and the transmitter (Tx) and receiver (Rx) antennas are needed to be placed in orthogonal polarization to each other.

Reader's digital and RF components for reading these frequency domain based chipless RFID tags are now available in the market. But yet no commercially suitable reader antenna has been developed for the reader with low profile, low cost, lightweight, high gain and wide bandwidth for reading these tags from a decent distance. Although, horn antennas are used widely by the researchers for proving their design concepts of new chipless tag and reader prototypes, they cannot be used as a commercial reader antenna due to their large profile, heavy weight and huge cost. Hence, an antenna with very low-profile is needed to be developed for the reader to read the chipless RFID tags. This requirement is marked in blue color in Fig. 1, where three different tag reading methods are demonstrated using single antenna and two antenna setups for reading the LP, OI, and CP tags.

A number of techniques have been analyzed for developing the UWB reader antenna. A symmetric approach using Vivaldi notch elements is presented in [8] and another design using log-periodic folded dipole is shown in [9]. But these types of end fire antennas possess a very large profile and cannot be used commercially. After analyzing several other low-profile wideband antenna design techniques, LP ACMPAs are found as wideband, low profile, low-cost, light weight, high gain antennas. This is because they have three additional parameters than the normal patch antennas for improving their performances, which are the tuning stub length, coupling aperture size and multilayer patches. In addition, its ground plane provides a shield to the patch radiation from the feed radiation and its small back radiation from the aperture can be reduced further by adding a back

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

reflector. Therefore, in this paper, an ACMPA with wide beam-width and high gain is proposed for achieving a decent reading range in the micro-wave frequency band, which can be used in commercial chipless RFID tag reader.

At first, a single ACMPA is designed and optimized in the micro-wave frequency band. After that, the optimized design is fabricated, assembled, and tested. Finally, the tag measurement performance is analyzed using the assembled antenna. Rest of the paper is organized as follows: Section II presents the design of the single element antenna. Analysis of the measured return loss, gain and radiation pattern are presented in Section III. Finally, tag measurement performance of the assembled antenna is evaluated in Section IV, which is followed by a conclusion.

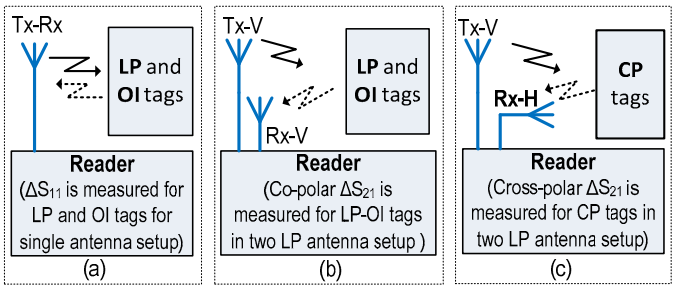


Fig. 1. Reading methods for the tags (a) single antenna and (b) two antenna setup for LP and OI tags and (c) two antenna setup for CP tags

II. Single Element LP ACMPA Design

A. LP ACMPA Design Principle

Plan view and 3-D perspective view of the single element LP ACMPA antenna are shown with design parameters in Fig. 2(a) and (b), respectively. It consists of two substrate layers with length, l and width, w. A rectangular patch with length, l_p and width, w_p is etched on the top substrate layer 2.

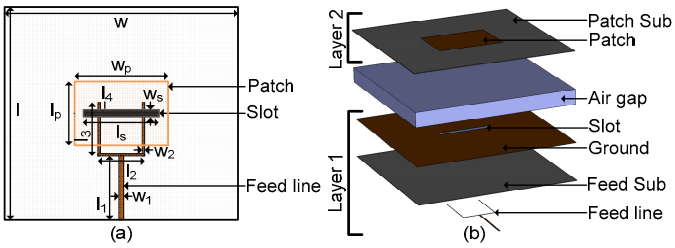


Fig. 2: Single element LP ACMPA (a) plan view and (b) expanded 3-D view

Table I: Single element LP ACMPA specifications

Antenna parameters	Targeted values for single element
Bandwidth	6.4-10.6 GHz (~50%)
Weight	100 g
Manufacturing cost	< $ 10
Beam-width	~ 65 degree
Gain	~ 7 dBi

The aperture slot with length, l_s and width, w_s is etched from the ground plane, which is on the top side of the bottom substrate layer 1, and the dual offset feed line is placed on the bottom side of the bottom substrate layer 1. The dual offset feed line starts with a 50 Ω line with length, l_1 and width, w_1, which is then divided equally in to two 100 Ω lines, each with length, $l_{f1} = 0.5l_2$ and width w_2. Then these 100Ω lines are bent at right angles to be coupled orthogonally with the aperture slot with a length, l_3. A portion of the feed with length, l_4 is extended over the slot for impedance matching, which is called the tuning stub. Besides, air gap of height, t_{fm} is kept between the two substrate layers 1 and 2 for bandwidth enhancement. Here, the center of the patch, aperture and feed line are perfectly aligned with each other to maximize the coupling.

B. Design and Simulation of a Single Element LP ACMPA

CST 2011 is used for the design, simulation and optimization of the single element LP ACMPA. First, a commercially available substrate is chosen for developing the antenna. Then, the antenna parameters and feed line widths on that substrate are calculated for the center frequency. After that, the slot is matched with the dual offset feed lines over the entire bandwidth in simulation. Next, the radiating patch is matched with the slot and feed line. Finally, the single element LP ACMPA is optimized further for fabrication. A parameterized single element LP ACMPA structure is created in CST, following the layout shown in Fig. 2. Commercially available Taconic TLX-8 substrate with permittivity, ε_r = 2.55, loss tangent, $tan\ \delta$= 0.0019, substrate thickness, h = 0.13 mm and copper thickness, t = 0.018 mm is used for the antenna design and fabrication. Dimensions of different parameters for achieving the target set in the previous section are: antenna substrate length and width, $l = w$ = 40 mm, patch length, l_p = 12 mm, patch width, w_p = 16 mm, slot length, l_s =13.0 mm, slot width, w_s =1.4 mm, 50 Ω feed line length, l_1 = 12.0 mm, separation between 100 Ω dual offset feeds, l_2 = $2\times l_{f1}$= 8.0 mm, 100Ω dual offset feed length, $l_3 = l_{f2}$ = 10.2 mm, matched stub length over the slot, $l_4 = l_{stub}$=1.5 mm, 50Ω feed width, w_1 = 0.4 mm, 100 Ω feed width, w_2 = 0.1 mm and air gap height between layer 1 and 2, t_{fm} =3.9 mm.

III. Measurement Results of the LP ACMPA

Two layers of the single element LP ACMPA, optimized on Taconic TLX-8 substrate are fabricated on separate PCBs. These two PCBs are then assembled by placing one over another with the help of a proper jig arrangement, which are made of Plexiglas. Proper air gap distance is kept between the two layers and finally, a K-type flange connector by Anritsu, having model number K-103F, is connected with the feed line through a glass bead to complete the antenna assembly. The bottom PCB layer with the feed line and ground plane can be fabricated on the RF board directly and over it, the patch layer can be placed without any complex jig requirement. This will make the reader system more compact, efficient and simpler. This will also reduce the manufacturing cost, as no extra connector and no separate feed-ground layer fabrication will be needed. Assembled antenna performances are measured using the VNA Agilent PNA E8361A. Return loss, gain, radiation pattern, main lobe, side lobe and back lobe level of the assembled antenna is presented in detail in the next section.

Fabricated two PCB layers and the assembled antenna for the single element LP ACMPA are shown in Fig. 3. Top view of the substrate layer-2 with a rectangular patch is shown in Fig. 3(a). Here, copper is removed from the bottom side of this substrate layer-2. Substrate layer-1, which contains the ground plane with a slot on the top side and a dual offset feed line on its bottom side, is shown respectively in Fig. 3(b) and (c). The patch layer-2 is placed over the ground-feed layer-1 inside the antenna jig made of Plexiglas by keeping an air gap and K type connector is added to get the complete antenna as shown in Fig. 3(d). The assembled antenna is then tested in

the VNA and the measured return loss is plotted in Fig. 4(a), along with the simulated S_{11} result. From the measured result, 10 dB return loss bandwidth is found as 6.5 to 10.6 GHz, which covers the targeted frequency band for the LP reader antenna. Measured realized gain is plotted with the simulated realized gain in Fig. 4(b) and around 7 dBi realized gain is achieved from the measured result.

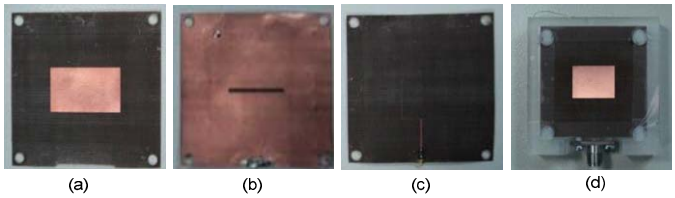


Fig. 3: Fabricated PCBs (a) patch on the top side of substrate layer 2, (b) ground with slot on the top side of substrate 1, (c) feed on the bottom side of substrate 1, (d) assembled antenna

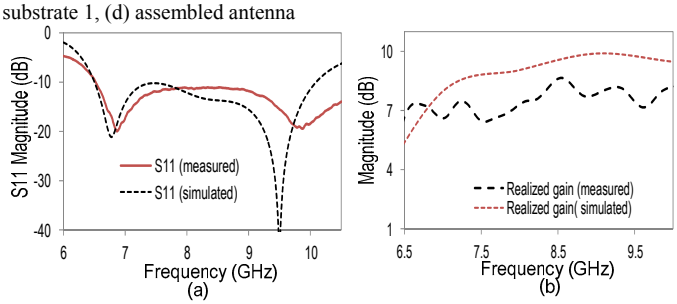


Fig. 4. Measured and simulated (a) S_{11}, and (b) realized gain

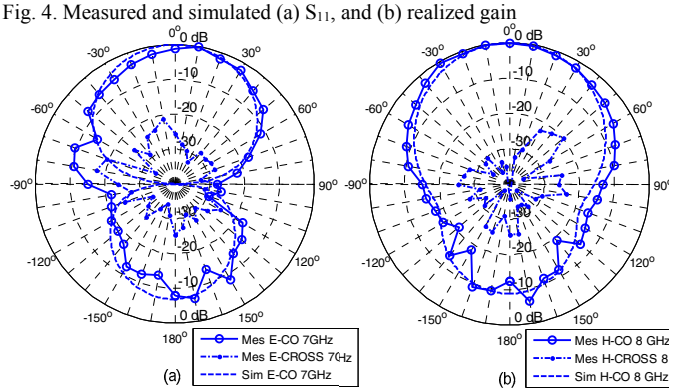


Fig. 5: Radiation pattern (a) E-field at 7 GHz, and (b) H-field at 8 GHz

Measured E-field and H-field radiation patterns are shown in Fig. 5 (a) and (b), respectively. A good matching is observed between the simulated and measured E-field and H-field values. Cross polar measured values are found around 20 dB below the co-polar values. The two layers of thin PCBs used in the single element LP ACMPA, which will not add more than 100 gm weight to the reader and their manufacturing cost in mass production will be lesser than $10. Hence, all the targeted specifications set at the beginning: bandwidth 6.5-10.6 GHz, beam-width 70°, gain ~7dBi, weigh <100g and cost<$10, has been achieved by the assembled single element LP ACMPA.

IV. Chipless Tag Reading Performance of LP ACMPA

Performance of the single element LP ACMPA as a radiating wideband, low profile, low cost antenna has been verified from the simulation and measurement results presented in the previous sections. However, its tag reading performance has not been verified yet. Hence, a detail experimental verification is presented in this section on the chipless RFID tag reading performance for the single element LP ACMPA in different measurement setups. Three different tag measurement setups, which are used in this antenna verification process, are shown in Fig. 6. OI tag reading technique using a single LP antenna is shown in Fig. 6(a). In Fig. 6(b), through reading technique is shown using two LP antennas, where the transmitter LP antenna 1 (TxV) is placed in the front side of the tag and the receiver LP antenna 2 (RxV) with the same polarity is placed at the back side of the tag. The third reading technique is shown in Fig. 6(c), where two LP antennas (antenna 1 and 2) with same polarities are placed at the same side of the tag.

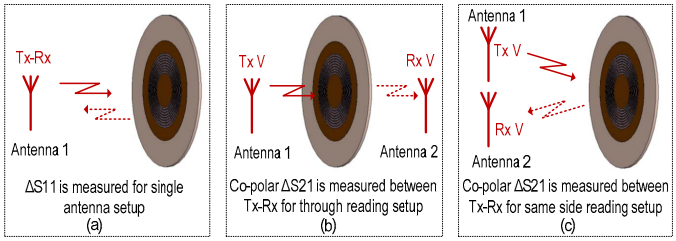


Fig. 6: Measurement setups (a) single antenna reading, (b) two antenna through reading and (c) two antenna same side reading setups.

Four 'O' slot loaded OI tags with different IDs, shown in Fig. 7, are used here for verifying the tag reading performance of the developed LP ACMPA. As the working principle of these tags are presented in detail in [10] with the proof of simulation and horn antenna measurement results, hence, in this paper, only the LP ACMPA measured results are presented. 'O' slot loaded four OI tags shown in Fig. 7(a), (b), (c) and (d) have tag IDs 'OI-11111111', 'OI-00000000', 'OI-10101010' and 'OI-01010101'.

Fig. 7. OI tags used for antenna performance verification (a) 'OI-11111111', (b) 'OI-00000000', (c) 'OI-10101010', and (d) 'OI-01010101'

A. Single Antenna Reading Using the Developed LP ACMPA

In this section, tag measurement performance of the single element LP ACMPA antenna is evaluated by measuring the tags using the single antenna measurement setup shown in Fig. 6(a). First, the single element LP ACMPA is connected with Agilent PNA8361A VNA and the transmission power is set to 0 dBm. Then, no tag is placed in front of the antenna and data ($S_{11(no\ tagV)}$) are saved in the memory. After that, the tags shown in Fig. 7 are placed one by one, at a distance of 5 cm away from the antenna and in a new window, difference between the return loss values of tag reading ($S_{11\ (tagV)}$) with the memory values with no tag ($S_{11(no\ tagV)}$) are plotted, where the plotted values are $\Delta S_{11(V)} = S_{11\ (tagV)} - S_{11(no\ tagV)}$.

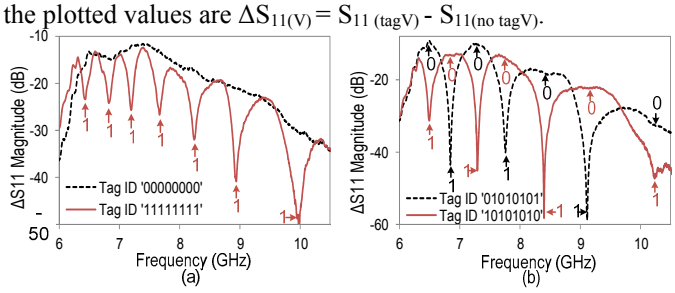


Fig. 8. 'O' slotted OI tags shown in Fig. 7 (a), (b), (c) and (d) measured changes in magnitude results for (a) 'OI-11111111' and 'OI-00000000', (b) 'OI-10101010' and 'OI-01010101'

Measured changes in magnitude results for the tags shown in Fig. 7 (a) and (b) with IDs 'OI-11111111' and 'OI-000000

00' are shown in Fig. 8(a). Here, seven notches are generated for the first tag with all 1's and no notches are generated for the second tag with all 0's in their IDs. Similarly, notches are found at the corresponding alternate positions of 1's and no notches are found at the corresponding alternate positions of 0's in the measured changes in the magnitude curves shown in Fig. 8(b) for the tags in Fig. 7 (c) and (d) with IDs 'OI-10101010' and 'OI-01010101'. Here, in all cases the, only the eighth notch is not captured properly by the single element LP ACMPA, as it falls a way outside its operating frequency bandwidth. However, in all other notch positions, significant notch depths (>10 dB) are observed in the measured results using this single element LP ACMPA, which confirms the correct detection of the encoded bits for this 'O' slot loaded OI tags, as well as the usability of the antenna in tag reader.

B. Through Reading Using Two LP ACMPAs

In this section, following the through reading measurement setup shown in Fig. 6(b), the 'O' slot loaded tags of Fig. 7 are measured. Similar procedure is followed here to save the tag measurement data, which is used for the single antenna setup. First, the two LP ACMPAs (antenna 1 and 2) are placed face to face at a distance of 10 cm with 0 dBm transmitted power from the VNA and co-polar transmission data between the two antennas without placing any tag ($S_{21(notagV)}$) is saved in the memory. After that, tags are placed in between the antennas one by one and the difference between the transmission coefficient values of tag reading ($S_{21\ (tagV)}$) with the memory values with no tag ($S_{21(notagV)}$) are plotted, where the plotted values are $\Delta S_{21(V)} = S_{21(tagV)} - S_{21(notagV)}$. All the measurement results for the 'O' slotted OI tags correspond well with their anticipated values using the single element LP ACMPA antenna in through reading setup as well.

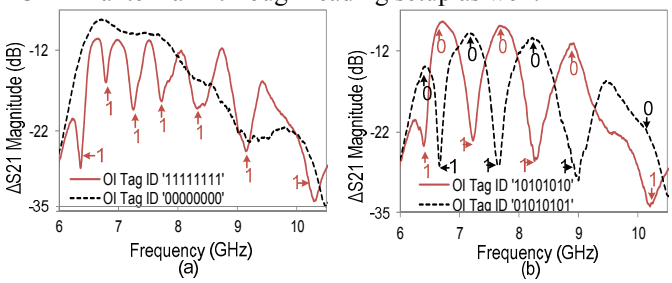


Fig. 9. 'O' slotted OI tags of Fig. 7(a), (b), (c) and (d) measurement results (a) 'OI-11111111' and 'OI-00000000', (b) 'OI-10101010' and 'OI-01010101'

C. Same Side Reading Using Two LP ACMPAs

The measurement setup shown in Fig. 6(c) is followed here to measure the 'O' slotted OI tags of Fig. 7. First, two single element LP ACMPAs (antenna 1 and 2) are placed side by side, targeting an imaginary tag positioning point at 5 cm away from them. Then, the co-polar transmission coefficient between the two antennas are saved in the memory without placing any tag ($S_{21(no\ tagV)}$). After that, tags are placed in front of the antennas one by one and the difference between the tag reading values ($S_{21(tagV)}$) with the memory values are plotted, where the plotted values are $\Delta S_{21(V)} = S_{21(tagV)} - S_{21(no\ tagV)}$.

All the measurement results for the 'O' slot loaded OI tags correspond well with their anticipated values in the three different measurement setups. Therefore, it can be inferred that, the developed low-profile, low-cost, lightweight, wideband single element LP ACMPA is capable of reading OI and LP tags and hence, can be used in commercial readers.

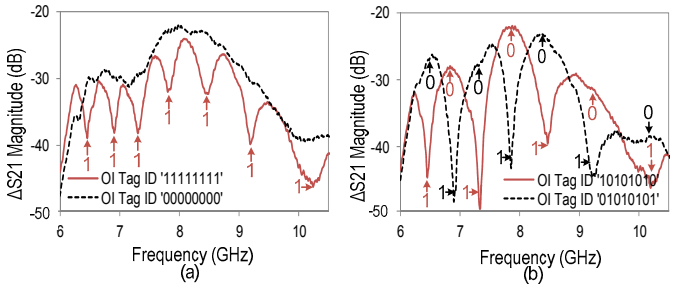


Fig. 10. 'O' slotted OI tags of Fig. 7(a), (b), (c) and (d) measurement results (a)'OI-11111111' and 'OI-00000000', (b) 'OI-10101010' and 'OI-01010101'

V. Conclusion

A commercially suitable single element LP ACMPA is presented in this paper for reading chipless RFID tags in microwave frequency band. First, a thorough review is done on existing low profile wideband antennas. From the literature analysis, ACMPA design technique is found as the best match for developing the reader antennas. Next, antenna specifications are listed for developing a single element LP ACMPA (bandwidth 6.5-10.6 GHz, beam-width 70°, gain ~7dBi, weight < 100g and cost < $10). Then, a single element LP ACMPA with dual offset 100 Ω feedline is designed and optimized to achieve the targeted values, which is fabricated and assembled next. After that, the ACMPA is measured in VNA, which fulfill the set targets. Next, the tag reading performances of the antenna from a distance of 5 cm are evaluated in three different measurement setups through measurement of 'O' slot loaded OI tags. These tag measurement results have confirmed the usability of the developed antennas. This low-profile, low-cost, lightweight, wideband, wide beam, high gain LP ACMPA can be used commercially in the frequency domain-based chipless RFID tag reader for reading LP, OI, and CP tags.

References

[1] U. Kaiser and W. Steinhagen, "A low-power transponder IC for high-performance identification systems," *IEEE Proc. of Solid-State Ckt,* vol.30, no.3, pp:306-10, March 1995.

[2] S. Botao, *et al.*, "An ultra-low-cost RFID tag with 1.67 Gbps data rate by ink-jet printing on paper," in *IEEE Asian Solid State Ckt. Conf.*, 2010.

[3] M. A. Islam, *et al.*, "A novel compact chipless RFID tag and near-field reader," *the Asia-Pacific Microwave Conference,* 2011, pp. 1518-1521,.

[4] A. Vena, *et al.*, "Novel Compact RFID Chipless Tag," in *Progress In Electromagnetics Research Symp, Morocco*, 2011, pp. 1062-1066.

[5] M. A. Islam, *et al.*, "Compact printable orientation independent chipless RFID tag," *Progress In Electromag.Research C,* vol. 33, pp. 55-66, 2012

[6] A. Vena, *et al.*, "High-Capacity Chipless RFID Tag Insensitive to the Polarization," *IEEE Trans on Ant & Prop,* vol. 60, pp. 4509-4515, 2012.

[7] S. Preradovic and N. C. Karmakar, "Design of fully printable planar chipless RFID transponder with 35-bit data capacity," *European Microwave Conference (EuMC),* pp. 13-16, 2009.

[8] P. Baldonero, *et al.*, "UWB Double polarised phased array," in *European Conference on Antennas and Propagation (EuCAP)*, 2009, pp. 556-560.

[9] R. Olsson, *et al.*, "Measurements of a 1 to 13 GHz model of a dual polarized low-profile log-periodic feed for US-SKA," in *IEEE Antennas and Propagation Society Symposium*, 2005, pp. 700-703 vol. 2B.

[10] M. A. Islam, *et al.*, "Orientation independent compact chipless RFID tag," presented at the IEEE International Conference on RFID-Technologies and Applications (RFID-TA), pp.137-141, 2012.

Discrimination of Focal and Non-Focal EEG Signals using Entropy-based Features in EEMD and CEEMDAN Domains

Anindya Bijoy Das*and Mohammed Imamul Hassan Bhuiyan
Department of Electrical and Electronic Engineering(Bangladesh University of Engineering and Technology, Bangladesh)
Email: *anindya149@yahoo.com

***Abstract*—In this paper, a comprehensive analysis for the discrimination of the focal and non-focal electroencephalography (EEG) signals is carried out in the ensemble empirical mode decomposition (EEMD) and complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) domains. A number of spectral entropy-based features such as the Shannon entropy, log-energy entropy, Renyi entropy and Tsallis entropy are calculated in the EEMD and CEEMDAN domain of the EEG signals. In lieu of using the signals from the EEG channels, the differences between two adjacent EEG channels are used due to its robustness to noise and interference. The ability of the entropy-based features in separating the focal and non-focal EEG signals is explored utilizing the one-way ANOVA analysis and the box-whisker plots. The results reveal that among the entropy measures computed in the EEMD and CEEMDAN domains, the quadratic Renyi entropy and log-energy entropy measures are most promising in discriminating the focal and non-focal EEG signals. The analysis may encourage the researchers to develop improved algorithms to classify these signals and would be helpful in locating the epileptogenic zones.**

Keywords*— *Electroencephalogram(EEG); Entropy; Ensemble Empirical Mode Decomposition(EEMD); Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN); Analysis of Variance (ANOVA)

I. Introduction

Epilepsy is a common neurological disorder that affects about 50 million people worldwide [1]. It is characterized by the recurrent seizures and has debilitating effects on epileptic patients, and sometimes may lead to death. An important problem in epilepsy treatment is the limited response to anti-epileptic drugs in case of a significant portion of the patients. For example, mesial temporal lobe epilepsy(MTLE), which is the most common type of epilepsy, is often refractory to the "conventional pharmacological treatment" [1]. These types of focal epilepsies are characterized with unprovoked and recurrent seizures from a certain section of the brain (in MTLE, hippocampus and its surrounding), known as epileptogenic focus [2]. In such case, often the treatment policy is resective surgery to remove the epileptogenic sites [3]. However, these types of surgeries are quite risky in case of multi-focal epilepsy, and may have overlapping of those sites with other eloquent areas of the brain which are responsible for primary motor or vision. Thus, it is extremely important to identify the epileptogenic zones prior to the surgery.

Identification of the epileptogenic sites can be carried out by employing conventional diagnostic systems such as MRI. However, these systems provide limited temporal information, leading to poor accuracy [1]. An effective approach to locate the epileptogenic zones is to classify focal and non-focal EEG data, where the EEG signals obtained from the epileptogenic foci are called focal EEG signals, and those originating from other areas are named as non-focal EEG signals [4]. Recently, empirical mode decomposition (EMD) has emerged as a highly promising tool for analyzing EEG signals [5]- [11], including the classification of focal and non-focal EEG signals. However, the EMD suffers from the mode-mixing problem. Basically, dissimilar oscillations may appear in the same mode or very similar oscillations may be found in different modes. This complicates feature extraction and subsequent classification processes. The ensemble EMD (EEMD) eradicates this problem by performing the signal decomposition over an ensemble of signals and Gaussian noise [12]. However, it has residual noise and can lead to different number of modes [13]. The complete ensemble EMD with adaptive Noise (CEEMDAN) gives an exact reconstruction of a signal with a superior spectral separation of the mode functions [13]. Although the EEMD and CEEMDAN have been used recently for the analysis of bomedical signals, to the best of our knowledge, they have not yet been employed for the analysis of focal and non-focal EEG signals. Our objective is to explore the potential of analyzing focal and non-focal EEG signals in EEMD and CEEMDAN domains using entropy based parameters. Various entropy measures, such as Shannon entropy, log-energy entropy, Renyi entropy and Tsallis entropy, are used due to their ability to characterize the complexity and measure the variability of the EEG signals [5], [8], [14]. Extensive analysis is carried out using a publicly available, comprehensive and benchmark database and employing one way ANOVA analysis.

II. Materials

The focal and non-focal EEG signals used in this study are obtained from a publicly available EEG database, called Bern-Barcelona EEG database [15]. It has been used in a number of related works [5], [8], [14] and is comprehensive enough to ensure statistical significance. These signals are collected from five epilepsy patients and are sampled at 512 Hz. Each

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

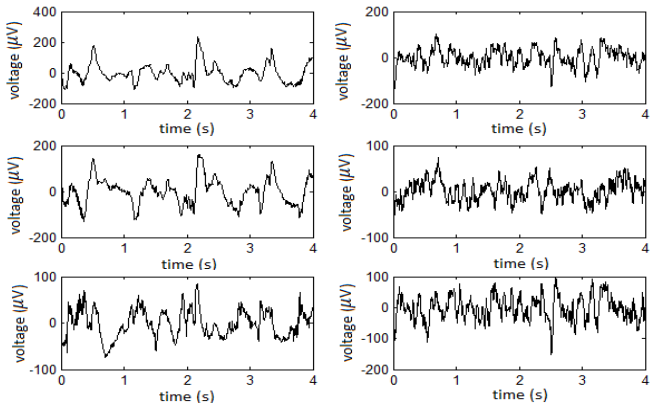


Fig. 1. Sample focal(left) and non-focal(right) EEG signals for channels x, y and $x-y$ in the first, second and third row respectively

of the focal and non-focal EEG records contain 3750 pairs of simultaneously recorded signals, denoted as x and y, where the former signals are collected randomly from any of the five patients from all those channels that detects the first ictal activity and the latter corresponds to those from one of the neighboring channels. All other channels are termed as non-focal EEG channels and the database contains those recordings for two adjacent channels also, named as x and y. In this paper, first 50 sets of signals for both of focal and non-focal EEG, named Data_F_50 and Data_N_50 respectively, are utilized for the analysis in the EEMD and CEEMDAN domains. Figure 1 shows four seconds of sample focal and non-focal EEG signals. In the left panel, focal EEG signals x, y and their differences $x-y$ are shown in the first, second and third rows, respectively. The right panel shows the same for the non-focal EEG signals. In this study, the difference signals, that is $x-y$, are used for analysis due to the robustness of EEG signals at adjacent channels to noise and interference (see [8], [16]).

III. EEMD AND CEEMDAN: A BRIEF REVIEW

The Empirical Mode Decomposition(EMD) provides amplitude and frequency modulated oscillatory patterns known as intrinsic mode functions (IMF), which are derived using the basis obtained from the signal subjected to EMD. As discussed before, a major problem with EMD is mode-mixing; recently, variants of EMD such as EEMD and CEEMDAN have been proposed to eradicate this problem. In the following, a brief review of these variants is provided.

A. Ensemble Empirical Mode Decomposition (EEMD)

The EEMD provides the EMD of ensemble average of Gaussian noise-assisted data. A low-level Gaussian noise is mixed with the original signal which is then decomposed by EMD [10]. The steps of EEMD are as follows:

- First, from the actual signal $x(n)$, $x^i(n)$ is calculated as $x^i(n) = x(n) + w^i(n)$ where $w^i(n)$ indicates the realizations of white Gaussian noise (while $i = 1, 2, 3, ..., R$) with zero mean and unit variance, and R is the number of realizations of the noise.

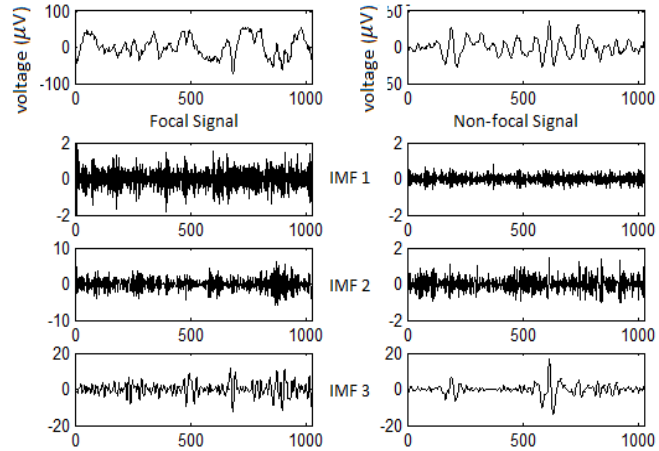


Fig. 2. Focal(left) and non-focal(right) EEG signals along with their first three mode functions in the EEMD domain in the respective column

- Next, each of the signals of $x^i(n)$ is subjected to EMD, thus giving the IMFs, denoted as $IMF_k^i(n)$, where $k = 1, 2, 3, ..., L$ represent the different modes.
- Finally, the mode functions in EEMD domain, $IMF_k(n)$, are calculated by averaging the $IMF_k^i(n)$ as

$$IMF_k(n) = \frac{1}{L}\sum_{i=1}^{R} IMF_k^i(n) \tag{1}$$

In Fig. 2, the first three mode functions of the focal and non-focal EEG signals in EEMD domain have been shown for the signal $x-y$ in both cases.

B. Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN)

Although EEMD successfully minimizes the mode-mixing problem, due to the use of different realizations, it can give different number of modes and residual noise. The CEEMDAN gives an exact reconstruction of the data while computationally less expensive as compared to EEMD [13]. In CEEMDAN, an operator M_j is defined that produces the jth mode of EMD. Unlike EEMD, the added white Gaussian noise has a certain standard deviation denoted as σ_0 . Thus, the steps of the CEEMDAN algorithm is given as below:

- First, $x^i(n)$ is obtained as $x^i(n) = x(n) + \sigma_0 w^i(n)$ and transformed in EMD to obtain the first mode functions.
- Next, the first mode IMF_1 and the first residue $r_1(n)$ in CEEMDAN are calculated as

$$IMF_1(n) = \frac{1}{L}\sum_{i=1}^{R} IMF_1^i(n) \tag{2}$$

$$r_1(n) = x(n) - IMF_1(n) \tag{3}$$

- If σ_1 is the standard deviation of the white Gaussian noise at this stage, then IMF_2 is determined as

$$IMF_2(n) = \frac{1}{L}\sum_{i=1}^{R} M_1(r_1(n) + \sigma_1 M_1(w^i(n))) \tag{4}$$

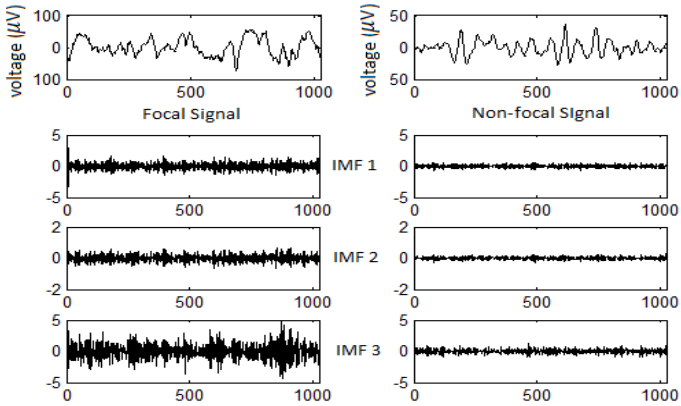


Fig. 3. Focal(left) and non-focal(right) EEG signals along with their first three mode functions in the CEEMDAN domain in the respective column

- Subsequently, the kth residue and the $(k+1)^{th}$ mode functions are obtained as

$$r_k(n) = r_{k-1}(n) - IMF_k(n) \tag{5}$$

$$IMF_{k+1}(n) = \frac{1}{L}\sum_{i=1}^{R} M_1(r_k(n) + \sigma_k M_k(w^i(n))) \tag{6}$$

- The previous step is repeated until the residue becomes a monotonic function such that further extraction of an IMF is impossible. If k is the total number of modes and $r_k(n)$ the final residue, the input x(n) can be reconstructed from all the IMFs as

$$x(n) = \sum_{i=1}^{k} IMF_k(n) + r_k(n) \tag{7}$$

In Fig. 3, the first three mode functions of the focal and non-focal EEG signals in CEEMDAN domain have been shown for the signal $x - y$ in both cases.

IV. Analysis of EEG Signals in EEMD and CEEMDAN Domains

Entropy is a widely used parameter which measures the randomness a chaotic system and the complexity of the non-linear signals such as the EEG signals [17]- [18]. Among the various types of entropy, spectral entropy-based parameters (e.g. Shannon entropy or log-energy entropy) have been successfully utilized to measure the irregularity and the amount of EEG disorders [8]. In case of the non-linear signals such as EEG signals, Shannon entropy measures the average information contained in the probability distribution function (pdf) [5]. On the other hand, log-energy entropy can describe the electro-physiological behavior of various types of EEG signals [8]. Renyi Entropy has been reported to have the capability to derive the continuous family of mutual information measures [9]. Tsallis wavelet entropy has the ability to extract some improved features by reducing the negative effects of wavelet aliasing [14]. For a given data-set, with

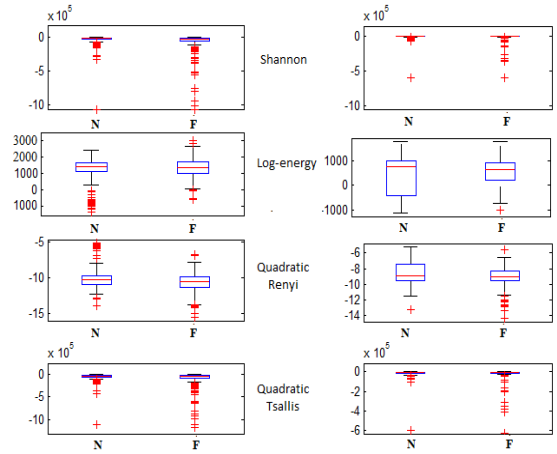


Fig. 4. Sample box plots for IMF 3 for EEG signal $x - y$ in for Shannon entropy, log-energy entropy, quadratic Renyi entropy and quadratic Tsallis entropy in EEMD domain(left) and CEEMDAN domain(right)

length N where the probability of occurrence of the kth component is p_k, , the corresponding Shannon-entropy(H_{SE}), log-energy entropy(H_{LE}), Renyi entropy(H_{RE}) and Tsallis entropy(H_{TE}) are expressed as

$$H_{SE} = -\sum_{k=1}^{N} p_i^2 \times log_2(p_k^2) \tag{8}$$

$$H_{LE} = \sum_{k=1}^{N} log_2(p_k^2) \tag{9}$$

$$H_{RE} = \frac{1}{1-\alpha} log \sum_{k=1}^{N} p_k^{\alpha} \tag{10}$$

$$H_{TE} = \frac{1}{\alpha - 1}(1 - \sum_{k=1}^{N} p_k^{\alpha}) \tag{11}$$

The Renyi entropy (H_{RE}) and Tsallis entropy(H_{TE}) are of order α where $\alpha > 0$ and $\alpha \neq 1$. If α is equal to 2, then the measurements equally emphasize the sub-Gaussian and the super-Gaussian components [8] and called as quadratic Renyi entropy and quadratic Tsallis entropy which are given by

$$H_{RE} = -log \sum_{k=1}^{N} (p_k)^2 \tag{12}$$

$$H_{TE} = 1 - \sum_{k=1}^{N} p_k^2 \tag{13}$$

These entropy-based parameters are calculated for the difference EEG signal $x - y$ in the EEMD and CEEMDAN domains for both focal and non-focal EEG signals. Next, one-way ANOVA analysis is carried out and the corresponding p-values are shown in tables I and II. Sample box plots for IMF 3 for various types of entropy-based parameters in EEMD and CEEMDAN domains are shown in fig. 4.

TABLE I
P-VALUES FOR ENTROPY-BASED FEATURES IN EEMD DOMAIN

IMFs	Shannon	Log-energy	Quad. Renyi	Quad. Tsallis
1	0.1013	0.0742	2.1269e-06	0.0998
2	0.0044	4.8437e-04	3.2639e-08	0.0062
3	0.0457	0.0085	1.5371e-07	0.0401
4	0.0776	0.0035	3.7581e-06	0.0568
5	0.0136	0.2831	2.7069e-04	0.0100
6	0.0146	0.0019	4.3050e-06	0.0098
7	0.0160	8.0231e-16	5.0625e-17	0.0092

TABLE II
P-VALUES FOR ENTROPY-BASED FEATURES IN CEEMDAN DOMAIN

IMFs	Shannon	Log-energy	Quad. Renyi	Quad. Tsallis
1	0.1084	0.0959	2.2076e-06	0.1063
2	0.0393	0.2990	8.8121e-07	0.0378
3	0.0306	0.3688	4.3565e-04	0.0249
4	0.0072	0.1546	8.4354e-06	0.0041
5	0.0321	1.0262e-08	1.0478e-10	0.0224
6	0.0813	1.8202e-14	1.2049e-14	0.0594
7	0.0618	2.1351e-11	1.6058e-11	0.0426

The following observations can be made based on the the p-values listed in tables I and II and the box plots:

- The p-values obtained are small enough to indicate the discrimination capability of the entropy-based features in EEMD and CEEMDAN domain.
- The p-values for quadratic Renyi entropy are significantly smaller than the those for the others in both EEMD and CEEMDAN domains which indicates its superior capability to distinguish between the focal and non-focal EEG signals.
- For some particular mode functions in EEMD and CEEMDAN domains, the p-values for log-energy entropy are also very small as compared to others, for example in the IMFs 5, 6 and 7 in CEEMDAN domain.
- The box plots further confirm that Renyi entropy and log-energy entropy perform a better discrimination than Shannon entropy and quadratic Tsallis entropy.

V. CONCLUSION

In this paper, a statistical analysis of the focal and non-focal EEG signals has been carried out to explore their discrimination in the EEMD and CEEMDAN domains. Various types of entropy measurs such as the Shannon entropy, log-energy entropy, Renyi entropy and Tsallis entropy have been used for their previous success to discriminate the EEG data. These parameters are calculated from the IMFs of the EEG signals in the EEMD and CEEMDAN domains. Subsequently, their capability in discriminating the focal and non-focal EEG signals is explored utilizing the one-way ANOVA analysis and the corresponding p-values. The small p-values obtained for the entropy-based features indicate their good discrimination capability in the EEMD and CEEMDAN domains. Among them, the p-values for quadratic Renyi entropy and log-energy entropy are significantly smaller than those for the Shannon entropy and quadratic Tsallis entropy which indicates their superior potential in classifying the EEG signals. The future goal of this study is to employ suitable classifiers, like support vector machine (SVM) or artificial neural network (ANN) or k-nearest neighbor (KNN), to classify the focal and non-focal EEG signals where these entropy-based measures could be utilized as the classification features. Thus, this analysis can aid the researchers to develop an improved method that would be helpful to identify the epileptogenic zones.

REFERENCES

[1] M. T. Salam, J. L. P. Velazquez and R. Genov,"Seizure Suppression Efficacy of Closed-Loop Versus Open-Loop Deep Brain Stimulation in a Rodent Model of Epilepsy," *IEEE Transactions on Neural Systems and Rehabilitation Engineering* , vol. 24, November, 2015.

[2] Eva Verhellen and Paul Boon, Chia-Hsiang Yang and Herming Chiueh,"EEG source localization of the epileptogenic focus in patients with refractory temporal lobe epilepsy, dipole modelling revisited," *Acta Neurologica Belgica*, vol. 107, pp. 71-77, September, 2007.

[3] S. Liang, Y. Chen, Y. Wang, P. Chen, C. Yang and H. Chiueh,"A hierarchical approach for online temporal lobe seizure detection in long-term intracranial EEG recordings," *Journal of Neural Engg.*, May, 2013.

[4] R. G. Andrzejak, K. Schindler and C. Rummel, "Nonrandomness, nonlinear dependence, and nonstationarity of electroencephalographic recordings from epilepsy patients", Phy. Rev. E, 86, 046206, 2012.

[5] R. Sharma, R. B. Pachori and U. R. Acharya, "Application of Entropy Measures on Intrinsic Mode Functions for the Automated Identification of Focal Electroencephalogram Signals", *Entropy*, vol. 2, no. 17, 2015.

[6] S M S. Alam and M. I. H. Bhuiyan, "Detection of Seizure and Epilepsy using Higherorder Statistics in the EMD Domain," *IEEE Trans. on Information Technology in BioMedicine*, vol. 17, March, 2013.

[7] V. Bajaj, R.B. Pachori, "EEG Signal Classification Using Empirical Mode Decomposition and Support Vector Machine", Proc. of *Int. Conf. on Soft Computing for Problem Solving*, December, 2011.

[8] A. B. Das and M. I. H. Bhuiyan, "Discrimination and classification of focal and non-focal EEG signals using entropy-based features in the EMD-DWT domain", *Elsevier Journal On Biomedical Signal Processing and Control*, vol. 29, pp. 11-21, 2016.

[9] Sharma, R.; Pachori, R.B.; Gautam, S., "Empirical mode decomposition based classification of focal and non-focal EEG signals", in Proc. of *Int. Conf. on Medical Biometrics*, pp. 135-140, China, 2014.

[10] A. R. Hassan and M. I. H. Bhuiyan, "Computer-aided sleep staging using Complete Ensemble Empirical Mode Decomposition with Adaptive Noise and bootstrap aggregating", in Elsevier Journal on *Biomedical Signal Processing and Control*, vol. 24, February, 2016.

[11] L. Orosco, E. Laciar, A. G. Correa, A. Torres and J. P. Graffigna,"An Epileptic Seizures Detection Algorithm based on the Empirical Mode Decomposition of EEG ", Proc. of 31st IEEE Conference on *Engineering in Medicine and Biology Society (EMBS)*, Minnesota, USA, 2009.

[12] W. Zhaohua and Norden Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method", *Advances in Adaptive Data Analysis* , vol. 1, January, 2009.

[13] M. Torres, M. Colominas, G. Schlotthauer, P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise", Proc. of *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4144-4147, 2011.

[14] R. Sharma, R. B. Pachori and U. R. Acharya, "An Integrated Index for the Identification of Focal Electroencephalogram Signals Using Discrete Wavelet Transform and Entropy Measures", *Entropy*, vol. 8, 2015.

[15] Bern-Barcelona database: http://ntsa.upf.edu/downloads/andrzejak-rg-schindler-k-rummel-c-2012-nonrandomness-nonlinear-dependence-and

[16] R. Yadav, M. N. S. Swamy and R. Agarwal "Rapid Identification of Epileptogenic Sites in the Intra-cranial EEG", proc. of 33rd Annual International Conference of *IEEE EMBS*, September, 2011.

[17] Guohun Zhu, Yan Li, Peng Paul Wen, Shuaifang Wang, and Min Xi, "Epileptogenic focus detection in intracranial EEG based on delay permutation entropy", Proc. of *AIP International Conf.*, vol. 31, 2013.

[18] C. C. Jouny and G. K. Bergey, "Characterization of early partial seizure onset: Frequency, complexity and entropy", Elsevier Journal on *Clinical Neuro-physiology*, vol. 123, pp. 658-669, 2012.

MX-MAC: A Multichannel Based Low Latency Asynchronous MAC Protocol For Wireless Sensor Networks

Md Mustafa Kamal,[1,*] Shafika Showkat Moni,[1,2] and Mohammad Shah Alam[1]

[1]IICT, Bangladesh University of Engineering and Technology

[2]CSE, Rajshahi University of Engineering and Technology

[*] mustafacomp@gmail.com

Abstract—**In this paper we propose a Multichannel based X-MAC protocol (MX-MAC) with the objective of eliminating hidden node problem of existing asynchronous X-MAC protocol, as well as reducing average packet delay. To address the hidden terminal problems, multi channel enables a receiver to send a preamble with next receiver ID on the control channel to notify its neighbors that an ongoing reception is in progress and alert intended next hop neighbor to wake up on time as well as to forward received data to next hop without sending any additional preamble. As a result, packet collisions are reduced with an increase of network throughput. We compare our algorithm with X-MAC and X-MAC/CA. Simulation results show that the proposed technique can significantly improve the throughput as well as can reduce average packet delay.**

Index Terms— **MAC; X-MAC; Multichannel MAC; Markov Model; Analytical Modeling;**

I. Introduction

In Wireless Sensor Networks (WSN), a critical resource is energy. Energy saving is an essential aspect to extend life time of the WSN [1]. The major Advantages of asynchronous MAC protocols for WSNs like B-MAC and WiseMAC are that the sender and receiver can be completely decoupled in their duty cycles. Despite all the benefits provided by the asynchronous MAC protocols, the long preamble in Low Power Listening (LPL) illustrates several disadvantages: it is subject to overhearing that causes excess energy exhaustion at non-target receivers, and it introduces extra waiting at each hop[2]. In order to alleviate these problems, a new approach has been proposed to LPL called X-MAC, which employs a short preamble instead of long preamble, to reduce energy consumption and shorten latency. The hidden terminal problem, one of the basic challenges in wireless MAC protocols, has still unresolved in X-MAC protocol.

This research work aims to eliminate the hidden node problem and minimize the latency of existing asynchronous X-MAC protocol for WSNs. The main contributions of this work are summarized as follows: (a) MX-MAC initiates multichannel concept over X-MAC protocol containing control channel and data channel, thereby avoiding hidden terminal problem. (b) MX-MAC introduces receiver preamble on the control channel to inform neighbors that an ongoing transmission is in progress for a multi-hops transmission which reduces the extra energy. (c) We developed an algorithm based on X-MAC and X-MAC/CA, which is energy-intensive in wireless sensor network. (d) Furthermore, we also developed Markov chain model of our proposed protocol MX-MAC to investigate the performance of proposed MAC protocol. (e) Finally, our proposed model has been simulated and simulation results has been compared with the simulation results of X-MAC and X-MAC/CA. From simulation it is evident that in our proposed model energy savings has been increased and latency has been decreased in contrast to X-MAC and X-MAC/CA.

The rest of the article is prepared as follows: Section II discusses related works. Proposed MX-MAC protocol is introduced in Section III. Section IV depicts the Markov model of proposed Multichannel X-MAC Protocol. Section V provides throughput analysis of MX-MAC protocol. Section VI describes energy consumption of proposed protocol. Section VII presents latency of proposed Multichannel protocol. In Section VIII performance of proposed MX-MAC protocol is evaluated in terms of latency, energy consumption and throughput. We conclude with final remarks in Section IX.

II. Related Works

X-MAC is Contention-based protocol, which solve the overhearing problem by dividing the one long preamble into a series of short preamble packets, each containing the ID of the target node, as indicated in Fig 1. When a node wakes up and receives a short preamble packet, it looks at the target node ID that is included in the packet. If the node is not the intended recipient, the node returns to sleep immediately and continues its duty cycling as if the medium had been idle. If the node is the intended recipient, it remains awake for the subsequent data packet. As seen in the figure, a node can quickly return to sleep, thus avoiding the overhearing problem [3].

X-MAC/CA protocol is the same as the X-MAC. In X-MAC/CA protocol when a sensor node wakes up and has some data to send in its queue, the sensor node performs preamble backoff in which it randomly draws an integer number i from a range (1, Wm) and waits for i backoff slots without polling the channel state unlike IEEE 802.11 [4].

III. Overview Of Mx-Mac

To achieve the higher throughput we design a Multichannel asynchronous X-MAC (MX-MAC) protocol for Wireless Sensor Network based on X-MAC protocol.

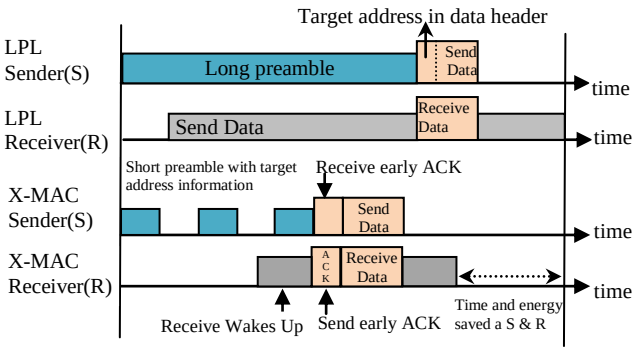


Fig. 1 A Visual Representation of X-MAC [3].

MX-MAC protocol uses a series of short preamble like X-MAC. In addition, our proposed protocol uses two separate channels (control and data channel) to avoid hidden terminal problem and to manage the limited resources efficiently in terms of node energy consumption and packet delay. The main idea is to eliminate the hidden node problem of X-MAC scheme by transmitting preamble packets in the control channel by the receiver node whenever it receives data from the sender node.

In MX-MAC protocol, when a node wants to send data, it first performs Clear Channel Assessment (CCA) in the control channel as well as in the data channel to keep away from any possible collision with the ongoing transmission. If both data and control channel are idle then sender send short preamble by control channel with pauses into the series of short preamble packets, creating a strobed preamble, which enables the target receiver to shorten the strobed preamble via an early acknowledgement, thereby achieving additional energy saving at both the sender and receiver, as well as a reduction in per-hop latency. If sender received early acknowledges (ACK), it instantly ends sending short preamble and switch to data channel for sending data packets. If the destination node gets any preamble packet with its own address in the 'Target address' field, it sends an ACK packet by control channel and turns on its data channel to receive data from sender node. During the whole receiving period the receiver switched on both of its data and control channel and receive data by data channel and sends preamble packet with its next hop receiver address by control channel at a time. In this way the receiver preamble can restricts all the neighbor nodes of the receiver to transmit any data packets to receiver. In this way, receiver can easily eliminate the hidden terminal problem. When a neighbor node of receiver awakes, it first scans the medium. If it detects any preamble then the node decides whether it is the target node or not. If it is target node then it sends early ACK and immediately goes to sleep. In this way, neighbor node of the receiver node can easily sense from the receiver's preamble that a data transmission is going on with the receiver end, even the nodes are not within the range of the sender node.

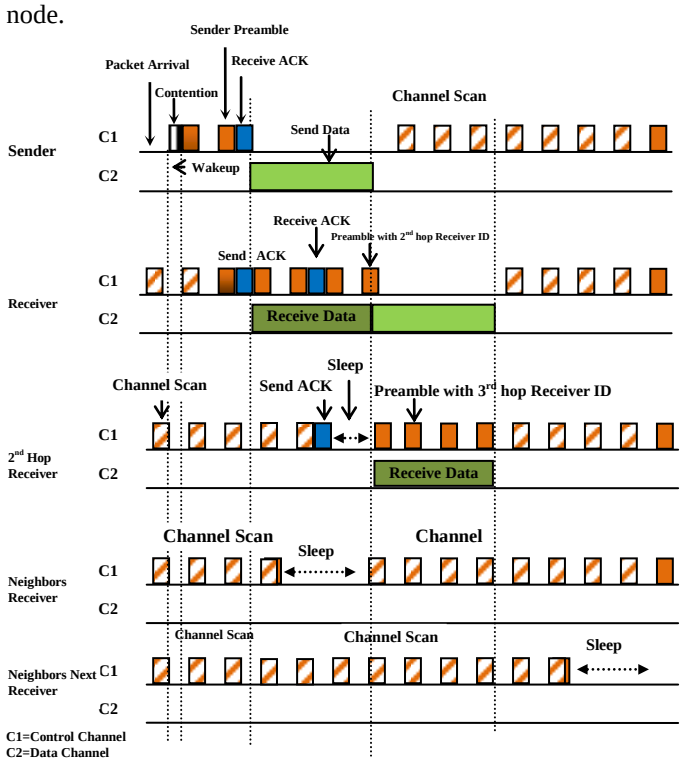


Fig. 2 A Visual Representation of MX-MAC.

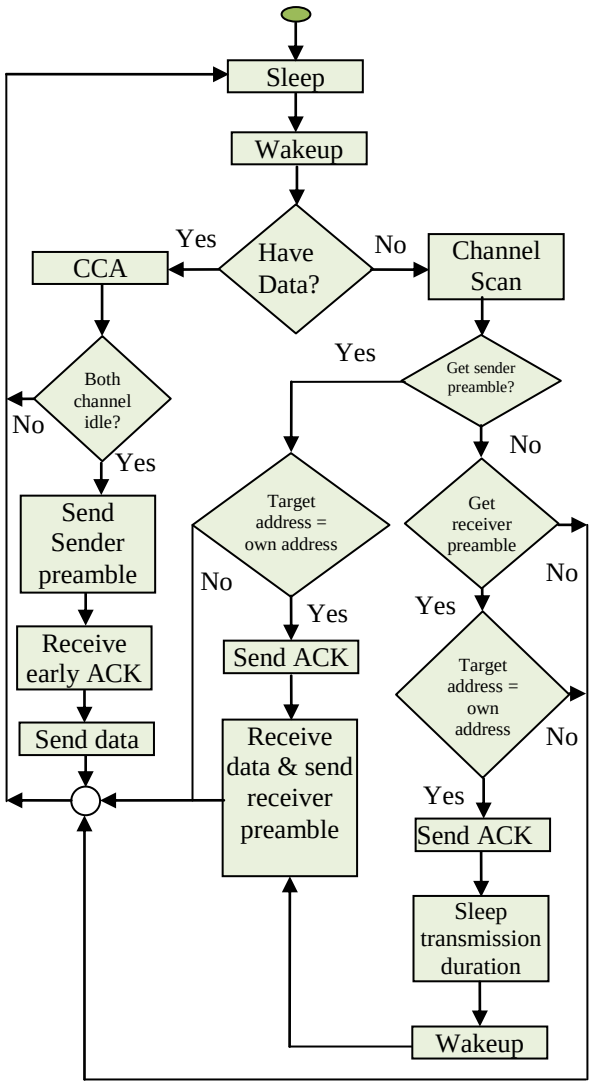


Fig. 3 Flow chart of MX-MAC protocol.

IV. Markov Model Of Mx-Mac

The Markov model has Q+1 states, first digit of the state indicates channels (c=control, d=data) and second digit from left to right corresponding to 0 to Q packets in the queue, Q indicates the queue full. In Fig. 4 the probability P_{ij} (i<j) indicates that j new data packets arrive at a sender node i. Probability P_{ij} (i>j) defines the probability of the transmission of only one data packet. Probability P_{ij} (i=j) implies non-decreasing queue. R_{ij} (i=j) means i new data packets arrive at sender node and j data packets is being received and forward to next hop without sending any additional preamble. R_{ij} (i<j) reveals i+1 new data packets arrived at receiver node. R_{ij} (i>j) exhibits i-1 data packets transmitted to final destination (sink). S_{ij} (i<=j) defines i data packets transmitted to sink and j new data packets arrive at the any sensor node.

V. Throughput Analysis Of Mx-Mac

Throughput is defined as the amount of data successfully delivered within a unit time [5]. For MX-MAC, the throughput can be calculated within a cycle time, since nodes work in a duty cycled fashion. Assuming there are N nodes in the network, a DATA packet has size S , a cycle has T slots, the length of a slot is τ , and the probability for each node to successfully deliver a DATA packet in a cycle is p_s , if the proposed Markov model has a unique stationary distribution $\pi = \pi_0,\ldots, \pi_Q$, the throughput of MX-MAC is

$$THR = \frac{N\cdot(1-\pi_0)\cdot P_s\cdot S}{T.\tau} \quad (1)$$

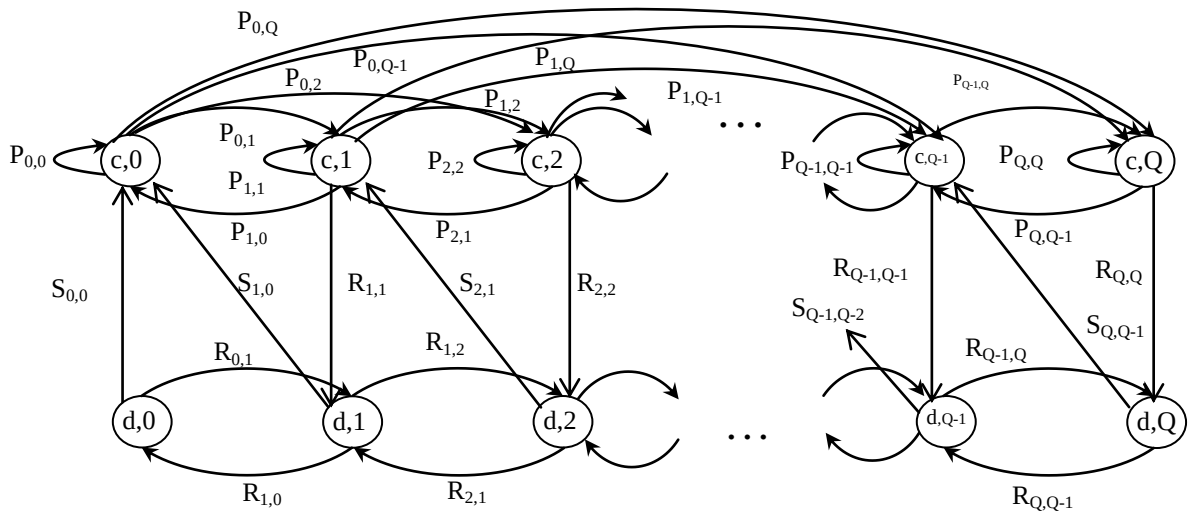


Fig. 4 Markov model for proposed MX-MAC.

N, S and T are known variable, only π_0 and P_s are need to be calculated. π_0 and P_s are determined by our proposed Markov model and the media access rule of MX-MAC. The Markov model has a state space 0,1….Q and a transition matrix. The stationary probability of the empty queue state can be obtain from Markov model which is shown in Fig. 4.

$$P=P_s + P_c \tag{2}$$

Data transmission probability P can be define as a addition of Probability of success P_s and collision P_c . The probability of success depend on status of the queue and channel free probability.

$$P_s=P_{avs} \,.\, P_r(free) \tag{3}$$

Average Probability of Success:

$$P_{avs} = \sum_{t=1}^{T} \frac{1}{T} \left(\sum_{w=1}^{CW} \left(\sum_{Ta=0}^{N-1} \left(\sum_{Wi=0}^{Ta} \left(\sum_{Sd=0}^{Wi} P_r(A|affect, \overline{empty}) \cdot P_r(B|wake, \overline{empty}) \cdot P_r(C|data, \overline{empty}) \right) \right) \right) \right) \tag{4}$$

$$P_r(A|affect, \overline{empty}) = \binom{N-1}{Ta} \left(\frac{Wm}{T}\right)^{Ta} \left(\frac{T-Wm}{T}\right)^{N-1-Ta} \tag{5}$$

$$P_r(B|wake, \overline{empty}) = \binom{Ta}{Wi} \left(\frac{1}{Wm}\right)^{Wi} \left(\frac{Wm-1}{Wm}\right)^{Ta-Wi} \tag{6}$$

$$P_r(C|data, \overline{empty}) = \binom{Wi}{Sd} (\pi_0)^{Wi-Sd} (1-\pi_0)^{Sd} \left(\frac{W_{succ}}{Wm}\right)^{Sd} \tag{7}$$

Here A is the event where Ta nodes out of (N-1) nodes wakeup at contention window *affected* area. B is the event where Wi number of nodes *wakeup* out of Ta nodes at the ith slot of contention window and C is the event where Sd nodes out of Wi nodes have some *data* but they do not hamper the transmission on this contention window. W_{succ} defines successful data transmission in data channel and also indicates that there is no collision due to hidden terminal problem, though collision will still happen in control channel. In case of control channel, the average probability of success will be $W_{succ}=Wt_i$, as we set random backoff time t_i.

The average free channel probability $P_r(free)$ at given time is the ratio of duration of channel free and entire channel time.

$$P_r(free) = \frac{E_{free}}{E_{free} + E_{busy}} \tag{8}$$

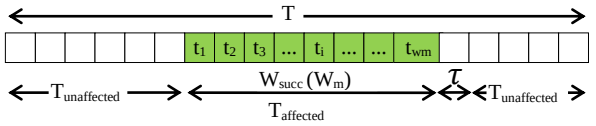


Fig. 5 A Snap-Shot of a Cycle [4].

VI. Energy Consumption Of MX-MAC

The average energy consumption can be calculated as average length of a successful transmission (T/2+L_{data} times), plus total energy requires to send and receive preamble. For simplicity, it is assumed that the total energy requires exchanging the data transmission period (*t*) only. Since MX-MAC eliminates the hidden node problem, no retransmission energy due to hidden node's collision is necessary. The only additional energy that is incurred in MX-MAC compared to X-MAC is its channel switching energy.

The total energy consumption for the proposed MX-MAC is:

$$E_{MX-MAC} = \sum_{n=0}^{\infty} \sum_{t=1}^{T} \left(\left(\frac{T}{2} + L_{DATA}\right) . P_{busy}^{succ} . E_{TRX-DATA} + \left(\frac{T}{2} + L_{DATA}\right) . P_{busy}^{succ} . E_{TRX-PREA} \right) \tag{9}$$

$$P_{busy}^{succ} = \sum_{i=0}^{N-1} \sum_{j=1}^{N-i} \sum_{k=1}^{j} \binom{N}{i} \left(\frac{t}{T}\right)^{i} \left(\frac{T-Wm}{T}\right)^{N-1-i} . \binom{N-i}{j} \left(\frac{1}{Wm}\right)^{j} \left(\frac{Wm-1}{Wm}\right)^{N-i-j} \binom{j}{k} (1-\pi_0)^{k} . \pi_0^{\,j-k} \left(\frac{t}{Wm}\right)^{j-k} \tag{10}$$

VII. Latency Of Proposed MX-MAC

The end-to-end delay is defined as the latency from the time when a source node has a data packet ready to send to the time the first packet is received at the target node. There are many factors affecting latency of the network. Factors such as overhearing, over-emitting, collisions, and control packets overhead affect energy saving as well as forwarding delay. On the other hand, low duty-cycling and sleep mode of a radio cause end-to-end delay degradation.

The expected latency for a single hop data exchange is given by [3]:

Lat = (duration of preamble + ACK listen) × (expected number of iterations required) + duration to send a data packet

$$= \frac{(R_l+R_s)}{(R_l-S_p)} \times (S_p + S_{al}) + S_d \tag{11}$$

For *n* packets with *h* hop count, the only sustained latency due to preamble packet will be only for once and the delay due to subsequent preambles will not be accounted for as the receiver node sends preambles during its packet reception and does not need to send any preamble before forwarding the received packet to the next hop receiver.

TABLE I.

MATHEMATICAL MODEL VARIABLES FOR MX-MAC

Symbol	Meaning	Unit
S_p	Duration of the sender's preamble	ms
S_{al}	Duration of the Acknowledgement listen	ms
R_l	Receiver listen period	ms
R_s	Receiver sleep period	ms
t_d	Data transmission period for a single packet	ms
R_a	Duration of the receiver's ACK period	ms
R_d	Duration of the receiver's data period	ms

Therefore, the total latency for the proposed scheme is derived as:

$$Lat_{MX\text{-}MAC} = t_{contention} + ENIR \times (t_p + t_{al}) + t_d \times n \times h \quad (12)$$

$$ENIR = \frac{(R_l + R_s)}{(R_l - S_p)} \quad (13)$$

ENIR=Expected Number of Iteration Required

VIII. PERFORMANCE ANALYSIS

In this section, we present the results of the simulation that we have done to evaluate our algorithm. The results have shown that our algorithm performs better than X-MAC and X-MAC/CA in most of the situations tested. Fig 6. depicts that MX-MAC decreases the latency with respect to X-MAC and X-MAC/CA protocol. Proposed protocol is more energy efficient which is shown in Fig 7.

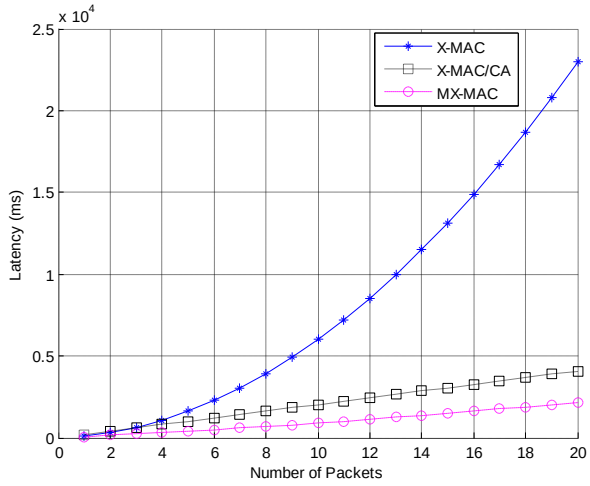


Fig. 6 Latency(ms) with respect ot No of Packets.

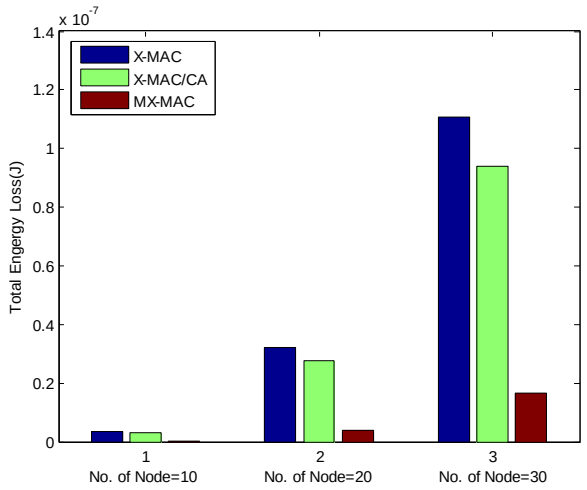


Fig. 7 Total Energy Loss with respect to Number of Node.

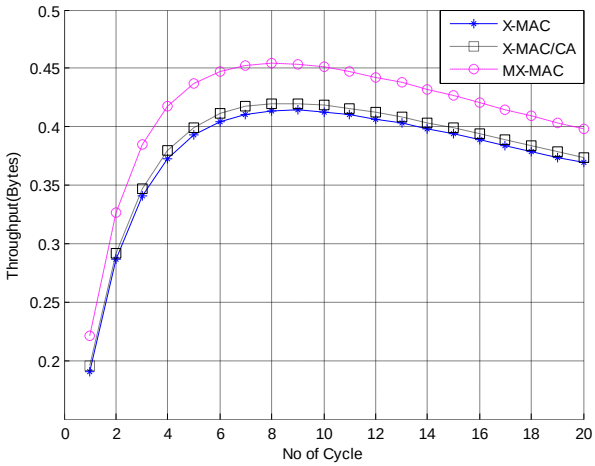


Fig. 8 Throughput with respect ot No of Cycle.

Finally, Fig 8. illustrates the throughput with respect to no of cycle for each algorithm. This figure observably shows that throughput is higher with respect to other algorithm. Comparing with X-MAC and X-MAC/CA protocol, MX-MAC reduces 30% packets retransmission due to hidden node.

IX. CONCLUSIONS

In this paper, we have proposed a Multichannel Asynchronous MX-MAC protocol to reduce the impact of collisions. The key intention of our algorithm is to extend the life span of the WSN by reducing latency and energy loss. From the simulation results it is evident that our proposed algorithm has an enhanced throughput compared to X-MAC and X-MAC/CA.

REFERENCES

[1] Qiang Ma, Kebin Liu, Zhichao Cao, Tong Zhu, Xin Miao and Yunhao Liu "Opportunistic Concurrency: A MAC Protocol for Wireless Sensor Networks", IEEE Transactions on Parallel and Distributed Systems, vol. 26, no. 7, july 2015.

[2] Shuguo Zhuo, Zhi Wang. "A Traffic Adaptive Multi-channel MAC Protocol with Dynamic Slot Allocation for WSNs", IEEE Transactions on Mobile Computing , 27 August 2015, Dubrovnik, Croatia.

[3] Michael Buettner, Gary V. Yee, Eric Anderson, Richard Han. "X-MAC: A Short Preamble MAC Protocol for Duty-Cycled Wireless Sensor Networks", SenSys'06, November 1–3, 2006, Boulder, Colorado, USA.

[4] Ayaz Ullah, Jong-Suk Ahn, and Gayoung Kim. "X-MAC Protocol with Collision Avoidance Algorithm", Ubiquitous and Future Networks (ICUFN), 2013 Fifth International Conference on Year: 2013 Pages: 228 - 233, IEEE Conference Publications.

[5] Ou Yang, and Wendi B. Heinzelman. "Modeling and Throughput Analysis for X-MAC with a Finite Queue Capacity", Global Telecommunications Conference (GLOBECOM 2010), Year: 2010 Pages: 1 - 5, IEEE Conference Publications.

[6] Antonio G. Ruzzelli, Gregory O'Hare, Raja Jurdak, and Richard Tynan. "Advantages of Dual Channel MAC for Wireless Sensor Networks", Communication System Software and Middleware, 2006. Comsware 2006. Pages: 1 - 3, Publisher: IEEE.

[7] Ou Yang and Wendi Heinzelman. "Modeling and Throughput Analysis for SMAC with a Finite Queue Capacity", Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2009 5th International Conference on Year: 2009 Pages: 409 - 414, IEEE Conference Publications.

[8] T. Dam and K. Langendoen, "An adaptive energy-efficient MAC protocol for wireless sensor networks", in Proceedings of the 1st Int. Conf. on Embedded Networked Sensor Systems, 2003, Pages. 171-180.

[9] Hyungkeun Lee and Hyukjoon Lee. "Modeling and Anaysis of an Energy-efficient MAC Protocol for Wireless Sensor Networks", Conference on Information Networking (ICOIN), 2016.

Miniaturized Body Implantable PIFA Operates in MICS-Band and SAR Analysis at Brain, Breast and Muscle of a Human Phantom Model

Sadia Sultana,[1,*] Raja Rashidul Hasan,[2] M. Mohiuddin Miran,[3] Tonmay Kumar Mondal,[4] Shafkat Hossain[5]

Department of Electrical and Electronic Engineering, American International University – Bangladesh
Dhaka, Bangladesh

[*]hemal_aiub@yahoo.com

Abstract—**This paper represents a standard design of a Planar Inverted-F Antenna (PIFA) for medical applications which operates in MICS band (402-405MHz). The antenna is designed very small in size and a better resonant performance is observed at different tissues of human phantom Model. Rogers R03010 is used as substrate and Copper as material. By employing meandering and including the superstrate, the dimension of the proposed antenna is to 12.5 ×12.5 mm^2.The antenna is operated at the desired frequency of 404 MHz and also the S_{11} parameter is obtained -25.81 dB maximum. Specific Absorption Rate (SAR) is calculated for Brain, Breast and Muscle of a human phantom model. At resonant the SAR value is found 0.5745 W/kg for brain, 0.5751 W/kg for muscle and 0.5772W/kg for breast tissue. In brain tissue the antenna gives better frequency response. Safety measurements were taken into account by stimulating the Specific Absorption Rate (SAR) distribution of the antenna to make it perfectly biocompatible.**

Index Terms—**PIFA, Specific Absorption Rate, MICS, Phantom, Resonant.**

I. Introduction

Nowadays, the use of Implantable Medical Devices (IMDs) as part of bio- medical procedures is highly attractive. All important elements of implantable antennas are embedded in those systems, which control the communication of data between implantable devices and external environment. So Miniaturized antennas for implantable medical devices are on prominent demand for nowadays. In wireless communication system a low profile antenna is required which supports multiband and wideband operations. Planar Inverted-F Antenna is met these requirements. This antenna is compact, has minor inverted radiation and has quarter-wavelength patch. Because of its miniature size, simply to integration, manufacture and low fabrication cost, PIFA can be needed for many application such as radio communication, mobile phone [1].On the other hand, Federal Communication Commission have approved various band for bio-medical applications such as Industrial Scientific and Medical (ISM) bands (5.725-5.875GHz, 2.4-2.48GHz, 902-928MHz and 433-434.8MHz), Medical Implant Communication Service (MICS) band (402-405MHz). This band was extended to 401-406 in 2009 and renamed Medical Device Radio Communication Service (Med Radio) [2].

In this paper the other characteristics of antenna like radiation pattern, gain, directivity, VSWR, radiation efficiency are also observed at different phantom tissue in the same band.

II. Experimental Design

To modify the theoretical antenna, single-layer structure and meandered line concepts are introduced here. The antenna performance in free space as well as on human phantom has the main significance here. In order to fulfill the radiation requirements within the miniaturized size and implanted into three layer phantom model antenna is the principle goal of the design procedure.[3]

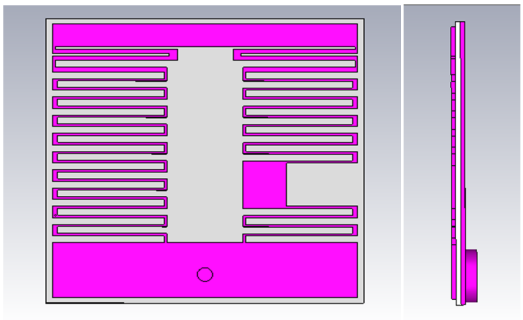


Fig. 1 Top view and side view of the proposed Planar Inverted-F Antenna in free space

The antenna is employed meandering strategy and designed 12.5 ×12.5mm^{2}ground plan. 12 ×12 mm^{2}is consisted quarter wavelength patch antenna. For basic designing purpose Rogers RO3010 (ε_r=10.2 and tanδ=0.003) is used as substrate and superstrate with 0.25mm thickness and copper is used as patch material. The other parameters of antenna are changed in order to improve antenna performance at 404 MHz centerfrequency.A feeding point radius of 0.3mm and a shorting pin radius of 0.03mm are inserted between the patch and groundplane.[3]

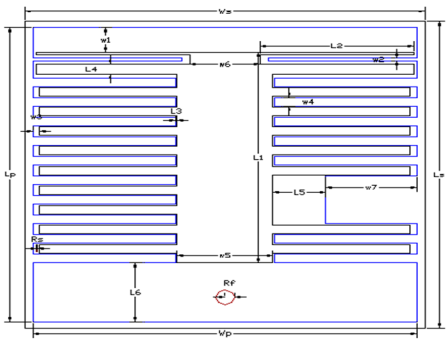


Fig. 2 Antenna designed with AutoCAD

The antenna has designed in free space firstly. CST Microwave Studio is used in different steps of the antenna

performance analysis. The specific absorption rate (SAR) and the resonance performance are verified from CST voxel human model [3] [6]. Table.I shows the final dimensions of the antenna by calculating other parameters.

TABLE I
List of Antenna Size Parameters

Parameters	Size (mm)	Parameters	Size (mm)
L_s	12.5	L_p	12
W_s	12.5	W_p	12
L1	8.6	w1	1
L2	4.8	w2	0.1
L3	0.05	w3	0.2
L4	0.4	w4	0.35
L5	1.65	w5	3
L6	2.4	w6	2.2
R_f	0.3	w7	2.85
R_s	0.03	T_{sub}	0.25

III. Simulated Result

The simulation procedure was performed in CST Microwave Studio. Designed antenna was tested in free space and also on different tissues of human phantom model for observing the performance onwards. The dielectric constant, conductivity and the depth of the human phantom model materials were affected the antenna characteristics. The input impedance was also set to 50 ohm to observe the performance.

A. Simulated Result for Free Space

In Figure.3 shows S_{11} parameter of the antenna is observed -13.17 dB at resonance frequency of 0.406 GHz. S-parameter shows the amount of power radiated and reflected from the antenna. For better performance S_{11} parameter should be less than -10 dB. The bandwidth of the antenna can also be calculated by drawing a straight line at -10dB.From the graph the bandwidth is found to be 7MHz (403MHz-410MHz) .

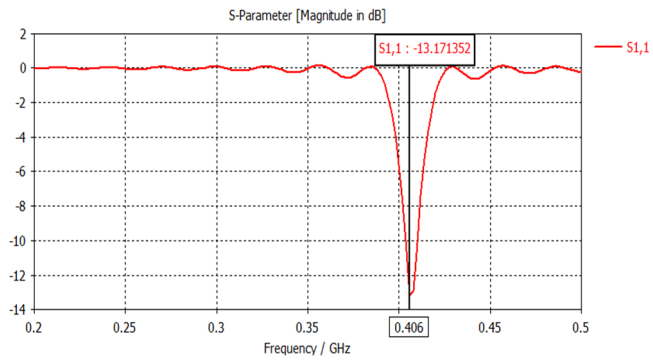


Fig. 3 S11 parameter of the antenna in free space

B. Simulated Result for Brain

After implementing the designed antenna on brain of a human phantom model it is observed that the return loss of the antenna is -25.81 dB at resonance frequency of 0.404 GHz shown in Figure.5 which means the loss of the reflected back signal is very less compared to the radiated signal. From the graph the bandwidth is found to be 18MHz (394MHz-412MHz).

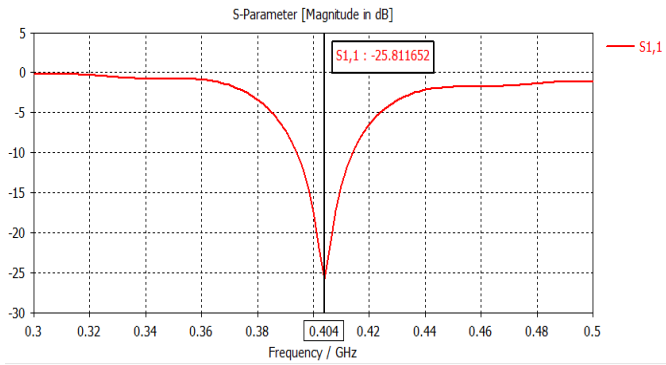


Fig.5 S_{11} parameter of the antenna on Brain of Human Phantom Model (Skin, Fat, Bone & Brain)

In figure.6 it can be found that the designed antenna in human phantom model (brain) has the directivity of 1.962 dBi and radiation efficiency of -47.08 dB.

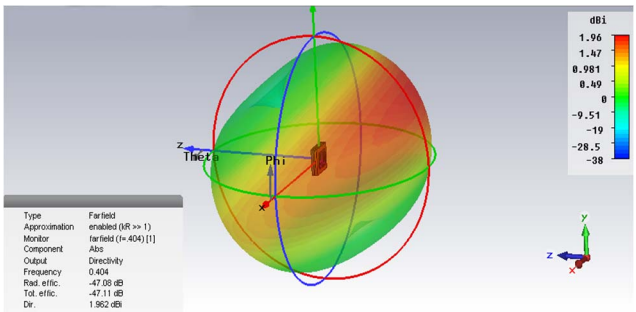


Fig. 6 3D view of Far-Field Region of the antenna on Brain of Human Phantom Model (Skin, Fat, Bone & Brain)

The SAR is measured on brain of Human Phantom Model in figure.7. The Standards for SAR calculation was regulated by IEEE C95.3-2002 standard, where the 1-g average SAR should not exceed 1.6 W/kg as given by FCC and ICNIRP guidelines[4][5]. For this antenna, the SAR was measured at the resonant frequency of 0.404 GHz when an input power of 1mW was provided and the maximum SAR value is found 0.5745W/kg from figure.7 which is below the mark to fulfill the ICNIRP and FCC guidelines regarding SAR.

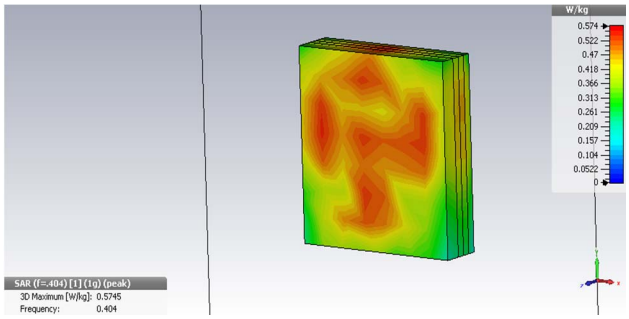


Fig. 7 SAR distribution for an input of 1mW on Brain of Human Phantom Model (Skin, Fat, bone & brain)

C. *Simulated Result for Muscle*

After implementing the designed antenna on muscle of a human phantom model it is observed that the return loss of the antenna is -24.10 dB at resonance frequency of 404 MHz shown in figure.8 and the bandwidth is found to be 20 MHz (394MHz-414GHz) from the graph

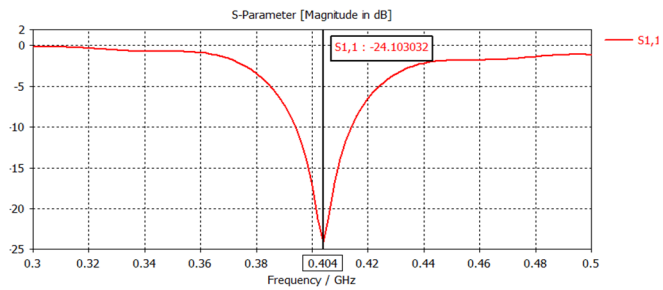


Fig. 8 S_{11} parameter of the antenna on Muscle of Human Phantom Model (Skin, Fat & Muscle)

In figure.9 it can be found that the designed antenna on Brain of Human Phantom Model (Skin, Fat &Muscle) has a directivity of 1.982 dBi and the radiation efficiency of it is -47.44 dB.

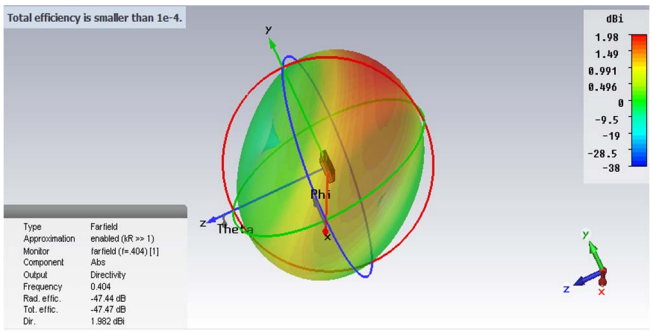


Fig. 9 3D view of Far-Field Region of the antenna on Muscle of Human Phantom Model (Skin, Fat & Muscle)

Input power of 1mW was provided and the maximum SAR value is found 0.5751W/kg for muscle from figure.10 which is below the mark to fulfill the ICNIRP and FCC guidelines regarding SAR.

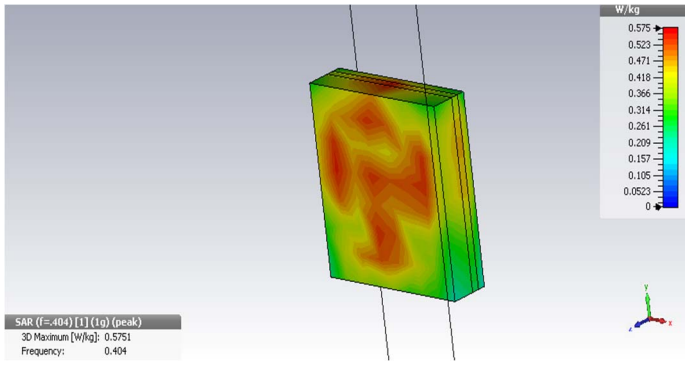


Fig. 10 SAR distribution for an input of 1mW on Muscle of Human Phantom Model (Skin, Fat &Muscle)

D. *Simulated Result for Breast*

After implementing the designed antenna on breast of a human phantom model it is observed from figure.11 that the return loss or S_{11} parameter of the antenna is -23.59 dB when the resonant frequency is 0.404 GHz. The bandwidth of the antenna is measured 20 MHz (394 MHz -414 MHz)from the graph.

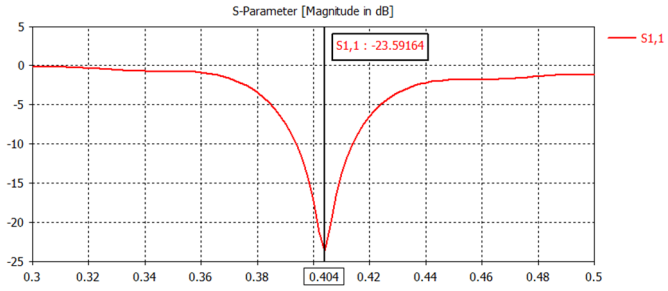


Fig. 11 S_{11} parameter of the antenna on Breast of Human Phantom Model (Skin, Fat, & Breast)

From the figure.12 it can be found that the designed antenna is on Human Phantom Model (Skin, Fat &Muscle) has a directivity of 2.009 dBi and the radiation efficiency of it is -47.62 dB.

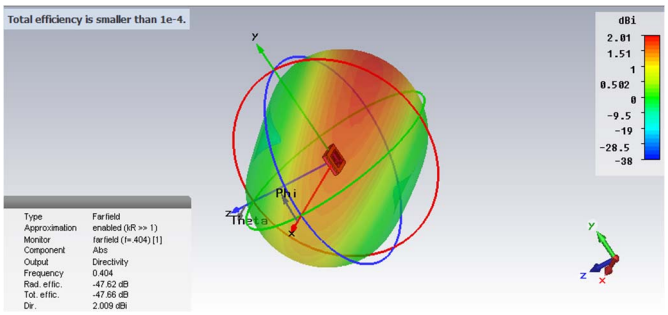


Fig.12 3D view of Far-Field Region of the antenna on Breast of Human Phantom Model (Skin, Fat &Breast)

From figure.12 the SAR value is obtained 0.5772W/kg for breast tissue which is below the mark to fulfill the ICNIRP and FCC guidelines regarding SAR.

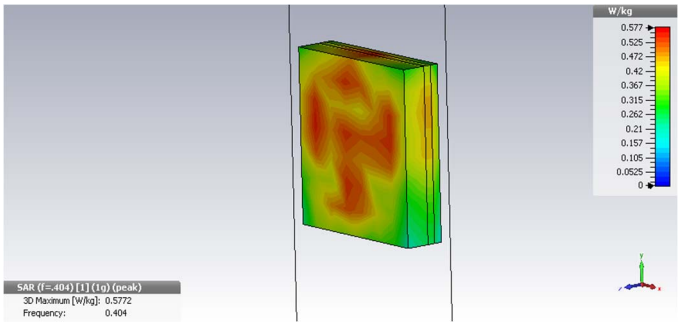


Fig. 13 SAR distribution for an input of 1mW on Breast of Human Phantom Model (Skin, Fat &breast)

Table. II shows the performances analysis of designed antenna for different tissues of Human Phantom Model.

TABLE III
Performance analysis of the designed antenna at different tissue

Parameters	Different tissues of Human Phantom Model		
	Brain	**Muscle**	**Breast**
Resonance Frequency (MHz)	404	404	404
Return Loss(dB)	-25.81	-24.10	-23.59
Directivity (dBi)	1.962	1.982	2.009
VSWR	1.107	1.132	1.141
Side lobe (dB)	-0.9	-1.0	-1.1
Main lobe Magnitude (dBi)	1.96	1.98	2.01
Direction (degree)	149	147	146
Radiation Efficiency (dB)	-47.08	-47.44	-47.62
Total Efficiency (dB)	-47.11	-47.47	-47.66

IV. Conclusions

Considering all the factors and with the help of theoretical calculation finally a planar inverted-f antenna was designed. By analysing the results it has seen that the designed antenna is operated in MICS band and the antenna resonant frequency is 404 MHz. The return loss was measured maximum at -25.81 dB which was also very good to support the criteria of the on body biocompatible antenna. The directivity of the radiation pattern of the antenna was also satisfactory. SAR of the antenna was measured at 1-g averaged SAR with an input power of 1mW maintaining IEEE C95.3 standard technique & the results were satisfactory according to FCC & ICNIRP. Future continuation of this thesis should cover the integration of this antenna with a complete biomedical system or device. Finally, by analysing all the results, the conclusion can be made that the proposed antenna is practically implantable & ready to give a good performance in biomedical application.

References

[1] Online Available :http://www.antennatheory.com//patches/pifa.php. [Accessed: July 15, 2016].

[2] Online Available :https://www.fcc.gov/general/medical-device-radiocommunications-service-medradio. [Accessed: August 5, 2016]

[3] Li, Hua, et al. "A miniature-implantable antenna for medradio-band biomedical telemetry." IEEE Antennas and Wireless Propagation Letters 14 (2015): 1176-1179.

[4] IEEE, “Standard for Safety Levels with Respect to Human to Radio Frequency Electromagnetic Fields, 3kHz to 300GHz,” 1999.

[5] IEEE, “Standard for Safety Levels with Respect to Human Exposure to Radio Frequency Electromagnetic Fields, 3kHz to 300GHz,” 2005.

[6] Online.Available:www.qsl.net/va3iul/Antenna/PIFA/PIFA_Planar_Inverted_F_Antenna.pdf. [Accessed: July 15, 2016]

On the Autocorrelation and Cross-correlation of Ternary Möbius Sequence over Finite Field

Fatema Akhter,[1,*] and Md. Selim Al Mamun,[2]
[1]Jatiya Kabi Kazi Nazrul Islam University, Bangladesh
[2]Okayama University, Japan
*fatema.kumu02@gmail.com

***Abstract*—This paper proposes a novel approach for pseudo random ternary sequence over finite field. The sequence is generated from a primitive polynomial over finite field applying trace function and möbius function. This paper further presents some important properties such as period, autocorrelation and cross-correlation of the generated sequence. The proposed approach is experimented on various bit length of odd characteristic and compared with some existing works. The comparison results shows the superiority of the proposed approach over existing ones.**

***Index Terms*—Primitive polynomial; finite field; trace function; möbius function; autocorrelation, cross-correlation**

I. Introduction

The pseudo random sequences have a wide range of applications in communication systems, cryptography and network security [1]–[4]. In recent years, the study of random sequences has made a significant progress. Several sequences such as Maximum Length Sequence [5], Legendre Sequences [6]–[8] and Sidelnikov Sequences [9]–[11] have been studied by many researchers for their flourishing applications in communications. Pseudo random sequences have some important properties such as period, autocorrelation, cross-correlation and linear complexity [12] that characterize the applications of the sequence in communication or security. Unfortunately, it is familiar that these important properties of the random sequences cannot be proved theoretically. If the security is the major concern, then the period, autocorrelation and linear complexity of the sequence are preferred to be large while cross-correlation is preferred to be low [13]. Therefore the study of these features have also become important while studying the generation of the sequences. This paper proposes a new approach for generation of pseudo random ternary sequence and presents the period, autocorrelation and cross-correlation of the generated sequence. Linear complexity is not studied in this paper, which will be in the future work.

In the extension field of $\mathbb{F}_{p^m}$ let, p be an odd characteristic and m be the degree of the primitive polynomial $f(x)$. $f(x)$ is primitive meaning that a maximum length sequence over $\mathbb{F}_{p^m}$ can be generated from the polynomial. Let ω be its zero, that becomes a primitive element in $\mathbb{F}^*_{p^m}$, then the sequence $\mathcal{S} = \{s_i\}$, $s_i = \mathrm{Tr}\left(\omega^i\right)$, $i = 0, 1, 2, 3, ..., (p^m - 2)$ becomes a maximum length sequence of period, $L = p^m - 1$, where $\mathrm{Tr}\left(\cdot\right)$ is the trace function over $\mathbb{F}_p$. In this proposed work, first a primitive polynomial is generated over finite field, then a maximum length vector sequence is generated from the primitive polynomial, then trace function is applied on the vector sequence to map the vectors to scalars and finally, möbius function is applied on the scalars to generate the pseudo random sequence. Throughout this paper, let random sequence mean three values $\{-1, 0, +1\}$ corresponding to möbius function, p denotes an odd prime number and $\mathbb{F}_{p^m}$ denotes the finite field over p, where $m \geq 2$ is a positive integer representing extension degree. $\mathbb{F}^*_{p^m}$ is the the multiplicative group of $\mathbb{F}_{p^m}$ i.e. $\mathbb{F}^*_{p^m} = \mathbb{F}_{p^m} - \{0\}$. This paper presents all the terminologies and theorems based on the odd characteristic prime field $\mathbb{F}_{p^m}$. However, any arbitrary field can be adaptable as the base field.

The paper is organized as follows: Section II describes the preliminary studies for proposed approach with the necessary definitions and algorithms needed for implementation. Section III describes the proposal of pseudo random sequence with some examples. Section IV presents the experimental results and describes the period, autocorrelation and cross-correlation properties of the generated sequence. Finally, section V concludes the paper with some future recommendations.

II. Preliminaries

This section briefly describes the basic terminologies and mathematical fundamentals such as primitive polynomial, trace function and möbius function, autocorrelation and cross-correlation of pseudo random sequence.

A. Primitive Polynomial

A primitive polynomial of a finite field is a polynomial that can generate all the elements of the field. The multiplicative group, $\mathbb{F}^*_{p^m}$ of non-zero elements of $\mathbb{F}_{p^m}$ is cyclic and can be generated from a primitive element. Let g be a generator of $\mathbb{F}^*_{p^m}$ with elements $\{1, g, g^2, ..., g^{p^m-2}\}$, then every non-zero element is represented as g^i for $i = 0, 1, .., p^m - 2$. The element g is primitive if and only if $\gcd(i, p^m - 1) = 1$. $f(x)$ is a primitive polynomial if and only if the field element x generates the cyclic group of all non-zero field elements of the finite field $\mathbb{F}^*_{p^m}$. Therefore, $f(x)$ is primitive polynomial if and only if

1) $x^{p^m-1} \equiv 1 (\mathrm{mod}\ f(x))$,
2) $x^k \not\equiv 1 (\mathrm{mod}\ f(x))$ for $1 \leq k \leq p^m - 2$.

Property 1: Let g be a generator of $\mathbb{F}^*_{p^m}$, $g^{(p^m-1)/(p-1)}$ becomes a non-zero element in the prime field $\mathbb{F}_p$ and is also a generator of $\mathbb{F}^*_p$.
(Proof) Since g is a generator of $\mathbb{F}^*_{p^m}$ with order $p^m - 1$, if i is a non-negative integer, the order of g^i is given by

$$\frac{p^m - 1}{\gcd(i, p^m - 1)}. \tag{1}$$

Then, the order of $g^{(p^m-1)/(p-1)}$ becomes $p - 1$. Therefore, $g^{(p^m-1)/(p-1)}$ is a generator of $\mathbb{F}^*_p$.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

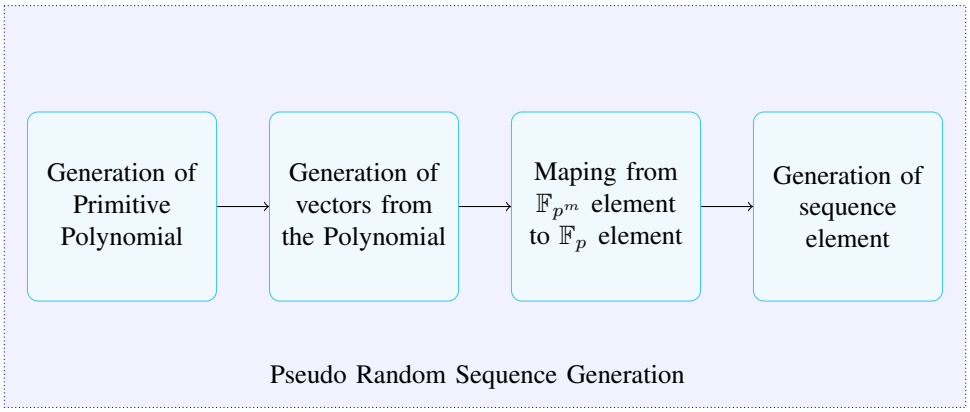


Fig. 1: Block Diagram of Proposed Pseudo Random Sequence

B. Trace Function

Trace function $\mathrm{Tr}\,(\cdot)$ is used to map an extension field element to prime field element. Let X be an extension field element of $\mathbb{F}_{p^m}$, this can be map to prime field element x of $\mathbb{F}_p$ as

$$x = \mathrm{Tr}\,(X) = \sum_{i=0}^{m-1} X^{p^i}. \tag{2}$$

The trace function has the linearity on $\mathbb{F}_p$ when $a, b \in \mathbb{F}_p$ and $X, Y \in \mathbb{F}_{p^m}$ as follows:

$$x = \mathrm{Tr}\,(aX + bY) = a\mathrm{Tr}\,(X) + b\mathrm{Tr}\,(Y). \tag{3}$$

Property 2: For each $i = 0, 1, 2, ..., p-1 \in \mathbb{F}_p$, the number of elements in $\mathbb{F}_{p^m}$ whose trace with respect to $\mathbb{F}_p$ becomes i is given by p^{m-1}.
(Proof) Elements in $\mathbb{F}_{p^m}$ are the roots of $x^{p^m} - x$. It is factorized over $\mathbb{F}_p$ as follows

$$x^{p^m} - x = \prod_{i=0}^{p-1} (\mathrm{Tr}\,(x) - i). \tag{4}$$

Since the degree of $\mathrm{Tr}\,(x)$ is p^{m-1} and $\mathrm{Tr}\,(x)$ does not have any duplicate root, the property is shown. Therefore, the number of elements in $\mathbb{F}^*_{p^m}$ whose trace becomes i is p^{m-1}.

C. Möbius Function

Möbius function, denoted as $\mu(\cdot)$ can translate any natural number to any of the values, $\{-1, 0, +1\}$. For any $n \in \mathbb{N}$, where $\mathbb{N}$ is natural number, möbius function $\mu(\cdot)$ is defined as:

$$\mu(n) = \begin{cases} 1, & \text{if } n = 1, \\ (-1)^k, & \text{if } n \text{ is a product of } k \text{ distinct primes}, \\ 0, & \text{if } n \text{ is divisible by square of a prime.} \end{cases} \tag{5}$$

This paper uses the möbius function to translate the prime field element x of $\mathbb{F}_p$ to the sequence element $s_i \in \{-1, 0, +1\}$. The möbius function takes only natural number, $n \in \mathbb{N} : n > 0$ and has no definition for $n = 0$. As 0 is an element in prime field $\mathbb{F}_p$, this paper modifies the definition of möbius function on prime field $\mathbb{F}_p$ as follows:

$$\mu(n) = \begin{cases} 1, \text{if } n = 1, \\ (-1)^k, \text{if } n \text{ is a product of } k \text{ distinct primes}, \\ 0, \text{if } n = 0 \text{ or } n \text{ is divisible by square of a prime.} \end{cases} \tag{6}$$

For $n \leq 11$, the möbius function values are:

n	0	1	2	3	4	5	6	7	8	9	10	11
μ	0	1	-1	-1	0	-1	1	-1	0	0	1	-1

D. Autocorrelation

This paper presents the pseudo random sequence as: $\mathcal{S} = \{s_i\}$, for $i = 0, 1, 2, 3, ..., (L-1)$, where $s_i \in \{-1, 0, +1\}$ and $L = p^m - 1$ is the period of the sequence, i.e. $s_i = s_{i+L}$. Then, the autocorrelation, sometimes called the periodic autocorrelation, $\mathcal{R}_{\mathcal{S}}(x)$ of the sequence $\mathcal{S}$ after shifted x position is defined as follows:

$$R_S(x) = \sum_{i=0}^{L-1} s_i s_{i+x}. \tag{7}$$

for all $x = 0, 1, ..., L-1$, where s_i and s_{i+x} are the ith element and the x position shifted element of the generated sequence $\mathcal{S} = \{s_i\}$.

E. Cross-correlation

Let $\mathcal{S} = \{s_i\}$ and $\mathcal{T} = \{t_i\}$ be two random sequences with same length, $L = p^m - 1$, then this paper defines the cross-correlation, $\mathcal{C}_{\mathcal{S},\mathcal{T}}(x)$ of $\mathcal{S}$ and $\mathcal{T}$ as follows:

$$\mathcal{C}_{\mathcal{S},\mathcal{T}}(x) = \sum_{i=0}^{L-1} s_i t_{i+x}. \tag{8}$$

for all $x = 0, 1, ..., L-1$, where s_i and t_{i+x} are the ith element of sequence $\mathcal{S}$ and x position shifted ith element of $\mathcal{T}$ respectively.

III. Proposed Pseudo Random Sequence

This section presents the proposed pseudo random sequence and briefly describes the implementation procedure with some examples.

A. Proposal of Pseudo Random Sequence

Let ω be a primitive element in the extension field $\mathbb{F}_{p^m}$. The proposed pseudo random sequence $\mathcal{S} = \{s_i\}$ is presented as:

$$\mathcal{S} = \{s_i\}, s_i = \mu\left(\mathrm{Tr}\left(\omega^i\right)\right) \text{ for } i = 0, 1, 2, 3, ..., (2^m - 2). \tag{9}$$

where $\mathrm{Tr}\,(\cdot)$ represents the trace function and $\mu(\cdot)$ represents the möbius function. The period of the generated sequence and the number of peaks in autocorrelation are given below:

- Period, $L = p^m - 1$
- Number of peaks in autocorrelation = $p - 1$

It is noted that the period of the generated sequence and the number of peaks of autocorrelation are given from the experimental results. The theoretical proof is not studied in this paper, which will be in our future work.

B. Implementation Procedure

Fig.1 shows the phases of the proposed pseudo random ternary sequence generation. This section briefly describes the implementation procedure of the proposed pseudo random sequence.

Pseudo Random Sequence Generation

---

① Generate primitive polynomials over finite field $\mathbb{F}_{p^m}$. There are $n = \dfrac{\varphi(p^m - 1)}{m}$ number of primitive polynomials, where $\varphi(\cdot)$ represents the *euler totient function*.

② Select one of the primitive polynomials generated in the above step. This is the generator primitive polynomial, $f(x)$.

③ From the primitive polynomial, $f(x)$, generate all the vectors over finite field $\mathbb{F}_{p^m}$. There are $2^m - 1$ vectors in the finite field $\mathbb{F}_{p^m}$.

④ For each vector $X^i \in \mathbb{F}_{p^m}$ for $i = 0, 1, 2, ..., 2^m - 2$ do the followings:

  a) Map vector element X^i to prime field element $x_i \in \mathbb{F}_p$ using the trace function defined in Eq. (2) as follows: $x_i = \mathrm{Tr}\left(X^i\right)$

  b) Translate prime field element $x_i \in \mathbb{F}_p$ to pseudo random sequence element $s_i \in \mathcal{S}$ using the möbius function defined in Eq. (6) as follows: $s_i = \mu(x_i)$.

⑤ Finally, output the pseudo random sequence $\mathcal{S} = \{s_i\}$ for $i = 0, 1, .., p^m - 2$.

C. Examples of the sequence

Let $p = 3$, $m = 2$ and $f(x) = x^2 + x + 2$ be the primitive polynomial over $\mathbb{F}_{3^2}$. In this case,

Tr	0	1	2
μ	0	1	-1

Then, the generated pseudo random sequence is:

$$\mathcal{S} = \{-1, -1, 0, -1, 1, 1, 0, 1\}. \tag{10}$$

Let $p = 5$, $m = 2$ and $f(x) = x^2 + x + 2$ be the primitive polynomial over $\mathbb{F}_{5^2}$. In this case,

Tr	0	1	2	3	4
μ	0	1	-1	-1	0

Then, the generated pseudo random sequence is:

$$\begin{aligned}\mathcal{S} = \{ &-1, 0, -1, 0, 1, 0, 0, -1, 0, 0, -1, -1, \\ &-1, 1, -1, 0, 0, 1, 1, -1, 1, 0, -1, -1\}.\end{aligned} \tag{11}$$

IV. Results and Discussions

This section describes the features of proposed pseudo random sequence such as period, autocorrelation, cross-correlation and compare with some popular existing works.

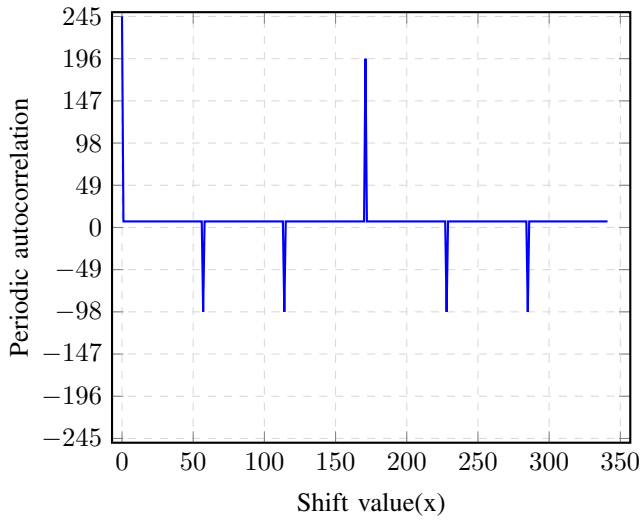


Fig. 2: Autocorrelation for the sequence, $p = 7$ and $m = 3$.

A. Period of Sequence

The proposed pseudo random sequence, $\mathcal{S} = \{s_i\}$ has the period of $L = p^m - 1$ for finite field $\mathbb{F}_{p^m}$. For $p = 5$ and $m = 3$, there are 20 primitive polynomials. In this paper, 2 primitive polynomials are used to generate squences. The generated sequence and the period of the sequence for each primitive polynomial is given below:

- $f_1(x) = x^3 + 6x^2 + 6x + 2$
  Period, $L = 342$

$$\mathcal{S}=\{\begin{array}{cccccccccccc}
-1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 \\
1 & 1 & -1 & 0 & -1 & 1 & 0 & 0 & 1 & -1 & 0 & 0 \\
-1 & 1 & -1 & -1 & 1 & 1 & 1 & 0 & 1 & -1 & 0 & -1 \\
-1 & 0 & 0 & 1 & 1 & -1 & 1 & 1 & -1 & 0 & 0 & 1 \\
0 & 1 & 1 & 0 & 0 & 1 & -1 & 1 & 1 & 1 & -1 & 1 \\
0 & -1 & 0 & -1 & -1 & -1 & 0 & -1 & -1 & -1 & -1 & 1 \\
- & - & - & - & - & - & - & - & - & - & - & - \\
- & - & - & - & - & - & - & - & - & - & - & - \\
- & - & - & - & - & - & - & - & - & - & - & - \\
-1 & -1 & 1 & -1 & 0 & 0 & -1 & -1 & 1 & -1 & -1 & 1 \\
-1 & 0 & -1 & 0 & -1 & 0 & 0 & -1 & 0 & -1 & -1 & 0
\end{array}\}$$

- $f_2(x) = x^3 + 6x^2 + 5x + 4$
  Period, $L = 342$

$$\mathcal{T}=\{\begin{array}{cccccccccccc}
-1 & 1 & -1 & -1 & 1 & 1 & 0 & 1 & -1 & 0 & 1 & 0 \\
-1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & -1 & -1 & 1 \\
-1 & 1 & 0 & 0 & 1 & -1 & -1 & -1 & -1 & 1 & -1 & 1 \\
1 & 1 & 0 & -1 & 1 & 0 & 0 & 0 & -1 & -1 & -1 & -1 \\
1 & 1 & -1 & -1 & -1 & -1 & 1 & 0 & -1 & -1 & -1 & 1 \\
1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\
- & - & - & - & - & - & - & - & - & - & - & - \\
- & - & - & - & - & - & - & - & - & - & - & - \\
- & - & - & - & - & - & - & - & - & - & - & - \\
-1 & 1 & -1 & -1 & -1 & 0 & 0 & -1 & 1 & 0 & 1 & -1 \\
1 & 0 & 1 & -1 & -1 & 0 & 1 & -1 & -1 & -1 & 1 & 0
\end{array}\}$$

B. Results of Autocorrelation

The periodic autocorrelation of the sequence represented at $f_1(x)$ is presented in Eq. (12) and shown in Fig. 2.

$$\mathcal{R}_{\mathcal{S}}(x) = \begin{cases} 245 & \text{if } i = 0 \\ -98 & \text{else if } i = 57, 114, 228, 285 \\ 196 & \text{else if } i = 171 \\ 7 & \text{otherwise} \end{cases}. \tag{12}$$

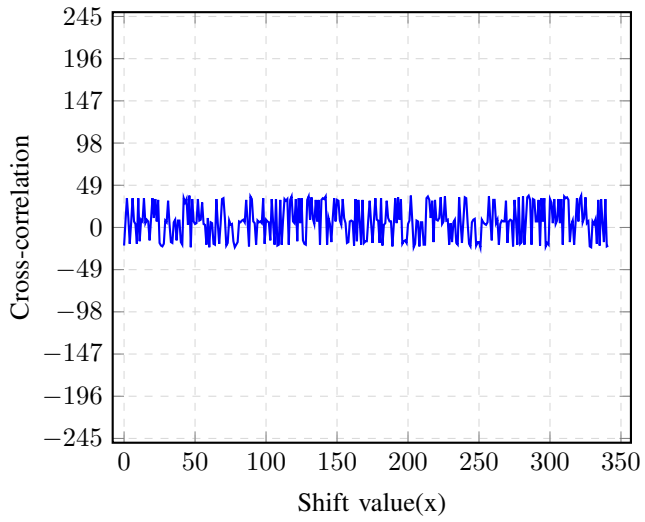


Fig. 3: Cross-correlation between sequence, $\mathcal{S}$ and $\mathcal{T}$.

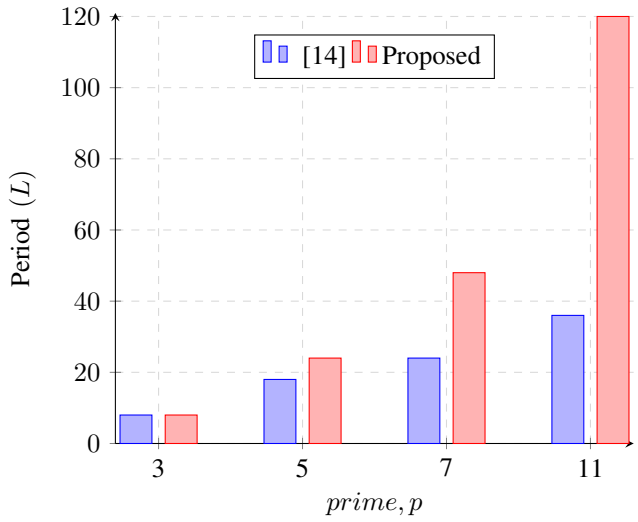


Fig. 4: Comparison of period between [14] and proposed sequence.

C. Results of Cross-correlation

To study the cross-correlation property of the proposed sequence, sequence $\mathcal{S} = \{s_i\}$ is generated from $f_1(x)$ while sequence $\mathcal{T} = \{t_i\}$ is generated from $f_2(x)$. The cross-correlation of the two sequence is shown in Fig. 3.

D. Comparisons with Existing Approaches

The proposed random sequence $\mathcal{S} = \{s_i\}$ has the period, $L = p^m - 1$ over odd characteristic $\mathbb{F}_{p^m}$. This is the maximum period that can be achieved on $\mathbb{F}_{p^m}$ and is equal to the length of maximum value sequence. The proposed sequence generated using möbius function has not been reported anywhere in the literature yet, where our goal is to generate random sequence with maximum period and autocorrelation value. The original Legendre sequence has the length of p which is extended by [13]–[15] later. Sequence generated in [14] has a period of length, $L = \dfrac{k(p^m - 1)}{p - 1}$[1]. A comparison result with [14] is shown in Fig. 4. The proposed sequence has larger period than that of [14] for various bit length of characteristic.

V. Conclusion

This paper proposes a novel approach for generation of random ternary sequence utilizing primitive polynomial, trace function and möbius function over finite field $\mathbb{F}_{p^m}$. The proposed approach is clarified through appropriate parameterization and examples. Then, this paper presents period, autocorrelation and cross-correlation of the generated sequence for $\mathbb{F}_{7^3}$. The results are finally compared with some popular approaches to establish the superiority of the proposed approach. Future works include study of other important properties such as linear complexity, distribution of bit pattern of the sequence.

Acknowledgment

The author would like to express her greatest gratitude to the anonymous reviewers for their constructive feedback and critical suggestions that helped significantly to elicit the utmost technical attribute of this research work.

[1]Proposed $\mathcal{S} = \{s_i\}$ takes only 3 values, $s_i \in \{-1, 0, 1\}$, this paper considers only the existing work with 3 values, i.e. $k = 3$ in [14] for comparison.

References

[1] M. K. Simon, J. K. Omura, R. A. Scholtz and B. K. Levitt, "Spread spectrum communications handbook," Comput. Sci. Press, McGraw-Hill, New York, Apr. 1994.
[2] P. Fan, and M. Darnell, "Sequence design for communications applications," John Wiley & Sons, New York, 1996.
[3] X. Du. and Z. Chen, "Linear complexity of quaternary sequences generated using generalized cyclotomic classes modulo $2p^*$," IEICE Trans. on Fundamentals of Electron., Commun. and Comput. Sci., vol. E94-A, no. 5, pp. 1214 - 1217, May 2011.
[4] Y. T. Kim, M. K. Song, D. S. Kim, and H. Y. Song, "Properties and crosscorrelation of decimated sidelnikov sequences," IEICE Trans. on Fundamentals of Electron., Commun. and Comput. Sci., vol. E97-A, no. 12, pp. 2562 - 2566, Dec. 2014.
[5] S. W. Golomb, "Shift register sequences," Holden-Day, San Francisco, 1967.
[6] N. Zierler, "Legendre Sequence," MIT, Lincoln Publications, 1958.
[7] J. S. No, H. K. Lee, H. Chung, H. Y. Song and K. Yang, "Trace representation of legendre sequences of mersenne prime period," IEEE Trans. on Inf. Theory, vol. 42, no. 6, pp. 2254 - 2255, Nov. 1996.
[8] C. Ding, T. Hesseseth, and W. Shan, "On the linear complexity of legendre sequences," IEEE Trans. on Inf. Theory, vol. 44, no. 3, pp. 1276 - 1278, May 1998.
[9] N. Y. Yu and G. Gong, "New construction of M-ary sequence families with low correlation from the structure of sidelnikov sequences," IEEE Trans. on Inf. Theory, vol. 56, no. 8, pp. 4061 - 4070, Aug. 2010.
[10] M. Su and A. Winterhof, "Autocorrelation of legendre sidelnikov sequences," IEEE Trans. on Inf. Theory, vol. 56, no. 4, pp. 1714 - 1718, Apr. 2010.
[11] Y. T. Kim, D. S. Kim and H. Y. Song, "New M -ary sequence families with low correlation from the array structure of sidelnikov sequences," IEEE Trans. on Inf. Theory, vol. 61, no. 1, pp. 655 - 670, Nov. 2014.
[12] T. W. Cusick, C. Ding and A. Renvall, "Stream ciphers and number theory," Elsevier, Amsterdam, Apr. 1998.
[13] Y. Nogami, K. Tada and S. Uehara, "A geometric sequence binarized with legendre symbol over odd characteristics field and its properties," IEICE Trans. on Fundamentals of Electron., Commun. and Comput. Sci., vol. E97-A, no. 12, pp. 2336 - 2342, Dec. 2014.
[14] H. Ino, Y. Nogami, N. Begum, S. Uehara, R. M. Zaragoza and T. Kazuyoshi, "A consideration on crosscorrelation of a kind of trace sequences over finite field," 3rd. Int. Symp. on Computing and Networking(CANDAR), pp. 484 - 486, Dec. 2015.
[15] Y. Nogami, K. Tada and S. Uehara, "A binarization of geometric sequences with legendre symbol and its autocorrelation," The 6th Int. Workshop on Signal Design and Its Applications in Commun., pp. 28-31, Oct. 2013.

Rational Krylov Subspace Method (RKSM) for Solving the Lyapunov Equations of Index-1 Descriptor Systems and Application to Balancing Based Model Reduction

M. Monir Uddin,[1,*] M. Sumon Hossain,[2] and Mohammad Forhad Uddin[2]
[1] Department of Mathematics and Physics, North South University
Dhaka - 39106, Bangladesh
[2] Department of Mathematics, Bangladesh University of Engineering and Technology
Dhaka - 1000, Bangladesh
*monir.uddin@northsouth.edu

***Abstract*—This paper focuses on the iterative method for solving Lyapunov equations to compute the low-rank Gramian factors. Such Lyapunov equations arise from large-scale sparse index-1 descriptor system. The technique is mainly based on rational Krylov subspace method (RKSM) which is introduced in [1]. However, there the proposed technique is applicable for the standard state space model. Here, we extend this idea for index-1 descriptor system to compute the low-rank Gramian factors. The Gramian factors are then applied to the balancing based model reduction to reduce the complexity of the underlying system. Several test examples are considered to show the efficiency of the proposed method numerically.**

***Index Terms*—Index-1 descriptor systems, Lyapunov matrix equations, rational Krylov subspace method, model reduction, balanced truncation.**

I. Introduction

This paper discusses an efficient method to solve the continuous-time algebraic Lyapunov equations

$$APE^T + EPA^T = -BB^T, \tag{1}$$

$$A^TQE + E^TQA = -C^TC, \tag{2}$$

where $E, A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times p}$, $C \in \mathbb{R}^{m\times n}$ and $p, m \ll n$. The solutions P and Q are known as the controllability and observability Gramians of the linear time invariant (LTI) dynamical system

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t); \quad x(t_0) = x_0, \quad t \geq t_0 \\ y(t) &= Cx(t) + Du(t), \end{aligned} \tag{3}$$

where the vectors $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^p$ and $y(t) \in \mathbb{R}^m$ are known as respectively, the state, input and output. This dynamical system is assumed to be obtained by eliminating the algebraic part of the *differential-algebraic equations* (DAEs) or descriptor system [2] of the form

$$\begin{aligned} E_1\dot{x}(t) &= A_1x(t) + A_2z(t) + B_1u(t) \\ 0 &= A_3x(t) + A_4z(t) + B_2u(t) \\ y(t) &= C_1x(t) + C_2z(t) + D_au(t), \end{aligned} \tag{4}$$

where $z(t) \in \mathbb{R}^{n_a}$ is the vector of algebraic variables and the matrices E_1, A_1, A_2, A_3, A_4, B_1, B_2, C_1, C_2, and D_a are *sparse matrices* with adequate sizes. The relation between the systems (3) and (4) is as follows

$$\begin{aligned} &E := E_1, \\ &A := A_1 - A_2{A_4}^{-1}A_3, \quad B := B_1 - A_2{A_4}^{-1}B_2, \\ &C := C_1 - C_2{A_4}^{-1}A_3, \quad D := D_a - C_2{A_4}^{-1}B_2. \end{aligned}$$

In the literature the system is known as *index-1* since the null space of the block matrix A_4 contains only zero (block) i.e., A_4 is nonsingular block matrix. In a matrix-vector form (4) can be represented as

$$\begin{aligned} \underbrace{\begin{bmatrix} E_1 & 0 \\ 0 & 0 \end{bmatrix}}_{\tilde{E}} \underbrace{\begin{bmatrix} \dot{x}(t) \\ \dot{z}(t) \end{bmatrix}}_{\dot{\tilde{x}}(t)} &= \underbrace{\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}}_{\tilde{A}} \underbrace{\begin{bmatrix} x(t) \\ z(t) \end{bmatrix}}_{\tilde{x}(t)} + \underbrace{\begin{bmatrix} B_1 \\ B_2 \end{bmatrix}}_{\tilde{B}} u(t) \\ y(t) &= \underbrace{\begin{bmatrix} C_1 & C_2 \end{bmatrix}}_{\tilde{C}} \begin{bmatrix} x(t) \\ z(t) \end{bmatrix} + D_au(t). \end{aligned} \tag{5}$$

The dynamical systems (3) and (4) (or (5)) are equivalent in a sense that their *transfer functions* and the *finite spectrums* are the same [3]. Transfer function [4] is essentially the input-output relations of the system. For our given system it can be defined as

$$G(s) = C(sI - A)^{-1}\ B + D = \tilde{C}(s\tilde{E} - \tilde{A})^{-1}\tilde{B} + D_a,$$

where $s \in \mathbb{C}$ (complex-plane). Such a descriptor system arises in different applications of science and engineering. The model that we consider for our numerical test is obtained by linearizing the power system model around the equilibrium point (e.g., see [5] and the references therein). The applications of this model include the design of controllers, optimizations, simulations (time response or frequency response) and so on. Either of the tasks become infeasible if the system becomes very large. A large-scale dynamical system often leads to not only the extra memory requirements but also additional computational complexity. Therefore, *model order reduction* is unavoidable. That is, the large-scale system is replaced by the reduced order system where the dynamics of the original and reduced systems are approximated in a possible context. The way to find the reduced order model from the large model is called *model order reduction* (MOR). See e.g., [4], [6] for motivations, applications, restrictions, and techniques of MOR.

In the literature there exist many techniques for computing the MOR. Recently, the system theoretic method balanced truncation (BT) is the most popular technique, since in this

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

method the reduced system preserves some required properties of the original system. The main demerit of this method is to solve two Lyapunov equations for computing the Gramian factors. Authors in [5] discussed an efficient balancing based methods for the model reduction of the system in (4). To solve the Lyapunov equations, they apply the low-rank alternating direction implicit (LR-ADI) iterations [7], [8]. If there are eigenvalues of the system near zero, the convergence of the LR-ADI method becomes worse. To overcome the problem, in [5], the authors use so-called α-shift approach for moving the eigenvalues closest to origin to the complex plane left-side.

In this paper we investigate another efficient method to solve the Lyapunov equations for the balancing based model reduction of the system (4). Our technique is based on rational Krylov subspace method (RKSM) which is originally introduced in [1]. There, the proposed algorithm is designed for standard (or generalized) state space systems. Here, we will extend this algorithm for index-1 DAEs (4) to compute the low-rank Gramian factors. The computed Gramian factors will then be used for the balanced truncation based MOR of the DAEs. Also, numerical results will be discussed to show the efficiency of our proposed method.

II. Preliminaries

In this section we briefly introduce the balanced truncation and related issues for the large-scale generalized linear time-invariant (LTI) continuous-time system

$$\mathcal{E}\dot{\xi}(t) = \mathcal{A}\xi(t) + \mathcal{B}u(t), \quad y(t) = \mathcal{C}\xi(t) + \mathcal{D}u(t), \tag{6}$$

where $\mathcal{E}, \mathcal{A} \in \mathbb{R}^{n_g \times n_g}$, $\mathcal{B} \in \mathbb{R}^{n_g \times p}$, $\mathcal{C} \in \mathbb{R}^{m \times n_g}$ and $\mathcal{D} \in \mathbb{R}^{m \times p}$. The matrix $\mathcal{E}$ is non-singular. To obtain a reduced order model of this system using the BT method, we need to perform three major steps (see, e.g., [4] in details). At first, compute the low-rank Gramian factors $\mathcal{R}$ and $\mathcal{L}$ by solving the two Lyapunov equations

$$\mathcal{A}\mathcal{P}\mathcal{E}^T + \mathcal{E}\mathcal{P}\mathcal{A}^T = -\mathcal{B}\mathcal{B}^T, \quad \text{and} \tag{7}$$

$$\mathcal{A}^T\mathcal{Q}\mathcal{E} + \mathcal{E}^T\mathcal{Q}\mathcal{A} = -\mathcal{C}^T\mathcal{C}, \tag{8}$$

respectively, where $\mathcal{P} \approx \mathcal{R}\mathcal{R}^*$ and $\mathcal{Q} \approx \mathcal{L}\mathcal{L}^*$. Then compute the singular value decomposition

$$\mathcal{R}^T\mathcal{E}\mathcal{L} = \mathcal{U}\Sigma\mathcal{V}^T = [\mathcal{U}_1, \mathcal{U}_2] \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} \mathcal{V}_1^T \\ \mathcal{V}_2^T \end{bmatrix}, \tag{9}$$

where $\mathcal{U}_1$ and $\mathcal{V}_1$ are composed of the leading $k(\ll n)$ columns of $\mathcal{U}$ and $\mathcal{V}$, respectively, Σ_1 is the first $k \times k$ block of the matrix $\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_k, \ldots \sigma_n)$, where σ_i, $i = 1, 2, \ldots, n$ are the *Hankel singular values* [4] of the system. Finally, construct the left and right *balancing and truncating transformations* by defining

$$\mathcal{T}_l := \mathcal{R}\mathcal{U}_1\Sigma_1^{-\frac{1}{2}}, \qquad \mathcal{T}_r := \mathcal{L}\mathcal{V}_1\Sigma_1^{-\frac{1}{2}}. \tag{10}$$

Applying these transformations into the system (6) find the reduced-order model (ROM)

$$\hat{\mathcal{E}}\dot{\hat{\xi}}(t) = \hat{\mathcal{A}}\hat{\xi}(t) + \hat{\mathcal{B}}u(t), \qquad \hat{y}(t) = \hat{\mathcal{C}}\hat{\xi}(t) + \mathcal{D}u(t), \tag{11}$$

where the reduced matrices can be formed as

$$\hat{\mathcal{E}} = \mathcal{T}_l^T\mathcal{E}\,\mathcal{T}_r,\ \hat{\mathcal{A}} = \mathcal{T}_l^T\mathcal{A}\mathcal{T}_r,\ \hat{\mathcal{B}} = \mathcal{T}_l^T\mathcal{B}, \quad \hat{\mathcal{C}} = \mathcal{C}\mathcal{T}_r.$$

The ROM obtained in this way satisfies the global error bound [4]

$$\|\mathrm{G}(.) - \hat{\mathrm{G}}(.)\|_{\mathcal{H}_\infty} \leq 2\sum_{i=k+1}^{n} \sigma_i, \tag{12}$$

where $\mathrm{G}(s)$ and $\hat{\mathrm{G}}(s)$ are called the transfer functions (defined above) of the full and reduced model, respectively, and $\|.\|_{\mathcal{H}_\infty}$ denotes the $\mathcal{H}_\infty$-norm. The relation (12) is indeed an *a priori* error bound. Hence, one can easily determine the required dimension of the ROM for a given error tolerance. In [9] this procedure for the BT method is summarized as the *square-root* (SR) algorithm.

In the SR-algorithm the main computational task is to solve two Lyapunov equations for computing the Gramian factors. During the recent decades several approaches have been developed that allow to exploit the fact that often all coefficient matrices are highly sparse and the number of inputs and outputs are very small compared to the number of DoFs (Degree of Freedoms). Then usually the Gramians $\mathcal{P}$ and $\mathcal{Q}$ can be expressed in low rank factored form, i.e., $\mathcal{R}$ and $\mathcal{L}$ are the thin rectangular matrices. In computing the low-rank factors $\mathcal{R}$ and $\mathcal{L}$, currently, the low-rank alternating direction implicit (LR-ADI) [7], [8] method is one of the most commonly used technique. Another prominent technique is the rational Krylov subspace method (RKSM) which is originally introduced in [10]. This technique is attractive since it is assumed to be cheap in computation. Since in this paper we mainly focus on this technique, in the following we present an algorithm from [10] to solve the Lyapunov equations (7-8) using the RKSM for computing the low-rank Gramian factors of the system (6). The Lyapunov equations (7) and (8) are dual of each other. Hence, we discuss the method elaborately only for the Lyapunov equation (7).

The rational Krylov subspace projection-type method [10] solves the Lyapunov equation (7), delivering a low rank approximate solution $\tilde{\mathcal{P}}$ in factored form i.e., $\tilde{\mathcal{P}} = \mathcal{R}\mathcal{R}^*$, so that only the tall matrix $\mathcal{R}$ needs to be stored. The projection is performed onto the rational Krylov subspace which is generated iteratively. For the given shift parameters $\mu_i \in \mathbb{C}$; $i = 1, 2, \cdots, m$, the m dimensional *rational Krylov subspace* is defined as

$$\mathcal{K}_m := \operatorname{span}\left\{(\mathcal{A} - \mu_1\mathcal{E})^{-1}\mathcal{B}, \cdots, \prod_{i=1}^{m}(\mathcal{A} - \mu_i\mathcal{E})^{-1}\mathcal{B}\right\}.$$

Let us consider that orthonormal columns of $\mathcal{V}_m$ span the approximation space imposing the Galerkin condition:

$$\mathcal{V}_m^*(\mathcal{A}\mathcal{P}\mathcal{E}^T + \mathcal{E}\mathcal{P}\mathcal{A}^T + \mathcal{B}\mathcal{B}^T)\mathcal{V}_m = 0.$$

Then we can write $\tilde{P} = \mathcal{V}_m Y \mathcal{V}_m^*$, for the matrix Y, which can be uniquely determined by solving the small Lyapunov equation

$$\mathcal{A}_m Y \mathcal{E}_m^T + \mathcal{E}_m Y \mathcal{A}_m^T = -\mathcal{V}_m^*\mathcal{B}\mathcal{B}^T\mathcal{V}_m, \tag{13}$$

where $\mathcal{E}_m = \mathcal{V}_m^*\mathcal{E}\mathcal{V}_m$ and $\mathcal{A}_m = \mathcal{V}_m^*\mathcal{A}\mathcal{V}_m$. We know that Y is symmetric and positive (semi-)definite matrix, hence it can be factorized as $Y = \mathcal{S}\mathcal{S}^*$. Therefore, we obtain $\tilde{\mathcal{P}} = (\mathcal{V}_m\mathcal{S})(\mathcal{V}_m\mathcal{S})^* = \mathcal{R}\mathcal{R}^*$, which implies $\mathcal{R} = \mathcal{V}_m\mathcal{S}$. Applying the *eigenvalue decomposition* to the matrix Y and truncating the negligible eigenvalues we can ensure that the computed $\mathcal{R}$ has smallest possible number of columns. Readers are referred to [10] for further information and the implementation details.

The method can be stopped if the norm of the residual

$$\mathcal{R} = \mathcal{A}\tilde{\mathcal{P}}\mathcal{E}^T + \mathcal{E}\tilde{\mathcal{P}}\mathcal{A}^T + \mathcal{B}\mathcal{B}^T \tag{14}$$

is sufficiently small. Note that following [10, Proposition 4.1] we can compute the Frobinious norm of $\mathcal{R}$ cheaply. Another

Algorithm 1: Computation of new shift

Input : $\{\lambda_k\}_{k=1:m}$, $\{\mu_k\}_{k=1:m}$,
$\{\nu_1 = \mu_0^{(1)}, \nu_2 = \mu_0^{(2)}, \cdots \nu_{m+1} = \mu_{m+1}$.
Output: μ_{j+2}.

for $j = 1 : l-1$ **do**
  $\sigma_j = \text{argmax}_{\sigma \in [\nu_j, \nu_{j+1}]} \frac{1}{|r_m(\sigma)|}$, for
  $r_m(z) = \prod_{j=1}^{m} \frac{|z-\lambda_j|}{|z-\mu_j|}$
Compute $\mu_{m+2} = \text{argmax}_{j=1,\cdots,l-1} \frac{1}{|r_m(\sigma_j)|})$

Algorithm 2: RKSM.

Input : $\mathcal{E}$, $\mathcal{A}$, $\mathcal{B}$, i_{max} (number of iterations),
$\{\mu_0^{(1)}, \mu_0^{(2)}\}$.
Output: $\mathcal{R}$, such that $\mathcal{P} \approx \mathcal{R}\mathcal{R}^*$.

Compute $v = (\mathcal{A} - \mu_1 \mathcal{E})^{-1}\mathcal{B}$, $V_1 = \frac{v}{\|v\|}$ for $\mu_1 = \mu_0^{(1)}$
while *(not converged) or* $m \leq i_{max}$ **do**
  Find $\lambda = eig(V_m^* \mathcal{E} V_m, V_m^* \mathcal{A} V_m)$ and select new shift μ_{m+1} using Algorithm 1.
  Find $v = (\mathcal{A} - \mu_{m+1}\mathcal{E})^{-1} V_m$
  Orthogonalize v against V_m to obtain v_{m+1}, $V_{m+1} = [V_m, v_{m+1}]$
  Solve the small Lyapunov equation
$$\mathcal{A}_{m+1} Y \mathcal{E}_{m+1}^T + \mathcal{E}_{m+1} Y \mathcal{A}_{m+1}^T = -\mathcal{B}_{m+1}\mathcal{B}_{m+1}^*,$$
  for Y where $\mathcal{A}_{m+1} = \mathcal{V}_{m+1}^* \mathcal{A} \mathcal{V}_{m+1}$, $\mathcal{E}_{m+1} = \mathcal{V}_{m+1}^* \mathcal{E} \mathcal{V}_{m+1}$ and $\mathcal{B}_{m+1} = \mathcal{V}_{m+1}^* \mathcal{B}$.
  Compute the norm of the residual as defined in (14)
Compute eigenvalue decomposition
$Y = T \Lambda T^* = \begin{bmatrix} T_1 & T_2 \end{bmatrix} \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix} \begin{bmatrix} T_1^* \\ T_2^* \end{bmatrix}$
Truncate Λ_2 if the eigenvalues are sufficiently small and construct $\mathcal{R} = V_{m+1} T_1 \Lambda_1^{\frac{1}{2}}$

computational task of this algorithm is to compute the shift parameter μ_i at the i-th iteration which is crucial for better convergency of the algorithm. The authors in [10] also show an adaptive shift parameter computation approach which is computationally efficient. In this paper the authors suggested to consider the initial shifts $\mu_0^{(1)}, \mu_0^{(2)}$ from the spectral region of $(\mathcal{A}, \mathcal{E})$. Then the next shift μ_{m+2} is selected by applying Algorithm 1. In this algorithm the values λ_k , $k = 1, 2, ..., m$ are the eigenvalues of pair $(\mathcal{A}_m, \mathcal{E}_m)$. More details and computational strategy can be found in [10]. A complete procedure to find the low-rank Gramian factor by solving the Lyapunov equation (7) is summarized in Algorithm 2. However, this algorithm can be used for solving the Lyapunov equation (7) for computing the low-rank Gramian factor $\mathcal{L}$. In that case, the inputs $\mathcal{E}$, $\mathcal{A}$ and $\mathcal{B}$ are replaced by respectively, $\mathcal{E}^T$, $\mathcal{A}^T$ and $\mathcal{C}^T$.

III. Solution of the Lyapunovs and Model Reduction of the index-1 DAEs

This section discusses the MOR of index-1 descriptor system (4). To perform the balanced truncation of this system we can first convert this into the generalized state space system (3) which gives exactly same structure of the system as in (6). Therefore, we can apply the procedure as discussed in Section II. We already have mentioned (in Section I) that system matrices in (3) are dense and hence, they lead to huge computational complexities in the implementation. Consequently, in the following we show that how to avoid this formulation explicitly to perform the BT method for the system (3).

We first need to solve the Lyapunov equations (1) and (2) to compute the low-rank controllability (R) and observability (L) Gramian factors, respectively. To solve (1) for computing R we can follow the Algorithm 2 replacing the inputs ($\mathcal{E}$, $\mathcal{A}$, $\mathcal{B}$) by (E, A, B). A close observation reveals that in each iteration of this algorithm we have to solve a shifted linear system like

$$(A - \mu E)v = V, \tag{15}$$

which is equivalent to the solution of the linear system

$$((A_1 - A_2 A_4^{-1} A_3) - \mu E_1)v = V.$$

Recalling the strategy in [5], instead of solving the above linear system we can solve the linear system

$$\begin{bmatrix} A_1 - \mu E_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} v \\ \Lambda \end{bmatrix} = \begin{bmatrix} V \\ 0 \end{bmatrix}, \tag{16}$$

for v. This is important, because, although the linear system in (16) has a larger dimension than that of (15), it is sparse (see, e.g, [3, Chapter 3]) and can efficiently be solved by suitable direct [11] or iterative solvers [12].

We can use the same algorithm for solving the Lyapunov equation (2) to compute the L. In that case input ($\mathcal{E}$, $\mathcal{A}$, $\mathcal{B}$) is changed to (E^T, A^T, C^T). Again in each iteration we solve a linear system $(A^T - \mu E^T)w = C^T$. Which is equivalent to solve the linear system

$$\begin{bmatrix} A_1^T - \mu E_1^T & A_3^T \\ A_2^T & A_4^T \end{bmatrix} \begin{bmatrix} w \\ \Lambda \end{bmatrix} = \begin{bmatrix} C_1^T \\ C_2^T \end{bmatrix}, \tag{17}$$

for w.

Once we have R and L, the balancing and truncating transformations T_l and T_r can be constructed following the procedures (9) - (10). Then applying these transformations we construct the ROM

$$\begin{aligned} \hat{E}\dot{\hat{x}}(t) &= \hat{A}\hat{x}(t) + \hat{B}u(t), \\ \hat{y}(t) &= \hat{C}\hat{x}(t) + \hat{D}u(t), \end{aligned} \tag{18}$$

where the reduced matrices can be formed as:

$$\begin{aligned} \hat{E} &:= T_l^T E_1 T_r, \\ \hat{A} &:= \hat{A}_1 - \hat{A}_2 A_4^{-1} \hat{A}_3, \quad \hat{B} := \hat{B}_1 - \hat{A}_2 A_4^{-1} \hat{B}_2, \\ \hat{C} &:= \hat{C}_1 - \hat{C}_2 A_4^{-1} \hat{A}_3, \quad \hat{D} := D_a - C_2 A_4^{-1} B_2, \end{aligned} \tag{19}$$

in which $\hat{A}_1 = T_l^T A_1 T_r$, $\hat{A}_2 = T_l^T A_2$, $\hat{A}_3 = A_3 T_r$, $\hat{B}_1 = T_l^T B_1$ and $\hat{C}_1 = C_1 T_r$.

IV. Numerical Results

To assess the performance of our proposed method, we apply this to several test-systems derived from Brazilian Interconnected Power System (BIPS) models introduced in [5]. The number of the differential and algebraic variables and the input-output are given in Table I.

Model	n	n_a	inputs/outputs
Mod-1	606	6 529	4/4
Mod-2	1 142	8 593	4/4
Mod-3	3 078	18 050	4/4

TABLE I: Number of the differential and algebraic variables and the inputs/outputs for the different models.

First we compute the low-rank controllability and observability Gramian factors for all the models by solving their respective Lyapunov equations (as given in (1) and (2)) applying the RKSM method. Then the reduced models are computed for all the data sets using the BT. The comparisons of the original and reduced models are depicted in Table II showing their absolute and relative deviations. Note that for all data the truncation tolerance is set to 10^{-2}. However, the dimension of the ROM can be decreased or increased by changing (increasing or decreasing) the tolerance if desired or required. Due to the lack of space we exemplary present only selected plots (e.g., for the largest model Mod-3).

Model	Dimension		Error	
	full	reduced	absolute	relative
Mod-1	7 135	86	2×10^{-2}	1.7×10^{-3}
Mod-2	9 735	89	9×10^{-4}	3×10^{-4}
Mod-3	21 128	93	7×10^{-4}	4.3×10^{-4}

TABLE II: Comparisons of the full and reduced models.

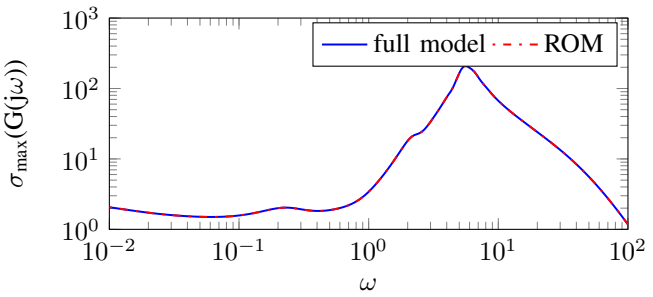


Fig. 1: Sigma plot (maximum singular values) of full and reduced order models.

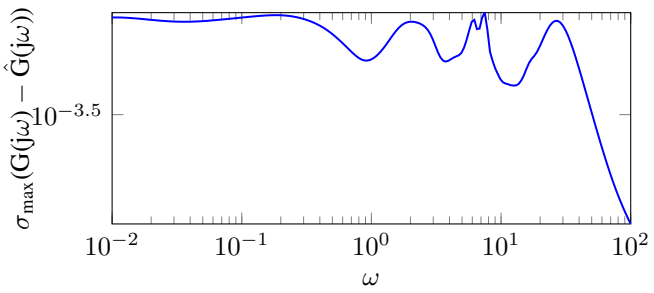


Fig. 2: Absolute error in the sigma plot of full and reduced order models.

Figure 1 shows the frequency responses (largest singular value of $G(j\omega)$) of full and the 93 dimensional reduced order models in frequency domain over the frequency (ω) range of 10^{-2} to 10^{2}. In Figure 2, the absolute error between the frequency responses of full and reduced models is shown. The relative deviation of lower dimensional models from full models are depicted in Figure 3 which is below 10^{-3}. From these figures we can conclude that the generated ROM ensures good quality and hence they can be used for controller design, simulation or optimization.

V. Conclusions

We have presented an efficient method for low-rank Gramian factors by solving the Lyapunov equations which

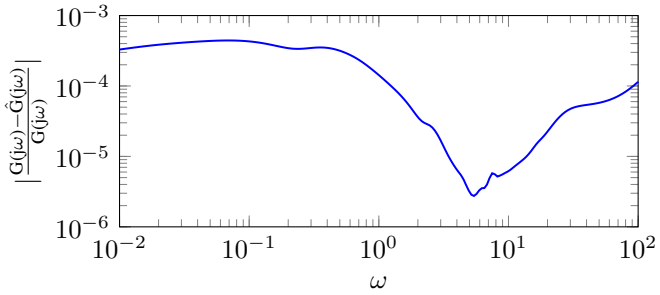


Fig. 3: Relative error in the sigma plot of full and reduced order models.

arise from the large-scale sparse index-1 descriptor systems. The technique was based on the iterative rational Krylov subspace method (IRKM) in [1] which is originally discussed for standard state space model. We have shown that eliminating the algebraic variables the descriptor system can be converted into the standard state space model which, however, is not possible for large-scale model due to memory restriction of the computer and additional computational complexity. In this paper we have shown that this dense formulation of the dynamical is not required to compute explicitly. The computed low-rank Gramian factors have been applied to balancing based model reduction of the index-1 DAEs. We have applied the method to several test-systems derived from Brazilian Interconnected Power System (BIPS) models found in [5]. Numerical results also have been discussed to show the efficiency and accuracy of our proposed method. As future works we would like to extend the idea for higher index DAEs. We also want to investigate how to compute the shift parameters used in the IRKM method efficiently.

References

[1] V. Druskin, L. Knizhnerman, and V. Simoncini, "Analysis of the rational Krylov subspace and ADI methods for solving the Lyapunov equation," *SIAM J. Numer. Anal.*, vol. 49, pp. 1875–1898, 2011.

[2] P. Kunkel and V. Mehrmann, *Differential-Algebraic Equations: Analysis and Numerical Solution*, ser. Textbooks in Mathematics. EMS Publishing House, 2006.

[3] M. M. Uddin, "Model reduction for piezo-mechanical systems using Balanced Truncation," Master's thesis, Stockholms University, Stockholm, Sweden, 2011.

[4] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*. Philadelphia, PA: SIAM Publications, 2005.

[5] F. Freitas, J. Rommes, and N. Martins, "Gramian-based reduction method applied to large sparse power system descriptor models," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1258–1270, Aug. 2008.

[6] P. Benner, V. Mehrmann, and D. C. Sorensen, *Dimension Reduction of Large-Scale Systems*, ser. Lect. Notes Comput. Sci. Eng. Springer-Verlag, Berlin/Heidelberg, Germany, 2005, vol. 45.

[7] T. Penzl, "A cyclic low rank Smith method for large sparse Lyapunov equations," *SIAM J. Sci. Comput.*, vol. 21, no. 4, pp. 1401–1418, 2000.

[8] A. C. Antoulas, D. C. Sorensen, and Y. Zhou, "On the decay rate of Hankel singular values and related issues," *Systems Control Lett.*, vol. 46, no. 5, pp. 323–342, 2002.

[9] M. S. Tombs and I. Postlethwaite, "Truncated balanced realization of a stable nonminimal state-space system," *Internat. J. Control*, vol. 46, no. 4, pp. 1319–1330, 1987.

[10] V. Druskin and V. Simoncini, "Adaptive rational Krylov subspaces for large-scale dynamical systems," *Systems and Control Letters*, vol. 60, no. 8, pp. 546–560, 2011.

[11] T. A. Davis, *Direct Methods for Sparse Linear Systems (Fundamentals of Algorithms 2)*. Philadelphia, PA, USA: SIAM, 2006.

[12] Y. Saad, *Iterative Methods for Sparse Linear Systems*. Philadelphia, PA, USA: SIAM, 2003.

Cascading of Dual Drive Mach Zehnder Modulator and Electro Absorption Modulator for producing an Optical Comb

Rifat Nazneen,[1,*] Ebad Zahir,[2]

[1,2]Department of Electrical and Electronic Engineering
American International University- Bangladesh (AIUB)
Dhaka, Bangladesh
[*]syeda.rifat18@gmail.com

Abstract—**A new scheme for generatinga flat optical frequency comb using three Laser sourcesbased on cascading of Dual Drive Mach Zehnder Modulator (DD-MZM) and Electro Absorption Modulator (EAM) through an optical nonlinear fiber loop is explored in this paper. A single RF signal of 25 GHz having different amplitudes feed into the two arms of DD-MZM and then feed the signal into EAM to modulate the pulse train. By controlling the DC bias of both modulators as well as the RF signal frequency, the scheme is successful in generating evenly spaced 15 flat comb lines among 25 lines with a spectral power variation of about 2dB.The paper compares the performance of the proposed scheme against other methods where only one modulator was used. The scheme is also analysed for producing an optimum number of spectral lines based on number of sources, channel spacing and RF signal frequency. Finally, the performance of the system was successfully evaluated in a preliminary WDM setup of five channels/modes at 10Gb/s for a channel of 1 km length. Results in the form of BER and Q factor are presented in this paper.**

***Index Terms*—Optical Frequency Comb; Four Wave Mixing; Highly Nonlinear Fiber; Dual Drive Mach Zehnder Modulator;Electro Absorption Modulator;Fiber optic Communication; Wavelength Division Multiplexing.**

I. Introduction

An optical comb having evenly spaced channels with equal flat spectrum power and large bandwidth is desired for optical communication systems as well as other applications. Optical frequency comb with evenly spaced and flat comb lines are applicable in metrology [1], arbitrary frequency optical millimeter wave generation [2], spectroscopy, generation of high pulse train, dense wavelength division multiplexing (DWDM) [3], etc. Crosstalk, inter-symbol interference, noise between channels can be reduced through equal spacing and flattening of the spectral lines. There are many purposes to generate an evenly spaced large number of frequency comb such as gain switching, mode locked laser, nonlinearity effect in optical fibers [4], re-circulating shifter loop [5] etc. but it is difficult to flatten the spectrum. Previously simulated designs have shown the generation of evenly spaces 27 comb lines [6] using phase modulators resulting in uneven gain. On the other hand, cascading of intensity modulators and phase modulators can generate a limited number of lines with poor flatness due to the limited amplitude of RF signals in the phase modulator and limited phase deviation [7]. Recent research has shown that by using a single drive dual parallel Mach Zehnder Modulator (MZM) with two different amplitudes, RF signals have been used to generate 5-7 flat lines in 18dbm [8]. Also two cascaded intensity modulators have generated 9 flat spectral lines in 2dbm power deviation [9]. And, by incorporating the optoelectronic oscillation loop with cascaded modulators in optical comb generator has produced 11 ultra flat lines [10].

This paper presents an alternative approach than the methods mentioned above which can generate 15 flat comb lines with only 2dbm power deviation using two25GHz RF drive signals. This scheme shows promising results applicable towards WDM systems by cascading a DD-MZM and an EAM in nonlinear optical feedback loop to generate the flat comb lines. As the DD-MZM can modify both the amplitude and the phase of the modulated signal, there is no phase modulator required in the scheme. The scheme was evaluated for number of comb lines and flatness of power deviation by varying the number of sources with different channel spacing. Performance evaluation was carried out by testing the BER and Q factor for each of the five modes after introducing a 10Gb/s signal for 1km transmission length which justifies the use of the modes as individual channels for a Wavelength Division Multiplexing (WDM) system. Finally, the system was optimized for the maximum transmission length.

II. DD-MZM & EAM

ADD-MZM is a two arm modulator where an electric field is applied to control the phase and if the contribution in each arm is properly controlled then it becomes an intensity modulator. It requires two electrical signals, each driving the modulator at half the voltage of a single-drive device [11]. Normal operation consists of applying half of the required electrical amplitude on each RF port, RF + and RF-, with a π phase shift between them;

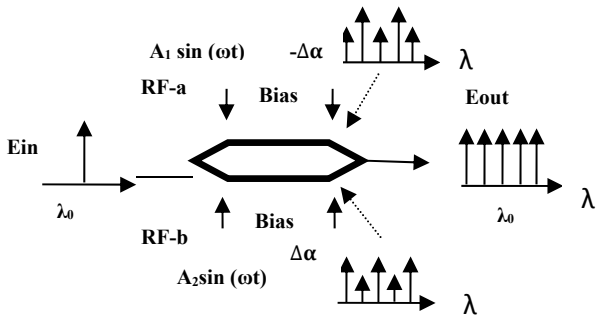


Fig.1. General Scheme of optical frequency comb generation using a MZM

In Fig. 1, the two electrical RF signals RF-a and RF-b are applied on the two arms of the MZM with two different amplitudes of $A_1\sin(\omega t)$, $A_2\sin(\omega t)$ along with a phase shift. These two arms are biased with $-\Delta\alpha$ and $+\Delta\alpha$ so that the amplitude can be controlled. Eout can be expressed as;

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

$$E_{out}(t)=\frac{E_{in}(t)}{2}\sum_{k=-\infty}^{+\infty}(J_k(A_1)\ell^{J(kwt+\alpha_1)}+J_k(A_2)\ell^{J(kwt+\alpha_2)} \quad \text{.... (1)}$$

By simplifying equation (1), it can define $2\Delta A$ as the peak-to-peak phase difference induced in each arm;

$$\Delta A = \frac{A_1 - A_2}{2} \quad \text{..... (2)}$$

Dc bias difference between the arms $2\Delta\alpha$ can be defined from equation (3)

$$\Delta\alpha = \frac{\alpha_1 - \alpha_2}{2} \quad \text{..... (3)}$$

Combining equations (2) & (3) can be used to generate a flat optical frequency comb;

$$\Delta A + \Delta\alpha = \frac{\pi}{2} \quad \text{..... (4)}$$

Also the MZM is based on electro-optic effect, which means that there fractive index, n, of the material changes with respect to the voltage, V, applied across the arms as well as the phase change in the signals [12].This relationship can be defined by equation (5):

$$\Delta\alpha = \frac{2\pi}{\lambda}\Delta nL \quad \text{..... (5)}$$

EAM is another single drive external modulator which drives the RF signal with a specific bias voltage on the plane of the waveguide [12]. The RF signal is injected on the EAM to absorb the unwanted signals and modulate the desired signal according to the operating wavelength. The general equation for EAM to modulate the output signal depends on bias voltage E(v) and can be defined by equation (6):

$$E(v) = I(v)^{(1+j\alpha_p(v)/2} \quad \text{..... (6)}$$

Four–wave mixing (FWM) is the nonlinear effect which is present in the nonlinear feedback loop. It is used to generate the optical frequency comb using a wavelength conversion method. If two or more light sources are injected in the fiber which has nonlinear effects, they generate a new wave at a new wavelength [13] as expressed below in equation (7).

$$f_{new} = f_1 + f_2 - f_3 \quad \text{..... (7)}$$

Where, f_1& f_2 = pumped waves, f_3= weak signal or probe light, andf_{new}= new photon at new wavelength or idler wavelength.

III. PROPOSEDMETHODOLOGY

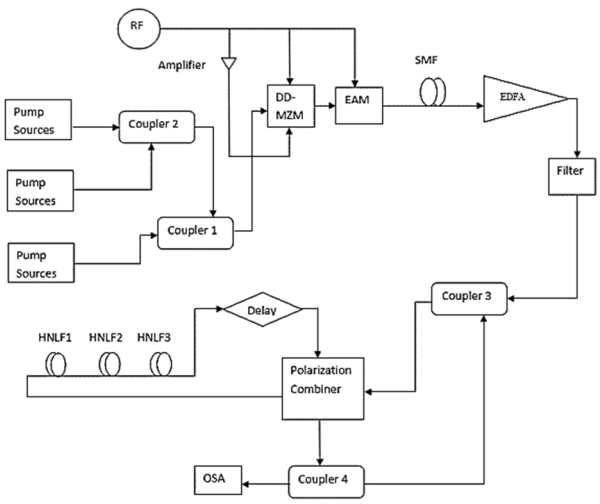


Fig.2. Proposed Methodology for flat optical comb generation

As shown in Fig.2, three pump sources of 1551nm, 1552nm and 1553nm are coupled using 3dm couplers of 10MHz line width and then passed through a cascade of both DD-MZM & EAM modulators. The DD-MZM is driven by a RF signal of 25GHz frequency which is split and introduced into the two arms of the modulator. Since the two RF signals have to be of different amplitudes, one of them is first amplified by -1dB gain before it is injected in one arm. The biasing voltages used in these two arms are -4.65V and -3.1V.As the DD-MZM operates as both an intensity and a phase modulator, there is no need to add any phase shifter. The RF signal is also injected into the EAM modulator which modulated the spectrum with a 1.6V peak to peak voltageto modify the sideband signals and flatten the spectrum.

After this spectrum travels through a 1km long single mode fiber (SMF) to suppress Brilloiuin scattering, it is amplified by the EDFA by 1W. Here the spectrum picks up amplified stimulated noise (ASE), so it is then passed through a filter to get rid of this unwanted noise. Couplers 3 and 4 are using 9% coupling ratio for feedback into the polarization combiner[13].

The filtered spectrum then finally passes through the feedback loop of 0.03km long highly nonlinear fibers(HNLF) with the polarization combiner. The three HNLFs have zero dispersion wavelengths (ZDW) at 1600nm, 1550nm, and 1500nm as well as an additional delay inserted. The effective area of all HNLFs are A_{eff}=64μm^2. Finally, the polarization combiner initiates the FWM phenomenon which is assisting in creating the proposed spectral lines.

IV. RESULTS AND DISCUSSION

Fig.3 shows the results while using only a phase modulator. The system is able to generate about 15 lines but with a high degree of fluctuation in the amplitudes of each line. Such a response would create a lot of problems in a communication system.

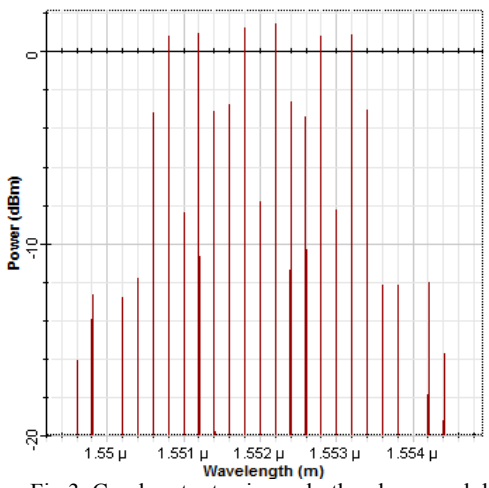


Fig.3. Comb output using only the phase modulator

Fig.4 illustrates the result using the cascading of single drive MZM and EAM. It is evident here that by cascading the two modulators, the resultant comb has less lines but the amplitude deviation is only about 2dBm.

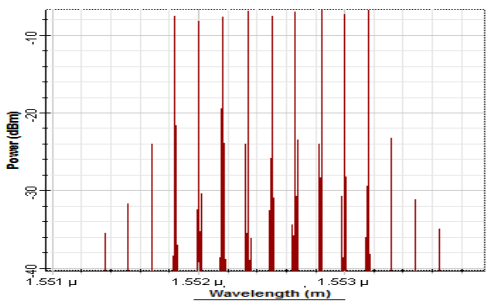


Fig.4. Response with cascading of single drive MZM and EAM

Fig.5 is displaying the output for the proposed scheme of cascading the DD-MZM and EAM. Here the number of comb lines is the same as Fig.3 while the amplitude deviation is only 2dBm which is the same as Fig.4. Clearly, it has the best part of both aforementioned scenarios.

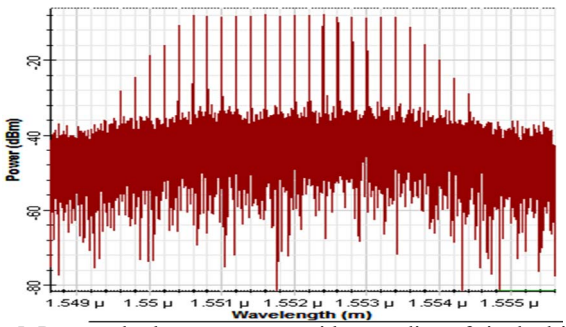


Fig.5. Proposed scheme response with cascading of single drive MZM and EAM

V. Optimization Of The Comb Response

While investigating the system, it is only natural to question the impact of the critical parameters such as number of pump sources and the amount of channel spacing.

When the number of pump sources were varied by keeping the channel spacing at 1nm, Fig.6 shows the correlation. It is evident that the system is yielding the best response when 5 pump sources are used. At this point, it is producing 23 flat spectrum lines.

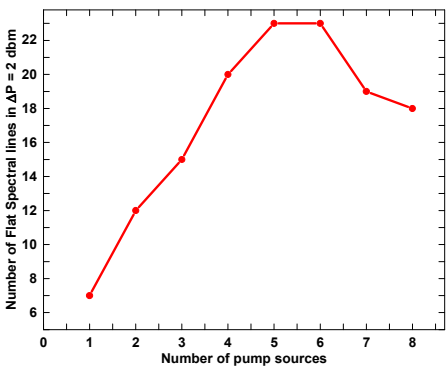


Fig.6. Correlation of number of pump sources on comb spectrum

As seen in Fig.7, varying the channel spacing between the sources increases the spectral lines as well as having an impact on the flatness of the power deviation. When 1nm channel spacing is used, 15 spectrum lines are generated with 2dBm power deviation.

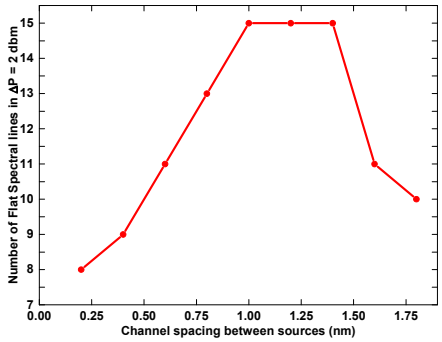


Fig.7. Correlation of flat spectrum lines and channel spacing

The driving RF signals bandwidth also has an impact on the flatness of the comb spectrum. Fig.8.shows that the maximum of 15 flat spectrum lines can be generated using 25GHz.

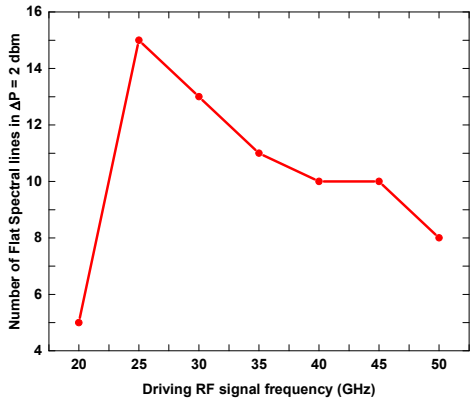


Fig.8. Correlation between driving RF signals and flat spectral lines.

VI. Using the Comb as Message Carriers

Previously the comb generator operated for 25 GHz RF signal to generate maximum and flat comb lines in small power deviation. Now, in the propose scheme when the 10Gb/s message signal is inserted, the shape of the response changes. Fig.9. shows all the comb lines carrying the same identical message.

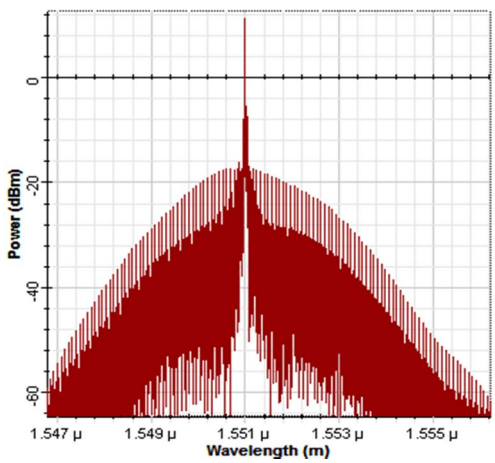


Fig.9. Response of proposed scheme with 10Gb/s message

The comb lines then tested as individual channels carrying the 10Gb/s signals evaluating the Eye Diagram, BER and Q factor. After a transmission of 1km distance, the results shown in Fig 10-12 are quite promising for the first 5 modes.

It is evident in Fig.10 that even though all tested channels have acceptable Eye Diagrams, the performance of each individual channel is somewhat different due to the nonlinear behaviour of the FWM.

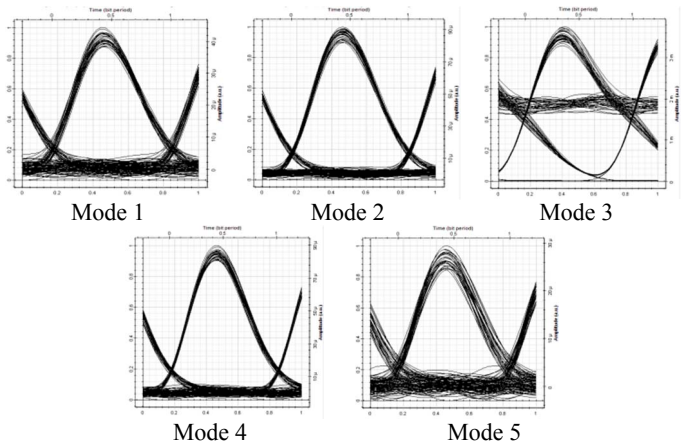


Fig.10. Eye diagrams of 5 modes for 1km transmission

Although the Q factor is inversely proportional to the length of the transmission line, the nonlinear behaviour of the FWM is showing an interesting response for the first 5 channels in Fig. 11. When the response of the first 5 modes are compared with an increase in transmission length, the best combined result is evident at 1km fiber length. This proves that the effect of FWM on each mode is different and thus the performance needs to be optimized based on the number of channels.

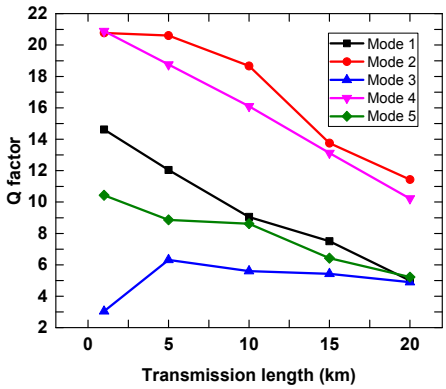


Fig.11. Variation of Q factor according to transmission length

A third performance parameter, BER, was also compared for the 5 modes and the results showed a similar behaviour in Fig. 12where the combined best results were at 1km transmission.

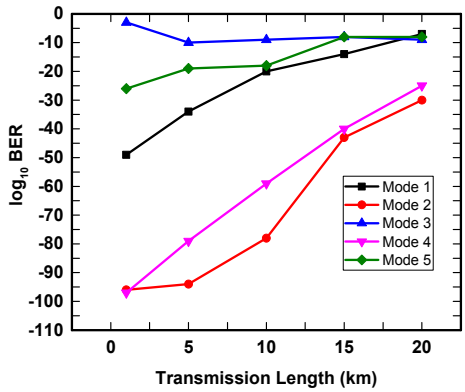


Fig.12. Variation of BER according to transmission length

VII. Conclusion

This paper has presented a novel scheme using cascaded DD-MZM and EAM to produce an optical comb spectrum applicable for a fiber optic communication systems using WDM. The scheme has been characterized based on critical parameters such as number of pump sources, channel spacing and RF bandwidth showing optimum performance of 15 flat comb lines, equally spaced with a minimal power deviation of 2dB. Promising performance results using the scheme in a basic WDM system has also been presented for a transmission length of 1km.

References

[1]. A. Antikainen and G. Agrawal, "Dual-pump frequency comb generation in normally dispersive optical fibers", Journal of the Optical Society of America B, vol. 32, no. 8, p. 1705, 2015

[2]. Yang, Yang et al. "Optical Frequency Comb Generation using two Cascaded Polarization Modulators". Photonic Network Communications (2015)

[3]. Yang, Ting et al. "Comparison Analysis of Optical Frequency Comb Generation with Nonlinear Effects In Highly Nonlinear Fibers". Opt.Express 21.7 (2013): 8508.

[4]. Jianrui Li, Jianrui Li et al. "Scheme for Generation of flat top and high signal-to-noise ratio optical frequency comb". Chinese Optics Letters 13.1 (2015): 010605-10610.

[5]. Nagarjuna, K.P, Shankar Kumar Selvarajaa, and V.R Supradeepa."Generation Of Tunable, High Repetition Rate Frequency Combs With Equalized Spectra Using Carrier Injection Based Silicon Modulators". Proc. of SPIE, vol. 9752, SPIE (2016)

[6]. R.Nazneen, E. Zahir, "Analysis and Optimization of an Optical Frequency Comb for Multiple Channel Transmission in a Fiberoptic System"3rd International Conference on Electrical Engineering and Information Communication Technology (iCEEiCT 2016) at MIST, Mirpur Cantonment, Dhaka

[7]. Zhang, Fangzheng et al. "Flat Optical Frequency Comb Generation and Its Application For Optical Waveform Generation". Optics Communications 290 (2013): 37-42.

[8]. Shang, Lei et al. "A Flexible And Ultra-Flat Optical Frequency Comb Generator Using A Parallel Mach–Zehnder Modulator With A Single DC Bias". Optics Communications 356 (2015): 70-73.

[9]. X. Zhou, X. Zheng, H. Wen, H. Zhang, Y. Guo, B. Zhou, All optical arbitrary waveform generation by optical frequency comb based on cascading intensity modulation, Opt. Commun. 284 (2011) 3706–3710.

[10] Dai, Jian et al. "Self-Oscillating Optical Frequency Comb Generator Based on an Optoelectronic Oscillator Employing Cascaded Modulators". Opt. Express 23.23 (2015): 30014.

[11] Morant, Maria et al. "Dual-Drive Linbo_3 Interferometric Mach-Zehnder Architecture with Extended Linear Regime For High Peak-To-Average OFDM-Based Communication Systems". Opt. Express 19.26 (2011): B452.

[12] Kumar, Shiva and M. Jamal Deen. Fiber Optic Communications. Hoboken: Wiley, 2014.

[13]. Dong, Qimeng et al. "Flat Frequency Comb Generation Based On Efficiently Multiple Four-Wave Mixing Without Polarization Control". Photonic Sens 6.1 (2015): 85-89.

Analysis of mechanical and micro-structural properties of trabecular bones with statistical review of osteoporosis among females in Bangladesh

Atia Rahman,[1,*] Shamimur Rahman Akanda,[2] Raihan Hussain,[3] Monon Mahboob,[4] Md. Ashiqur Rahman[5]

[1]Student, Department of Mechanical Engineering, BUET, Dhaka-1000, Bangladesh.
[2]Student, Department of Mechanical Engineering, BUET, Dhaka-1000, Bangladesh.
[3] Professor, National Institute and Nuclear Medicine & Allied Sciences, BSMMU, Dhaka, Bangladesh.
[4]Assistant Professor, Department of Mechanical Engineering, BUET, Dhaka-1000, Bangladesh.
[5]Assistant Professor, Department of Mechanical Engineering, BUET, Dhaka-1000, Bangladesh.
atiarahman1110170@gmail.com

Abstract—**Osteoporosis is one of the most notable causes of bone fracture and approximately 200 million women are affected by osteoporosis worldwide. According to a report of International Osteoporosis Foundation, at least one in five men and one in three women in Bangladesh suffer a bone fracture due to osteoporosis in their remaining lifetime. Limited data is available in the open literature on osteoporosis among women in Bangladesh and more so on the micro-structural analysis of osteoporosis. The present study particularly aims at identifying the trend of osteoporosis of women of in Bangladesh, with special emphasis on the mechanical micro-structural analysis of female bones. A comprehensive statistical analysis has been conducted by collecting data of 250 Bangladeshi females with age ranging from 20 to 80 years. Trabecular bones of Tibia of female have been analysed with computed tomography for micro-structural assessment. To characterize several parameters of trabecular bones, the computed tomography scans have been analysed and parameters like bone volume fraction, fractal dimension etc. have been evaluated. A correlation between structural properties of the bones has been formed by multiple regression analysis. It shows a positive correlation between fractal dimension and bone volume fraction of the specimens.**

Index Terms—**Osteoporosis, BMD, Multiple Regression Analysis, Micro-structural Properties, Women In Bangladesh.**

I. Introduction

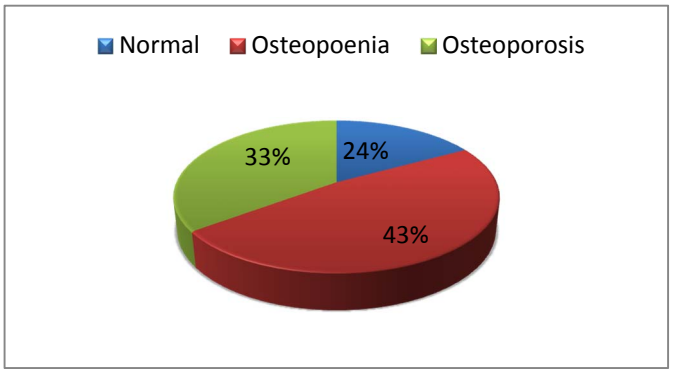
Osteoporosis is an under-rated and often overlooked disease from a long time, especially in our country. It is a bone disease that occurs when the body loses too much bone or makes too little bone or both. It is a skeletal disorder caused by compromised bone strength leading to an increased risk of fracture. Bone density is a parameter which is used in World Health Organization (WHO) operational definition of osteoporosis [1]. The WHO uses T-scores to define normal bone mass, low bone mass (or osteopenia), and osteoporosis. T-score compares bone density to the average bone density of young healthy adults of same gender [2]. Normal and healthy condition of bone is ensured if T-score lies between +1 and -1. Low bone mass is reported and is termed as osteopenia when T-score falls between -1 and -2.5. A T-score greater than -2.5 indicates the possibility of osteoporosis.

Although being a common disease, it had gained less attention all over the world and has become a public health issue in the recent past [1,3]. Bone mineral density is expressed by T-score [4]. Multivariable regression models based on 16-45 and 46-65 year old women showed that age and BMI are were associated with osteopenia/osteoporosis at lumbar spine, femoral neck or either of the 2 sites [5]. S. Reza et al. [6] reported statistical analysis of bone mineral densities at the lumbar spine and femoral neck along with hip axis for 336 females where 5 years age band has been used. In this study it was found that BMD and mean hip axial length were both started declining at the band 45-49. For accurate prediction of osteoporosis quantification and analysis of bones is necessary .Although bone mineral density is mainly used for prediction of osteoporosis, only bone mineral density cannot predict mechanical properties fully [7]. Three dimensional micro structural properties provide better information for understanding bone quality [8-12]. Trabecular bone is a three dimensional solid foam situating at the epiphysis of long bone [13-14]. Computed Tomography is a useful tool on non-destructive technique and may be used for quantification of the three dimensional structure of bone. Different mechanical tests can be performed to determine mechanical behaviours of trabecular bones [14]. In the present study, a comprehensive statistical analysis has been performed showing relationship between T-score of trabecular bones and parameters like age, weight etc. An attempt has been made to find out correlations between structural properties, mechanical properties and age of six trabecular bones.

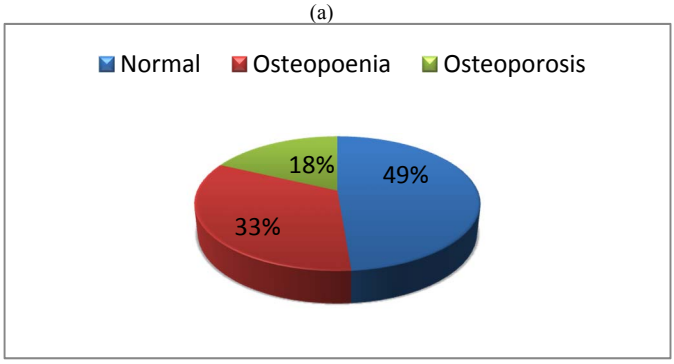
II. Statistical Analysis

DEXA scans of 250 female patients have been analyzed which were collected from The Nuclear Medicine Department of BSMMU. The data consisted of age, weight, height, Bone mineral Index and T-score of Lumbar Spine. The model of the DEXA machine is Norland 4.5. From Fig. 1(a), distribution of all the test subjects according to T-score shows that most of the patients had osteopenia (43%) then osteoporosis (33%) and rest were found normal (24%).

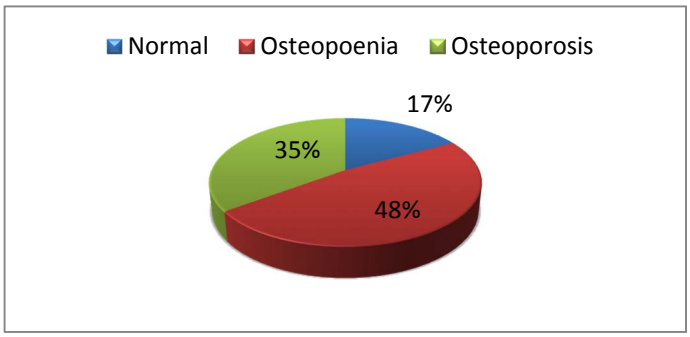
978-1-5090-2963-1/16/$31.00 ©2016 IEEE



(a)



(b)



(c)

Fig. 1 Distribution of test subjects according to T-score (a) in overall lifetime, (b) in premenopausal stage and (c) postmenopausal stage.

From Fig. 1(b) it can be found that the probability for osteoporosis in premenopausal phase (age below 45) is around 18% and in postmenopausal phase (age above 45) is around 35% which is almost double the value of the previous one.

This is due to lack of estrogen and after menopause bone resorption (breakdown) overtakes the building of new bone.

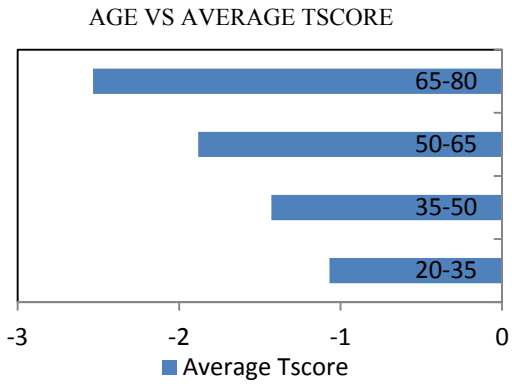
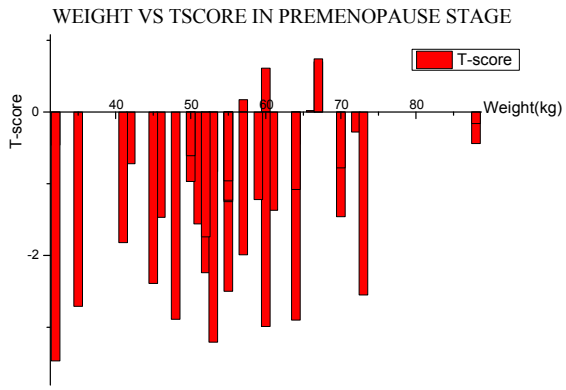


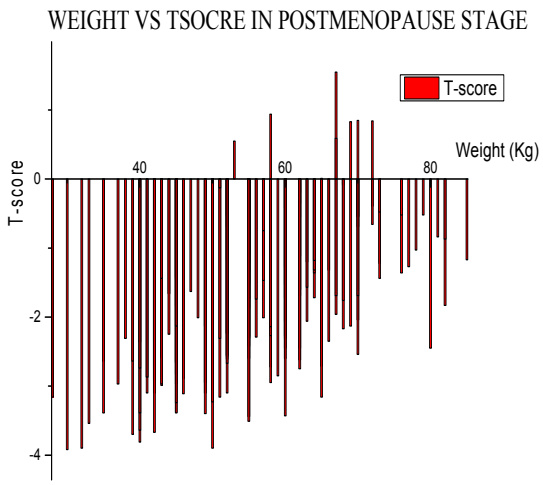
Fig. 2 Average T-score in different age groups(20-35,35-50,50-65 and 65-80).

It is clearly evident that aging accelarates the loss of bone from Fig. 2. During development and aging, bone shape changes in response to load and the changes are due to morphological changes which are related to genetics, the loading of the bones, and the activity of the cells [15]. These morphological changes widely affect the content of BMD in elderly people and reduce the BMD level and there fore decreasing T-score.

For both the case of premenopausal and postmenopausal stages from weight vs T-score column chart in Fig. 3(a) and 3(b) it has been seen that for that of with the increase of weight T-score is increasing. Weight, BMI and BMD had a positive correlation, which means low weight is much more likely to be related to osteoporosis with no other factor considered [16].



(a)



(b)

Fig. 3 Distribution of test subjects according to T-score (a) in premenopausal stage and (b) in postmenoapausal stage.

III.CT SCAN AND IMAGE PROCESSING

CT scan of six samples of female tibia bones of different age groups has been performed in a high resolution PET-CT scan device. The intensity was of 39 mA, voltage of 120 kV and slice thickness of 3.75mm.

Around 102 slices were made for each of the bone. After Ct-scan evaluation the images were processed by ImageJ Software. Structural properties of tibia bones like Structural

Model Index (SMI), Bone Volume fraction (BV/TV), Bone Surface Fraction (BS/TV) and Fractal Dimension (FD) have been resolved successfully using BoneJ plug-in from ImageJ software. The slices were cropped equally having a fixed origin (X=159.18 & Y=309.57).The stacked slices were successfully reconstructed using 3d viewer plug-in to obtain a full framed 3d image. The SMI value for most of the specimen are about 2.5-3.0 ensuring rod type structure [17]. The 2D fractal dimension method was interpreted using box-counting method [18-19]. Descriptive statistics have been formed by mean value, standard deviation and range.

Fig. 4 Set up of a CT-scan image of a trabecular tibia bone with an intensity of 39 mA, 120 kV.

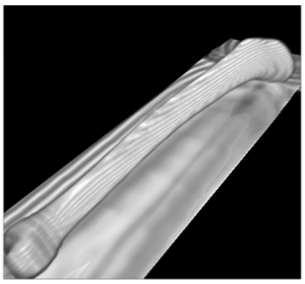


Fig. 5 A 3D image of a trabecular tibia bone from CT-scan in ImageJ software

Fig. 6 A cropped image of a slice of tibia (131*84) from CT-scan in ImageJ Software.

Fig. 7 Three point bending test of trabecular bones (Tibia) is performed by universal testing machine with a stroke rate between 1mm/min and 5 mm/min.

IV.RESULT

Differences between the parameters were assessed by two tailed test by Originlab-2016. The strength of the correlations between the respective structural properties and ages of the bones has been determined using a multiple analysis by stepwise analysis. For each specimen Pearson Correlation Coefficient was determined for examining correlation between each structural group and age of the bone.

The values presented here in the table are the structural properties obtained for 6 samples of tibia bone tissue. From this table, Fractal dimension (FD) is positively correlated with BV/TV with the value of coefficient of correlation=0.94614. The ages of the specimens are negatively correlated with BV/TV and FD. Modulus of elasticity of bones is strongly associated with fractal dimension with a value of coefficient of correlation of about 0.839. There is no other significant correlation found among other parameters. For evaluation mechanical properties modulus of elasticity of the bones have been determined by three point bending test. Moment of inertia of the bones have been deduced by using weighted average values obtained from ImageJ with the aid of BoneJ plugin. The scatter plot between fractal dimension and bone volume fraction has been shown in Fig. 8 which clearly shows that there is a positive correlation between fractal dimension and bone volume fraction.

TABLE I
VALUES OF THE STRUCTURAL PARAMETERS DESCRIBING GEOMETRY FOR EACH OF THE GROUPS TESTED

	Number of Specimen	**Mean**	**Standard Deviation**	**Range**
SMI	6	2.82883	0.70648	2.224-4.128
BV/TV	6	2.267	1.863	1.2-6
BS/TV	6	0.5	0.268	0.2-1
FD	6	1.87733	1.87733	1.819-2.002
Age	6	55.33	15.35	35-74
E (MPa)	6	3891.66	1639.84	1925-6406

TABLE II
THE PEARSON PRODUCT MOMENT CORRELATION COEFFICIENTS R, AMONG INDIVIDUAL STRUCTURAL PROPERTIES, $p<0.05$

	AGE	**BV/TV**	**BS/TV**	**SMI**	**FD**	**E**
AGE	1	-0.6643	0.10192	0.37754	0.73355	-0.984
BV/TV	-0.6643	1	0.17604	0.05084	0.94614	0.776
BS/TV	0.10192	0.17604	1	0.66615	0.32804	-0.184
SMI	0.37754	0.05084	0.66615	1	0.11219	0.3519
FD	0.73355	0.94614	0.32804	0.11219	1	0.839
E	-0.984	0.776	-0.184	0.3519	0.839	1

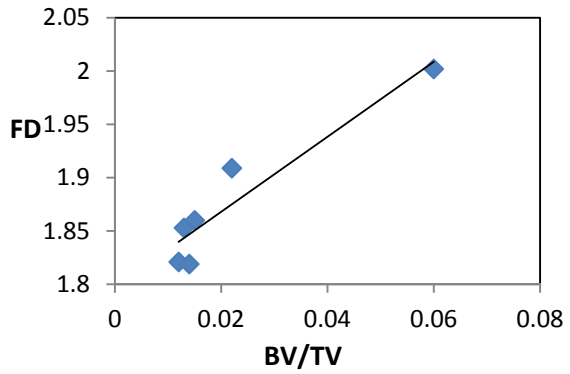


Fig. 8 Scatter plots for the parameters between FD (fractal dimension) and BV/TV (bone Volume Fraction) showing positive correlation.

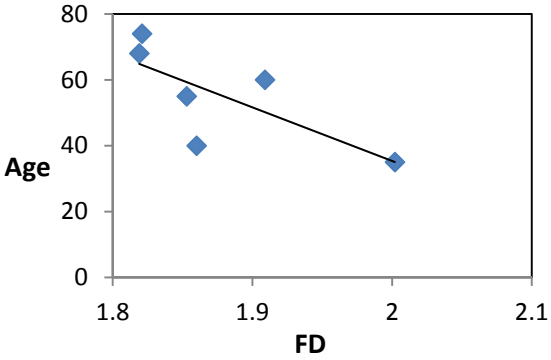


Fig. 9 Scatter plots for the parameters between age and fractal dimension showing negative correlation between them.

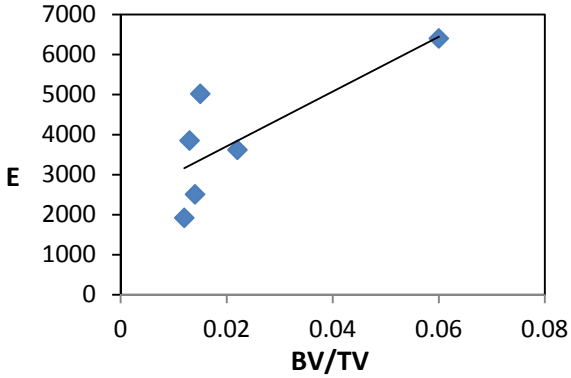


Fig. 10 Scatter plots for the parameters between modulus of elasticity and bone volume fraction showing positive correlation between them

Fig. 9 clearly indicates negative correlation between fractal dimension and age. Fig. 10 exhibits that modulus of elasticity is strongly associated with fractal dimension.

V. Conclusions

This work is aimed at working an indirect relation with the trend of age, weight and T-score and structural properties of trabecular bone. CT evaluation has been taken place in place of Micro CT evaluation. So, some errors are generated by image processing while revealing values of structural properties. Age and bone volume fraction are positively associated with fractal dimension. No other significant correlation has been found from this work. The main limitation is that some intrinsic properties of bone tissue like trabecular thickness, trabecular number of bones, trabecular spacing were not taken into account which was studied by other research teams [7-8]. There was scope to establish relationship between the structural properties and mechanical properties of respective bones.

Acknowledgement

The authors would like to express their gratitude to the Institute of Nuclear Medicine & Ultrasound, BSMMU, Dhaka for permitting us to perform the CT scanning of bones. We also sincerely acknowledge the guidance and valuable suggestions of Dr. Khondkar Siddique Rabbani, Chairperson, Department of Biomedical Physics & Technology, University of Dhaka.

References

[1] R Handa, A Ali Kalia ,G Maalouf , "Osteoporosis in developing countries",Elsevier,vol.22, p.693-670, Aug. 2008.

[2] S. A. Ferdoucy, M. A. Hussain M. N. Akter, M. S. Islam, M. A. Sadek,"Correlation between Degenerative Changes and Osteoporosis in Lumber Spine of Elderly women in Dhaka city, Bangladesh", City Dental College Journal, vol. 10, no.10, Jan. 2013.

[3] B. L. Riggs, L. J. Melton III, "The Prevention and Treatment of Osteoporosis", New Engl J Med, vol. 327, p. 620-627, Aug. 1992.

[4] T. R. Morrison, S. A. Russell, "The score index and bone mineral density assessments", BCMJ, vol. 42, p. 302-303, Aug. 2000.

[5] R. A. Begum,L. Ali, J. Akter, O. Takahashi, T. Fukui, R, M. Rahman, " Osteopenia and Osteoporosis among 16-65 year old women attending optpatient clinics", J Community Health, vol. 39, p. 1071-1076, March, 2014.

[6] S. Reza, M. Rahman, S. Hossain, S. Afroz, "Bone mineral density in normal Bangladeshi women", Iran. J. Radiat. Res. vol.6, No. 3, p. 157-160, winter 2008.

[7] H. Gong, M. Zhang, H. Y. Yeung, L. Qin, "Regional variations in micro structural properties vertebral trabeculae with aging", J Bone Miner Metab, vol. 23, p. 174-180, 2005.

[8] RW Goulet, SA Goldstein, MJ Ciarelli , JL Kuhn , MB Brown , LA Feldkamp, "The relationship between the structural and orthogonal compressive properties of trabecular bone", J Biomech, vol. 27,p. 375–389, April 1994.

[9] J. Homminga, BR. Mccreadie, H. Weinans, R Huiskes, "The dependence of the elastic properties of osteoporotic cancellous bone on volume fraction and fabric", J Biomech, vol. 36, p. 1461–1467, Oct. 2003.

[10] J. Kabel, B. van Rietbergen, A Odgaard, R Huiskes, "Constitutive relationships of fabric, density, and elastic properties in cancellous bone architecture", Bone, vol. 25, p. 1159–1164, Oct. 1994.

[11] T. M. Keaveny, E. F. Wachtel, CM. Ford, W. C. Hayes , "Differences between the tensile and compressive strengths of bovine tibial trabecular bone depend on modulus", J Biomech, vol. 27,p. 1137–1146, Sep. 1994.

[12] B. Van Rietbergen , A. Odgaard, J. Kabel, R. Huiskes, "Relationships between bone morphology and bone elastic properties can be accurately quantified using high-resolution computer reconstructions" J. Orthop Res, vol. 16, p. 23–28, Jan. 1998.

[13] J. GIbson, M.F. Ashby, *Cellular solids: Structure and Properties*, 2nd Ed, Cambridge University Press, Cambridge, 1997.

[14] A. Catarina Vale , M.F.C. Pereira, A. Maurício, P. Amaral, L.G. Rosa, A. Lopes, A. Rodrigues, J. Caetano-Lopes, B. Vidal, J. Monteiro, J.E. Fonseca, H. Canhão, M. Fátima Vaz, "Micro Computed Tomography and compressive characterization of Trabecular bone", Elsevier, vol. 438,p. 199-205, Dec. 2013.

[15] A.L. Boskey and R. Coleman, "Aging and Bone", Journal of Dental Research, vol. 89,p. 1333-1348, Dec. 2013.

[16] S. J. Kim, W.-G. Yang, E. Cho and E.-C. Park, "Relationship between weight, body mass index and bone mineral density of lumber spine in women", J Bone Metab, Vol. 19, p. 95-102, Nov. 2012.

[17] T. Hilderbrand, P. Ruegsegger, "Quantification of Bone Microarchitecture with the structural model index", Computer Methods in Biomechanics And Biomedical Engineering, vol. 1, p. 15-23, 1997

[18]]. A. NIkodem, "Correlations between structural and mechanical properties of human trabecular bone", Acta Bioeng Biomech Vol. 14, p. 37-46, 2012.

[19]]. N.L. Fazzalari, I.H. Parkinson, "Fractal Dimension and Architecture of Trabecular Bone", Journal of Pathology, vol. 178, p.100-105, Jan. 1996.

A hybrid deep model with HOG features for Bangla handwritten numeral classification

S M A Sharif, Nabeel Mohammed, Nafees Mansoor, Sifat Momen
Department of Computer Science and Engineering
University of Liberal Arts Bangladesh (ULAB)
Satmasjid Road, Dhaka, Bangladesh
Email:{sma.sharif.cse,nabeel.mohammed,nafees.mansoor,sifat.momen}@ulab.edu.bd

***Abstract*—Considering the practical significances, handwriting recognition is getting an intense interest to the research community. Through, several studies have been conducted for Bengali handwriting recognition, a robust model for Bengali numerals classification is still due. Therefore, a hybrid model is presented in this paper, which aims to classify the Bengali numerals more precisely. The proposed model bridges hand crafted feature extraction based approaches with the automatically learnt features of Convolutional Neural networks (CNN). It is observed that the proposed model outperforms existing models with lesser epochs. The proposed model is trained and tested with the ISI numeral dataset and also cross-validated with the CAMTERDB numeral dataset. For both scenarios, proposed model shows consistency and demonstrate the maximum accuracy of 99.02% and 99.17%, respectively. For the CMATERDB collection, the proposed model achieves the best accuracy rate reported till date.**

I. Introduction

Due to the various application potentials, recognition of handwritten numerals has been receiving a profound interest to the computer vision researchers over the past few years. However, Bangla, the seventh most widely spoken language in the world and the second most used language in the Indian subcontinent, is lagging far behind in handwritten Bangla numeral research. Meanwhile, with progressive affordability of technology and incentives by the concerned Governments, Bangla document analysis is more relevant now than every before.

Deep learning methods, specifically convolutional neural network (CNN) based solutions exhibit notable performance in handwriting recognition, image classification, image annotation and various other fields. CNNs are multi-layered networks that learns, extracts and classifies features automatically. In contrast, other popular approaches such as SIFT [1],HOG [2] and LBP [3] are"hand crafted"', as these methods extract salient image information through techniques which were devised through experimentation and ingenuity of the researchers. Inspired by the performance of deep learning models a few recent studies have reported on the performance of CNN-based models for Bangla handwriting recognition.

The main contribution of this paper is the proposal of a hybrid model which combines the use of a automatically learnt CNN-based features and the hand crafted Histogram of Oriented Gradients (HOG) feature. The proposed model is trained on a data set of Bangla isolated numerals. It is observed that the proposed hybrid model outperforms other CNN-only methods in terms of accuracy and the number of epochs required to achieve the best accuracy. Moreover, through cross dataset testing, very good performance is observed by the proposed model even on untrained collections.

The rest of the paper is organized as follows. In §II, existing CNN-based Bangla numeral recognition methods are discussed. §III-D provides the details of the different experiments done for this research as well as the motivations. §IV presents the results that are obtained from the experiments. Next, conclusion and future works are presented in §V.

II. Background

In a breakthrough study [4] showed that CNNs can successfully be applied to large and complex image classifications tasks. Such CNNs are neural networks which accept gray scale or colour images as their input, and have successive layers of filters, which are the trainable weights, and pooling/subsampling layers [5]. Non-linearities, such as the sigmoid or ReLU functions, are typically applied after the convolutional layers. Such sequences of filtering and pooling layers are often followed by fully connected(FC) layers for classification purposes. These networks are effective because they are both feature learners and extractors.

Prior to the widespread adoption of deep learning, hand crafted features such as Histogram of Oriented Gradients (HOG) [2], Scale Invariant Feature Transform (SIFT) [1], Local Binary Patterns (LBP) [3] were among the most effective tools for image classification. These features were devised through experimentation and ingenuity of the researchers involved, and were considered to be state of the art. [6] published a detailed analysis of five different image features, their encoding schemes and classification performance on standard data sets. They identified HOG, which is based on image gradients being accumulated in localised histograms, as one of the most effective features. However, in a follow up paper, the authors compared the performance of such hand crafted approaches with deep learning techniques [7] and concluded that deep architectures outperform the hand crafted methods by a large margin [7]

A trend towards deep learning techniques can be observed in recent studies concerning Bangla handwriting recognition. [8] used a LeNet-like five layer CNN for character classification. This network has two sets of convolutional and subsampling layers in sequence, successively learning and extracting higher level features, with the last layer (F5) extracting a 300 dimensional feature vector, which is then used in fully connected layers during training. The motivation of this work was to train a single network on a single dataset and then re-use the network features. They initially trained their network on a data set with 50 classes (characters) and then applied the same network to other data sets, using the higher level features to train an SVM classifier. This method gave them a generalised

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

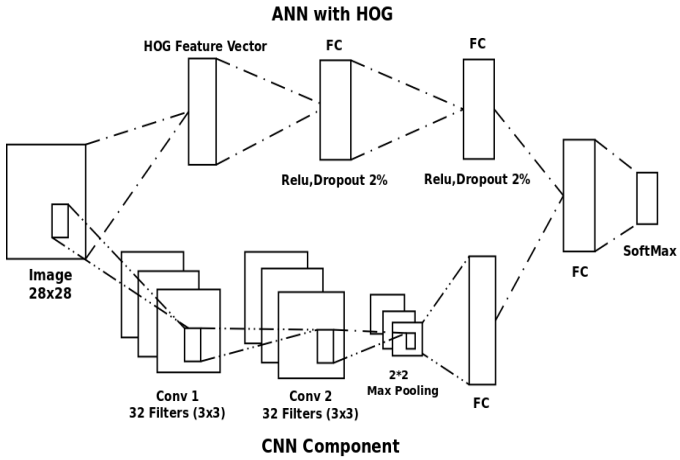


Fig. 1: Hybrid Model Overview

feature extractor (the CNN) tuned for characters which could then be used to train classifiers like SVM which do not demand the same volume of training data. Using this scheme they achieved 98.375% accuracy on the ISI Bangla numeral dataset. Recently [9] used a smaller CNN to train on the ISI numeral dataset and achieved more impressive results of 98.98% accuracy. In this case the network was specifically trained for the numeral dataset. As the training set is quite small ($19,392$ images), they augmented the training set by creating rotated versions of the training images. They trained different networks for rotation angles of $5°, 10°, 20°$ and $30°$. Best validation accuracy of 98.98% was achieved when the training set was augmented by $10°$ rotations. Unfortunately the merit of the augmentation method is questionable as for each experiment the images were rotated by a fixed angle, and then they reported the best observed results from multiple such experiments. This is equivalent to hand fitting for the testing set. Our attempts to reproduce the results on the same network, but when using random rotations, resulted in sub 98% validation accuracy.

None of these studies have explored the efficacy of combining CNN's with handcrafted features for Bangla OHR. In this paper we present the use of a hybrid network which combines the use of features learnt and extracted by CNNs, with the very popular histogram of oriented gradients (HOG) image feature.

III. Experimental Setup

A. Proposed Hybrid model

The hybrid model proposed in this paper has two components, as shown in Figure 1. The first component is an Artificial Neural Network with two hidden layers of size 32 and 64 respectively. The input to this network is the HOG feature vector extracted from an image. The second component of the hybrid model is a traditional multi-layered CNN which has two convolutional layers in sequence, each with 32 3×3 filters, followed by a 2×2 max-pooling layer. The output of the sub-sampling layer is flattened to a vector and forwarded to a fully connected layer.

The output of the two components are combined into a larger feature vector, which then acts as the input to another network comprising of two fully connected layers, the last of which is a softmax layer which provides probability distributions of class predictions.

(a) ISI (b) CMATERDB

Fig. 2: Example images from the ISI Character database (a) and CMATERDB3.1.1(b)

The ANN component and the CNN component were also trained individually by forwarding their output to a 10 node softmax layer. We will be referring the the Hybrid model with the "Hybrid-" prefix. The CNN only and ANN only models will be refered by the "CNN" and "ANN-" prefixes.

B. Data Sets

This study utilises two independently collected data sets – the *Indian Statistical Institute (ISI) handwritten character database* [10], specifically the Bangla handwritten numeral data set, and the *CMATERDB3.1.1* collection [11]. Both data sets contain images of isolated handwritten Bangla numerals. The ISI images have variable size and contains noise. The CMATERDB3.1.1 images are all 32×32 pixels, mostly noise free but the character edges seem to be more blocky. Figure 2 shows an example image from each of the two collections.

For deep learning, the recommended approach is to have three partitions: a training set, a validation set and a testing set. The ISI collection comes with a partition of $19,392$ training images and $3,986$ testing images. The CMATERDB3.1.1 collection has 6000 images, which has been split into testing and training sets by different studies, usually 5000 training images and 1000 testing images. In this study, the ISI collection's training and testing partitions are used for training and validation respectively. The entire CMATERDB3.1.1 collection is used as the testing set.

C. Dataset preparation

A smoothing operation with a small kernel is performed at the very beginning to remove any noise from the images. The images are then converted to a binary format, followed by a *not* operation to have black background and white foreground. An edge thickening filter is applied next, and finally all images are resized to 28×28 pixels.

Following the recommendation of [7] each training image is rotated clockwise and anticlockwise by a random angle between $0° - 50°$, resulting in a final training set of $58,176$ training images. The difference with [9] is that rotation is performed by random angles, not fixed ones.

D. List of experiments and their motivation

Table I lists the different models trained to compare the performance of three different networks and their inputs described in §III-A. CNN-64 is very similar to the CNN described previously, except the first layer has 64 filters instead of 32. This network has been trained to observe the effect of increasing the number of network parameters, without introducing hand crafted features. A comparison between the performance of CNN and the Hybrid-* models can demonstrate the efficacy of including HOG features in the Hybrid model.

Experiment List			
Experiment Label	Network Type	Input	No. of Parameters
CNN	CNN	Grey scale 28×28 image	600810
CNN-64	CNN	Grey scale 28×28 image	610346
ANN-Hog-PPC4	ANN	HOG feature(392 dimensions)	15338
ANN-HOG-PPC8	ANN	HOG feature(72 dimensions)	5098
Hybrid-HOG-PPC4	Hybrid Model	HOG feature(392 dimensions) and 28×28 image	616138
Hybrid-HOG-PPC8	Hybrid Model	HOG feature(72 dimensions) and 28×28 image	605898

TABLE I: Details of the different models used in the experiments

E. Model training

All models were trained on the augmented training set of the ISI Bangla numeral collection. HOG features were extracted from all the images of the augmented set. The Adadelta optimiser was used with categorical cross entropy as the loss function. Each model was trained for 100 epochs.

IV. Results and Analysis

Fig. 3 illustrates how the loss values change over the training period, whereas Fig. 4 shows the how the accuracy values (training and validation) change over the epochs. Training results for different models are summarized in Table II, where the accuracies are evaluated over $3,986$ validating images. CNN, CNN-64 and Hybrid-HOG-PPC4 models achieve accuracy values around 98.9% in 100, 87 and 98 epochs, respectively. Higher number of parameters for the Hybrid-HOG-PPC4 is a result of the higher input dimensionality of the HOG feature vector (392). With lower number of parameters compared to Hybrid-HOG-PPC4 and CNN-64, Hybrid-HOG-PPC8 still shows better performance by achieving a validation accuracy of 99.02% in only 55 epochs.

Experiment Name	Accuracy(%)	Epoch
HOG-PPC4	96.61	98
HOG-PPC8	89.76	87
CNN	*98.87*	100
CNN-64	*98.87*	86
Hybrid-HOG-PPC4	98.90	98
Hybrid-HOG-PPC8	**99.02**	**55**

TABLE II: Best validation accuracy and the number of epochs required

Therefore, from the validation set results, it is observed that hand crafted feature that works well individually may not be suitable with learnt CNN features (Hybrid-HOG-PPC4). Moreover, it is also observed that a well-chosen hand crafted feature can improve the accuracy of a deep model and the training speed.

Performances of the proposed Hybrid-HOG-PPC8 and other existing models are presented in Table III for ISI Bangla numeral data set and CMATERDB3.1.3 data set. Table III shows that Hybrid- Hog-PPC8 achieves an accuracy of 99.17% on the CMATERDB3.1.1 data set, where a collection of 6000 images is considered for the test. Till date, this is the most precise recognition rate for the CMATERDB3.1.3 collection. Moreover, it is worth to mention that the 6000 test images are not at all used in the training phase, which explains the robustness of the proposed model.

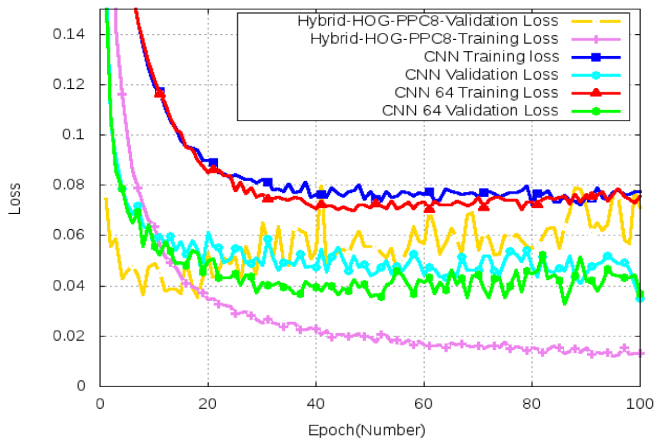


Fig. 3: Loss values over the training period

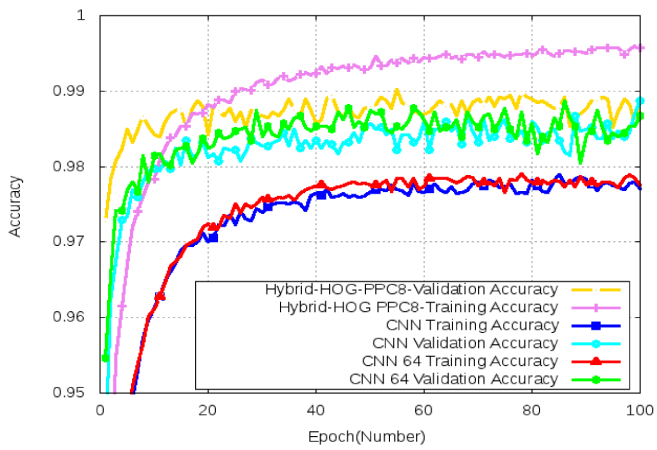


Fig. 4: Accuracy values over the training period)

V. Conclusion

This paper presents a hybrid deep learning model that combines CNN with HOG features. Bangla handwritten isolated numeral data set is considered to train the proposed model. From the experimental results, it is observed that the proposed hybrid model achieves better accuracy with lesser epochs as compared to other CNN only models. Moreover, in cross data set testing, the proposed model exhibits the most precise recognition rate for handwritten Bengali numerals. Moreover, the proposed model outperforms CNNs with greater number of parameters, thereby demonstrating the efficacy of using the HOG features. In future, the proposed hybrid model will be integrated with multiple handcrafted features and behavior of that particular system will be studied.

ISI Numeral Dataset		CMATERDB 3.1.1	
Work	Accuracy	Work	Accuracy
Bhattacharya and Chaudhuri [10]	98.20%	Haider Adnan Khan et al. [12]	94%
Wen and He [13]	96.91%	Hassan et al. [14]	96.7%
Das et al. [15]	97.70%	Das et al. [16]	98.55%
Nasir and Uddin [17]	96.80%	Sarkhel el at. [18]	98.23%
Akhnad et al. [19]	97.93%	Basu et al. [20]	97.15%
CNNAP, [9]	98.98 %	Basu et al. [21]	95.1%
Hybrid-HOG-PPC8 (proposed)	**99.02%**	Hybrid-HOG-PPC8 (Proposed)	**99.17%**

TABLE III: Performance comparison of Hybrid-HOG-PPC8 with previously published results

REFERENCES

[1] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[2] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1. IEEE, 2005, pp. 886–893.

[3] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105. [Online]. Available: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

[5] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," *The handbook of brain theory and neural networks*, vol. 3361, no. 10, p. 1995, 1995.

[6] K. Chatfield, V. S. Lempitsky, A. Vedaldi, and A. Zisserman, "The devil is in the details: an evaluation of recent feature encoding methods." in *BMVC*, vol. 2, no. 4, 2011, p. 8.

[7] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman, "Return of the devil in the details: Delving deep into convolutional nets," *arXiv preprint arXiv:1405.3531*, 2014.

[8] D. S. Maitra, U. Bhattacharya, and S. K. Parui, "Cnn based common approach to handwritten character recognition of multiple scripts," in *Document Analysis and Recognition (ICDAR), 2015 13th International Conference on*. IEEE, 2015, pp. 1021–1025.

[9] M. M. H. R. M. A. H. Akhand, Mahtab Ahmad, "Convolutional neural network training with artificial pattern for bangla handwritten numeral recognition," *ICIEB*, vol. 1, no. 1, pp. 1–6, 2016.

[10] U. Bhattacharya and B. B. Chaudhuri, "Handwritten numeral databases of indian scripts and multistage recognition of mixed numerals," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, no. 3, pp. 444–457, 2009.

[11] R. Sarkar, N. Das, S. Basu, M. Kundu, M. Nasipuri, and D. K. Basu, "Cmaterdb1: a database of unconstrained handwritten bangla and bangla–english mixed script document image," *International Journal on Document Analysis and Recognition (IJDAR)*, vol. 15, no. 1, pp. 71–83, 2012.

[12] H. A. Khan, A. Al Helal, and K. I. Ahmed, "Handwritten bangla digit recognition using sparse representation classifier," in *Informatics, Electronics & Vision (ICIEV), 2014 International Conference on*. IEEE, 2014, pp. 1–6.

[13] Y. Wen and L. He, "A classifier for bangla handwritten numeral recognition," *Expert Systems with Applications*, vol. 39, no. 1, pp. 948–953, 2012.

[14] T. Hassan and H. A. Khan, "Handwritten bangla numeral recognition using local binary pattern," in *Electrical Engineering and Information Communication Technology (ICEEICT), 2015 International Conference on*. IEEE, 2015, pp. 1–4.

[15] N. Das, R. Sarkar, S. Basu, M. Kundu, M. Nasipuri, and D. K. Basu, "A genetic algorithm based region sampling for selection of local features in handwritten digit recognition application," *Applied Soft Computing*, vol. 12, no. 5, pp. 1592–1606, 2012.

[16] N. Das, J. M. Reddy, R. Sarkar, S. Basu, M. Kundu, M. Nasipuri, and D. K. Basu, "A statistical–topological feature combination for recognition of handwritten numerals," *Applied Soft Computing*, vol. 12, no. 8, pp. 2486–2495, 2012.

[17] M. K. Nasir and M. S. Uddin, "Hand written bangla numerals recognition for automated postal system," *IOSR Journal of Computer Engineering*, vol. 8, no. 6, pp. 43–48, 2013.

[18] R. Sarkhel, N. Das, A. K. Saha, and M. Nasipuri, "A multi-objective approach towards cost effective isolated handwritten bangla character and digit recognition," *Pattern Recognition*, vol. 58, pp. 172–189, 2016.

[19] S. I. P. S. Md. Mahbubar Rahman, M. A. H. Akhand and M. M. H. Rahman, "Bangla handwrittencharacter recognition using convolutional neural network," *I.J.Image, Graphics and Signal Processing(IJIGSP*, vol. 7, no. 3, pp. 42–49, 2015.

[20] S. Basu, N. Das, R. Sarkar, M. Kundu, M. Nasipuri, and D. K. Basu, "A novel framework for automatic sorting of postal documents with multi-script address blocks," *Pattern Recognition*, vol. 43, no. 10, pp. 3507–3521, 2010.

[21] S. Basu, R. Sarkar, N. Das, M. Kundu, M. Nasipuri, and D. K. Basu, "Handwritten bangla digit recognition using classifier combination through ds technique," in *Pattern Recognition and Machine Intelligence*. Springer, 2005, pp. 236–241.

Wavelet Coherence Based Channel Selection for Classifying Single Trial Motor Imagery

Simanto Saha*, Khawza I. Ahmed, and Raqibul Mostafa
Biomedical, IMage and Signals (BIMS) Research Group
Department of Electrical and Electronic Engineering
United International University
Dhaka-1209, Bangladesh
*simantosaha@gmail.com

***Abstract*—Multi-channel electroencephalography (EEG) recordings require excessive computation and sometimes engender outliers, which make brain computer interface (BCI) systems inefficient. Thus, optimal channel selection becomes a key factor for developing a more comfortable BCI. This study emphasized on a time-frequency (T-F) coherence method, called as Wavelet Coherence (WC), for selecting lesser number of channels. The selected sets of channels were then used to classify two motor imagery (MI) tasks, i.e., right hand (RH) and right foot (RF). The data was collected from publicly available dataset IVa from BCI Competition III. Common spatial pattern (CSP) with and without regularization were applied as preprocessing techniques. While the classification accuracy is 90% using available 118 channels for subject *ay*, we have achieved higher classification accuracy of 93% using only 24 channels using CSP with regularization. Interestingly, the achieved classification accuracy for subject *av* is 67% using 4 channels only, that outperform the classification accuracy (i.e., 61%) achieved using 118 channels.**

***Index Terms*—Wavelet Coherence, Motor Imagery, Brain Computer Interface, Electroencephalography, Channel Selection.**

I. Introduction

Brain computer interface (BCI) provides an unorthodox communication pathway between the brain and the computer or other external devices, while disregarding the communication channel via the central nervous system (CNS). In addition to cognitive rehabilitation, recent studies have proposed many applications of BCI for healthy users. These applications include, but are not limited to lie detection [1], brain fingerprinting [2], mood assessment [3] and gaming [4], etc. In some applications, portability is an important criterion for user comfort. However, electroencephalography (EEG) records electrocortical activities of the brain using high dimensional electrode montages. Simultaneous processing of EEG sources requires high computational resources, thus makes BCI less portable and less comfortable. Furthermore, lying within uninvolved scalp areas for specific cognitive tasks, the excessive channels do not produce distinguishable event related oscillations, rather these unnecessary channels sometimes generate outliers and, thus, debase BCI performance. In this paper, the proposed hypothesis is that optimally selected channels can reduce the number of channels (computational burden) and can minimize the effect of outliers.

Many studies, involving MI-EEG based BCI development, have proposed various signal processing and machine learning tools for selecting reduced number of channels. Fewer channels are selected by eliminating redundant channels, based on the Laplacian Derivative (LAD) of power averaged across frequency bands features [5]. A novel channel selection approach using common spatial pattern (CSP), in which the actual EEG sources were correlated with CSP channels, has been proposed [6]. Within a constraint of optimal classification accuracy, optimal channels have been selected based on CSP with sparsity [7]. Another study has proposed filter bank CSP (FBCSP) and head model geometry together for channel selection [8] within specific regions of interest. Similarly, a work has been proposed to employ lesser number of channels, featuring spatial brain dynamics corresponding to MIs [9], i.e., sensorimotor areas are mostly responsive during MIs [10]. However, it is evident from literature that only one EEG source can be used for classifying MIs [11]. In this paper, we have proposed wavelet coherence (WC) based channel selection method during MI-EEG based BCI development, for the first instance. As a popular method of measuring coherence in time-frequency (T-F) space, WC had been previously applied to assess associativity of inter-personal brain activity [12] and to quantify coherence between two neural signals [13].

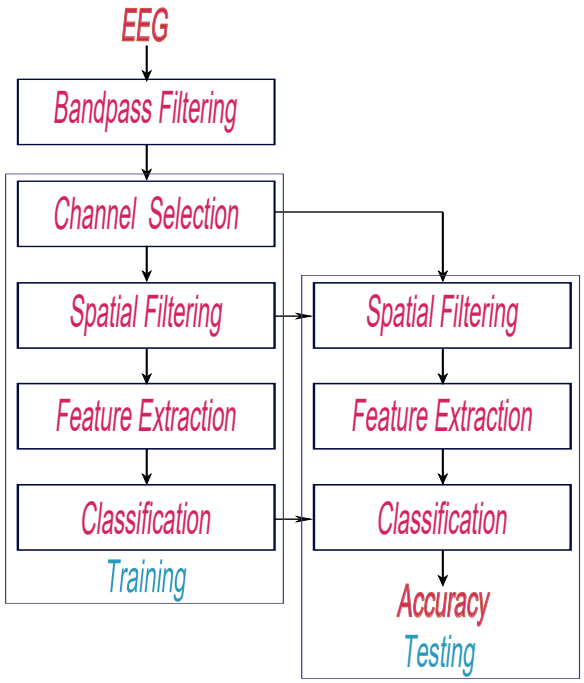


Fig. 1. Basic Block Diagram of Proposed Paradigm

II. Proposed Experimental Paradigm

The basic block diagram of the proposed BCI paradigm has been shown in Fig. 1. At first, the single trial EEG signals

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

corresponding to MIs were bandpass filtered (4 Hz to 32 Hz) using Butterworth filter of order 10. Then, wavelet coherence (WC) analysis based channel selection method was proposed to select lesser number of channels. As spatial filtering technique, common spatial pattern (CSP) or regularized common spatial pattern (RCSP) had been applied to optimally project selected EEG sources, with a view to achieving distinguishable features. Afterwards, wavelet decomposition (wavelet basis: db3, level: 3) based subband energy and subband entropy features were used to model a Two-Layer Feed-Forward Neural Network [14] and classification accuracies (testing phase) were finally measured.

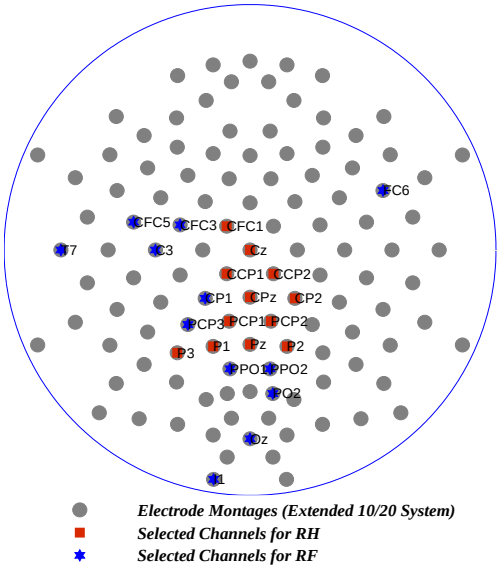


Fig. 2. **Electrode Montages: Extended 10/20 system:** selected 24 channels corresponding to RH and RF MIs for subject *ay*

A. Dataset and Experimental Settings

We have used dataset IVa from BCI competition III [15] to validate the proposed channel selection approach. The data were recorded from five healthy subjects while sitting on a comfortable chair with arms resting on armrests. Total 118 channels (Extended 10/20 System) were employed to capture the multichannel EEG signals (see Fig. 2). The sampling rate was set at 1 KHz during recording session. The available data are downsampled to 100 Hz. The subjects were given visual cues to perform each of three MIs (left hand (LH), right hand (RH) and right foot (RF)). However, the data were provided for only two classes, i.e., RH and RF. The data consist of 280 trials for each subject and each trial consists of 3.5s of recording for any of the two MIs. It should to be noted that we have chosen $2.5s$ of data for each trial after $1s$ of the corresponding visual cues.

Out of total 280 trials for each subject, there are 140 trials for each class. In this experiment, the first 120 trials were used to train the algorithm. The consecutive 60 trials were then used to validate the classifier for avoiding overfitting. Finally, the remaining 100 trials were used to test the classifier model.

B. Channel Selection based on Wavelet Coherence

WC reveals associativity of two time signals in time-frequency space. This similarity measure finds the time-frequency locations where the signals significantly co-vary, but doesn't necessarily have high power. Suppose x_n and y_n are two time signals (where, n=1,2,...,N), then the WC is defined as follows [16]:

$$R_n^2(s) = \frac{|S\left(s^{-1}W_n^{xy}(s)\right)|^2}{S\left(s^{-1}|W_n^x|^2\right).S\left(s^{-1}|W_n^y|^2\right)} \tag{1}$$

where S is a smoothing operator and defined as

$$S(W) = S_{scale}(S_{time}(W_n(s))) \tag{2}$$

S_{scale} and S_{time} represent smoothing along scale/frequency and time axes respectively. W_n^{xy} is the Cross Wavelet Transform (XWT) which finds the time-frequency regions where two signals co-vary, but have high power in contrast to WC.

$$W^{xy}(s) = W^x(s).W^{y^*}(s) \tag{3}$$

where W^x and W^y denote the Continuous Wavelet Transform (CWT) and represents the signals x and y in time-frequency space. The CWT of a time signal $(x_n, n = 1, 2, ..., N)$ can be written as

$$W_n^x(s) = \sqrt{\frac{\delta t}{s}} \sum_{n'=1}^{N} x'_n \psi_0\left[(n'-n)\frac{\delta t}{s}\right] \tag{4}$$

Here, δt and s represent uniform time shifts and scaling respectively. ψ_0 is the mother wavelet (wavelet basis function) and Morlet wavelet with dimensionless frequency $\omega_0 = 6$ has been selected in this specific case.

$$\psi_0(\eta) = \pi^{-\frac{1}{4}} e^{i\omega_0\eta} e^{-\frac{1}{2}\eta^2} \tag{5}$$

where η is the dimensionless time. For more details about WC calculation, see [16].

The training trials were categorized according to classes, i.e., 60 trials for RH and 60 trials for RF. Let's consider $M = 60$. Then, the WC had been applied to measure similarities of first $M/2$ trials with rest $M/2$ trials one by one. The WC power is normalized to 1 and thus the average WCP ranges from 0 to 1. The accumulated WCP over trials are averaged and denoted as WCP_N.

$$WCP_{Nc} = \frac{2}{M} \sum_{i=1}^{M/2} \left(\left(R_n^2(s)\right)_c\right)_i \tag{6}$$

where $c \in (RH, RF)$. The $R_n^2(s)$ in the range from 0.6 to 0.9 had been considered significant and used to measure associativity of inter-subject channels, i.e., $R_n^2(s) \rightarrow [0.6, 0.9]$. An underlying assumption is that high coherence power ($>$ 0.9) might manifest resting EEG dynamics, irrespective of any information regarding MIs.

Finally, the estimated WCP_N is used to rank the channels in descending order. Several cases were studied concerning different number of channels, i.e., 4 channels to 24 channels with an increment of 4. Two channels were selected for each class every time and continued until total number of channels becomes 24, based on higher WCP_N. If any channel falls within the selection criteria for both classes, we have selected another channel with higher WCP_N either for RH or RF. However, the visualization of WCP_N calculation corresponding to RH MI for subject *ay* has been shown in Fig. 3.

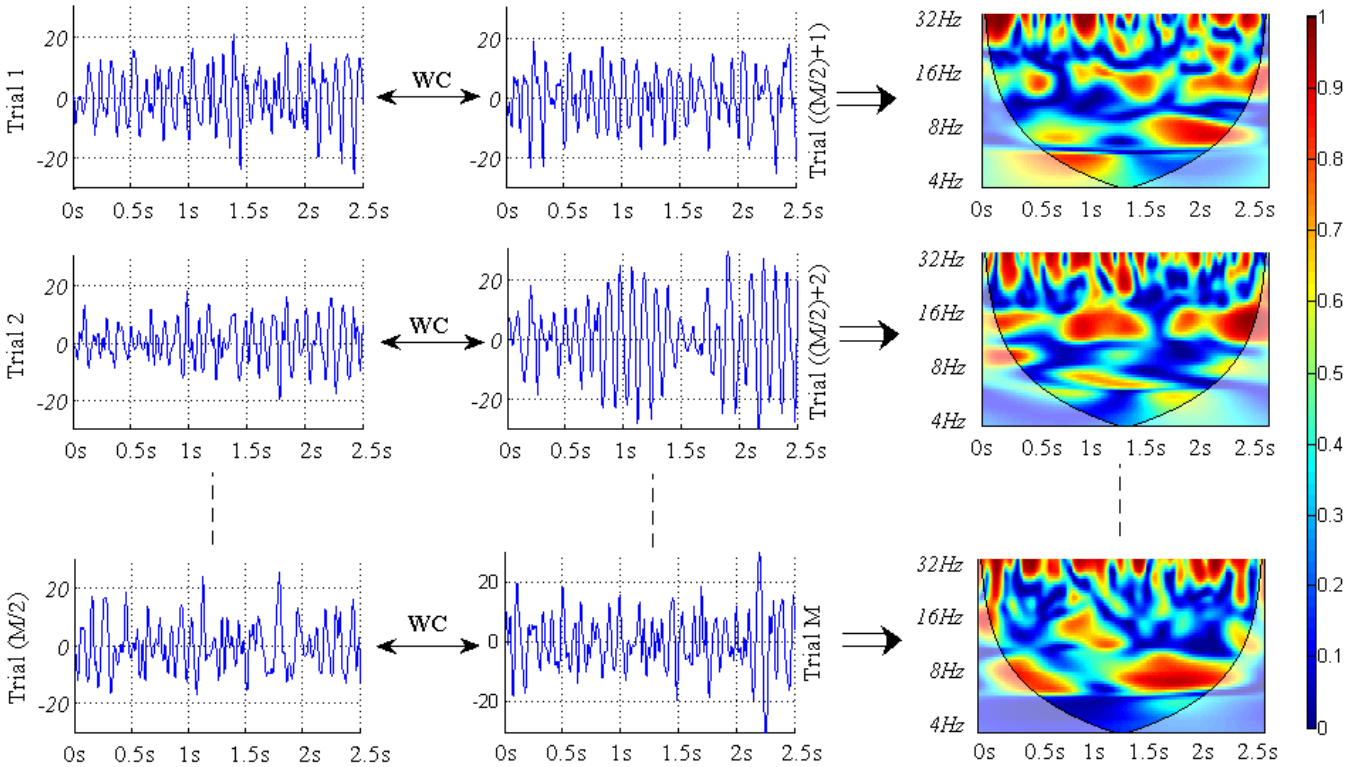


Fig. 3. **Calculation of Wavelet Coherence Power (WPN) on Single Trial Motor Imagery (MI)**: the depicted trials are for right hand (RH) MIs from the channel CP5 corresponding to subject *ay*.

C. Common Spatial Pattern with and without Regularization

In 2000, Ramoser et al. proposed CSP, as a preprocessing method, for classifying single trial MIs for the first instance [17]. In this study, CSP with and without regularization parameters have been used [18]-[19]. Usually, CSP gives projection of multi-channel EEG signals, so that the difference between two classes is maximized.

A brief description of CSP with and without regularization has been given as following. For further details about the regularized CSP, refer to [18]-[19]. Suppose $C1$ and $C2$ represent two MIs, i.e., RH and RF respectively. Then, CSP estimates a set of spatial filters W whose first components have maximum variance for $C1$ and minimum for $C2$, while last components have maximum variance for $C2$ and minimum for $C1$. The training trials are used to estimate sample based covariance matrix.

Let's consider, N is the number of channels and P is the number of samples per trial, then E represents EEG signals of size $N \times P$.

$$E_{csp} = W^T E \tag{7}$$

E_{csp} refers to the spatially filtered EEG signals. Simultaneous diagonalization of two covariance matrices, which represent two classes (i.e., $\sum_{C1}$ and $\sum_{C2}$), is performed while estimating optimal spatial filters.

$$\begin{cases} W^T \sum_{C1} W = \Lambda_{C1} \\ W^T \sum_{C2} W = \Lambda_{C2} \end{cases} \tag{8}$$

where Λ_{C1} and Λ_{C2} represent the diagonal matrices, comprising the eigenvalues. W has been estimated such that $\Lambda_{C1} + \Lambda_{C2} = I$. I is the identity matrix.

As a data driven method, CSP is prone to be affected by outliers, thus be overfitted [20]. To get rid of this problem, regularization of covariance estimation sometimes becomes useful. Two regularization parameters, i.e., γ and β had been used to bias the covariance matrix. γ controls the shrinkage of the training samples covariance matrix estimate to the pooled estimate and β controls the degree of shrinkage towards a multiple of the identity matrix. Different combination of values were used during regularization and indicated as follows: $\beta = (0, 0.001, 0.01, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9)$ $\gamma = (0, 0.01, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9)$.

It should be noted that total four CSP components, two for each class, were selected for extracting features.

III. Results and Discussions

The intra-subject classification accuracies achieved using different number of channels are shown in Table I. The classification accuracies differ significantly from subject to subject. However, the performances showcase the usefulness of WC as a novel channel selection method. For subject *ay*, the achieved classification accuracies are 93% (RCSP) and 91% (CSP) using 24 channels and 20 channels respectively, while 90% accuracy has been achieved using available 118 channels from the whole brain. The classification accuracies increase, although lesser number of channels have been employed. For example, the classification accuracies for subject *av* are 67% (RCSP) and 64% (CSP) using only 4 channels, which have also outperformed the performances achieved using 118 channels. For subject *aa*, 70% (RCSP) accuracy has been achieved using only 20 channels, which is just 1% below the accuracy achieved using all the channels. In summery, results

TABLE I
CLASSIFICATION ACCURACIES (%): INTRA-SUBJECT BRAIN COMPUTER INTERFACE (BCI) WITH SETS OF DIFFERENT NUMBER OF CHANNELS.

Number of Channels	aa		al		av		aw		ay	
	CSP	RCSP	CSP	RCSP	CSP	RCSP	CSP	RCSP	CSP	RCSP
4	60	63	57	68	**64**	**67**	52	60	80	82
8	57	62	76	81	59	63	69	74	86	89
12	56	63	75	79	50	57	82	76	79	84
16	66	69	83	90	59	58	79	82	89	83
20	59	72	80	90	53	57	80	75	**91**	92
24	63	70	88	88	51	58	80	81	90	**93**
:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:
118	**72**	**73**	**97**	**98**	60	61	**91**	**91**	90	90

represent the usefulness of the proposed channels selection approach.

Fig. 2 shows the selected channels corresponding to RH and RF MIs respectively. The selected channels for RH densely lie within the sensorimotor areas. On the contrary, the selected channels for RF are relatively sparse, but mostly exist around the sensorimotor regions of the brain. Literature suggests that sensorimotor areas are mostly responsible for MIs [10]. However, although WC has already been applied to measure couplings between brain regions using sensory-evoked potentials [21] and to find out associative inter-subject brain activity [12], this work demonstrate how this method can be adopted to select optimal channels.

IV. CONCLUSIONS

Computational cost is a key factor for developing and optimizing a portable and simple BCI system for various commercial applications, since simultaneous processing of multi-channel EEG data imposes high computational burden. As evident from previous literature, specific brain areas are responsible for specific cognitive tasks, i.e., sensorimotor areas robustly respond during MIs. Thus, localization of optimal sources corresponding to specific cognitive task becomes crucial, that lessens computational cost and effect of outliers. This study has demonstrated the feasibility of WC based channel selection approach for implementing computationally more efficient BCI systems. As this experiment is limited to binary classification, further work can be extended to multiclass classification problem.

ACKNOWLEDGMENT

The authors would like to thank the organizers of BCI Competition III for availing the dataset.

REFERENCES

[1] H. Wang, W. Chang and C. Zhang, "Functional brain network and multichannel analysis for the P300-based brain computer interface system of lying detection", Expert Systems with Applications, vol. 53, pp. 117-128, 2016.

[2] L. Farwell, D. Richardson, G. Richardson and J. Furedy, "Brain fingerprinting classification concealed information test detects US Navy military medical information with P300", Front. Neurosci., vol. 8, 2014.

[3] F. Ortiz Carreon, J. Gonzalez Serna, A. Montes Rendon, N. Gonzalez Franco, A. Pena Jimenez and J. Hernandez Gomez, "Induction of emotional states in people with disabilities through film clips using brain computer interfaces", IEEE Latin America Transactions, vol. 14, no. 2, pp. 563-568, 2016.

[4] B. van de Laar, H. Gurkok, D. Plass-Oude Bos, M. Poel and A. Nijholt, "Experiencing BCI Control in a Popular Computer Game", IEEE Trans. Comput. Intell. AI Games, vol. 5, no. 2, pp. 176-184, 2013.

[5] H. Yang, C. Guan, C. C. Wang and K. K. Ang, Maximum dependency and minimum redundancy-based channel selection for motor imagery of walking EEG signal detection, 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, 2013, pp. 1187-1191

[6] Y. Wang, S. Gao and X. Gao, Common Spatial Pattern Method for Channel Selelction in Motor Imagery Based Brain-computer Interface, 2005 IEEE Engineering in Medicine and Biology 27th AnnualConference, Shanghai, 2005, pp. 5392-5395.

[7] M. Arvaneh, C. Guan, K. K. Ang and C. Quek, Optimizing the Channel Selection and Classification Accuracy in EEG-Based BCI, in IEEE Transactions on Biomedical Engineering, vol. 58, no. 6, pp. 1865-1873, June2011

[8] A. Zaitcev, G. Cook, W. Liu, M. Paley and E. Milne, Feature extraction for BCIs based on electromagnetic source localization and multiclass Filter Bank Common Spatial Patterns, 2015 37th Annual International Conferenceof the IEEE Engineering in Medicine and Biology Society (EMBC), Milan,2015, pp. 1773-1776

[9] S. Saha and K. I. Ahmed, Efficient event related oscillatory pattern classification for EEG based BCI utilizing spatial brain dynamics, Electrical and Computer Engineering (ICECE), 2014 International Conference on,Dhaka, 2014, pp. 707-710

[10] C. Zich, S. Debener, C. Kranczioch, M. Bleichner, I. Gutberlet and M. DeVos, Real-time EEG feedback during simultaneous EEG/MRI identifies the cortical signature of motor imagery, NeuroImage, vol. 114, pp. 438-447,2015

[11] S. Ge, R. Wang and D. Yu, "Classification of Four-Class Motor Imagery Employing Single-Channel Electroencephalography", PLoS ONE, vol. 9, no. 6, p. e98019, 2014.

[12] X. Cui, D. Bryant and A. Reiss, "NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation", NeuroImage, vol. 59, no. 3, pp. 2430-2437, 2012.

[13] J. Lachaux, A. Lutz, D. Rudrauf, D. Cosmelli, M. Le Van Quyen, J. Martinerie and F. Varela, "Estimating the time-course of coherence between single-trial brain signals: an introduction to wavelet coherence", Neurophysiologie Clinique/Clinical Neurophysiology, vol. 32, no. 3, pp. 157-174, 2002.

[14] D. Svozil, V. Kvasnicka and J. Pospichal, "Introduction to multi-layer feed-forward neural networks", Chemometrics and Intelligent Laboratory Systems, vol. 39, no. 1, pp. 43-62, 1997.

[15] Blankertz, B. (2005). BCI competition III webpage. Webpage. URL:http://www. bbci. de/competition/iii.

[16] A. Grinsted, J. Moore and S. Jevrejeva, "Application of the cross wavelet transform and wavelet coherence to geophysical time series", Nonlin. Processes Geophys., vol. 11, no. 56, pp. 561-566, 2004.

[17] H. Ramoser, J. Muller-Gerking and G. Pfurtscheller, Optimal spatial filtering of single trial EEG during imagined hand movement, in IEEE Transactions on Rehabilitation Engineering, vol. 8, no. 4, pp. 441-446, Dec2000.

[18] F. Lotte and C. Guan, Regularizing Common Spatial Patterns to Improve BCI Designs: Unified Theory and New Algorithms, in IEEE Transactions on Biomedical Engineering, vol. 58, no. 2, pp. 355-362, Feb. 2011.

[19] A. Buccino, H. Keles and A. Omurtag, "Hybrid EEG-fNIRS Asynchronous Brain-Computer Interface for Multiple Motor Tasks", PLOS ONE, vol. 11, no. 1, p. e0146610, 2016.

[20] C. Sannelli, C. Vidaurre, K. Mller and B. Blankertz, "Ensembles of adaptive spatial filters increase BCI performance: an online evaluation", J. Neural Eng., vol. 13, no. 4, p. 046003, 2016.

[21] A. Klein, T. Sauer, A. Jedynak and W. Skrandies, "Conventional and Wavelet Coherence Applied to Sensory Evoked Electrical Brain Activity", IEEE Transactions on Biomedical Engineering, vol. 53, no. 2, pp. 266-272, 2006.

Non-Isolated Single Stage PFC based LED Driver with THD Minimization using Cúk Converter

S.M.Ferdous,[1] Mohammad Abdul Moin Oninda,[2,*] Golam Sarowar,[2] Kazi Khairul Islam,[2] and Md. Ashraful Hoque[2]

[1]Department of Electrical, Electronics, Computers and Systems Engineering, University of Oviedo, Gijon, Asturias, Spain-33504.

[2]Department of Electrical and Electronic Engineering, Islamic University of Technology, Board Bazar, Gazipur-1704.

*mamoin@iut-dhaka.edu

Abstract—**This paper proposes a hysteresis current controller based PFC on Cúk converter with improvement in total harmonic distortion (THD), power factor (PF) and efficiency. Cúk is a dc-dc converter which by varying duty cycle can be used for step up and step down of voltage, and has low output ripples. Cúk converter integrated with PFC technique operates in discontinuous mode and provides zero current switching (ZCS) of the power switches causing power factor to approach near unity. The proposed converter supplies a constant current to the load, designed for LED driver application. Theoretical analysis, simulation and experimental result confirm that the proposed control system improves THD and is better compared to conventional control system.**

***Index Terms*—Hysteresis Current Control; Cúk Converter; THD; LED driver; Average Current Control.**

I. Introduction

For the development of viable and productive solution to the power quality problem, increasing demand for better power quality in power utility has captivated the consciousness of the power engineers [1]. Generally a ac-dc converter contains large amount of harmonics and decreased power factor due to distorted input line current. The source current gets distorted to a non-sinusoidal current due to these harmonics in the line current. AC-DC converters must comply with harmonic standards such as IEC61000-3-2.

In almost all the major areas such as communication, computer, biomedical, automotive etc., active power factor correction techniques are extensively used in power supplies, fulfilling harmonic regulations as well as other standards.

The dynamics of a Cúk converter changes for different input voltage and output load conditions. In order to restore the power factor to as close as possible to 1, power factor correction (PFC) technique is employed. DC-DC converter used with PFC technique helps to reduce the distortion in the input line current.

Several power factor correction (PFC) techniques [2]–[9] have already been discussed in many papers earlier. PFC for AC-DC converter can be classified into two types – two stage ac-dc converter [10], [11] and single stage ac-dc converter [12]-[21]. PFC provides advantages of improved efficiency, maximization of real power drawn from the mains, reduction in cost, I^2R loss, voltage drop in long cables and electric burden on cables and electrical components [22]. This paper proposes to improve THD of Cúk converter by integrating the converter with hysteresis current controller based PFC.

Section II of this paper deals with the conventional average current control of Cúk converter, proposed design of the current controller is shown in Section III, simulation and result of the proposed converter is given in Section IV of this paper.

II. Conventional Average Current Control of Cúk Converter

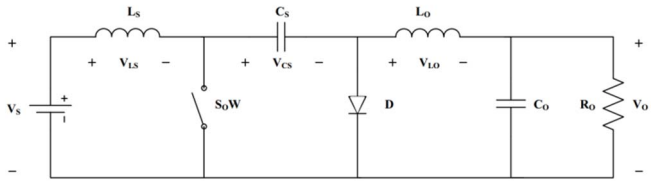


Fig.1 Circuit Diagram of Cúk Converter

Average voltage across capacitor C_S over 1 switching cycle, T:-

$$V_S - V_{LS} - V_{CS} - V_{LO} - V_O = 0$$

$$V_S - V_{CS} - V_O = 0$$

$$V_{CS} = V_S - V_O \quad (1)$$

At steady state,

Average voltage across inductors over one full switching cycle is zero. Hence,

$$V_{LS} = V_{LO} = 0$$

Now,

$$v_{LS(ON)} = V_S DT$$

$$v_{LS(OFF)} = (V_S - V_C)(1-D)T$$

So, average voltage across L_S for one full switching cycle,

$$\overline{V_{LS}} = \frac{\left(v_{LS(ON)} + v_{LS(OFF)}\right)}{T}$$

$$\overline{V_{LS}} = V_S D + (V_S - V_S + V_O)(1-D) \quad \text{[From eq}^{n}\text{ (1)]}$$

$$\overline{V_{LS}} = V_S D + V_O(1-D)$$

$$\overline{V_{LS}} = D(V_S - V_O) + V_O$$

From which Duty Cycle, D can be determined as:-

$$D = \frac{\overline{V_{LS}} - V_O}{V_S - V_O} \quad (2)$$

Voltage & current gain for Cúk converter are given as

$$V_O = \frac{D}{1-D} V_S$$

$$I_O = \frac{1-D}{D} I_S$$

It can be seen that I_S, which is the input current of the converter is equal to the input inductor current i.e. current through inductor, L_S.

$$i_{LS} = \left(\frac{D}{1-D}\right) I_O$$

So, the control block diagram can be shown as:-

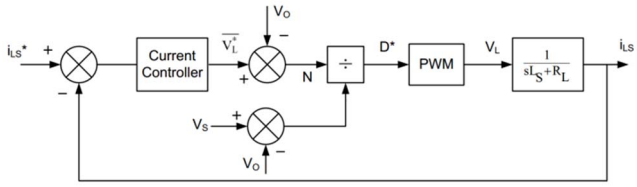


Fig.2 Control block diagram for close loop current control of the Driver.

III. Proposed Design of Current Controller

The proposed controller is a hysteresis current controlled based PFC with improvement in THD. The controller is designed for LED driver application where it is desired to supply a constant current to the load. So, LED driver circuit is basically a constant current load where the current control is done by a hysteresis controller. The advantage of the controller is very fast response, independent of system dynamics and improvement in performance.

The short-comings are – reduced accuracy at the constant output current and variable switching frequency which is the most common drawback of using hysteresis controller.

LED Driver load is rated as 30V, 20W DC.

$$I_{LOAD} = \frac{20W}{30V} = 0.667A = 667mA$$

Equivalent load resistance,

$$R_{LOAD} = \frac{30}{0.667} = 45\Omega$$

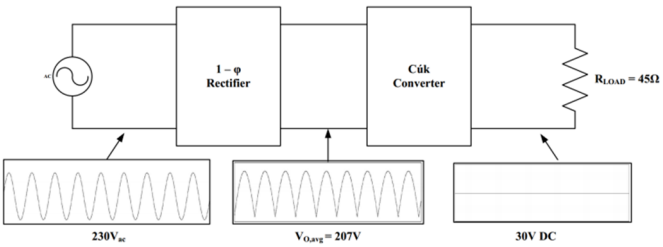


Fig.3 Overall Converter Topology with desired Voltage Output.

At unity power factor operation, the input to the Cúk converter is pulsating DC (Full wave rectified DC).

Considering a 230V AC utility supply system, the maximum value can be calculated as:-

$$V_{max} = 230 \times \sqrt{2} = 325 V_{ac}$$

So, the e value of the rectified DC voltage:-

$$V_{O,avg} = \frac{2V_{max}}{\pi} = \frac{2 \times 325}{\pi} = 207 V_{dc}$$

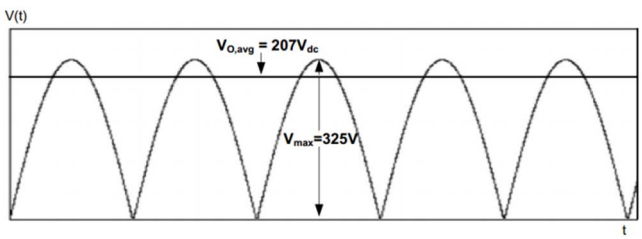


Fig.4 Determination of Average Voltage at the Cúk Converter Input.

Now, using the voltage gain equation of Cúk converter,

$$V_O = \frac{D}{1-D} V_S$$

$$V_O - DV_S = DV_S$$

$$D = \frac{V_O}{V_O + V_S} = \frac{30}{30+207} = 0.1266$$

So, the steady state duty cycle value = 0.1266.

Hence the converter input current can be calculated as:-

$$I_S = \frac{D}{1-D} I_O = \frac{0.1266}{1-0.1266} \times 0.667$$

$$I_S = 0.0967\ A$$

So, the average input inductor current is 0.0967A.

For unity PF operation the maximum value of the current is needed to be calculated. Using the basic relation of maximum value and average value of Full wave rectified DC wave –

$$I_{S,max} = \frac{\pi}{2} \times I_{S,avg}$$

$$I_{S,max} = \frac{\pi}{2} \times 0.0967$$

$$I_{S,max} = 0.152\ A$$

So, the commanded input inductor current of the controller should be –

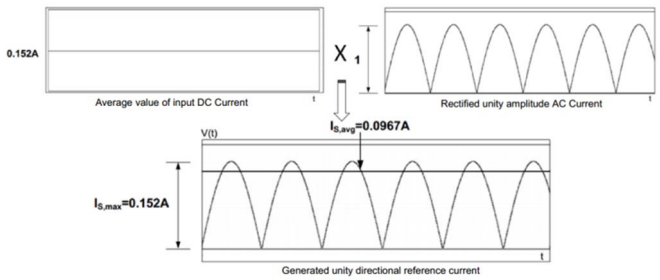


Fig.5 Generation of Reference Current i_L^* of the Converter for Unity Power Factor operation and desired output current.

If an ideal scenarion with 100% efficiency is assumed, Input Power,

$$P_{in,(avg)} = V_{s,(avg)} \times I_{S,(avg)}$$

$$P_{in,(avg)} = 207 \times 0.0967$$

$$P_{in,(avg)} = 20W$$

which is equal to the output power.

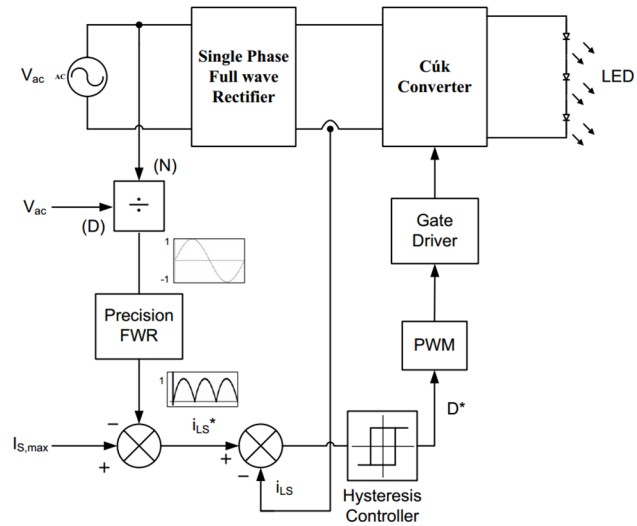


Fig.6 Block Diagram representation of the Proposed Converter for Unity Power Factor Operation.

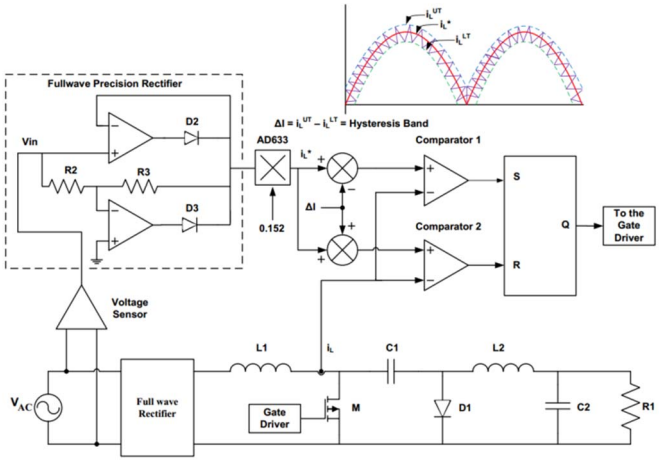


Fig.7 Proposed Cúk Converter with Hysteresis Current Controller.

IV. SIMULATION AND RESULT

The following diagram shows the source voltage (Vs), the source current (Is) and the input power factor (PF).

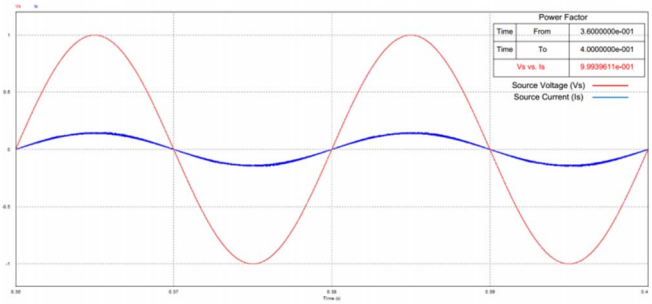


Fig.8 Plot of Source Voltage (V_s), Input Current (I_s) with a Power Factor of 0.99936. This corresponds to a Unity Power Factor Operation of the Converter.

The following diagram indicates the input source current and the associated THD. It may be noted here that the value of THD is significantly reduced compare to the commercially available LED Drivers.

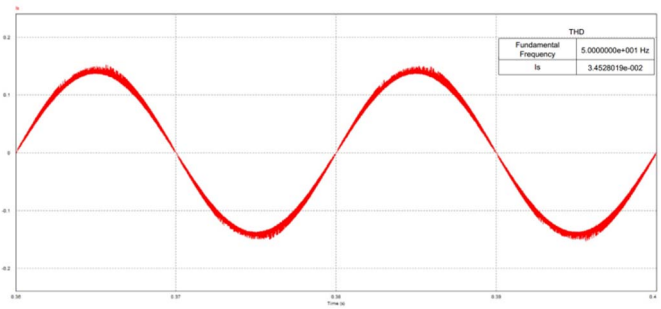


Fig.9 Plot of Input Current (I_s). The calculated value of THD is 3.45% as obtained from the simulation in PSIM.

The following diagram shows the reference current i_L^* and inductor current i_L controlled by the Hysteresis Controller. The response is very fast because there is no time constant dynamics involved in the operation of the controller.

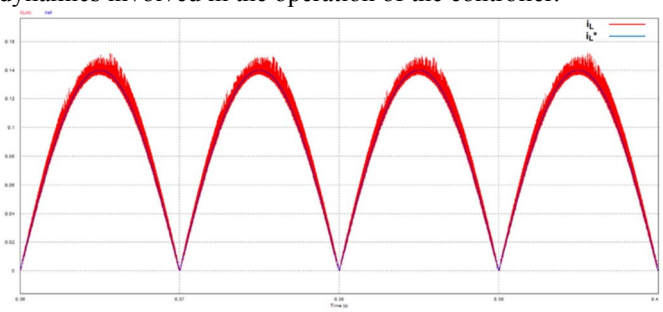


Fig.10 Plot of reference i_L^* and controlled inductor current i_L.

The following diagram shows the output voltage, current and the power of the converter for LED Driver Application.

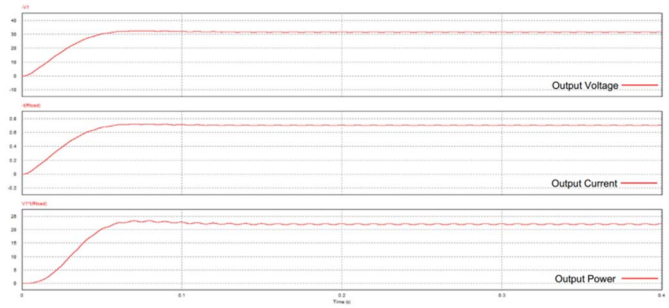


Fig.11 Plot of Output Voltage (31.64V), Output Current (0.703A) and Output Power (22.25W). The values shown here for Voltage, Current and Power are the approximate average value at the output of the Converter.

Now, a capacitor of 1μF is added at the input of the rectifier to further reduce the THD. The addition of the capacitor improves the THD but makes the converter to operate at leading Power Factor, which can be acceptable as the normal Power factor of the Utility supply system is lagging. This may somewhat serve the purpose of improvement of the Power Factor in a way.

Adding a passive LC filter for THD improvement and EMI reduction is very common for most of the commercial LED Drivers. So instead of using any inductor a single capacitor is only used.

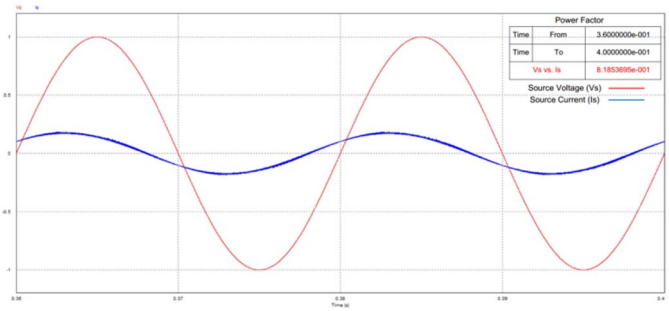


Fig.12 Plot of Source Voltage (V_s), Input Current (I_s) with a Power Factor of 0.8185. It is to be noted that, this corresponds to a leading Power Factor Operation of the Converter.

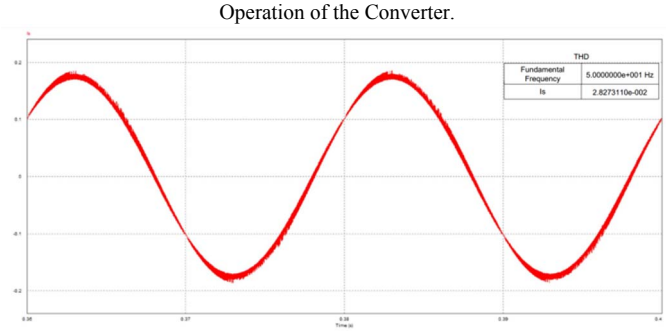


Fig.13 Plot of Input Current (Is) showing THD of 2.83%

The maximum allowable range of THD for Voltage and Current are given by IEEE Standard 519. According to IEEE 519, harmonic voltage distortion on power systems 69 kV and below is limited to 5.0% total harmonic distortion (THD) with each individual harmonic limited to 3%. The current harmonic limits vary based on the short circuit strength of the system they are being injected into. Essentially, the more the system is able to handle harmonic currents, the more the customer is allowed to inject. So the value of current THD obtained from the proposed design are in the acceptable range to meet the IEEE Standard given for Utility supply system.

The following diagram shows the reference current i_L^* and inductor current i_L.

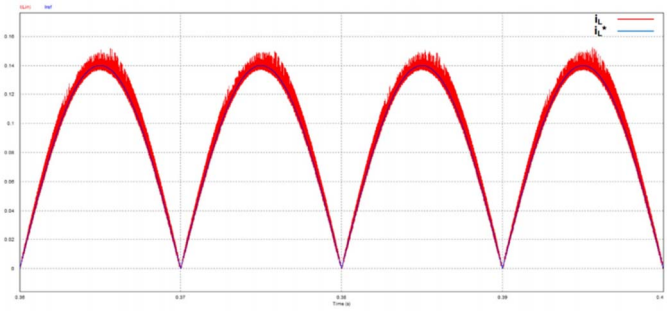


Fig.14 Plot of reference i_L^* and controlled inductor current i_L.

The following diagram shows the output voltage, current and the power of the converter for LED Driver Application.

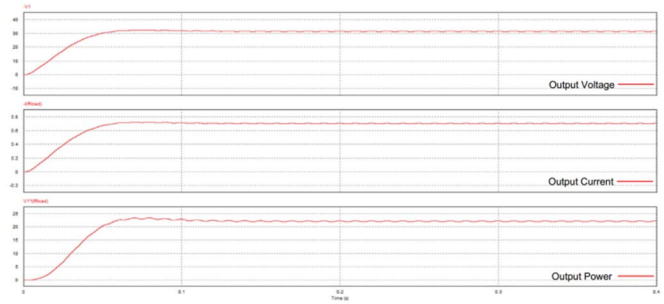


Fig.15 Plot of Output Voltage (31.53V), Output Current (0.7A) and Output Power (22.1W)

V. Conclusion

The proposed Cúk Converter based LED driver serves all the requirements as prescribed by the IEEE Standards. The Converter provides Unity Power Factor Operation with 3.45% of THD and in other case a 0.818 Leading Power Factor with 2.83% of THD in input current. The only shortcoming that is needed to be overcome is the output quantities. The Output Voltage, current and Power are not the exact values as calculated as there is no close loop control at the output of the converter. This is not too important for a LED driver where in most of the cases it is used for domestic lighting purpose. All the output parameters are in acceptable range provided that they are aimed for domestic application. However if a more precise output is desired from the controller the current reference of the converter i_L^* must be generated by a close loop controller and an additional control loop is to be added at the output of the controller. Nevertheless, it will also be more complicated to control both the loops together using a single stage converter operation as the numbers of controlling parameters (i.e. Duty Cycle/Input Voltage) are limited. Hence, a double stage converter with more degrees of freedom will be needed to serve the purpose.

References

[1] P.C. Todd, "UC3854 Controlled Power Factor Correction Circuit Design", Unit-rode Application Note, 1998

[2] B. Singh, B. N. Singh, A. Chandra, K. Al-Haddad, A. Pandey, and D. P. Kothari, "A review of single-phase improved power quality AC–DC converters, "IEEE Trans. Ind. Electron., vol. 50, no. 5, pp. 962–981,Oct. 2003.

[3] B. Singh, B. N. Singh, A. Chandra, K. Al-Haddad, A. Pandey, and D. P. Kothari, "A review of three-phase improved power quality ac–dc converters, "IEEE Trans. Ind. Electron., vol. 51, no. 3, pp. 641–660, Jun.2004.

[4] H. S. Kim, M. H. Ryu, J. W. Baek, and J. H. Jung, "High-efficiency isolated bidirectional AC–DC converter for a DC distribution system,"IEEE Trans. Power Electron., vol. 28, no. 4, pp. 1642–1654, Apr. 2013.

[5] M. Arias, D. G. Lamar, J. Sebastian, D. Balocco, and A. A. Diallo, "Highefficiency LED driver without electrolytic capacitor for street lighting, "IEEE Trans. Ind. Appl., vol. 49, no. 1, pp. 127–137, Jan./Feb. 2013.

[6] B. Tamyurek and D. A. Torrey, "A three-phase unity power factor singlestage AC–DC converter Based on an interleaved flyback topology, "IEEETrans. Power Electron., vol. 26, no. 1, pp. 308–318, Jan. 2011.

[7] S. C. Moon, G. B. Koo, and G. W. Moon, "A new control method of interleaved single-stage flyback AC–DC converter for outdoor LED lighting systems, "IEEE Trans. Power Electron., vol. 28, no. 8, pp. 4051–4062,Aug. 2013.

[8] M. Narimani and G. Moschopoulos, "A new single-phase single-stagethree-level power factor correction AC–DC converter, "IEEE Trans. Power Electron., vol. 27, no. 6, pp. 2888–2899, Jun. 2012.

[9] C. Y. Oh, D. H. Kim, D. G. Woo, W. Y. Sung, Y. S. Kim, and B. K. Lee, "A high-efficient nonisolated single-stage on-board battery charger for electric vehicles, "IEEE Trans. Power Electron., vol. 28, no. 12, pp. 5746–5757, Dec. 2013.

[10] J. M. Kwon, W. Y. Choi, and B. H. Kwon, "High-performance plasma display panel sustain power supply," inProc. IEE Elect. Power Appl.,vol. 152, pp. 1381–1388, Nov. 2005

[11] K. Y. Lee and Y. S. Lai, "Novel circuit design for two-stage ac/dc converter to meet standby power regulations," IET Power Electron, vol. 2, no. 6,pp. 625–634, Nov. 2009.

[12] D. D.-C. L. Lu, H. H.-C. Iu, and V. Pjevalica, "A single-stage ac/dc converter with high power factor,regulated bus voltage, and output voltage,"IEEE Trans. Power Electron., vol. 23, no. 1, pp. 218–228, Jan. 2008.

[13] T. S. Kim, G. B. Koo, G. W. Moon, and M. J. Youn, "A single-stage powerfactor correction ac/dc converter based on zero voltage switching fulbridge topology with two series-connected transformers,"IEEE Trans. Power Electron., vol. 21, no. 1, pp. 89–97, Jan. 2006.

[14] Y. C. Li and C. L. Chen, "A novel single-stage high-power-factor ac-todc LED driving circuit with leakage inductance energy recycling, "IEEE Trans. Ind. Electron., vol. 59, no. 2, pp. 793–802, Feb. 2012.

[15] Y. Ji and Y. Xu, "Design and analysis of single-stage power factor correction converter with a feedback winding," IEEE Trans. Power Electron.,vol. 25, no. 6, pp. 1460–1470, Jun. 2010.

[16] D. D. C. Lu, H. H. C. Iu, and V. Pjevalica, "Single-stage ac/dc boostforward converter with high power and regulated bus and output voltage," IEEE Trans. Ind. Electron., vol. 56, no. 6, pp. 2128–2132, Jun. 2009.

[17] T. H. Hsia, H. Y. Tsai, D. Chen, M. Lee, and C. S. Huang, "Interleaved active-clamping converter with ZVS/ZCS Features, "IEEE Trans. PowerElectron., vol. 26, no. 1, pp. 29–37, Jan. 2011.

[18] T. F. Wu, J. C. Hung, S. Y. Tseng, and Y. M. Chen, "A single-stage fast regulator with PFC based on an asymmetrical half-bridge topology,"IEEE Trans. Ind. Electron., vol. 52, no. 1, pp. 139–150, Feb. 2005.

[19] M. Arias, M. F. Diaz, D. G. Lamar, D. Balocco, A. A. Diallo, andJ. Sebasti´ an, "High-efficiency asymmetrical half-bridge converter without electrolytic capacitor for low-output-voltage ac–dc LED drivers,"IEEE Trans. Power Electron., vol. 28, no. 5, pp. 2539–2550, May 2013.

[20] R. T. Chen, Y. Y. Chen, and Y. R. Yang, "Single-stage asymmetrical halfbridge regulator with ripple reduction technique,"IEEE Trans. Power Electron., vol. 23, no. 3, pp. 1358–1369, May 2008.

[21] S. Ou and H. Hsiao, "Analysis and design of a novel single-stage switching power supply with half-bridge topology,"IEEE Trans. Power Electron., vol. 26, no. 11, pp. 3230–3241, Nov. 201

[22] John Ware, "Power Factor Correction" IEE Wiring Matters; Issue 18; spring 2006.

An Efficient MAC Protocol for Provisioning Fairness in Vehicle to Roadside Communications

Md. Abubakar Siddik[1,*], Shafika Showkat Moni[2] and Mohammad Shah Alam[1]

[1] IICT, Bangladesh University of Engineering and Technology
[2] CSE, Rajshahi University of Engineering and Technology
[*]siddik.ece@hstu.ac.bd

Abstract—**In this paper, we present an efficient medium access control (MAC) protocol based on IEEE 802.11 DCF, which minimizes the unfairness problem of two way vehicle ad hoc networks. Geographically highly dynamic topology in vehicle ad hoc networks lead to unfairness problem. Unfairness problem has two aspects: higher velocity vehicles cannot transmit a minimum number of packets and lower velocity vehicles cannot transmit above a minimum number of packets. To address the above mentioned issues, our proposed scheme adjusts the transmission probability for each vehicle according to its residence time by changing the value of MAC parameters dynamically. An analytical model is developed to analyze the performance of the proposed protocol in a non-saturated state. We derive the relationship between the transmission probability and the residence time. Furthermore, relation between the transmission probability and the minimum contention window size of a vehicle is also derived to correlated residence time with minimum contention window size. Analytical results show that our proposed scheme overcomes the limitations of existing MAC protocols by ensuring that packet transmission rate remains proportional to the residence time of the vehicles.**

Index Terms—**IEEE 802.11 DCF; MAC; unfairness; residence time; non-saturated traffic.**

I. Introduction

Vehicular ad hoc network (VANET) is a special type of mobile ad hoc network (MANET) which is an important component of intelligent transport system (ITS). VANET is introduced to enable vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication, thereby facilitating a variety of safety and non-safety applications [1, 2]. It has some unique characteristics such as geographically highly dynamic topology, predictable two direction mobility etc. which constitute VANET as a distinct research field in MANET [3]. An important challenge to design MAC protocols of VANET is unfairness [4]. To address these problems, an efficient MAC protocol is needed to coordinate the channel access of multiple vehicles.

Vehicles having two moving direction with different velocities have different residence time in the coverage area of an RSU. Due to small residence time of high velocity vehicles, IEEE 802.11 DCF protocol [5], a random access mechanism based on carrier sense multiple access with collision avoidance (CSMA/CA), provides less opportunity to communicate with RSU and consequently transfers less amount of data as compared to the low velocity vehicles. To address unfairness problem, E. Karamad and F. Ashtiani [6] proposed a modified IEEE 802.11 DCF based MAC protocol, which provides same opportunities all the vehicles by dynamically changing the MAC parameters. Although their protocol provides minimum opportunity of data transmission to vehicles with high velocity, it restricts the low velocity vehicles to a certain limit of data transmission. The former protocol is applicable for non-safety applications and later protocol for safety applications.

In this paper, we propose a new MAC protocol based on IEEE 802.11 DCF which has taken into account the residence time of the vehicles. According to residence time, all vehicles are grouped into three batches. Each vehicle calculates and updates their own batch number by considering their position, direction and velocity after every certain time interval. Every vehicle must have to chance to enter the high priority batch. If the batch number of a vehicle is changed then the vehicle adjust the MAC parameters to the respective batch dynamically. Based on the analytical model, we set the MAC parameters to each bath individually to ensure that high velocity vehicles can transmit a minimum number of packets and packet transmission rate of other vehicles is proportional to their residence time. The main idea behind our proposed protocol is that each vehicle dynamically updates and adapts their transmission rate according to their residence time by resetting MAC parameters after a certain time interval.

The remainder of this paper is organized as follows. Section II reviews related works. Section III describes the system model considered in this work. Section IV derives the performance models for IEEE 802.11 based MAC protocol. Section V shows the results of the derived analytical models. Section VI concludes this paper.

II. Related Works

Proposal of MAC protocol for V2I communications in VANET based on 802.11 DCF to ensure fair access and its performance Analysis has recently been studied in the literature [6-7, 9-10]. In [7], W. Alasmary and W. Zhuang proposed two dynamic CW based mechanisms to alleviate the performance degradation caused by different velocity of vehicles. However, they do not describe the exact procedure for the selection of optimal contention window size to avoid unfairness problem. In [6], to address unfairness problem, E. Karamad and F. Ashtiani proposed a modified IEEE 802.11 DCF based MAC protocol which dynamically assign minimum contention window to each vehicle batch at the beginning of entering time of the network for whole residence time by adjusting transmission probability with its speed to achieve fair access considering the network is in a saturated state. Moreover, they derived the relations between average velocity and the minimum contention window through analytical approximations. Performance of this protocol is analysed based on Bianchi model [8]. In [9], Q. wu and J. Zheng represented a MAC protocol proposed by E. Karamad and F. Ashtiani and analysed the performance by considering

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

the network is in a non-saturated state. Both authors [6, 9] consider a network model with following assumptions:

1) The vehicles arrive in the network in batches according to Poisson process with constant arrival rate.
2) The velocity of each batch's vehicles is same and constant for whole residence time of a particular RSU.
3) Moving direction and batch number of the vehicle do not change in residence time.

However, both authors do not ensure that packet transmission rate is proportional to the residence time of the vehicles.

In [10], V. P. Harigovindan, A. V. Babu and L. Jacob are derived the analytical expression for optimum CW required to ensure fairness, in the sense of equal opportunity to communicate with RSU among the vehicles of different average velocity in the network.

III. System Model

A. Network Architecture

We consider a simple traffic model for vehicles which are moving along a straight road with both directions covered by one AP or RSU at a fixed position on the road divider. We assume that the vehicles arrive in the network according to a Poisson process and no transmission error due to defective channel. Each vehicle has a queue independently contending for transmission and the links between vehicle and RSU both are symmetric. We ignore the effect of hidden terminals, exposed terminals and channel capture. In the coverage area, the vehicles may change the moving direction and velocity. Each vehicle must have capacity to measure its position, moving direction and velocity. The network has three classes of vehicles which are grouped according to residence time.

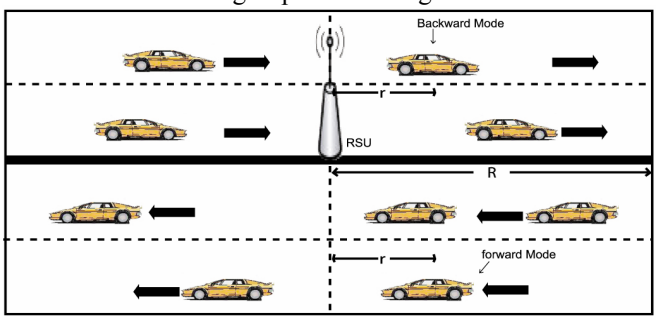


Fig. 1 A Vehicle to Infrastructure (V2I) based VANET model.

B. Proposed IEEE 802.11 DCF Based MAC Protocol

Our proposed IEEE 802.11 DCF based MAC protocol is a channel access mechanism designed for avoiding unfairness problem at the MAC layer. It defines three classes of vehicle in whole network covered by a RSU and denoted by batch 1, batch 2 and batch 3. Each batch vehicle has a queue independently contending for transmission with its own MAC parameters, including minimum contention window (CW_{min}), minimum contention window (CW_{max}), maximum backoff stage(m) and the retransmission limit value ($m+x$). In our proposed mechanism, after entering to the network, every vehicle computes their velocity, moving direction and position in the network. After that, vehicle calculates their residence time and then based on their residence time; vehicles are divided into three batches. Each vehicle updates their batch number after a successful transmission, after a packet drop or after a predefine time interval when the vehicle has empty queue. When the residence time is reaches to zero, the vehicle exit from network and continue the above procedures under the next RSU.

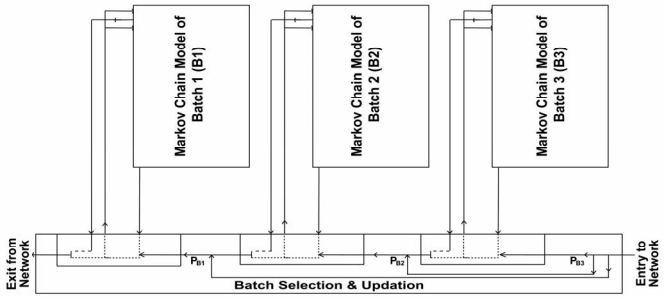


Fig. 2 Channel access mechanism of proposed MAC protocol.

IV. Performance Model for IEEE 802.11 based MAC Protocol

A. Batch Selection

We consider the network model described in figure 1. The instantaneous residence time, T_r of a vehicle can be express as $T_r = \frac{R \pm r}{v}$, where, R is the radius of the coverage range of RSU, r is the instantaneous distance of a vehicle from RSU, v is the instantaneous velocity of a vehicle, '+' indicates the moving direction of vehicle toward to RSU and '-' indicates the moving direction of vehicle backward from RSU. Let, the velocity of the vehicles is uniformly distributed in $[v_{min}, v_{max}]\, ms^{-1}$. Selection of batch number of the vehicles is determined by the table I. The packet transmission rate, R_i of a vehicle in batch i ($i = 1,2,3$) can be expressed as, $R_i = H \frac{R_{bit}}{N_{bit}} \frac{P_{t(i)}}{\sum_{k=1}^{3} P_{t(k)}}$; where, H is the normalized throughput of the network, N_{bit} is the average number of bits in the packet, R_{bit} is the bit rate over the channel and $P_{t(i)}$ is the transmission probability of the vehicle in batch i. To avoid unfairness by ensuring minimum access, the total number of packets transmitted within the average residence time of vehicle of each batch in the network is equal to N_P. Thus we have $R_i T_{r(i)} = N_P$; where, N_P is the minimum number of packets transmitted by every vehicle.

TABLE I
Batch Selection and Updation

Instantaneous Residence Time (T_r)	Batch No (Bi)	Maximum Residence Time($T_{r(i)}$)	Priority
$0 < T_r \le T_{r(min)}$	$B1$	$T_{r(min)}$	High
$T_{r(min)} < T_r \le T_{r(in)}$	$B2$	$T_{r(in)}$	Medium
$T_{r(in)} < T_r \le T_{r(max)}$	$B3$	$T_{r(max)}$	Low

B. Markov Models

We construct a 2-D markov chain to describe the post-backoff procedure and backoff procedure of vehicle queue for non-saturated and saturated state respectively, which is shown in fig. 3. According to the 2-D markov chain, all stationary probabilities $b_{i,j,k}$ are expressed by $b_{i,0,0}$ and which is finally determined by imposing the normalization condition, as follows:

$$\frac{1}{b_{i,0,0}} = 1 + \frac{(1-P_{eqat(i)})(W_{i,0}-1)}{2P_{idle(i)}} + \sum_{j=1}^{m_i+x_i} P_{coll(i)}^{j} \left(1 + \frac{W_{i,j}-1}{2P_{idle(i)}}\right) + \frac{P_{eqat(i)}}{W_{i,0}} \left[\frac{1-P_{eq(i)}{}^{W_{i,0}}}{1-P_{eq(i)}} \left\{\frac{1-P_{idle(i)}}{2P_{idle(i)}} \left(\frac{2P_{eq(i)}}{1-P_{eq(i)}} + (W_{i,0}-1)\right) - \frac{1}{P_{eq(i)}}\right\} + \left\{\frac{2(2-P_{eq(i)})}{P_{idle}(1-P_{eq(i)})} + \frac{1-P_{eq(i)}}{P_{eq(i)}}\right\}\right] \quad \text{(i)}$$

Packet transmission occurs if backoff counters becomes zero, regardless of the backoff stage. So, packet transmission probability of a vehicle in random time slot express as:

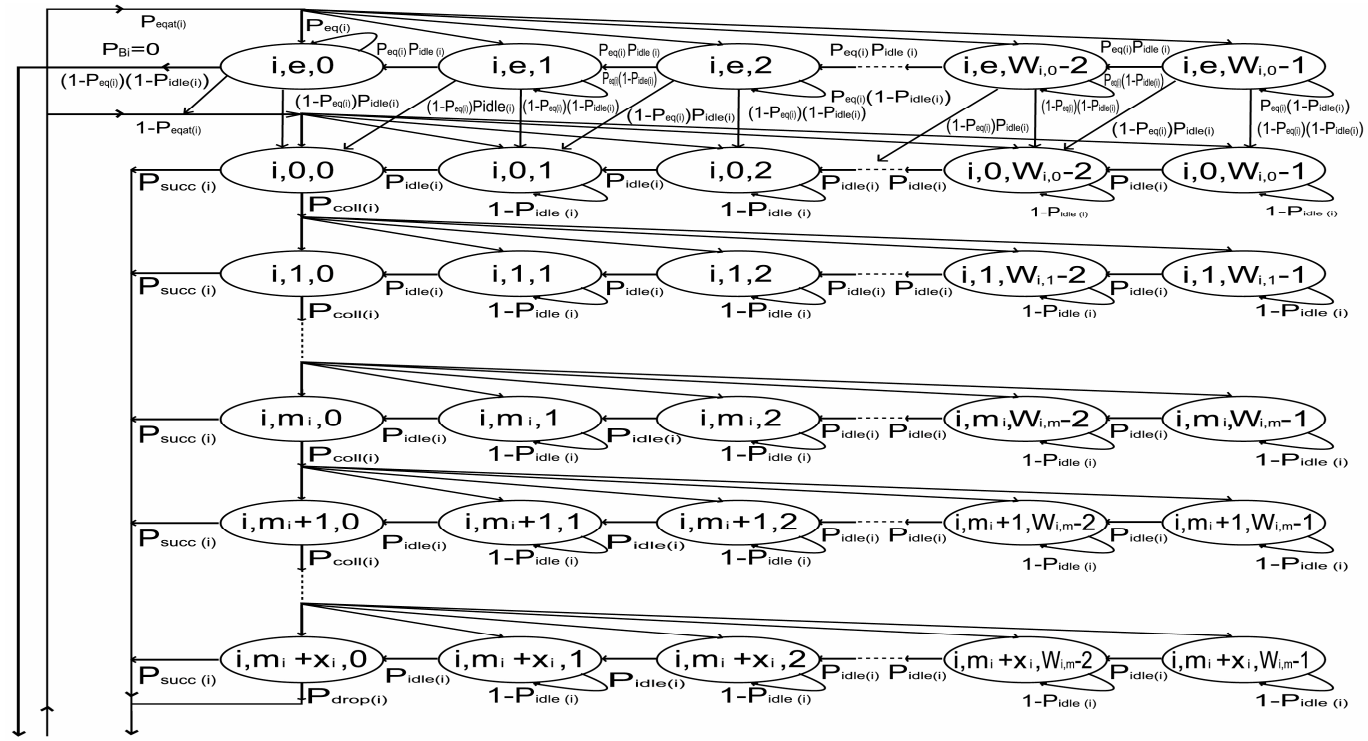


Fig. 3 Markov Chain model for the backoff procedure of a vehicle with both saturation and non-saturation state.

$$P_{t(i)} = b_{i,0,0}\,\frac{1-P_{coll(i)}^{m_i+x_i+1}}{1-P_{coll(i)}} \quad \text{..... (ii)}$$

Substituting eq. (i) and eq. (ii), we get the packet transmission probability of a vehicle in a randomly chosen time slot. However, $P_{t(i)}$ depends on the conditional collision probability, $P_{coll(i)}$ and channel idle probability, $P_{idle(i)}$ which is still unknown. A vehicle of batch i is sensed idle channel if there is no other vehicle of batch i and other batch transmits simultaneously. A successful transmission occurs if only one vehicle transmits packet in a time slot and collision occurs when more than one vehicle transmits packet in a time slot. So we can calculate $P_{idle(i)}$, $P_{succ(i)}$ and $P_{coll(i)}$ as follows:

$$P_{idle(i)} = \prod_{i=1}^{3}(1-P_{t(i)})^{n_i}$$

$$P_{succ(i)} = \sum_{i=1}^{3}\frac{n_i P_{t(i)}}{1-P_{t(i)}}\prod_{k=1}^{3}\left(1-P_{t(k)}\right)^{n_k}$$

$$P_{coll(i)} = 1-\sum_{i=1}^{3}\frac{n_i P_{t(i)}}{1-P_{t(i)}}\prod_{k=1}^{3}(1-P_{t(k)})^{n_k} - \prod_{i=1}^{3}(1-P_{t(i)})^{n_i}$$

According to Bianchi model [8], the normalized throughput of batch i of the network can be calculated as:

$$S_i = \frac{T_p P_{succ(i)}}{T_{succ}P_{succ(i)} + T_{coll}P_{coll(i)} + \sigma P_{idle(i)}}$$

Where, T_p average packet length in time, σ is the time duration of an empty slot, T_{coll} is the average time containing two or more colliding packet and T_{succ} is the average time containing a successfully transmitted packet.

V. Results

In this section, we carry out the performance analysis of the proposed MAC protocol based on the analytical model developed in section IV. For simplicity, we assume that all types of vehicles transmit packets of equal lengths (216 bytes) and have an equal probability (0.01) that a queue has packets waiting for transmission. Fig. 4 shows the relationship between the transmission probability of the vehicle and the minimum contention window size when the velocity of the vehicle is [5, 45] ms^{-1} and the diameter of the coverage area is 500 meters. It is seen that minimum contention window size increases with the transmission probability decreasing. According to figure 4, we set the MAC parameters including minimum contention window size, maximum backoff stage and retransmission limit which are listed in table II.

Fig. 5 shows the throughput of each batch vehicles under different number of vehicles in the network. Throughput of each batch reaches a maximum value with the number of vehicles varying. This is because the small number of vehicles leads to a low throughput for each vehicle. With the number of vehicles increasing within a certain range, it would not cause many collisions, which resulting in an increase of the throughput for each batch vehicles. With the number of vehicles further increasing, however, more vehicles will arrive in the network and will content for transmission simultaneously, which would result in more collisions and thus a reduction in the throughput for each batch vehicle.

TABLE III
Parameters Used in Numerical Calculations

Parameter	B1	B2	B3
CW_{min}	3	9	30
CW_{max}	12	112	1280
m	2	4	7
x	4	2	0

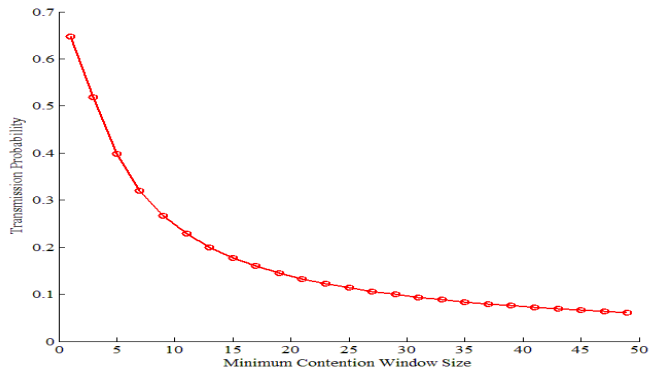


Fig. 4 Relation between transmission probability and minimum window size.

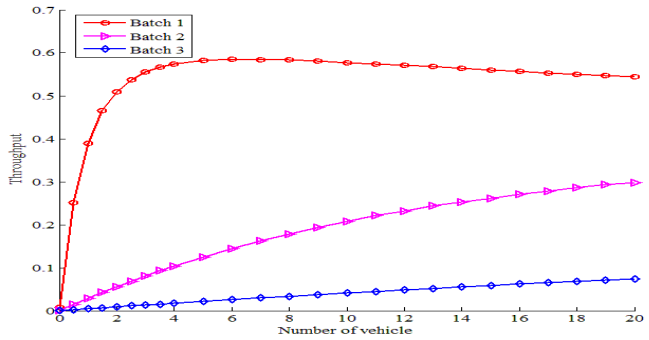


Fig. 5 Throughput of the proposed scheme for different number of vehicles.

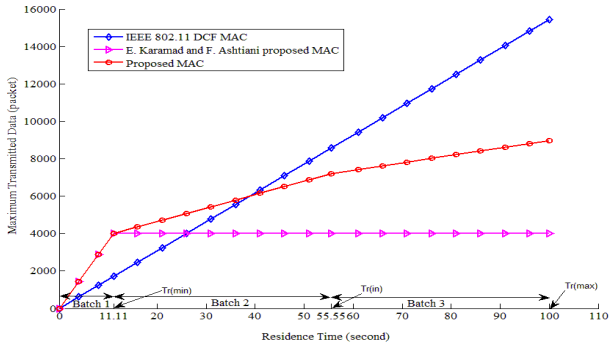


Fig. 6 Relationship between maximum transmitted data and residence time.

Fig. 6 shows the amount of maximum transmitted data (packet) under different MAC protocols. It is seen that IEEE 802.11 DCF MAC protocol offers a maximum transmitted data which is proportional to the residence time of the vehicles because a common MAC parameter is used by all the vehicles. Moreover, Karamad proposed MAC protocol offers an equal maximum transmitted data regardless of residence time because each batch vehicles use different MAC parameters which are determined by considering total number of packets transmitted within the residence time of each batch in the network should be equal.

The proposed protocol offers an opportunity to transmit a minimum number of packets for all vehicles and amount of maximum transmitted data is proportional to their residence time. This is because different batch vehicles use different MAC parameters and after certain time interval, vehicles update their batch number by resetting the higher MAC parameters.

Fig. 7 shows maximum number of packets transmission under different MAC protocols with various velocities. It is observed that in IEEE 802.11 DCF MAC, variation in data transmission amount between high and low velocity vehicles increases with velocity increment without ensuring a minimum access for high velocity vehicles. For Karamad protocol, the amount of packet transmission is constant regardless of velocity variation. However, our proposed protocol clearly demonstrates that data transmission amount remains proportional to residence time as well as a minimum channel access for high velocity vehicles is ensured .

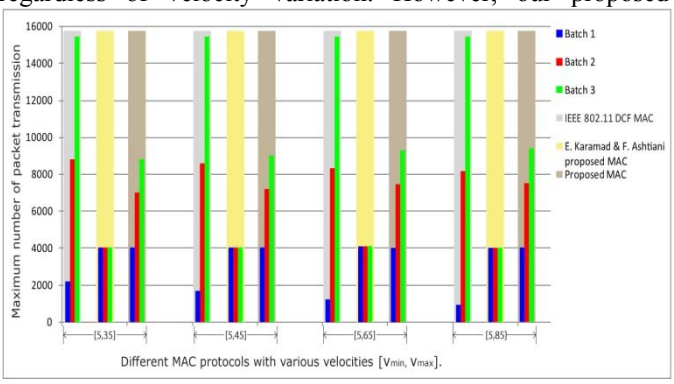


Fig. 7 Performance comparison of different MAC protocols (in terms of maximum number of packets transmitted at various velocities).

VI. Conclusions

In this paper, we considered the unfairness problem in vehicular ad hoc networks arisen due to the varying residence time of the vehicles. We proposed an efficient MAC protocol based on 802.11 DCF which adapts the MAC parameters of each vehicle according to its residence time and updates the MAC parameters after certain time interval. Every vehicle has chance to adapt at least higher priority MAC parameters. As a result, fairness in the sense of minimum chance of communicating with RSU for the vehicles having minimum residence time and time proportional chance of communicating with RSU for the other vehicles is achieved. Furthermore, we developed analytical models for analyzing the performance of the proposed MAC protocol in a non-saturated state take into account all major factors that could affect the access performance of the proposed MAC protocol, including the saturation condition, channel idle, packet drop and retransmission limit . Based on the analytical model, we derived the relation between the transmission probability and minimum contention window size, and the relationship between number of vehicles and throughput. Moreover, we can determine the value of MAC parameters for a given velocity in order to achieve fair access. Finally, we show the relationship between maximum transmitted data and residence time of the vehicles to compare the performance of the proposed MAC protocol with basic access mechanism IEEE 802.11 DCF MAC and Karamad MAC protocol. Accordingly, the proposed MAC protocol overcomes the unfairness problem by ensuring minimum access all the vehicles for both safety and non-safety applications. For future work, we will study the unfairness problem of the IEEE 802.11p EDCA mechanism.

References

[1] "Standard specification for telecommunications and information exchange between roadside and vehicle systems – 5 GHz band dedicated short range communications (DSRC) medium access control (MAC) and physical layer (PHY) specifications", ASTM E2213-03, 2010.

[2] "IEEE draft guide for wireless access in vehicular environments (WAVE)-architecture", IEEE P1609.0TM/D7.0, Aug. 2013.

[3] B. Yu, C. Xu, "Vehicular Ad-Hoc Networks: An information-centric perspective", ZTE communications, Vol. 8, No. 3, pp.42-49, 2010.

[4] N. Gupta, A. Prakash and R. Tripathi, "Medium access control protocols for safety applications in Vehicular Ad-Hoc Network: A classification and comprehensive survey" Vehicular Communications, Vol. 2, No. 4, pp. 223–237, Oct. 2015.

[5] "IEEE Standard 802.11, part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications", Nov. 1997.

[6] E. Karamad and F. Ashtiani, "A modified 802.11-based MAC scheme to assure fair access for vehicle-to-roadside communications", Computer Communication, Vol. 31, No. 12, pp. 2898–2906, July 2008.

[7] W. Alasmary and W. Zhuang,"Mobility impact in IEEE 802.11p infrastructureless vehicular networks", Vehicular Technology Conference Fall, 2010 IEEE 72nd, pp. 1-5, Sept. 2010.

[8] G. Bianchi, "Performance analysis of 802.11 distributed coordination function," IEEE Journal on Selected Areas in Communications, Vol. 18, No. 3, pp. 535-547, Mar. 2000.

[9] Q. Wu and J. Zheng, "Performance modeling of IEEE 802.11 DCF based fair channel access for vehicular-to-roadside communication in a non-saturated state", 2014 IEEE International Conference on Communications (ICC) , pp. 2575–2580, June. 2014.

[10] V. P. Harigovindan, A. V. Babu and L. Jacob, "Ensuring fair access in IEEE 802.11p-based vehicle-to-infrastructure networks", EURASIP Journal on Wireless Communications and Networking, May 2012.

Loadability Improvement of an Existing AC Line by Transmitting Simultaneous AC and DC Power

Mohammad Tawhidul Alam,[1] R. Sadnan,[1*] Irtiza Haque,[1] Mushfequr Rahman,[1] and Q. Ahsan[1]

[1] Dept. of EEE, Bangladesh University of Engineering and Technology, Dhaka, Bangladesh

[*] rabayet.sadnan@gmail.com

***Abstract*—To increase the power flow through an existing AC line, different compensator devices are used. Flexible alternating current transmission system (FACTS) are very common for the enhancement of power transmission capability of the line. Due to stability reasons, these compensation methods have limitations. Another alternative is to construct new AC transmission lines, which is not only expensive, but also time consuming. And it is not viable for every situation. However, transmitting both AC-DC power through the same line can increase the power flow without impairing the stability of the network. This paper investigates the increased power flow through an AC line when it is used to transmit both AC and DC power simultaneously. A model of a long transmission line carrying simultaneously both AC and DC power is studied for this purpose. The model uses a system of transformer and rectifier to produce and carry such power from a generating source to an infinite bus through the line. The study compares the power flow of the traditional AC line and the proposed simultaneous AC-DC power transmission.**

***Index Terms*—Zig-Zag transformer, loadability improvement, Simulink, long transmission line, simultaneous AC-DC power transmission.**

I. Introduction

Prevalent technology has resulted in most transmission lines around the world to be high voltage 3-phase alternating current (AC) lines. Towers have been constructed and tower insulation adopted with appropriate dimensions according to the voltage levels of transmission lines. Increasing loads in city centres indicate that even more power now needs to be transported along these corridors. In order to meet the rapidly increasing electrical demands in many load centres all over the world, detailed studies have been undertaken to evaluate the best method of upgrading power corridors. To transfer power more efficiently, traditional systems used high voltage AC (HVAC), or sometimes to increase the power flow controllability and capability FACTS compensators are used. However, many new ones utilise the benefits of high voltage DC (HVDC). HVDC transmission has the advantage of being more stabilized system. With the presence of current technology, HVDC transmission lines show cheaper costs for cable systems having a length greater than 25km [1] and overhead lines longer than 500km [2]. In [3], Larruskain discusses the various ways in which an upgrade can be made in order to increase the transmission capacity of the existing AC transmission lines. These methods include increasing voltage (substituting old lines with new compact AC lines), increasing current density (using high temperature - low sag conductors) and using the AC lines to carry DC power.

In [4], Clerici describes the upgradation of HVAC lines to HVDC lines, which provides a power upgrade of about multiple times than the conventional transmission. The differences in AC and DC towers are compared and feasibility of conversion is discussed. However, substantial changes must be made to the towers leading to significant cost and outages. [5] and [6] also discuss the conversion of AC lines to DC and the efficiency of such conversions with case studies, where conversion has been carried out. But for this conversion tower modifications and change of insulators are required.

Flexible AC Transmission System, a state-of-the-art of series compensation, uses power electronics technology to increase the power transmission capability of the existing AC line. Although series compensation improves the loading capability of a transmission line, it generates two protection problem. One is the protection of the capacitor banks against over voltages and another is the protection of the generator shaft system against torsional stress under subsynchronous resonance(SSR) condition. These two problems are considered as limiting factors for designing the amount of series compensation in electrical network.

To avoid the limitations of the above mentioned approaches of loadability improvement of the existing transmission line, H. Rahman and B. H. Khan proposed a novel approach in [7], [8] and [9] for simultaneous AC and DC power transmission, where the magnitude of DC and AC voltage were fixed while the power angle was varied. In this method, the single circuit line is able to carry more power without any changes to the tower structure This method has the clear benefit of requiring only changes to be made at the ends of the line and incur minimum outages in the conversion process. The study was undertaken in PSCAD/EMTDC simulation software.

This paper presents, a new method to assuage the high requirement problem of electrical power transmission by mixing the DC voltage with qualified AC voltage, without altering the tower structure or cable. In order to reach the thermal limit of the conductors, percentage of injected DC voltage is increased. Additionally, to keep the peak voltage constant, percentage of AC voltage is reduced. DC power is superimposed on AC power which ultimately increases total power flow through the line. Again varying the transmission angle, that power loadability can also be improved is shown here. The model is applied to realistic systems and the ranges of the variables are evaluated. The operation of a simultaneous AC-DC system is beneficial to these variables in terms of higher power transmission compared to pure AC system. This evaluation will help to develop the operational strategies of a simultaneous AC-DC transmission system.

In this paper, the proposed model along with the necessary mathematical expressions are provided in section II. In section III, the simulated results of the proposed system are presented along with the details. The observations are described and concluding remarks are drawn in section IV.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

II. Simultaneous AC-DC Power Transmission Model

A simultaneous AC-DC power transmission system transmits both AC and DC power over the same conductor. It only needs converters and zig-zag transformers to make the line compatible to carry both power. The conceptual schematic diagram of such system is shown below in Fig. 1.

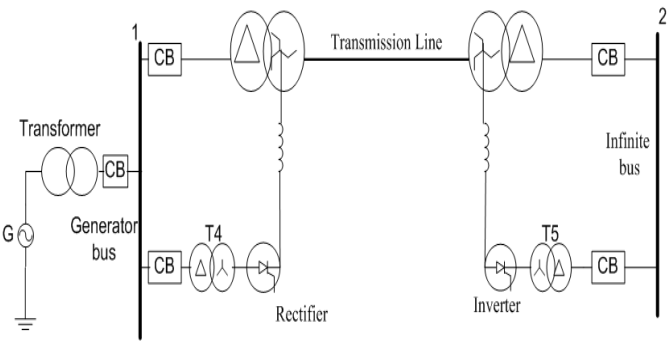


Fig.1 AC-DC composite power transmission system

In this system, a portion of generated AC power is converted into DC by controlled rectifier, and then injected into the long transmission line through the neutral point of a zigzag transformer at the sending end. At the receiving end of the transmission line, injected DC power is collected from the neutral point of zigzag transformer and converted into AC using inverter before feeding to the infinite bus.

Considering the transmission line of Fig.1 as a composite AC-DC system of transmission, the combined power flow between two buses, 1and 2, can be written as,

$$P_{comb} = \varepsilon P \tag{1}$$

Where,

$$\varepsilon = [\beta' + \beta''] \tag{2}$$

$\beta'$and β'' are constant and may be expressed as –

$$\beta' = (1-k)\beta \tag{3}$$

And β can be expressed in terms of sending and receiving end phase voltage, $E_{\emptyset s}$ and $E_{\emptyset r}$ respectively and the percentage DC voltage mix k and the reactance between sending and receiving end bus, X_{sr} as -

$$\beta = \frac{(1-k)\dfrac{E_{\emptyset s}\angle\delta_{ac} - E_{\emptyset r}\angle 0}{jX_{sr}}}{\dfrac{E_{\emptyset s}\angle\delta_{sr} - E_{\emptyset r}\angle 0}{jX_{sr}}} \tag{4}$$

Another constant β'' may be expressed as,

$$\beta'' = \sqrt{2}\, k\gamma \tag{5}$$

Where, $\gamma = \frac{I_{DC}}{I_L}$

I_L is the current flowing through the pure AC line. And I_{DC} is the DC current flowing through the conductor of a simultaneous AC-DC system

III. Simulated Result

To study the loadability of an existing AC line by transmitting AC-DC power simultaneously, two different transmission lines are considered. One is 804 km line Colstrip to Taft of USA, and another one is 500 kV 350 km line. Montana 500kV transmission line [10] is used to carry power from East, Colstrip, to west, Taft. The necessary parameters of the Montana 500kV transmission system are presented in Appendix-A. The maximum power transfer through the transmission line is evaluated considering,

a) Pure AC transmission and
b) Simultaneous AC-DC transmission.

The simultaneous AC-DC system of power transmission for different percentage of DC voltage mix is evaluated to determine the maximum power transmission. In this case, the transmission angle, relating to AC power flow of AC-DC system is equal to 44.47^0. And, the current flow is considered to be equal the thermal limit of the conductor. Though in pure AC system, the current – extended to the thermal limit of conductor, is not allowed to flow due to the stability constraint. Note that in the evaluation of the maximum power flow in a pure AC system, 30% stability margin is considered.

The possible maximum power transfer through the stated transmission system of Montana, for both simultaneous AC-DC and pure AC transmission system, are compared in Table I.

TABLE I
Comparison Of Maximum Power Flow At Different DC Voltage Mix

V_{DC} Mix(%)	Maximum power flow(MW)		Percentage of increased power flow in AC-DC system
	AC-DC system	Pure AC system	
20	2141	1028	108
30	2703	1028	162
40	3274	1028	218
49.5	3828	1028	272

Then, the maximum power transfer through the same 500kV Montana transmission line is estimated by varying the transmission angle at a fixed percentage of DC voltage mix of 49.5%. In this case, the DC current flow is attuned to extent the total current flow to the thermal limit of the conductor. The results are compared in Table II with that of pure AC system for this case.

TABLE II
Comparison of Maximum Power Flow at Different Transmission Angle

Transmission angle(Degree)	Maximum power flow(MW)		Percentage of increased power flow in AC-DC system
	AC-DC system	Pure AC system	
20	3675	1028	257
30	3745	1028	264
40	3805	1028	270
50	3854	1028	275
60	3892	1028	279

It may be stated that, DC voltage mix and transmission angle are two control variables to govern the power flow in a simultaneous AC-DC transmission system. Tables I and II show that with the increment of DC voltage mix or the transmission angle the maximum power flow in a composite AC-DC system increases.

It is certain that the scope of the increment of power flow through the line exists by controlling the transmission angle, though the mixing of DC voltage stretches to the limit. From the table, it is observed that the increment of power transfer is not linear with respect to percentage of DC voltage. It is perceived that, the average rate of the power transfer increment in composite AC-DC system is almost 10 times higher for DC voltage mix than controlling transmission angle.

It is also observed that, the composite AC-DC system transmits 2.79 times more power than existing pure AC system when the two controlling variables reach their respective limits. The two separate observations for two different controlling parameter indicate that in a simultaneous AC-DC transmission system, higher increment of power flow can be acquired by mixing the DC voltage with AC voltage, than by the control of transmission angle.

Another transmission system has been studied for reaching the maximum power transfer through the existing transmission line by changing the percentage of DC voltage. The model is simulated using MATLAB/Simulink software. Like the first case, the power transfer increased commensurately with the level of DC voltage mix. While the transmission angle varied from 35^0 to 27^0. The transmission system for this simulation is described in Appendix-B. The power (AC, DC, and total power) transfer with respect to different level of DC voltage are compared in Table III.

TABLE III
POWER MEASUREMENTS FOR DIFFERENT LEVEL OF DC VOLTAGE

DC (%)	P_{DC} (MW)	P_{AC} (MW)	P_{TOTAL} (MW)	Increase in P (%)
0.0	0.00	111.00	111.00	0.00
11.0	15.22	96.50	111.72	0.74
14.5	26.59	89.00	115.59	4.23
18.0	40.85	81.20	122.05	10.05
22.0	58.36	75.70	134.06	20.89
27.0	91.20	62.40	153.60	38.50
30.0	107.12	57.00	164.12	47.99
34.5	130.28	49.60	179.88	62.20
49.5	221.04	24.90	245.94	121.79

It is seen from table III that as the percentage of DC voltage mix increases, total power flow increases. The relationship is non-linear and the increment of power is higher as the percentage of DC increases.

Now, if the percentage of DC voltage vs Powers (P_{AC}, P_{DC} and P_{TOTAL}) are plotted the DC power increases quickly with the increasing DC voltage percentage whereas the AC power decreases at the same time in a non-linear manner. However, the total power flow increases. The result is similar to the first case. Here, for this case, the control variable is limited to the level of DC voltage, the transmission angle is not controlled for this simulation model.

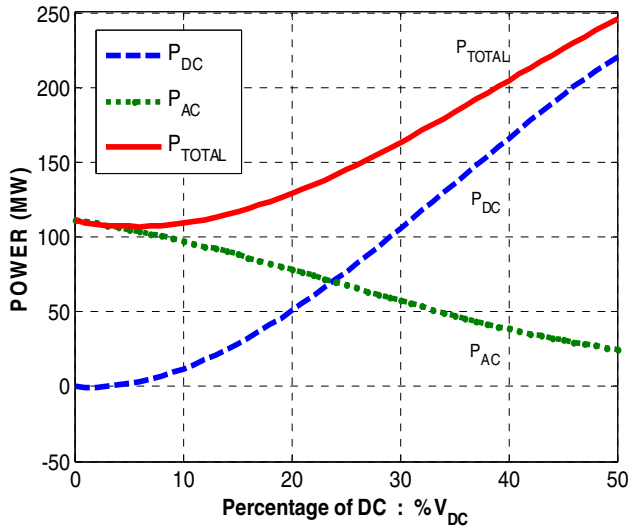


Fig. 2 Relation between percentage of V_{DC} and Power

Fig. 2 clearly shows that, the total power transmitted through the transmission line increases with the increment of the percentage of DC voltage mix as expected.

IV. CONCLUSION

The result shows that while maintaining the existing tower configurations, substantial power capacity upgrade is possible. The only change required is the setup of converter stations, which significantly reduces the outage time for upgrading work. The application of the mathematical model in a realistic system evidently reveals that the maximum power flow through the simultaneous AC-DC system using the existing AC line is much higher than that of existing pure AC system. The combined AC-DC system has the advantage of allowing the magnitude of current flow through the transmission line very close to the conductor's thermal limit, which is not possible for pure AC system. The application additionally states that, in the composite AC-DC transmission system, superior variation of power flow can be achieved through the existing AC line by controlling the DC voltage mix, when the technique has been compared with the scheme of controlling transmission angle. The advantage of such hybrid AC and DC transmission is the huge savings in cost due to lack of tower upgrade requirements and furthermore it saves the costs of compensation devices.

APPENDIX-A

Montana 500kV transmission system transmits power from the East (Colstrip) to west (Taft). The eastern Montana generation system has four coal-fired thermal units of a total capacity of 2272MW [10]. For the application of the proposed model only 500kV double circuit transmission line from Colstrip to Taft, excluding other portion of the power system, is considered. Colstrip with generation capacity of 2272MW and Taft are considered as generator and infinite bus, respectively. The simulation also does not consider Broadview and Garrison buses in between Colstrip and Taft. The

parameters of different components of this system are presented in Table A.

TABLE A
PARAMETERS OF THE DIFFERENT COMPONENTS OF MONTANA 500KV TRANSMISSION SYSTEM

Sl. No.	Component	Parameter
01	Line	(i) x = j253.21 Ω/phase/circuit (iii) Double circuit (iv) Three phase (v) 60Hz (vi) 804km (vii)500kV
02	Generator	358*2, 778*2(MW) 24kV,Reactance=0.3pu
03	Generator Transformer	24/230kV, Leakage reactance = 0.15pu.
04	Transformer (At the sending end of the line)	Δ-Y, 230/500kV, leakage reactance = 0.1pu. (pure AC) Δ-Z, 230/252.5kV, Leakage reactance = 0.1pu. (AC-DC)
05	Transformer (At the receiving end of the line)	Z-Δ, 252.5/500kV, Leakage reactance= 0.1pu, (AC-DC)
06	DC system	DC system rated voltage (maximum) and current are 202kVand 9kA, respectively.

APPENDIX-B

The system's frequency is 60Hz. The generator was of 1000MVA which is the nominal power for the system. The generator has a stator resistance of 2.8544×10^{-3} Ω pu. The nominal voltage for the system is selected 500 kV. The receiving end of the transmission line is connected with an infinite bus.

Transmission line parameters are calculated by the Compute RLC Line parameter tool of Matlab/Simulink. The parameters are given below in the Table B

TABLE B
TRANSMISSION LINE PARAMETERS FOR SIMULATION MODEL

Impedence	R (Ω/km)	L (H/km)	C (F/km)
Positive sequence	0.011688	0.0008683	1.3434×10^{-08}

REFERENCES

[1] V. Sood, HVDC and Facts Controllers - Applications of Static Converters in Power Systems. Kluwer Academic Publishers, 2004. ISBN 1-4020-7891-9.

[2] M. P. Bahrman, "Overview of HVDC transmission," *PSCE*, pp. 18–23, 2006.

[3] Larruskain, D. M., et al. "Power transmission capacity upgrade of overhead lines." *International Conference on Renewable Energy and Power Quality (ICREPQ), Spain*. 2006.

[4] Clerici, L. Paris, and P. Danfors, "HVDC conversion of HVAC line to provide substantial power upgrading," *IEEE Trans. Power Del.*, vol. 6, no. 1, pp. 324–333, Jan. 1991.

[5] Häusler, Michael, Gernot Schlayer, and Gerd Fitterer. "Converting AC power lines to DC for higher transmission ratings." *ABB REVIEW* (1997): 4-11.

[6] Matabaro, Mushagala, and Gary Atkinson-Hope. "Efficiency of an AC transmission line converted into a DC system."

[7] Rahman, Hafizur, and B. H. Khan. "Enhanced power transfer by simultaneous transmission of AC-DC: a new FACTS concept." *Power Electronics, Machines and Drives, 2004.(PEMD 2004). Second International Conference on (Conf. Publ. No. 498)*. Vol. 1. IET, 2004.

[8] Rahman, H., and B. H. Khan. "Power upgrading of transmission line by combining AC–DC transmission." *Power Systems, IEEE Transactions on* 22.1 (2007): 459-466.

[9] Rahman, H., and B. H. Khan. "Stability improvement of power system by simultaneous AC–DC power transmission." *Electric Power Systems* Research78.4 (2008): 756-764.

[10] Kosterev, D. N., W. A. Mittelstadt, R. R. Mohler, and W. J. Kolodziej. ac"An application study for sizing and rating controlled and conventional series compensation." In Northcon 95. I EEE Technical Applications Conference and Workshops Northcon95, pp. 280-286. IEEE, 1995.

Optimization of CNFET Op Amp for High Frequency Operation in Sub-10-nm Node

Md Zubair Ebne Rafique,* Alam Mahmud, and Sharif Mohammad Mominuzzaman

Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology (BUET)
Dhaka-1000, Bangladesh

*zirshovon@gmail.com

Abstract—**This paper presents optimization of various parameters of carbon nanotube field effect transistor (CNFET) based op amp in sub-10 nm node for high frequency operation. Parameters such as, channel length, gate oxide thickness, gate dielectric constant, spacer doping, spacer dielectric constant and pitch, are optimized. For high frequency operation of op amp, optimum values of CNFET parameters are found to be 10 nm channel length, 10 dielectric constant of gate oxide, 3 nm gate oxide thickness, 5 nm pitch, zero source/ drain doping concentration, 4 dielectric constant of spacer oxide. Performance of CNFET op amp incorporating these optimum values is investigated and compared with previous work. Obtained results reveals significant differences among previously designed Si CMOS op amp, 32 nm CNFET and the sub-10-nm CNFET op amp investigated here.**

Index Terms—**Carbon nanotube field effect transistor (CNFET), channel length, gate oxide thickness, spacer dielectric.**

I. Introduction

Carbon Nanotubes (CNT) is already acknowledged as a potential candidate to replace silicon in high performance electronics with their exceptional electrical properties. Carbon nanotube provides longer mean free-path (MFP) (~1 μm) than many other prevailing materials, which in turn provides very high mobility (10^5 cm^2 v^{-1} s^{-1}) and high conductivity (10^9 A cm^{-2}) [1], [2]. To put these properties in use, researchers have fabricated field effect transistors (CNFET) using CNT as channel material [3]–[5]. The need for proper CNFET model for research purpose both in academia and VLSI industry, led to the successful development of a compact SPICE model such as Stanford CNFET model [6], [7], virtual source CNFET model [18], [19] and many other. Design and performance analysis of analog circuits [8]–[10], digital logic circuits [11], [12] and memory circuits [13] using CNFET are also reported in literature.

In this work we have investigated application of CNFET in analog circuit, such as op amp. Optimized design of inverting amplifier and comparison using CMOS, CNFET and hybrid technologies [8], architecture optimization of analog circuits in mixed mode ICs [10], design and performance evaluation and comparison of elliptical filter with Si-OPAMPs and 32 nm CNT-OPAMPs [14], design of carbon nanotube FET based Operational Amplifier (Op Amp) [15], energy efficient op amp design using CNFET for biomedical application [16], design of SB-CNFET op amp and comparison with conventional CNFET op amp [17] are all reported in literature. But op amp designed for high frequency operation in sub-10 nm node is not reported. In an earlier work, op amp design using 32 nm CNFET and performance comparison with CMOS counterpart [9] was reported. Here, in this work, CNFET parameters such as, channel length (L_g), gate oxide (K_{ox}), oxide thickness (t_{ox}), pitch (s), source/drain extension doping profile (E_{fsd}), source/drain spacer oxide (K_{spa}) are optimized in such a manner that, the CNFET op amp provides high frequency operation in sub-10-nm node. Optimizations presented here are done with respect to amplifier gain and -3 dB cut off frequency. These optimized parameters are then used to analyze the performance of op amp.

The last section in this work presents overall performance analysis of optimized op amp and the comparison between the performances of CMOS op amp [9], 32 nm CNFET op amp [9] and sub-10-nm CNFET op amp.

For this work Virtual Source CNFET [18]–[20] model is used. Virtual Source (VS) CNFET model is a data-calibrated compact model of carbon nanotube (CNT) field effect transistors (CNFETs) based on the virtual-source (VS) approach. This model captures dimensional scaling effects and is suitable for technology benchmarking and performance projection at the sub-10-nm technology nodes [18], [19].

The CNFET in this work is considered to have a gate-all-around (GAA) architecture, shown in Fig.1 [18]. In this device, quasi ballistic transport is assumed and inversion gate capacitance, carrier mobility, threshold voltage, inverse subthreshold slope factor, drain-induced barrier lowering (DIBL) coefficient, and various other parameters are considered diameter dependent [18]. Various imperfections that arise in practical carbon nanotube based FETs, are rigorously investigated and the corresponding effects are incorporated in this model [19]. These considerations help to predict results close to experiment.

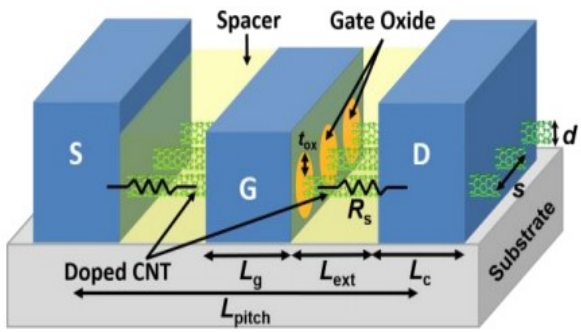


Fig. 1 GAA device structure of CNFET used in op amp. Visual representation of various parameters, critical dimensions, are also indicated in the figure. S -source, G -gate, D -drain [18].

II. CNFET Parameter Optimization for High Frequency Operation of Op Amp

Op amp design, which is used in all the simulations performed for this work is presented Fig.2. This circuit provides better performance in terms of gain, -3dB cutoff frequency, power, and noise rejection [9].

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

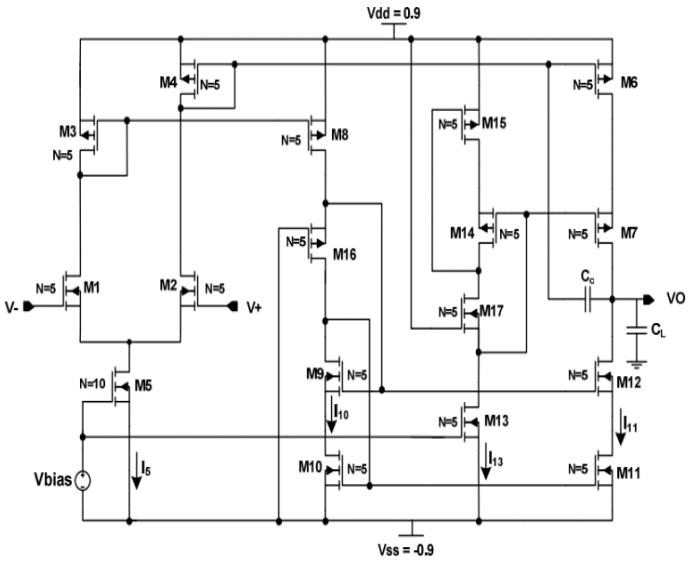


Fig. 2 Op amp design used in simulation. This design Provides high gain at -3dB frequency [9].

To optimize channel length, gain and frequency variation as a function of channel length is calculated. The obtained result is provided in Fig.3. At lower L_g, channel is ballistic so device speed improves; intrinsic capacitances decreases and drive current increase. Also with decreasing L_g, off current (I_{off}) and static power consumption increases. These effects are seen from Fig.3, with increasing channel length (L_g) gain increases and then after a certain value (here, Lg = 32nm) gain starts to decrease. So, there is an optimal value of L_g for which device speed and power consumption is optimal. Cut off frequency at 10 nm channel length is 17.5 GHz and at 7 nm channel length it is 18 GHz. But gain is higher for 10 nm channel length. So, optimum channel length is 10 nm.

To optimize gate oxide dielectric constant (K_{ox}), change of gain and frequency as a function of K_{ox} is calculated. The obtained result is provided in Fig.4. For gate all around (GAA) structure, gate capacitance is given by [18],

$$C_{OX} = \frac{2\pi K_{OX}\varepsilon_O}{\ln[(2t_{OX}+d)/d]} \quad (1)$$

Equation (1) indicates that gate capacitance (C_{ox}) increases with increasing gate dielectric constant (K_{ox}). As seen from Fig.4, gain decreases with increasing K_{ox} but for K_{ox} greater than 4, cut off frequency is relatively same, around 25 GHz. At K_{ox} 10, gain is 18.3582 dB, which is higher than the gain obtained for later values of K_{ox}. So, 10 is taken as the optimum gate dielectric constant.

Gate oxide thickness (t_{ox}) is also optimized in terms of gain and frequency. The obtained results are illustrated in Fig.5. It can be seen from (1) that gate capacitance decreases with increasing gate oxide thickness. For decreasing gate capacitance (C_{ox}) it can be inferred from (1) that, gate dielectric constant (K_{ox}) would have to decrease also. That means transconductance (g_m) will also decrease and gain will be reduced with it. This trend is very clear in Fig.5. The cut off frequency also follow the same pattern. At 3 nm gate oxide thickness, cut off frequency is the highest, 26 GHz, and gain is relatively high compared to the rest. So, 3 nm is the optimum gate oxide thickness.

Pitch (s) is the centre to centre distance between adjacent CNT in channel length. Through simulation, gain and frequency variation for various pitch is observed. These results are provided in Fig.6. As seen from Fig.6, gain is same, 16.7283 dB, for all pitch but -3dB cut-off frequency decreases with pitch. At 3 nm pitch, cut off frequency is 75 GHz, which is higher than the rest. But screening effect is very prominent at this distance. So, 5 nm, cut off frequency 26 GHz, is chosen to be the optimum value for pitch.

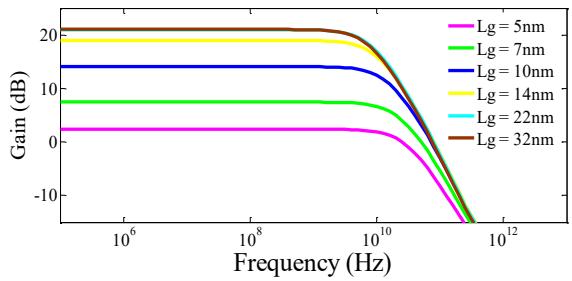


Fig. 3 Illustration of CNFET channel length optimization with respect to gain and cut off frequency.

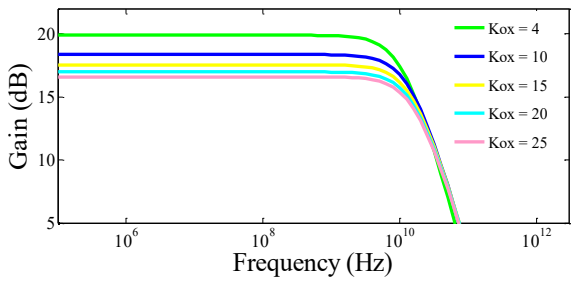


Fig. 4 Variation of gate dielectric constant to find optimum condition.

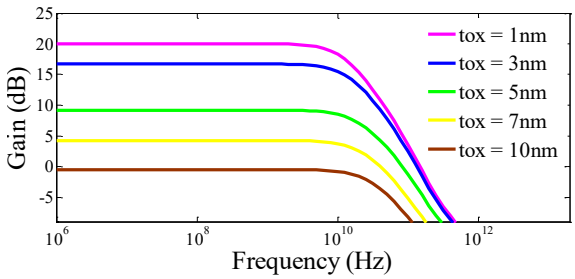


Fig. 5 Optimization of gate oxide thickness with respect to gain and frequency.

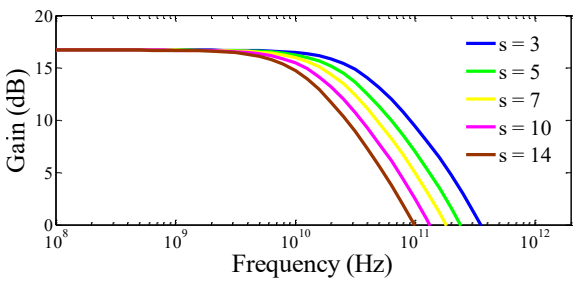


Fig. 6 Pitch variation for optimum operation in terms of gain and frequency.

To realize the effect of doping concentration of the source/drain extension or spacer doping (E_{fsd}) and obtain the optimum doping concentration, gain and frequency variation for various doping profile is investigated. Obtained results are illustrated in Fig.7. In general, better performance is obtained from MOSFET-like CNFETs with Ohmic metal-CNT contacts by heavily doping the source/drain (S/D) extensions [3], [4]. But here, as seen from Fig.7, gain decreases and cut off frequency stays almost same with increasing doping concentration (E_{fsd}) of the S/D extension. At zero (E_{fsd} = 0)

doping concentration gain is relatively higher with same operating frequency compared to the rest. So, for optimum performance S/D extension should not be doped.

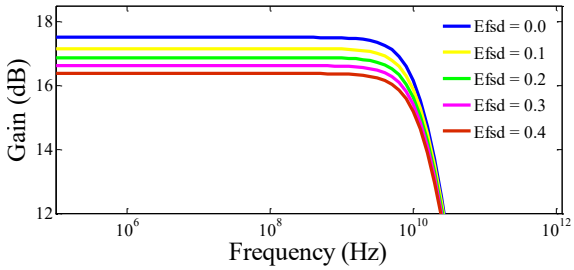


Fig. 7 Optimizing source/ drain spacer doping for high performance.

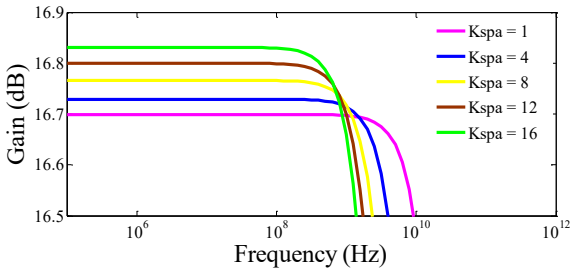


Fig. 8 Varying source/ drain spacer oxide dielectric constant for optimum performance.

To investigate the effect of K_{spa} on op amp performance, gain and frequency variation is simulated for various spacer oxide. The obtained results are provided in Fig.8. Spacer oxide with high dielectric constant (K_{spa}) results in wider energy barrier and stronger fringe field [6]. For GAA structure, the following empirical equation can be derived to understand the effect of K_{spa} [19].

$$L_{of} = (0.263K_{spa} + 0.056)(t_{ox}) \quad (2)$$

Here, L_{of} is an empirical parameter functioning like an extension of the L_g that captures the finite debye length and the gate fringing field [19]. Increasing K_{spa} decreases I_{SDT} (source/drain tunnelling current) [19]. Again increasing K_{spa} increases parasitic capacitance and degrades circuit speed [21]. It is seen from simulation results, Fig. 8, that with increasing K_{spa}, -3 dB cut-off frequency decreases but gain stays relatively same. At 4 dielectric constant, operating frequency is, 26 GHz, with relatively high gain. So, oxide having dielectric constant 4, should be used as spacer for optimum performance.

III. Performance Analysis and Comparison

Optimized parameter values, found in the previous section, are incorporated in CNFET to observe the overall performance of sub-10nm CNFET op amp. The obtained result, change of gain and phase as a function of frequency, is provided in Fig.9. 20.05 dB gain, 25 GHz operating frequency and 148.7 GHz unity gain frequency is obtained, which are very convenient compared to previous work.

For further analysis, performance of optimized op amp is analyzed regarding four performance parameters, such as, ICMR, PSRR, transient response, and flicker noise. ICMR is defined as the input common mode range. It specifies a range of common mode voltage over which the differential amplifier continues to sense and amplify the difference signal with the same gain. Fig.10 illustrates ICMR profile of op amp. Power Supply Rejection Ratio or Power Supply Ripple Rejection (PSRR) provides a measure of how well a circuit rejects ripple, of various frequencies, injected at its input. For any circuit PSRR is measured as circuit's power supply's rejection expressed in a log ratio of output noise to input noise. PSRR profile of sub-10-nm CNFET op amp is illustrated in Fig.11. ICMR and PSRR profile illustrated in Fig.10 and Fig.11, indicates large range of common mode amplification and high capability of rejecting input power supply ripple.

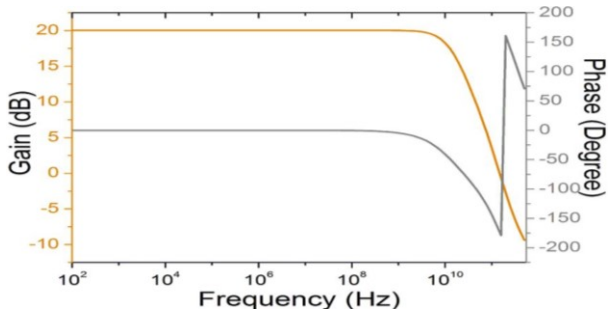


Fig. 9 Optimized sub-10-nm CNFET op amp gain and phase is illustrated as a function of frequency. Obtained gain is 20.05 dB with cut off frequency 25 GHz and unity gain frequency 148.7 GHz. Phase margin is 45°.

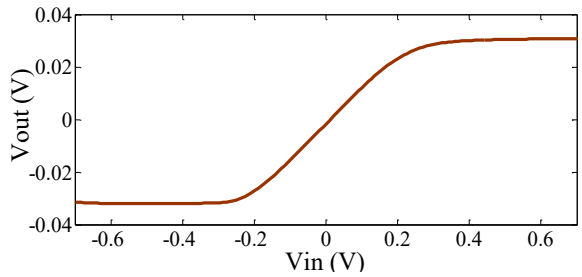


Fig. 10 Illustrating ICMR profile of Op amp.

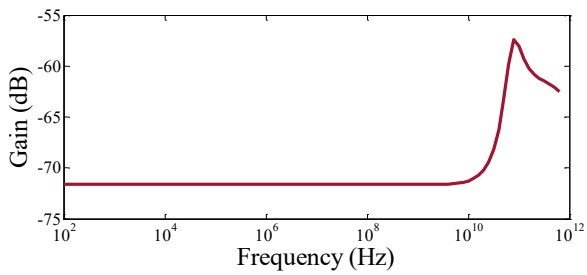


Fig. 11 Illustrating power supply rejection ratio of optimized Op amp.

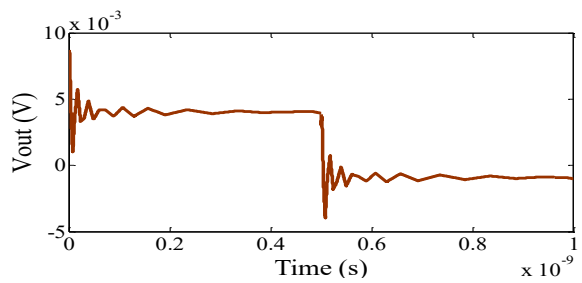


Fig. 12 Illustrating transient behaviour profile of optimized Op amp.

The transient response, i.e., the response to a step or pulse signal, of an amplifier is determined by the frequency response of the op-amp. Transient response of the op amp is shown in Fig.12. The overshoots observed in Fig.12, indicates the presence of various capacitive effects considered in

CNFET. Due to various defects, flicker noise arises in the low frequency region of operation. Flicker noise is calculated by Hooge's law [22]. Simulation result for flicker noise is provided in Fig. 13.

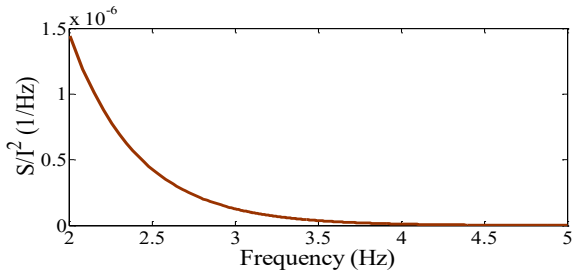


Fig. 13 Illustrating flicker noise profile of optimized Op Amp.

Here, as seen from Fig.13, flicker noise is prominent in optimized op amp. This is due to the fact that CNFET in very narrow device. In narrow devices, a single defect can cause high fluctuation in carrier mobility [23]. This in turn causes enormous current fluctuation in devices. Although, flicker noise is prominent in CNFET devices it is still lower than the CMOS counterpart, as reported in earlier work [9]. Maximum flicker noise of sub-10 nm CNFET op amp is 1.5μHz^{-1} which is lower compared to 2.6μHz^{-1} of CMOS op amp [9].

Our previous work demonstrated a comparison between CNFET op amp and CMOS op amp [9]. To clearly illustrate the difference in performance among previous work and the work presented here, various important figures of merit of op amp, with corresponding values, are summarized in Table I.

TABLE I
OVERALL PERFORMANCE COMPARISON OF SI CMOS OP AMP [9], 32 NM CNFET OP AMP [9] AND SUB-10 NM CNFET OP AMP WITH SAME GATE AREA

Parameter	Si CMOS Op Amp	32 nm CNFET Op Amp	10 nm CNFET Op Amp
Gain (dB)	9.8	32.43	20.05
Phase Margin (°)	52	45	45
-3 dB frequency (GHz)	9.24	3.08	25
Unity gain frequency (GHz)	25	134	148.7
CMRR (dB)	34.19	40.5	26.5938
PSRR (dB)	14.09	38.02	70.5
ICMR (mV)	-200 to 320	-380 to 210	-503 to 351
V_{out} swing (mV)	-623 to 425	- 350 to 320	-275 to 351
Offset voltage (mV)	-0.4	0.002	-0.0017
R_{out}(KΩ)	174	660	470
Settling time (ps)	118	17	173.94
Slew rate (V/μs)	477	1986	1687
Power (mw)	0.173	0.48	0.504

IV. CONCLUSIONS

This work is performed considering various defects, leakage currents, resistance, and parasitic capacitance which usually appear in experimental settings and are excluded from consideration for simple simulations. This is why obtained gain is lower than the previous works. Although the gain obtained in this work is lower than 32 nm CNFET op amp, it is still double of the gain obtained from the CMOS counterpart. Frequency of operation is in 15 to 26 GHz range for this work which is very convenient compared to 1 to 5 GHz for 32 nm CNFET op amp and 7 to 10 GHz for CMOS op amp reported in the earlier work.

REFERENCES

[1] T. Du1rkop, S. A. Getty, Enrique Cobas, and M. S. Fuhrer, "Extraordinary Mobility in Semiconducting Carbon Nanotubes", *Nano letters*. vol. 4, no. 1, pp.35-9, Jan. 2004.

[2] M. J. O'Connell, Carbon Nanotubes: Properties and Application; Tayor & Francis: 2006.

[3] A. Javey, J. Guo, Q. Wang, M. Lundstrom and H. Dai, "Ballistic carbon nanotube field-effect transistors," *Nature*, vol. 424, pp. 654-657, Aug. 2003.

[4] J. Chen, C. Klinke, A. Afzali and P. Avouris, "Self-aligned carbon nanotube transistors with charge transfer doping," *Appl. Phys. Lett.*, vol. 86, pp. 123108, 2005.

[5] B. C. Paul, S. Fujita, M. Okajima, and T. Lee, "Modeling and Analysis of Circuit Performance of Ballistic CNFET", In *Proceedings of the 43rd annual Design Automation Conference*, 2006, pp. 717-722.

[6] J. Deng and H.-S. P. Wong, "A compact SPICE model for carbon nanotube field effect transistors including non-idealities and its application—Part I: Model of the intrinsic channel region." *IEEE Trans. Electron Devices*, vol. 54, no. 12, pp. 3186–3194, Dec. 2007.

[7] J. Deng and H.-S. P. Wong, "A compact SPICE model for carbon nanotubefield effect transistors including non-idealities and its application—Part II: Full device model and circuit performance benchmarking," *IEEE Trans. Electron Devices*, vol. 54, no. 12, pp. 3195–3205, Dec. 2007.

[8] F. A. Usmani, and M. Hasan, "Carbon nanotube field effect transistors for high performance analog applications: An optimum design approach", *Microelectronics Journal,* vol. 41, no. 7, pp. 395-402, Jul. 2010.

[9] Md. N. Ferdous, H. Ahmad, A. K. Reza, S. M. Mominuzzaman, and A. B. M. Harun-Ur-Rashid, "Design optimization of high frequency op amp using 32 nm CNFET." In *International Conference on Electrical and Computer Engineering (ICECE), 2010*, pp. 230-234.

[10] J. S. Ajit, Y. B. Kim, and M. Choi, "Performance assessment of analog circuits with carbon nanotube FET (CNFET)." In *Proceedings of the 20th symposium on Great lakes symposium on VLSI*, 2010, pp. 163-166.

[11] S. Lin, Y. B. Kim, and F. Lombardi, "CNTFET-based design of ternary logic gates and arithmetic circuits." *IEEE Trans. on Nanotechnology*, vol. 10, no. 2, pp. 217-225, Mar. 2011.

[12] G. S. Tulevski, A. D. Franklin, D. Frank, J. M. Lobez, Q. Cao, H. Park, A. Afzali, S. J. Han, J. B. Hannon, and W. Haensch "Toward High-Performance Digital Logic Technology with Carbon Nanotubes", *ACS Nano*, vol. 8, no. 9, pp. 8730–8745, Aug. 2014.

[13] N. S. Bhat, "Design and modelling of different SRAM's based on CNTFET 32nm technology", *arXiv preprint arXiv: 1203.1971*, Mar. 2012.

[14] F. Rahman, I. Ilyas, N. Fatema, R. T. Shama, "Performance evaluation of a 32-nm CNT-OPAMP in analog circuits: Design and comparison of Elliptical filters", In *11th International Conference on Solid-State and Integrated Circuit Technology (ICSICT),* 2012, pp. 1-3.

[15] W. Makni, M. Najari, H. Samet, M. Masmoudi, "Operational Amplifier Circuit Design Using Carbon Nanotube Transistors", *Nanoscience & Nanotechnology-Asia*, vol. 3, no. 1, pp. 106-113(8), July 2013.

[16] A. Puri, A. Rana, "Performance analysis of CNTFET based low power operational amplifier in analog circuits for biomedical applications", In *International Conference on Electronics, Computing and Communication Technologies (CONECCT)*, 2015 pp. 1-5.

[17] W. Makni, M. Najari, H. Samet, M. Najari, "Schottky barrier carbon nanotube transistors Op-Amp circuit", In *12th International Multi-Conference on Systems, Signals & Devices (SSD)*, 2015, pp. 1-5.

[18] C.-S. Lee, E. Pop, A. Franklin, W. Haensch, and H.-S. P. Wong, "A Compact Virtual-Source Model for Carbon Nanotube Field-Effect Transistors in the Sub-10-nm Regime—Part I: Intrinsic Elements," *IEEE Trans. Electron Devices*, 2015.

[19] C.-S. Lee, E. Pop, A. Franklin, W. Haensch, and H.-S. P. Wong, "A Compact Virtual-Source Model for Carbon Nanotube Field-Effect Transistors in the Sub-10-nm Regime—Part II: Extrinsic Elements and Performance Assessment," *IEEE Trans. Electron Devices*, 2015.

[20] Stanford University Virtual Source CNFET Model, Version 1.0.1. Available: https://nano.stanford.edu/stanford-cnfet2-model

[21] J. Deng, K. Kim, C.-T. Chuang and H.-S. P. Wong, "The Impact of Device Footprint Scaling on High-Performance CMOS Logic Technology," *IEEE Trans. Electron Devices*, vol. 54, no. 5, pp. 1148-1155, May 2007.

[22] F. N. Hooge, "1/*f* noise is no surface effect," *Physics Letters A*, vol. 29 (3), pp. 139-140, Apr. 1969.

[23] F. Liu, K. Wang, D. Zhang, and C. Zhou, "Noise in carbon nanotube field effect transistors," Applied Physics Letters, vol. 89, 2006.

Ultra-High Nonlinear Photonic Crystal Fiber with GeO_2 Doped Core

Shekh Farhad Uddin Ahmed*, Sk. Nayemuzzaman, and Mohammad Faisal
Department of Electrical and Electronics Engineering
Bangladesh University of Engineering and Technology, Dhaka-1205, Bangladesh
*sk.farhad.eee@gmail.com

***Abstract*—This paper articulates our recent investigation on the properties of highly nonlinear photonic crystal fiber (HN-PCF) with concentrated GeO_2 doped core using Finite Element method (FEM). The proposed PCF shows extremely high nonlinearity with very small effective mode area and convenient dispersion characteristics in telecommunication window. Suitable nearly zero dispersion profile at operating wavelength for nonlinear applicationscan be achieved by tuning the global design parameters. Obtained values of nonlinear coefficient γ with this design are ~ 347.58 $W^{-1}km^{-1}$and ~ 232.53 $W^{-1}km^{-1}$ at 1.3 µm and 1.55 µm respectively with very low confinement loss suggesting this PCF as an attractive candidate for supercontinuum generation, ultrafast signal processing and fiber sensors etc.**

***Index Terms*—FEM, Photonic crystal fibers (PCFs), Fiber nonlinearity, Effective mode area.**

I. Introduction

PHOTONIC crystal fibers (PCFs) [1] also known as micro structured optical fibers have drawn huge research interest and attraction since there are lot of possibilities to shape their design in order to achieve desired optical properties. HN-PCF designs comprised of GeO_2 doped core region surrounded by multiple layers of air holes running through the fiber length have been one of the most promising recent improvements in the paradigm of nonlinear fiber optics [3-6].

To achieve high nonlinearity with suitable dispersion properties, different designs of HN-PCFs have been proposed over the recent years, such as, dispersion flattened PCF with 10 layers of air holes and defected core [2] has been demonstrated with nonlinear coefficient of only 34 $W^{-1}km^{-1}$ at 1.55 µm wavelength, which also has high confinement loss and comparatively large mode area as well. Uniform hexagonal distribution of air-holes throughout the cladding with core defected by germanium has been reported [3], which has offered a low dispersion and nonlinear coefficient of 67.23 $W^{-1}km^{-1}$at 1.55 µm. Hexagonal PCF with GeO_2 doped core and 5 layers of irregular air holes in the cladding also has been suggested which has offered a high nonlinearity of around 82 $W^{-1}km^{-1}$ at 1.55 µm and around 146 $W^{-1}km^{-1}$at 1.3 µm [6]. Furthermore, Ge-doped core with cladding having triangular lattice of air holes has been proposed [7]. It has provided flat dispersion over S+C+L wavelength bands and nonlinear coefficient of 46.6 $W^{-1}km^{-1}$ was achieved at 1.55 µm. Recently Tiang *et. al*. [8] has found nonlinear coefficient of 68 $W^{-1}km^{-1}$ for y-polarization at 1.55 µm using all circle air-holes in the PCF.

There is some more recent advancement in the field of nonlinear fiber optics, such as, very high nonlinearity has been reported using bismuth glass holey fibers [9], soft glass equiangular spiral PCF [10], chalcogenide PCFs [11] and PCF with nano-scale slot core [12]. But there are some design issues in the cases of bismuth and soft glass fibers to be considered. Confinement loss is very high in bismuth fibers. Confinement loss property is not analysed in [10]. In chalcogenide glass fibers the negative dispersion is very high which might be an issue while utilizing the nonlinear properties of the fiber. Fabrication of rectangular nano-scale slot in core is much difficult.

In this paper, we propose a promising design with highly concentrated GeO_2 doped core [4-5] surrounded by large, compact air holes near the core and smaller, variable size air holes in the outer cladding. The structure is simple with all circle air-holes. With this design, we have made an acceptable trade-off between the flatness [3], [6-7] of the dispersion curve and the effective mode area for attaining ultra-high nonlinearity. We obtained extremely large values of nonlinear coefficient in the operating wavelengths of 1.55 µm (~ 232.53 $W^{-1}km^{-1}$) and 1.3 µm (~ 347.58 $W^{-1}km^{-1}$) by tuning the global design parameters. We have used finite element method (FEM) with the help of COMSOL Multiphysics 5.0 to study the optical properties of this PCF within the range of 1.2 µm to 1.65 µm and obtained encouraging results.

II. Fiber Construction

Fig. 1 shows the cross-sections of proposed PCF with concentrated GeO_2 doped circular core in the center. The cladding structure has irregular air holes and the host material is silica. Seven layers of circular air holes in the cladding are arranged in a triangular array.

The air holes with diameters d_1, d_2, d_3, d_4, d_5, d_6, and d_7 positioned at different layers along with the pitch (Λ) are adjusted to meet certain design requirements. The diameters of the air holes of first 2 innermost layers and the circular core in the middle doped by GeO_2 with 97% [12] molar concentration have been kept the same ($d_0 = d_1 = d_2 = 0.89\Lambda$ at 1.3 and $d_0 = d_1 = d_2 = 0.97\Lambda$ at 1.55) and as large as possible. This novel design inspired by previous works [4-5], has resulted in very small effective mode area as well as extremely high nonlinearity. To reduce the backscattering and absorb the

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

leakage at the boundary, a circular perfectly matched layer (PML) is provided around the structure.

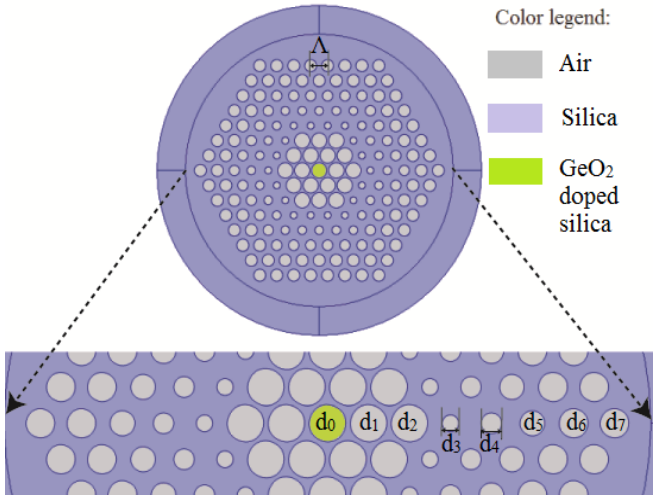


Fig. 1. Geometric cross section of the proposed PCF having concentrated GeO_2 doped core with 7 layers of air holes surrounded by a PML.

III. Theoretical Background

Simulations have been carried out using FEM to obtain the optical properties of the proposed structure. The results obtained from COMSOL Multiphysics 5.0 are exported to MATLAB for the calculation of dispersion, confinement loss, effective mode area and nonlinear coefficient. Dispersion was calculated by the equation as follows [13]

$$D(\lambda) = -\frac{\lambda}{c}\frac{d^2 Re[n_{eff}]}{d\lambda^2} \quad (1)$$

where $D(\lambda)$ is the chromatic dispersion, $Re[n_{eff}]$ is the real part of complex effective mode index n_{eff} , and c is the velocity of light in vacuum, and λ is operating wavelength. Wavelength dependent refractive index of SiO_2 can be found using Sellmeier's equation [14]. Refractive index of GeO_2 doped core can be found from [14]

$$n^2 = 1 + \sum_{i=1}^{3} \frac{[SA_i + X(GA_i - SA_i]\lambda^2}{\lambda^2 - [Sl_i + X(Gl_i - Sl_i)]^2}, \quad (2)$$

where SA_i, Sl_i, GA_i, Gl_i are the Sellmeier coefficients for the SiO_2 and GeO_2 glasses, respectively, and X is the mole fraction of GeO_2. Confinement loss is the leakage loss due to finite number of air holes in the fiber cross section. It is calculated as [13]

$$L_c = \frac{40\pi}{ln(10)\lambda} Im[n_{eff}] \quad (3)$$

where $Im[n_{\text{eff}}]$ is the imaginary part of n_{eff} . We also calculate the effective mode area by [15]

$$A_{eff} = \frac{\left\{\iint_{-\infty}^{\infty} |E(x,y)|^2 dxdy\right\}^2}{\iint_{-\infty}^{\infty} |E(x,y)|^4 dxdy} \quad (4)$$

where $E(x, y)$ is the electric field in the medium. Using the effective area in (3), we calculate the nonlinear coefficient as [16]

$$\gamma = \frac{2\pi n_2}{\lambda A_{eff}} \quad (5)$$

where the nonlinear refractive index, n_2 for GeO_2 doped SiO_2 can be calculated by [16-17],

$$n_2(GeO_2 - SiO_2) = n_2(SiO_2) + KX(GeO_2) \quad (6)$$

where, $X(GeO_2)$ is the GeO_2 concentration in mol.%, $n_2(SiO_2)$ = 2.16×10^{-20} m^2/W, and K = 0.033×10^{-20} $m^2/(\%W)$ and $n_2(GeO_2\text{-}SiO_2)$ is the effective nonlinear refractive index of GeO_2 doped structure in m^2/W.

IV. Results and Discussions

We have investigated the properties of the proposed structure with optimized parameters for nonlinear applications targeted at 1.3 µm and 1.55 µm. First, the electric field confinement plots of the structure are shown below in Fig. 2 indicating very small effective area, i.e., tight confinement of field in the core for both cases.

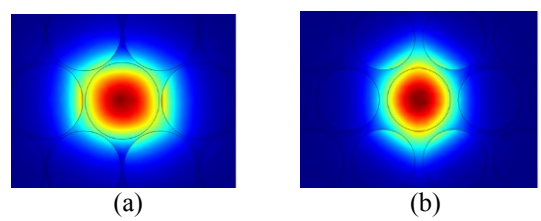


Fig. 2. Field confinement of the proposed PCF having concentrated GeO_2 doped core at the center for optimized parameters at 1.55 µm with effective mode area 0.934 μm^2 (a) and at 1.3 µm with 0.745 μm^2 (b).

We obtain a low dispersion profile around the operating wavelength of 1.3 µm with optimized pitch and other global design parameters as shown in Fig. 3. We find very small change of dispersion profile with ±1% pitch variation around 1.2 µm to 1.36 µm range.

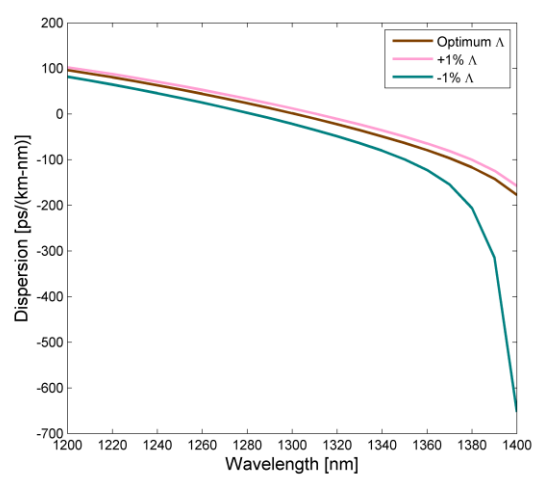


Fig. 3. Variation of dispersion as a function of wavelength resulted from optimized design parameters at 1300 nm with $d_0 = d_1 = d_2 = 0.89\Lambda$, $d_3 = 0.4\Lambda$, $d_4 = 0.5\Lambda$, $d_5 = 0.6\Lambda$, $d_6 = d_7 = 0.7\Lambda$ where $\Lambda = 0.9$ µm with ±1% change.

Next we present the confinement loss of the proposed PCF as a function of operating wavelength with variation in pitch as shown in Fig. 4. Observations indicate very low loss (~ 10^{-8} dB/km) at the operating wavelength of 1.3 µm for optimum pitch (Λ = 0.9 µm) which is convenient for our desired applications. The pitch variation of −1% yields more loss, but it is still very low (~ 10^{-7} dB/km) at 1.3 µm.

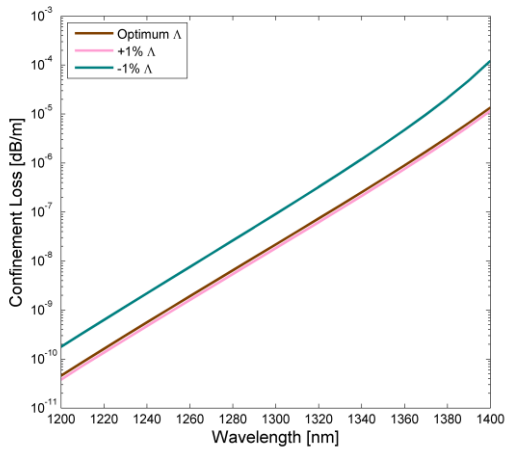


Fig. 4. Variation of confinement loss as a function of wavelength with optimized parameters for nonlinear applications at 1300 nm.

Now we demonstrate the nonlinear properties of the proposed structure in Fig. 5. Here the values of nonlinear coefficient and effective mode area are shown together along with ±1% variation in pitch. The effective mode area is found 0.73 μm^2 and nonlinearity is 353.99 $W^{-1}km^{-1}$ at 1.3 µm for optimum pitch.

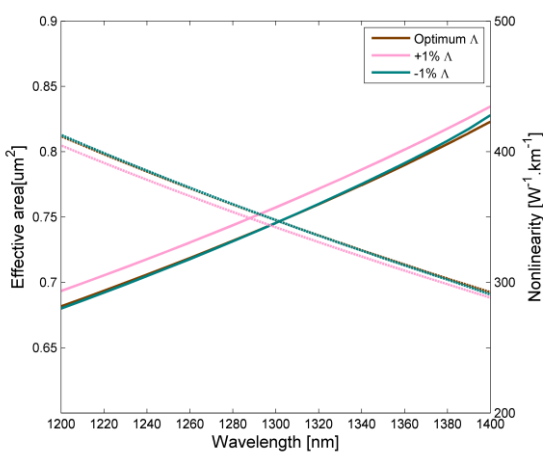


Fig. 5. Variation of effective area along with nonlinear coefficient γ as a function of wavelength with optimized parameters for nonlinear applications at 1300 nm.

Next, the same investigations were performed for operating wavelength of 1.55 µm as demonstrated in Fig. 6, Fig. 7 and Fig. 8. In each case the results of ±1% variation in pitch (Λ) together with optimum values are presented. These observations indicate the tolerable variation in nonlinear properties of the proposed PCF under the probable structural defect caused by manufacturing process. Fig. 6 shows the dispersion profile of the proposed PCF with optimized design parameters at the operating wavelength of 1.55 µm. The value of dispersion is low around the operating wavelength. Fig. 7 shows the confinement loss of the PCF. It is observed that the loss is much low (~ 10^{-4} dB/km) at the operating wavelength. The slight dip in the loss curve is due to the dependence of the loss on 3rd and 4th layer of the air holes in the cladding whose diameters are relatively small than other layers.

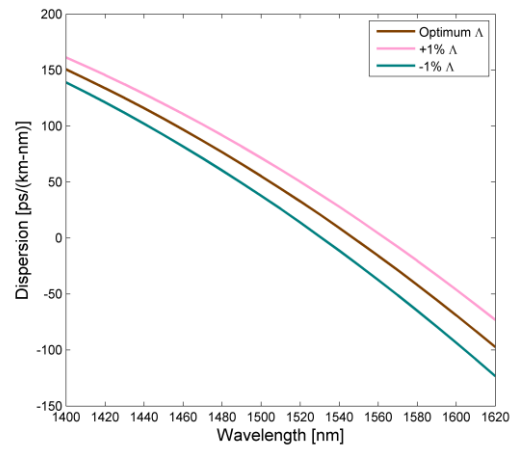


Fig. 6. Plots of dispersion versus wavelength resulted from optimized design parameters at 1550 nm with $d_0 = d_1 = d_2 = 0.97\Lambda$, $d_3 = 0.9\Lambda$, $d_4 = d_5 = d_6 = d_7 = 0.7\Lambda$ where Λ = 1.01 µm with ±1% change.

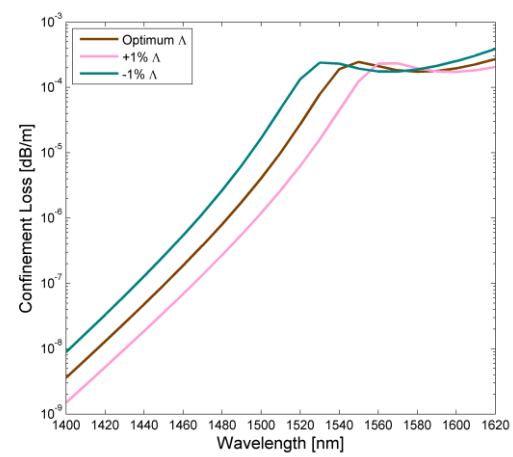


Fig. 7. Variation of confinement loss as a function of wavelength with optimized parameters for nonlinear applications at 1300 nm.

TABLE I
PCF PROPERTIES WITH PITCH VARIATIONS

% change in pitch (Λ)	λ [µm]	D(λ) [ps/(km-nm)]	γ [$W^{-1}km^{-1}$]	Eff. area [μm^2]	Conf. Loss [dB/km]
-1%	1.3	-21.7346	347.6295	0.745	9.111×10^{-7}
Optimized	1.3	1.5666	347.5775	0.745	2.17×10^{-8}
+1%	1.3	12.2568	342.1347	0.757	1.821×10^{-8}
-1%	1.55	-24.2281	234.1424	0.928	1.934×10^{-4}
Optimized	1.55	-3.3612	232.5331	0.934	2.453×10^{-4}
+1%	1.55	16.2816	230.8768	0.941	1.235×10^{-4}

Table I summarizes the observations with the values of dispersion, confinement loss, effective mode area and nonlinear coefficient along with pitch variations at two mostly desired operating wavelengths. Fig. 8 demonstrates the nonlinear properties of our proposed PCF for desired nonlinear applications targeted at 1.55 µm.

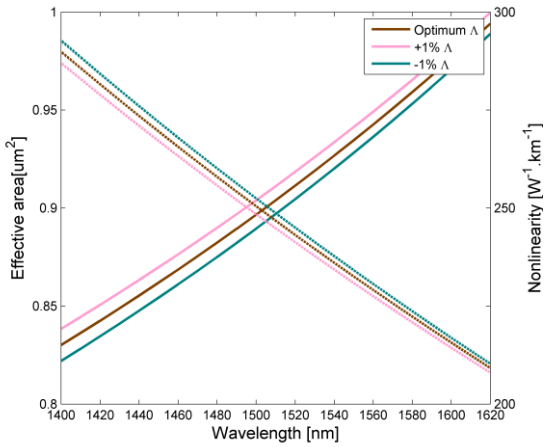


Fig. 8. Variation of effective area along with nonlinear coefficientγas a function of wavelength with optimized parameters for nonlinear applications at 1550 nm.

To determine the nonlinearity of the PCF, we calculated the effective mode area A_{eff} and then computed the nonlinear coefficient γ. From Fig. 8, it is found that the effective mode area is ~ 0.93 µm^2 and the nonlinearity is ~ 232.53 W^{-1}km^{-1}. The effective mode field distribution in Fig. 2(a) and 2(b) show the high density of field in the core which gives the lower effective mode area, consequently higher nonlinear effect. Moreover, the nonlinear refractive index n_2 for concentrated GeO_2 doped core is higher than pure silica and it also enhances the nonlinear effect.

V. Conclusions

The suggested novel PCF design containing highly GeO_2 doped core with nearby compact air holes having large diameters compared to the outer cladding presented in this paper has resulted in ultra-high nonlinearity with very small effective area in the operating wavelengths of 1.3 µm (~347.58 W^{-1}km^{-1}) and 1.55 µm. (~232.53 W^{-1}km^{-1}). Although much higher nonlinearity has been achieved in bismuth or chalcogenide glass fibers, or using silicon rectangular nano slot in core, however, these results are much promising and praiseworthy as the design is much simpler with all circular air-holes and only with doped GeO_2 core. These attractive features can be utilized in nonlinear applications such as supercontinuum generation, ultrafast signal processing in telecommunication window and sensors etc.

References

[1] J. C. Knight, T. A. Birks, P. St. J. Russell, and D. M. Atkin, "All-silica-single-mode optical fiber with photonic crystal cladding,"*Opt. Lett*. 21, 1547-1549 (1996).

[2] K. Saitoh and M. Koshiba, "Highly nonlinear dispersion-flattened photonic crystal fibers for supercontinuum generation in a telecommunication window", *Opt. Express*, vol. 12, no. 10, pp. 2027-2032, May 2004.

[3] M. S. Hossain, K. Neupane, M. S. Bin Hafiz, and S. P. Majumder, "Dispersion and nonlinear characteristics of a photonic crystal fiber (PCF) with defected core and various doping concentration," proc. in *ICECE 2014*, pp. 500-503, Dhaka, Bangladesh, 20-22 Dec., 2014.

[4] A. Nair, S. K. Sudheer, and M. Jayaraju, "Comparative Study of Octagonal Ring PhotonicCrystal Fiber with GeO2 Doped Structure", proc. in *ICICES 2014* Chennai, Tamil Nadu, India, 27-28 Feb. 2014.

[5] A. Labruyère, P. Leproux, V. Couderc, V. Tombelaine, J. Kobelke, K. Schuster,H. Bartelt, S. Hilaire, G. Huss, and G.Mélin, "Structured-Core GeO_2-Doped Photonic-Crystal Fibers for Parametric and Supercontinuum Generation", *IEEE Photon. Technol. Lett*. Vol. 22, no. 16, pp. 1259-1261, August 15, 2010.

[6] M. A. Hossain, Y. Namihira, M. A. Islam, S. M. A. Razzak, Y. Hirako, K. Miyagi, S. F. Kaijage, and H. Higa, "Tailoring supercontinuumgeneration using highly nonlinear photonic crystal fiber," *J. Opt. & Laser Technol.*, vol. 44, pp. 1889-1896, 2012.

[7] F. Hong, L. Shuqin, G. Tieying, Y. Lei, L. Honglei, and J. Shisheng, "Design of highly nonlinear photonic crystal fibers with flattened dispersion over S+C+L wavelength bands," Proc. of *SPIE*, vol. 6781, 67810P, 2007.

[8] T. Yang, E. Wang, H. Jinag, Z. Hu, and K. Xie, "High birefringence photonic crystal fiber with high nonlinearity and low confinement loss," *Optics Exp*., vol. 23, no. 7, p. 8329, 2015.

[9] H. Ebendorff-Heidepriem, P. Petropoulos, S. Asimakis, V. Finazzi, R. C. Moore, K. Frampton, F. Koizumi, D. J. Richardson, and T. M. Monro, "Bismuth glass holey fibers with high nonlinearity," Opt. Express, vol. 12, no. 21, pp. 5082–5087, 2004.

[10] A. Agrawal, N. Kejalakshmy, B. M. Azizur Rahman and Kenneth T. V. Grattan, "Soft Glass Equiangular Spiral Photonic Crystal Fiber for Supercontinuum Generation", IEEE Photon. Technol. Lett. Vol. 21, no. 22, pp. 1722 - 1724, November, 2009.

[11] D. Méchin, L. Brilland, J. Troles, Q. Coulombier, P. Houizot, A. Monteville, T. N. Nguyen., D. M. Nguyen, S. D. Le, M. Thual, T. Chartier, J. Adam, "Recent advances in very highly nonlinear chalcogenide photonic crystal fibers and their applications," in *Proc. Photonics Society Summer Topical Meeting Series, IEEE,* 2010.

[12] J. Liao and T. Huang, "Highly nonlinear photonic crystal fiber with ultrahigh birefringence using nano-scale slot core," *Opt. Fiber Technol*., vol. 22, pp. 107-122, 2015.

[13] K. Saitoh and M. Koshiba, "Ultra-flattened chromatic dispersion controllability using a defected core photonic crystal fiber with low confinement loss." *Optics Exp*, p. 8365, vol. 13, no. 21, 2005.

[14] J. W. Fleming, "Dispersion in GeO_2-SiO_2 glasses," *Appl. Opt*., vol. 23, no. 24, pp. 4486–4493, Dec. 1984.

[15] M. Pourmahyabadi* and S. M. Nejad, "Numerical Analysisof Index-Guiding Photonic Crystal Fibers with Low Confinement Loss and Ultra-Flattened Dispersion by FDFD Method", *Iranian Journal of Electrical & Electronic Engineering*, Vol. 5, No. 3, Sep. 2009.

[16] A. Boskovic, L. Gruner-Nielsen, O. A. Levring, S. V. Chernikov, and J. R. Taylor, "Direct continuous-wave measurement of n_2 in various types of telecommunication fiber at 1.55 µm," *Opt. Lett*., vol. 21, no. 24, pp. 1966–1968, Dec. 1996.

[17] Y. Yatsenko and A. Mavritsky, "D-scan measurement of nonlinear refractive index in fibers heavily doped with GeO2,"*Opt. Lett.*, vol. 32, no. 22, pp. 3257–3259, Nov. 2007.

A Lightweight PAPR Reduction Scheme using Greatest Common Divisor Matrix Based SLM Technique

Md. Moshiur Rahman,[1,*] Md. Nur Al Safa Bhuiyan,[2†]Muhammad Sajjadur Rahim,[3] and Sabbir Ahmed[4]
Dept. of Information and Communication Engineering, University of Rajshahi
Rajshahi-6205, Bangladesh
*mintu_apee@yahoo.com

***Abstract*—High Peak to Average Power Ratio (PAPR) is one of the major drawbacks for Orthogonal Frequency Division Multiplexing (OFDM) transmitted signal. Among the PAPR reduction techniques, Selected Mapping (SLM) technique is one of the promising one but results in added computational complexity. In this paper, we applied Greatest Common Divisor (GCD) sequence based SLM method using our proposed reverse searching technique. The proposed scheme is named lightweight as it is very simple to implement and offers improved performance through significant reduction in PAPR and computational complexity. Additionally, we explored side information free transmission with GCD default index that achieves higher throughput but sacrifices slight PAPR reduction.**

***Index Terms*—OFDM, PAPR, SLM, GCD Matrix, Computational Overhead and Side Information.**

I. Introduction

Recently, new multimedia wireless applications have resulted in an increased demand for technologies that support very high speed transmission rates, mobility and efficiently utilize the available spectrum and network resources. OFDM is one of the best solutions to achieve this goal and it offers a promising choice for future high speed data rate systems [1]. Additionally, it can combat multi-path fading efficiently and provides improved spectrum efficiency compared to single carrier systems. Because of all these advantages, OFDM has already been adopted as a standard technique for various wireless communication systems such as IEEE 802.11a standard for Wireless Local Area Networks (WLAN), IEEE 802.16 standard for Wireless Metropolitan Area Network (WMAN), Terrestrial Digital Video Broadcasting (DVBT) [2] and so forth.

However, the use of large number of sub-carriers in OFDM system points to the problem of high peak-to-average power ratio (PAPR). Thus, it is very important to reduce PAPR of the signal before transmission. Over the past decade various PAPR reduction techniques have been proposed [3]. These techniques can be broadly categorized into two: signal scrambling techniques and signal distortion techniques. Block coding, Selected Mapping (SLM) and Partial Transmit Sequences (PTS) fall into signal scrambling techniques [4], [5], whereas amplitude clipping, filtering and companding etc. represent signal distortion techniques [3].

Among them, SLM scheme is relatively attractive due to the achievement of better PAPR reduction without causing signal distortion. In this scheme, the choice of proper phase sequences to achieve good PAPR reduction is very important; hence researchers have proposed several ways to generate such sequences. For example, random values from the set (±1,$\pm j$) was chosen in the very first work on SLM [4]. Later a significant number of phase sequences were proposed. These includes Riemann, Hadamard and Hilbert matrix based phase sequence and so on [6], [7], [8].

In our previous work GCD matrix based SLM technique were introduced which showed good PAPR reduction with low computational overhead compared to the existing research work [9]. However, there is still room to reduce the computational complexity even further. In our previous work it is observed that to find out the proper phase sequence, searching is performed in usual manner which is in forward direction and the preferred phase sequences were almost always appeared in the last quadrant, i.e., from indices 41 to 62. Hence, in this work, we adopted an idea that if searching is started from the ending phase sequence of 64 and finished at starting phase sequence of 1, there is a higher probability of finding out proper phase sequences with less computational complexity. It is named as reverse searching technique in our current research. Moreover, to reduce the computational complexity, only specific band of phase sequences are considered for transmission as side information. Our simulation outcome revealed significant reduction of computational complexity for the proposed reverse searching technique and for this reason we named it lightweight.

Recently, side information free SLM method also has attracted significant research interest [10]. With the objective of using GCD sequence in conjunction with side information free transmission, we explored for the existence of "Default index" of GCD sequence. By default index, we indicated the GCD sequence that satisfied minimum PAPR for maximum number of OFDM data blocks. If such sequence is known then the receiver can be informed about it before the transmission and there is no need to send the side information with every data block. This will save valuable data throughput. In our research, 61st GCD phase sequence is found to be the default index. Additionally, PAPR of 2.5 dB is observed with this default index which is only 0.5 higher than the minimum PAPR.

II. PAPR IN OFDM

Let $X(0), X(1), ..., X(N-1)$ represent the data sequence to the transmitted in an OFDM symbol with N sub-carriers.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

The baseband representation of the OFDM symbol is given by:

$$x(t) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} X(n) e^{j2\pi nt/N} \qquad 0 \le t \le T-1, \quad (1)$$

where T is the duration of the OFDM symbol. According to the central limit theorem, when N is large, both the real and imaginary parts of $x(t)$ become Gaussian distributed, each with zero mean and a variance of $E[|x(t)|^2]/2$, and the amplitude of the OFDM symbol follows a Rayleigh distribution. Consequently it is possible that the maximum amplitude of OFDM signal may will exceed its average amplitude. Practical hardware (e.g. A/D and D/A converters, power amplifiers) has finite dynamic range; therefore the peak amplitude of OFDM signal must be limited. PAPR is mathematically defined as:

$$\text{PAPR} = 10log_{10} \frac{\max_{0 \le n \le N-1} [|x(t)|^2]}{E[|x(t)|^2]} (dB). \quad (2)$$

It is easy to see from equation (2) that PAPR reduction may be achieved by decreasing the numerator $max[|x(t)|^2]$, increasing the denominator $E[|x(t)|^2]$, or both.

The effectiveness of a PAPR reduction technique is measured by the Complementary Cumulative Distribution Function (CCDF), which is the probability that PAPR exceeds some threshold, i.e.:

$$\text{CCDF} = Probability(PAPR > \text{p}_0) \quad (3)$$

where p_0 is the threshold.

III. Phase Sequence Generation

Selection of proper phase sequences to achieve good PAPR reduction is very important in SLM. The PAPR reduction technique proposed in this paper is based on the technique introduced in [11]. Let $S = \{x_1, x_2, ..., x_n\}$ be a set of distinct positive integers. The $n \times n$ matrix $[S] = (s_{ij})$, where $(s_{ij}) = (x_i, x_j)$, the greatest common divisor of x_i and x_j is called the GCD matrix on S. Suppose $D = [d_1, d_2, ...d_m]$ is an factor-closed (FC) set containing S. Let the matrix $A = (a_{ij})$ be defined as follows:

$$a_{ij} = e_{ij}(\lambda_j)^{1/2} \quad (4)$$

Where

$$\text{e}_{ij} = \begin{cases} 1 & \text{if } d_j \text{ divides } x_i \\ 0 & \text{otherwise} \end{cases}$$

and $\lambda_j = \phi(d_j)$. Hence A is the $n \times m$ and A^T is $m \times n$. Furthermore,

$$\begin{aligned} (AA^T)_{ij} &= \sum_{k=1}^{m} \text{a}_{ik} \;\; a_{ij}. \\ &= \sum_{dk/xi, dk/xj} \sqrt{\phi(d_k)} \sqrt{\phi(d_k)}. \\ &= \sum_{dk/(xi,xj)} \phi(d_k) = (x_i, x_j) \\ &= B_{ij} \end{aligned}$$

Thus, $[S] = AA^T$

If the GCD matrix (B) is of size $N \times N$, the entries in the normalized GCD matrix (B) will be $B = gcd/N$. Using above condition, GCD Matrix (B) of order 4 can be writen as:

$$B = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 2 \\ 1 & 1 & 3 & 1 \\ 1 & 2 & 1 & 4 \end{bmatrix} \quad (5)$$

IV. Proposed Method

The following process are repeated for the proposed method with columns of a matrix are selected as phase sequence vectors for the SLM techniques.

- The input data streams are mapped as a vector symbols $X_0, X_1, X_2.....X_N - 1$ into N sub-blocks.
- Each block is multiplied point wise by B GCD phase sequence vector, each of length u. Mathematically, $B^{(u)} = [b_{u,0}, b_{u,1}, ...b_{u,N-1}]$ and the modified data block for the uth phase sequence can be written as $X^{(u)} = [X_0 b_{u,0}, X_1 b_{u,1}, ...X_{N-1} b_{u,N-1}], u = 64, 63, ..1$. This generates a modified data block and its PAPR is calculated.
- A particular data block is selected for transmission, and the phase index value is used as side information, if $PAPR <= Threshold$.
- The process is repeated for the same data block with the different phase vector at the decremented index, if $PAPR > Threshold$,

V. Determination of Default Index

One of our additional interests was to investigate the possibility of side-information free PAPR reduction mechanism using the GCD Sequence. For this, we tried to single out one particular GCD sequence that would yield maximum number of OFDM symbols achieved at one phase sequence. We name the corresponding GCD sequence index as the Default index. The underlying motivation is, if such sequence could be found, we can inform the receiver about the default index sequence only once at the start of the transmission. Thus, it is not required to send the sequence number used as side information for every OFDM symbol and there by improve the system efficiency. However, on the downside, we may have to sacrifice PAPR reduction capability. Following steps are taken in this method.

- The input data streams are mapped as a vector symbols $X_0, X_1, X_2.....X_N - 1$ into N sub-blocks.
- Each block is multiplied point wise by B GCD default phase sequence vectors of length u. and the modified data block is selected for side information free transmission after PAPR calculation.

VI. Simulation Results

At first, we performed simulations to compare the performance of SLM technique using GCD phase sequence with original OFDM system. The simulation parameters are listed in Table I. The result is shown in Fig. 1 that depicted number of OFDM symbols against the corresponding phase sequences with minimum PAPR when the GCD phase sequence is used. From this figure, it is seen that the least PAPR generating

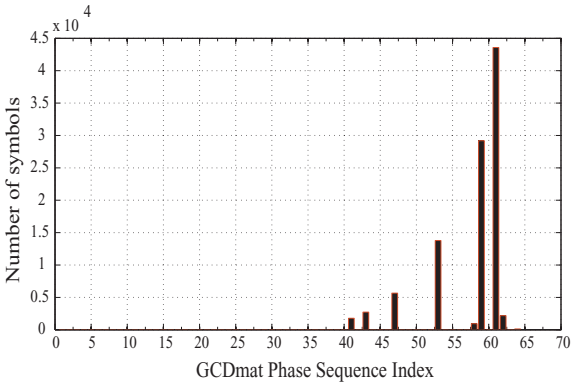


Fig. 1. Complexity Diagram for SLM using GCD Phase sequences.

symbols appeared from phase sequence 41 to 62 which can be called a band of phase sequence. Hence, searching can be focused only within this band requiring 22 iterations. As there are total of 64 iterations, computation complexity is reduced up to 66%. The comparative study of Complementary Cumulative Distribution Function (CCDF) of PAPR for GCD phase sequence based SLM technique with that of original OFDM is shown in Fig. 2. It can be seen from this figure that, for a probability of 10^{-3}, the corresponding PAPR of GCD and original OFDM are 2 dB and 10.1 dB, respectively. Thus, in GCD with SLM technique, PAPR reduction of 8.1 dB is achieved compared to the PAPR of original OFDM signal. Coming back to Fig. 1, it can also be seen that at phase sequences 59 and 61, most of the symbols with least PAPR appeared. Moreover, there is a clear tendency of low PAPR yielding sequences being present in the last quadrant i.e. in the region of sequence no. 41 to 62. Hence, in our proposed method, to find out the proper phase sequences with low computational complexity, we adopted a reverse searching technique which starts from index 64 and end at 1. Usually, searching is done from sequence 1 to 64 and can be called searching with forward direction. Now, if reverse searching is done with minimum PAPR, the result will be same as Fig. 1. Hence, with the target of minimizing computational complexity, some PAPR threshold values are selected which are slightly higher than the minimum PAPR (2 dB). In this research, the selected PAPR threshold values are 2.1 dB and 2.2 dB. Using each case of selected threshold values of PAPR for GCD sequence with SLM, searching for proper phase sequences is analyzed for both forward and reverse direction.

The number of symbols that achieved PAPR threshold of 2.1 dB versus the number of iterations required in forward and reverse searching is shown in Fig. 3 (a) and Fig. 3 (b), respectively. From Fig. 3 (a), it is observed that for forward searching, about 98,000 out of total OFDM symbols yields the PAPR threshold in 53 iterations (from 1 to 53). However, it is also observed that as 98,000 number of symbols appeared with in a band of phase sequence from 29 to 53. Hence, searching can be focused in this range. As, within this band, there are 25 number of phase sequences, complexity reduction of 61% is found. In contrast, for reverse searching, almost 98,200 symbols satisfied the PAPR threshold requirement within a band of phase sequence from 64 to 61. In this way, only 4 iterations are required for reverser searching and computational complexity reduction of about 94% is achieved. Proposed reverse searching required only 4 phase sequences to be transmitted as side information compared to around 25 for forward one.

TABLE I
SIMULATION PARAMETERS

Number of Subcarriers (N)	64
Modulation Technique	$BPSK$
Phase Sequences	$GCD\ matrix$
Number of Phase Sequences	64
Length of OFDM Symbols	10^5
Simulation Tool and Version	MATLAB [(R2009a)]

Similar simulation is repeated while increasing the PAPR threshold value slightly. In Fig. 4 (a) and (b), the simulation results for forward and reverse searching is given, respectively when the PAPR threshold is set at 2.2 dB. Almost similar situation is observed for Fig. 4 (a) and (b) for forward and reverse searching and complexity reduction of 70% and 94% is found, respectively. Hence, from the above discussions it is evident that the proposed reverse searching technique outperforms the forward one by a large margin and significantly reduces the computational complexity.

Irrespective of the searching with usual forward one or proposed reverse one, selected number of phase sequences must be sent as side information along with the OFDM signal. Now, it will be highly efficient, if it is possible to develop a transmission system that will support side information free transmission along with the PAPR reduction mechanism. This can be achieved using a special phase sequence called default index which yields minimum PAPR for maximum number of OFDM symbols. Going back to fig. 1 that most of the symbols satisfying minimum PAPR appeared at phase sequence 61st. Hence, index with 61st phase sequence is chosen as the default GCD sequence. However, some sacrifice in PAPR compared to the minimum one is found in case of default index which is shown if Fig. 2. From Fig. 2, it is seen that CCDF of PAPR for GCD phase sequence with default index 61st is 2.5 dB. As minimum PAPR for GCD phase sequence is 2 dB, PAPR sacrifice of only 0.5 dB is observed for 61st default index which is within acceptable range.

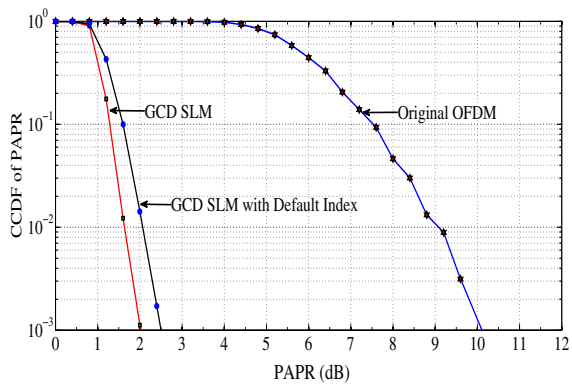


Fig. 2. CCDF vs. PAPR for SLM with GCD phase sequences, SLM with default GCD phase sequence and Original OFDM system with N=64, BPSK modulation.

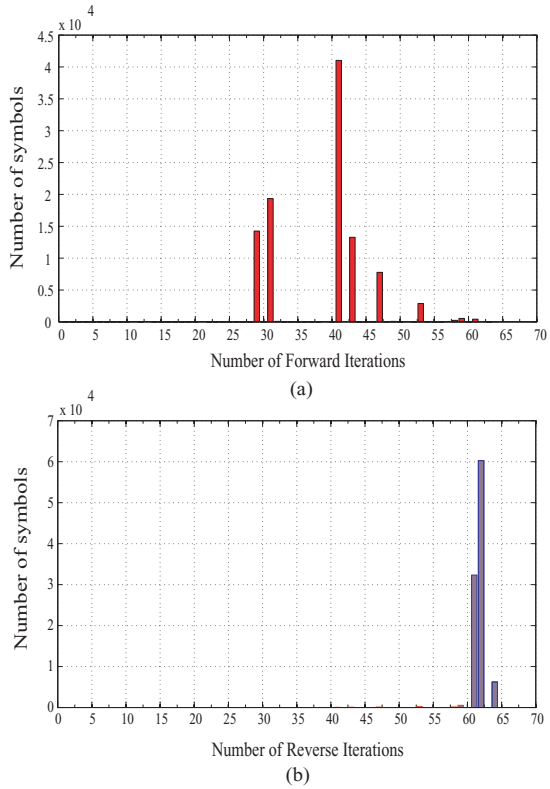


Fig. 3. Searching in GCD Matrix with PAPR threshold of 2.1 dB (a) Searching in Forward direction (b) Searching in Reverse direction.

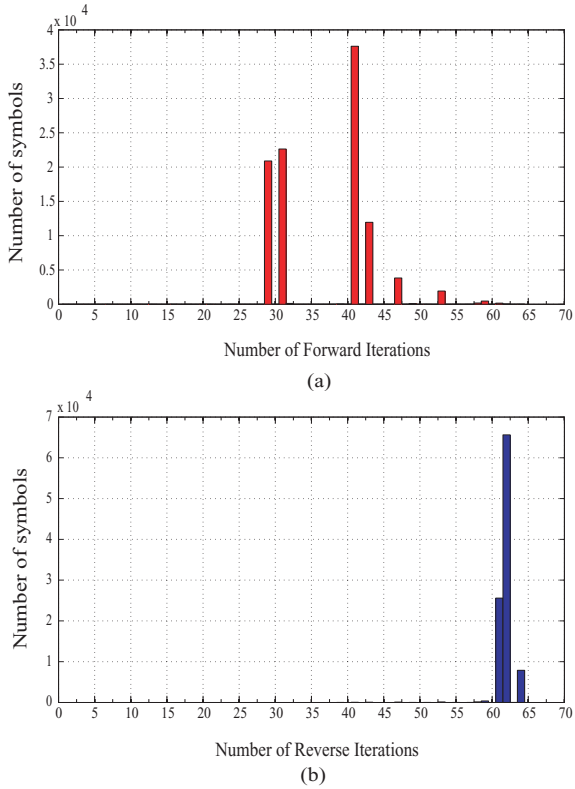


Fig. 4. Searching in GCD Matrix with PAPR threshold of 2.2 dB (a) Searching in Forward direction (b) Searching in Reverse direction.

Finally, it can be said that, depending on the application requirement and affordability of PAPR sacrifice, the selection of the proposed reverse searching technique with small number of side information or transmission with default index without any side information can be adopted.

VII. Conclusions

In OFDM system, compensating the adverse effect of high PAPR is one of the most challenging issues. Among the existing PAPR reduction methods, none of them can achieve all the desirable outcomes, namely, large reduction in PAPR, no spectral leakage, low computational complexity, no or little loss of data rate etc. In this paper, we focused on the complexity reduction issue along with PAPR reduction. Our proposed technique showed a new approach of using GCD sequence in SLM with reverse searching technique. In the proposed approach it is possible to simultaneously achieve significant reduction in PAPR along with very low computational overhead. The results showed that PAPR reduction of 8.1 dB can be obtained and computational complexity can be reduced up to 94%. It is also found that side information free transmission with default GCD index is achieved with only 0.5 dB sacrifice in PAPR compared to the minimum one. Hence, both the issues of reducing the computational complexity with reduced PAPR and side information free transmission using default GCD index, with small sacrifice in PAPR is accomplished.

References

[1] L. J. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing,"*IEEE Trans. Commun., vol. 33, no. 7, pp. 665-675, July 1985.*

[2] J. G. Andrews, A. Ghosh, and R. Muhamed, *Fundamentals of WiMAX, Prentice Hall, 2007.*

[3] Y. Rahmatallah and S. Mohan, "Peak-to-average power ratio reduction in OFDM systems: A survey and taxonomy," *IEEE Commun. Surveys and Tutorials, vol. 15, no. 4, pp. 1567-1592, 2013.*

[4] R. W. Bauml, R. F. H. Fischer, and J. B. Huber, "Reducing the peak-to-average power ratio of multicarrier modulation by selected mapping,"*IEEE Elect. Lett., vo1. 32, no. 22, pp. 2056-2057, Oct. 1996.*

[5] S. H. Muller and J. B. Huber, "OFDM with reduced peak-to-average power ratio by optimum combination of partial transmit sequences,"*IEEE Elect. Lett., vo1. 33, no. 5, pp. 368-369, Feb. 1997.*

[6] N. V. Irukulapati, V. K. Chakka, and A. Jain, "SLM based PAPR reduction of an OFDM signal using new phase sequence," *IEEE Elect. Lett., vol. 45, no. 24, pp. 1231-1232, Nov. 2009.*

[7] D. W. Lim, J. S. No, C. W. Lim, and H. Chung, "A new SLM OFDM scheme with low complexity for PAPR reduction," *IEEE Sig. Proc. Lett., vol. 12, no. 2, pp. 93-96, Feb. 2005.*

[8] S. L. Marple Jr., "Computing the discrete-time analytic signal via FFT," *IEEE Trans. on Sig. Proc., vol. 47, no. 9, pp. 2600-2603, Sep. 1999.*

[9] M. M. Rahman, M. S. Rahim, M. N. A. S. Bhuiyan, and S. Ahmed, "Greatest common divisor matrix based phase sequence for PAPR reduction in OFDM system with low computational overhead,"*in Proc. 1st Int. Conf. on Elect. and Electronic Eng. (ICEEE), RUET, Rajshahi, Bangladesh, Nov. 2015, pp. 97-100.*

[10] L. Chen and X. Hu, "A modified SLM scheme for PAPR reduction,"*in Proc. 3rd Int. Congress on Image and Sig. Proc. (CISP), Yantai, Shandong, China, Oct. 2010, pp. 4299 - 4302.*

[11] R. V. Nee and A. D. Wild, "Reducing the peak to average power ratio of OFDM,"*in Proc. 48th IEEE Semiannual Veh. Tech. Conf., Ottawa, Ontario, Canada, vol. 3, May. 1998, pp. 2072-2076.*

Impact of Phase Offset and Fading in Full Duplex MIMO System Under Antenna Cancellation

Md. Rayhan Khan and Md. Forkan Uddin*
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka-1205, Bangladesh
*mforkanuddin@eee.buet.ac.bd

***Abstract*—Full duplex MIMO communication is one of the key techniques for 5G wireless communication. The main challenge in full duplex operation is the cancellation of the self-interference (SI). In this paper, *Transmit Antenna Cancellation* technique is considered with symmetric placement of antennas to cancel SI in full duplex MIMO systems. The performance of a full duplex system with this cancellation technique may degrade due to phase offset and channel fading. Analytical model of signal to interference plus noise ratio (SINR) of a full duplex MIMO system is developed considering phase offset and channel fading for studying their impact on the performance of the system. Uplink capacity as well as outage probability of the full duplex MIMO system is determined numerically by using the developed analytical model. The results show that the performance of full duplex MIMO system significantly degrades due to channel fading and phase offset. It is also found that the performance of the system degrades with increasing the number of transmit antennas and decreasing the number of receive antennas at the full duplex transceiver.**

***Keywords*—Full duplex, self-interference, MIMO, phase offset, fading, signal to interference plus noise ratio.**

I. Introduction

A basic limitation of wireless communication is that the traditional wireless transceiver cannot transmit and receive on the same frequency at the same time, i.e., a transceiver operates in a half duplex mode. Half duplex systems are not optimal in use of time and frequency resource. As a result, full duplex transceiver design has become attractive to the researchers for 5G wireless communication [1]. In theory, full duplex data transmission is capable of doubling the spectral efficiency compared to that of half duplex transmission [2]. The key challenge in full duplex operation in a wireless transceiver is to detect the desired signal drowned in so called self-interference (SI) signal generated by proximately located transmitters which might be billion times stronger than the desired signal.

Multiple-input, multiple-output (MIMO) communication which uses multiple antennas at transmitter and receiver has been recognized as a promising technology to effectively handle the newer challenges and meet the flourishing requirements [3]. MIMO offers higher capacity and better transmission quality due to it's array gain, spatial diversity gain, spatial multiplexing gain and interference reduction capability [4], [5]. Although it is already in use with half duplex antennas, it is expected that (large scale) MIMO will be an integral part of 5th generation wireless communication network with full duplex antennas [6].

A significant amount of researches have already been done on designing full duplex system with SI cancellation [7]- [15]. In [7], the authors use two transmit antennas and one receive antenna to cancel SI. However, inability to work under wideband systems makes this system less practical [8]. In [8], the authors propose balanced and unbalanced (BALUN) cancellation that uses signal inversion through a BALUN circuit. According to [9], difficulties in tuning and quadratic growth of complexity for MIMO systems makes this system difficult to realize. The authors in [10] propose a single antenna SI cancellation scheme for single-input single-output (SISO) systems composed of a circulator, analog cancellation circuit and tuning algorithm. The same authors modify the version of SISO system for MIMO system [11]. Bandwidth constraint of ferrite circulators as well as the possibility of damaging sensitive receiver circuitry due to high transmit power limits the application of the proposed technique in [11]. In [14], the researchers develop two estimation algorithms for two-stage SI cancellation. In [15], the authors propose a linear model in frequency domain for SI channel in indoor environment. Digital SI cancellation is addressed in [12] and [13] which might not be feasible due to noise floor.

In [9], the authors propose SI cancellation by antenna cancellation which is simple, effective and scalable for (massive) MIMO system compared to the other proposed SI cancellation techniques. Antenna cancellation technique can be two types: transmit antenna cancellation and receive antenna cancellation. In transmit antenna cancellation, two identical transmit antennas separated by some distance transmit the same stream. If one unaltered and another passed through a π phase shifter, for a receive antenna the potential null points would be located on the perpendicular bisector on the line connecting the two transmit antennas. In receive antenna cancellation, two identical antennas are used to receive signals and a composite signal is created by passing one signal directly and the other signal through a π phase shifter.

Wireless channel has some intrinsic properties like fading and shadowing that hamper wireless communication severely. On the other hand, phase shifters are not ideal and they have phase offset. The performance of antenna cancellation techniques may degrade significantly due to channel fading and phase offset. However, so far, there is no clear indication in the literature on how much the performance of antenna cancellation technique will degrade due to channel fading and phase offset. So it is essential to study the effect of fading and phase offset on antenna based SI cancellation.

In this paper, a full duplex MIMO system is considered with transmit antenna cancellation technique. Linear minimum mean square error (LMMSE) MIMO detection technique is considered at the MIMO full duplex transceiver. The main contributions in this paper are as follows. Analytical model of signal to interference plus noise ratio (SINR) of the full duplex MIMO system is developed considering phase offset and channel fading. Uplink capacity as well as outage probability is computed numerically for the full duplex MIMO system using the developed analytical model and demonstrated that performance of a full duplex MIMO system degrades significantly due to channel fading and phase offset. The impact of the number of the transmit and receive antennas of the full duplex transceiver on the performance of

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

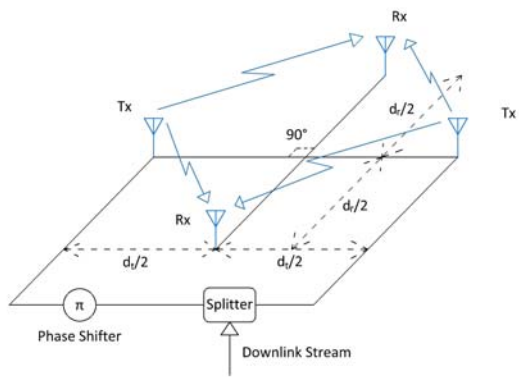


Fig. 1. Transmit antenna cancellation technique.

full duplex MIMO system is also studied.

The rest of the paper is organized as follows. The MIMO full duplex system is described in detail in Section II. Analytical model of SINR is developed in Section III. In Section IV, the numerical results are presented and various engineering insights are provided. The paper is concluded in Section V.

II. System Model

A. Self-Interference Cancellation Scheme

Transmit antenna cancellation is considered for SI cancellation [9]. In this cancellation technique, a pair of transmit antennas transmit the same symbol with equal transmit power. The phase difference between the carrier signals of the two transmit antennas are set to π radians by using a phase shifter. Since the potential null points are located on the perpendicular bisector of the line joining the transmit antennas, the receive antennas are placed on the perpendicular bisector of the line joining the two transmit antennas. A typical scheme is shown in Fig. 1 where, transmit and receive antennas are marked by Tx and Rx, respectively.

B. Network Architecture

We consider a single-cell full duplex MIMO system consisting of one base station (BS) and K active users. The BS is equipped with M pairs of transmit antennas and N receive antennas with full duplex capability, while each user terminal is equipped with a single-antenna. In general, a user can be equipped with multiple antennas. For simplicity of analysis, system with single-antenna users is assumed. We assume that all the users share the same time-frequency resource. Furthermore, we assume that the BS has perfect channel state information (CSI) of all the antennas.

C. Propagation Model

Only path loss is considered for the wireless propagation from the users to the BS. Since we are interested to study the fading effect on the performance of SI cancellation, path loss as well as channel fading is considered for the wireless propagation from the transmit antennas to the receive antennas of the BS. The channel gain between the kth user and the nth receive antenna is given as

$$g_{k,n} = \sqrt{\left(\frac{s_{k,n}}{d_0}\right)^{-\eta}} \tag{1}$$

where, $s_{k,n}$ is the distance between the kth user and the nth receive antenna, d_0 is the reference distance in the far field, η is the path loss exponent. The channel gain between the first antenna of the mth pair of antennas and n^{th} receive antenna is given as

$$a_{m,n} = h_{m,n}\sqrt{\left(\frac{d_{m,n}}{d_0}\right)^{-\eta}} \tag{2}$$

where, $d_{m,n}$ and $h_{m,n}$ are the distance and fast fading coefficient, respectively, from the first antenna of the mth pair transmit antennas to the nth receive antenna at the BS. The coefficient $h_{m,n}$ is assumed to be Gaussian random variable with zero mean and unit variance. The channel gain between the second antenna of the mth transmit antenna pair and the nth receive antenna, $a'_{m,n}$ can be obtained from (2) by replacing $h_{m,n}$ with $h'_{m,n}$. We assume $h_{m,n}$ and $h'_{m,n}$ are independent random variables.

D. MIMO Full Duplex Transmission and MIMO Detection Technique

Let $\mathbf{X}^u$ is the vector of the data symbols of the K users and $\mathbf{X}^d$ is the vector of the data symbols of the M transmit antenna pairs. The transmit power of all the users is assumed to be equal and denoted by P_u. The transmit power at the transmit antennas of the BS is also assumed to be equal and denoted by P_d. Due to better performance against noise, LMMSE detection algorithm is considered at the BS [16].

E. Phase Offset Model

Phase offsets of different pairs of transmit antennas are assumed to be independent. Phase offset of a pair of transmit antennas is assumed to be a random variable following the standard normal distribution having zero mean and σ^2 variance. The probability density function (PDF) of phase offset of the mth pair of transmit antennas is given by $f_{\theta_m}(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\theta^2}{2\sigma^2}\right)$.

III. Analytical Formulation

We are particularly interested in modelling uplink capacity because the main challenge for applying full duplex mode is to detect the signals from the users drowned in SI signal. Downlink signal is not affected by SI as single antenna is considered at each user terminal.

A. Received Signal Model

The signal from the kth user to the nth receive antenna can be written as

$$y_{k,n} = \sqrt{P_u} g_{k,n} x_k^u e^{i\left(\omega t + \frac{2\pi s_{k,n}}{\lambda}\right)} + n_{k,n} \tag{3}$$

where, x_k^u is the symbol transmitted by the kth user, ω is the angular velocity of carrier, λ is the carrier wavelength and $n_{k,n}$ is the additive white Gaussian noise (AWGN) at the receiver. Similarly, the SI signal from the mth transmit antenna pair to the nth receive antenna can be written as

$$\begin{aligned} y_{m,n} =& \sqrt{P_d} a_{m,n} x_m^d e^{i\left(\omega t + \frac{2\pi d_{m,n}}{\lambda}\right)} \\ &+ \sqrt{P_d} a'_{m,n} x_m^d e^{i\left(\omega t + \frac{2\pi d_{m,n}}{\lambda} + \pi + \theta_m\right)} + n_{m,n} \\ =& \sqrt{P_d}(a_{m,n} - a'_{m,n} e^{i\theta_m}) x_m^d e^{i\left(\omega t + \frac{2\pi d_{m,n}}{\lambda}\right)} + n_{m,n} \end{aligned} \tag{4}$$

where, x_m^d is the symbol transmitted from the mth transmit antenna pair and $n_{m,n}$ is the AWGN.

Let $\tilde{g}_{k,n} = g_{k,n} e^{\frac{2\pi i s_{k,n}}{\lambda}}$ and $\tilde{a}_{m,n} = (a_{m,n} - a'_{m,n} e^{i\theta_m}) e^{\frac{2\pi i d_{m,n}}{\lambda}}$. Let $\mathbf{G}$ and $\mathbf{G}_{SI}$ are the $K \times N$ and $M \times N$ matrices for the $\tilde{g}_{k,n}$ and $\tilde{a}_{m,n}$ variables, respectively. The

signal received by the nth receive antenna can be expressed as $y_n = \sum_{k=1}^{K} y_{kn} + \sum_{m=1}^{M} y_{mn}$. Let $\mathbf{Y} = [y_1, y_2, \ldots, y_N]^T$ is the received signal vector for all the receive antennas. Using vector format, $\mathbf{Y}$ can be expressed as

$$\mathbf{Y} = \sqrt{P_u}\mathbf{G}^H\mathbf{X}^u + \sqrt{P_d}\mathbf{G}_{SI}^H\mathbf{X}^d + \mathbf{n} \tag{5}$$

where, $\mathbf{n}$ is the $K \times 1$ noise vector.

B. LMMSE Detection

LMMSE MIMO detection minimizes the mean square error by determining a matrix $\mathbf{W}$ to obtain the post detection signal at the BS as

$$\mathbf{r} = \mathbf{W}^H\mathbf{Y}. \tag{6}$$

According to LMMSE MIMO detection [16], for the uplink system under consideration, $\mathbf{W} = \mathbf{R}_{yy}^{-1}\mathbf{R}_{yx}$, where, $\mathbf{R}_{yy} = E\left(\mathbf{Y}\mathbf{Y}^H\right)$ and $\mathbf{R}_{yx} = E\left(\mathbf{Y}\mathbf{X}^{uH}\right)$ are the covariance matrices. For the system under consideration, we obtain

$$\mathbf{R}_{yy} = P_u^2\mathbf{G}^H\mathbf{G} + P_d^2\mathbf{G}_{SI}^H\mathbf{G}_{SI}\sigma_n^2\mathbf{I} \tag{7}$$

and

$$\mathbf{R}_{yx} = P_u^{3/2}\mathbf{G}^H \tag{8}$$

where, σ_n^2 is the variance of noise. Thus, LMMSE MIMO detection matrix can be written as

$$\mathbf{W} = \left(P_u^2\mathbf{G}^H\mathbf{G} + P_d^2\mathbf{G}_{SI}^H\mathbf{G}_{SI}\sigma_n^2\mathbf{I}\right)^{-1} P_u^{3/2}\mathbf{G}^H. \tag{9}$$

C. SINR and Capacity Model

The post detection signal at the BS can be written as

$$\mathbf{r} = \sqrt{P_u}\mathbf{W}^H\mathbf{G}^H\mathbf{X}^u + \sqrt{P_d}\mathbf{W}^H\mathbf{G}_{SI}^H\mathbf{X}^d + \mathbf{W}^H\mathbf{n} \tag{10}$$

where, $\mathbf{r}$ is a $K \times 1$ matrix and the kth element representing the post detection received signal from the kth user. The SINR of the kth user can be obtained as

$$SINR_k = \frac{P_S}{P_I + P_N} \tag{11}$$

where, the post detection signal power P_S of the kth user is given as $P_u\left|w_{n,k}^H\tilde{g}_{n,k}\right|^2$, the interference power P_I is given as

$$P_I = P_u \sum_{i=0, i\neq k}^{K} \left|w_{n,i}^H\tilde{g}_{n,i}\right|^2 + P_d\|\mathbf{W}^H\mathbf{G}_{SI}\|^2 \tag{12}$$

and the post detection noise power P_N is given as

$$P_N = \sum_{i=1}^{K} \frac{\sigma_n^2\sigma_i^2}{(P_u^2\sigma_i + P_d^2\sigma_i\sigma_{SI_i} + \sigma_n^2)}. \tag{13}$$

where, σ_i and σ_{SI_i} are the singular values of a user i for $\mathbf{G}$ and $\mathbf{G}_{SI}$, respectively. In (12), $P_u\sum_{i=0,i\neq k}^{K}\left|w_{n,i}^H\tilde{g}_{n,i}\right|^2$ is the post-detection interference from other active users and $P_d\|\mathbf{W}^H\mathbf{G}_{SI}\|^2$ is the self-interference signal power. The ergodic capacity the kth user can be obtained from the Shannon capacity formula as follows:

$$C_k = E\left[log_2(1 + SINR_k)\right]. \tag{14}$$

IV. Simulation Results

In this section, the effect of phase offset and fading on the performance of the full duplex MIMO system is studied.

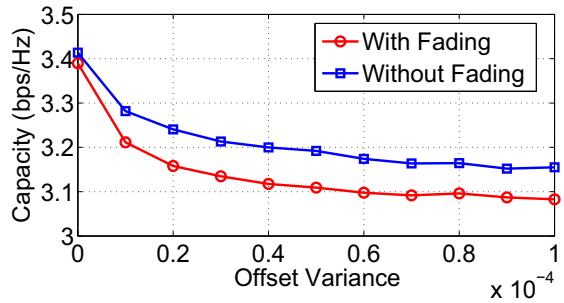


Fig. 2. Impact of phase offset and channel fading on the capacity of a user.

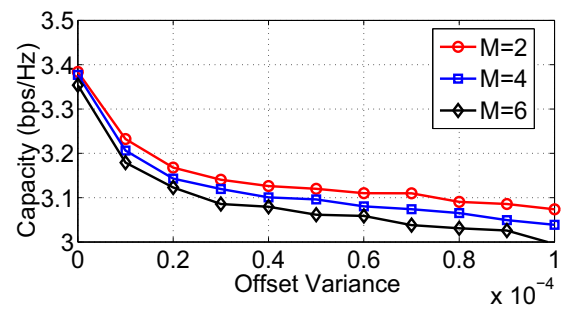


Fig. 3. The effect of the number of transmit antennas on the capacity of a user.

A. System Parameters

The transmit power levels at the transmit antennas of the users and the BS are taken to be 0.5 W and 5 W, respectively. The path loss exponent η is taken to be 3 assuming outdoor environment. Unless stated otherwise, the number of transmit antenna pairs is taken to be 2 and the number of receive antennas is taken to be 8. The total number of users in the system is assumed to be 4. The 4 users are located at 200 m, 300 m, 250 m and 500 m away from the BS. The transmit antennas are placed in a horizontal line and the receive antennas are placed in the perpendicular bi-sector horizontal line. The antennas are placed symmetrically keeping 20 cm distance between two adjacent antennas of same type. The carrier frequency is assumed to be 1900 MHz. The far field cross-over distance d_0 is taken to 10 cm. The noise variance σ_n^2 is assumed to be 10^{-25}.

B. Computation of Capacity

To compute the ergodic capacity of a user, the SINR of the user is determined 4000 times by generating the elements of the matrices $\mathbf{G}$ and $\mathbf{G}_{SI}$ according to their random distributions in each iteration and then average of Shannon capacity over the 4000 iterations is taken.

C. Impact of Phase Offset and Fading on Capacity

The uplink capacity of the 3rd user among the 4 users is computed for different values of phase offset variance with and without fading and presented in Fig. 2. The results show that the capacity of the user decreases for including channel fading and the impact of fading is more significant when phase offset is considered. The capacity loss is found to be around 0.7% for fading when phase offset is zero. On the other hand, the capacity losses are found to be 2.5% and 2.2% for fading for offset variances 2×10^{-5} and 8×10^{-5}, respectively. Further, it is seen that the capacity of the user decreases rapidly with increasing the phase offset. The capacity loss is found to be around 9% when offset variance is 10^{-4} with respect to the capacity without phase offset.

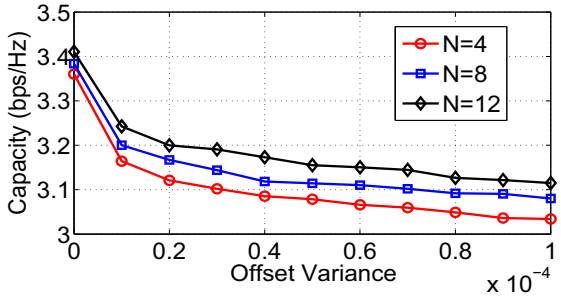


Fig. 4. The effect of the number of the receive antennas on the capacity of a user.

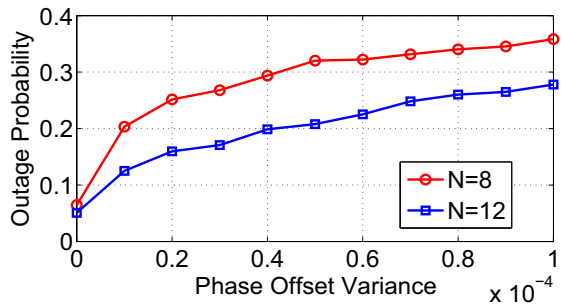


Fig. 5. The effect of phase offset on the outage probability of a user.

D. Impact of the Number of Transmit Antenna Pairs on Capacity

The uplink capacity of the 3rd user is computed under fading and phase offset for different number of transmit antenna pairs in the BS and shown in Fig. 3. The results show that the capacity of the user decreases with increasing the number of transmit antenna pairs. We attribute this to the fact that the SI increases with increasing the number of transmit antenna pairs at the BS due to phase offset and results in low SINR as well as low capacity.

E. Impact of the Number of Receive Antennas on Capacity

The capacity of the 3rd user is also computed under fading and phase offset for different number of receive antennas in the BS. The capacity results are illustrated in Fig. 3. The results show that the capacity of the user increases with increasing the number of receive antennas. We attribute this to the fact that noise and interference decreases with increasing the number of receive antennas due to the inherent properties of LMMSE detection..

F. Impact of Fading and Phase Offset on Outage Probability

The outage probability of a user is the probability that the received SINR at the receiver of the user is less than a threshold. To compute the outage probability of a user, the SINR of the user is determined 4000 times as mentioned above and then the outage probability is calculated as $N_f/4000$, where N_f is the number of iterations that the SINR of the user fails to meet the threshold.

The outage probability of the 3rd user is computed under fading and offset variance with a threshold of 7 dB and presented in Fig. 5. For $N = 8$, the outage probability is found to be around 0.07 when phase offset is zero and it becomes around 0.30 and 0.37 for offset variances 4×10^{-5} and 8×10^{-5}, respectively. Further, for $N = 12$ the outage probabilities are found to be around 0.20 and 0.28 for offset variances 4×10^{-5} and 8×10^{-5}, respectively. Thus, the outage probability increases significantly with increasing phase offset and decreasing the number of receive antennas.

V. Conclusions

The effect of phase offset and channel fading on the performance of full duplex MIMO system with antenna cancellation technique has been studied. By means of numerical results using the developed analytical model of SINR, it is demonstrated that (i) the capacity of a user degrades due to channel fading, (ii) the capacity of a user decreases with increasing phase offset and the number of transmit antenna pairs, (iii) the outage probability of a user increases significantly for increasing phase offset and (iv) the capacity and outage probability performance of a user can be improved by increasing the number of receive antennas. Since phase offset and channel fading are inherent properties of full duplex wireless MIMO system with such antenna cancellation scheme, to overcome the problem arises by these two properties, an effective digital cancellation technique should be provided along with the antenna cancellation technique.

References

[1] D. Korpi, T. Huusari, Y.-S. Choi, L. Anttila, S. Talwar, and M. Valkama, "Full-duplex mobile device-pushing the limits," *IEEE Communication Magazine (submitted)*, 2014.

[2] S. Hong, J. Brand, J. I. Choi, M. Jain, J. Mehlman, S. Katti, and P. Levis, "Applications of self-interference cancellation in 5G and beyond," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 114-121, 2014.

[3] H. Bölcskei, *Space-time wireless systems: from array processing to MIMO communications*. Cambridge University Press, 2006.

[4] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless personal communications*, vol. 6, no. 3, pp. 311-335, 1998.

[5] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585-595, 1999.

[6] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. M. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 122-130, 2014.

[7] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel, full duplex wireless communication," in *Proc. the sixteenth annual ACM international conference on Mobile computing and networking*, pp. 1-12, 2010.

[8] M. Jain, J. I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proc. the 17th annual ACM international conference on Mobile computing and networking*, pp. 301-312, 2011.

[9] M. A. Khojastepour, K. Sundaresan, S. Rangarajan, X. Zhang, and S. Barghi, "The case for antenna cancellation for scalable full-duplex wireless communications," in *Proc. the 10th ACM Workshop on Hot Topics in Networks*, pp. 17, 2011.

[10] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *ACM SIGCOMM Computer Communication Review*, vol. 43, no. 4, pp. 375-386, 2013.

[11] D. Bharadia and S. Katti, "Full duplex mimo radios," in *11th USENIX Symposium on Networked Systems Design and Implementation (NSDI14)*, pp. 359-372, 2014.

[12] D. Korpi, L. Anttila, and M. Valkama, "Reference receiver based digital self-interference cancellation in MIMO full-duplex transceivers," in *IEEE Globecom Workshops (GC Wkshps)*, pp. 1001-1007, 2014.

[13] D. Wu, C. Zhang, S. Gao, and D. Chen, "A digital self-interference cancellation method for practical full-duplex radio," in *Proc. IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, pp. 74-79, 2014.

[14] A. Masmoudi and T. Le-Ngoc, "Self-interference cancellation for fullduplex mimo transceivers," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 141-146, 2015.

[15] A. Nadh, J. Samuel, A. Sharma, S. Aniruddhan, and R. K. Ganti, "A linearization technique for self-interference cancellation in full-duplex radios," arXiv preprint arXiv:1605.01345, 2016.

[16] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM wireless communications with MATLAB*, John Wiley & Sons, 2010.

Stability Analysis of Grid Connected PV Array Under Maximum Power Point Tracking

Md. Hasan Shahriar, Md. Jawwad Sadiq and Md. Forkan Uddin*
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka-1205, Bangladesh
*mforkanuddin@eee.buet.ac.bd

***Abstract*—In this paper, the stability of a grid connected photovoltaic (PV) array is studied under two different maximum power point tracking (MPPT) algorithms. For the study, a 100 kW PV array connected to a grid via a voltage source inverter and a boost converter is considered. The system is modelled in MATLAB Simulink including the incremental conductance (InC) and perturb & observe (P&O) MPPT techniques. By Simulink simulation, the grid disturbances are studied under normal as well as transient conditions. The simulation results show that at normal condition the InC MPPT algorithm provides better response time and efficiency than the P&O MPPT algorithm. However, at transient condition P&O MPPT algorithm provides better recovery time than the InC MPPT algorithm. This study advocates to design hybrid MPPT algorithm such that the algorithm works effectively under both transient and normal conditions.**

***Index Terms*—Electricity grid, distributed generation, power system stability, maximum power point tracking.**

I. Introduction

Smart grid (SG) is recognized as the platform for the future of the power industry [1], [2]. The rapid rise of this issue is also leading to the fast expansion of distributed generation technology markets, such as fuel cells (FC), photovoltaic (PV) and wind turbine (WT). This trend will have a significant impact on future electricity technology. In the next few years, the world is supposed to face several problems related to the exhaustion of some energy sources, mainly those regarding fossil fuels. It is also well known that some aspects concerning the increase of oil price due to economic and political issues have been the cause of economic crisis in the last decades [3].

The search for renewable energy sources becomes more and more intense as a prominent alternative for the mitigation of the world energy crisis. PV solar energy solves the issue as a clean and green power source [4]. The significant cost reduction of PV modules in the last few years has made its usage popular among the small single-phase residential systems connected to the utility grid and stand-alone applications [5]. Despite all the advantages, the conversion efficiency is low and the initial cost is still high, as it is necessary to use maximum power point tracking (MPPT) techniques in order to extract the maximum power from the PV cell as the algorithms continuously track the maximum power point at which maximum power prevails [6], [7]. It is important to emphasize that there is only one maximum power point for each curve at given temperature and irradiance levels.

Due to the popularity of the PV solar energy, grid connected PV array has been studied in different aspects [6], [8], [9], [10]. In [6] and [8], the authors present experimental efficiency comparison of some MPPT algorithms on PV array. In [9] and [10], the authors perorm modeling and simulation of PV arrays. Stability is one of the most important issue in power system. Since PV array is connected to the grid through voltage source inverter and boost converter, it is very important to study the transient stability on this system. A little research has been carried out to investigate the stability issue when grid is integrated with PV array. Overall, research scenarios in the past have investigated the technical impact of distributed generation on existing grid only for normal condition, not for grid disturbances caused by different types of faults. Moreover, none of the research scenarios dealt with the issues from smart grid perspectives, i.e., grid connected distributed generation. Therefore, further investigation is required to analyze the impact of distributed generation on smart grid stability.

In this paper, a 100 kW PV array connected to the grid via a voltage source inverter and a boost converter is considered. The incremental conductance (InC) and perturb & observe (P&O) MPPT techniques are considered for improving the performance of the system. For transient analysis three phase line to ground fault is considered. The contributions in this paper are as follows:

- The grid connected PV array is modelled in MATLAB Simulink including the InC and P&O MPPT techniques.
- Simulink simulation is performed to study the grid disturbances under normal as well as fault conditions.
- Via simulation results, it is demonstrated that the InC technique provides better performance under normal condition and the P&O technique provides better response under fault condition.

The rest of the paper is organized as follows. The MATLAB Simulink model of grid connected PV array is presented in detail in Section II. In section III, the simulation results are presented and the various engineering insights are provided. The paper is concluded in Section IV.

II. Simulink Model of Grid Connected PV Array

The basic block diagram of a grid connected PV array is shown in Fig. 1, where the DC output of the PV array is converted to grid suitable AC by boost converter and inverter. In the figure, B_1, B_2 and B_3 represent three buses of the system. A MPPT technique is integrated with the boost converter for maximum power. The utility grid power is connected to the output of the inverter through step-up transformer. For transient study, line as well as load faults are considered in the output of the inverter. In the following, the Simulink model of the different components of the block diagram is described.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

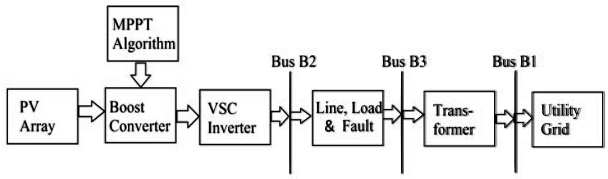


Fig. 1. The block diagram of a grid connected PV array.

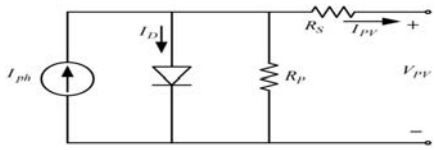


Fig. 2. Equivalent circuit of an ideal PV cell.

A. PV Cell Model

The I-V characteristic of an ideal PV cell is given as [11]

$$I_{PV} = I_{ph} - I_0[\exp(\frac{qV_{PV}}{akT}) - 1] \quad (1)$$

where, I_{ph} is the current generated by the incident light, I_0 is the reverse saturation or leakage current of the diode, q is the electron charge, k is the Boltzmann constant, T is the temperature in kelvin in the p-n junction and a is the diode ideality constant. The equivalent circuit of an ideal PV cell with series resistance R_s and parallel resistance R_p is shown in Fig. 2 [11], where I_D is the doide current. The mathematical model of PV array is developed in MATLAB Simulink as a current source where the current of the source depends on irradiance and temperature and is shown in Fig. 3. Various parameters of the PV array are determined by trial and error basis with $T = 298^o$ K. The I-V and P-V characteristics of the developed PV module at constant temperature 25^oC are presented in Fig. 4 for various irradiances (IR), where for a given voltage a higher irradiance provides higher current as well as higher power.

B. MPPT Algorithms

InC and P&O algorithms are considered for MPPT. In the InC method, the controller measures incremental changes in PV array current and voltage to predict the effect of a voltage change [12]. On the other hand, in P&O method, the controller adjusts the voltage by a small amount from the array and measures power and if the power increases, further adjustments in that direction are tried until power no longer increases [13], [14].

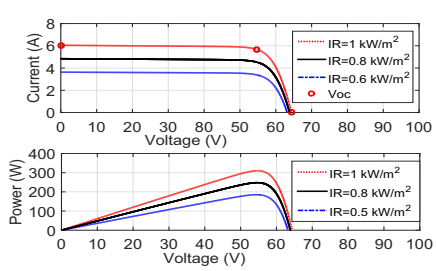
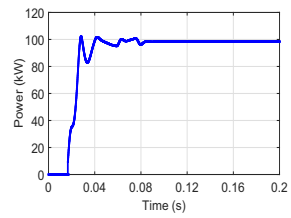
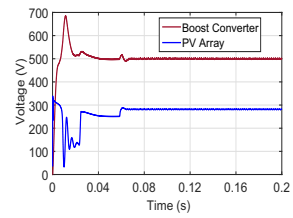


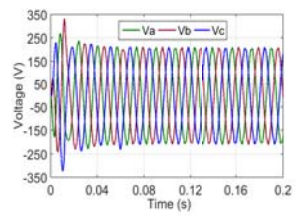
Fig. 4. I-V and P-V characteristics of the PV module at various irradiances and constant temperature.



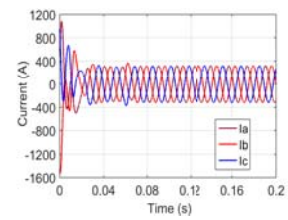
(a) Output power of PV array



(b) Output voltage of boost converter and PV array



(c) Output voltage of inverter



(d) Output current of inverter

Fig. 6. Voltage, current and power at different stages of the system under normal condition using InC algorithm.

C. Simulink Model

The Simulink model of 100 kW PV array with dc-dc boost converter and InC MPPT algorithm is shown in Fig. 5. The same conditions of the InC algorithm are applied for the P&O algorithm. The PV array is connected to a 25-kV grid via the DC-DC boost converter and a 3-phase 3-level voltage source converter (VSC). The PV array consists of 66 strings of 5 series-connected modules connected in parallel which delivers a maximum of 100 kW at 1 kW/m^2 sun irradiance. 5-kHz DC-DC boost converter increases voltage from PV natural voltage (273 V DC at maximum power) to 500 V DC. Switching duty cycle is optimized by the MPPT controller. In the system, 10-kVAR capacitor bank is used for filtering the harmonics produced by VSC and 100-kVA 260V/25kV three-phase coupling transformer is used to increase voltage level to the grid voltage level. Utility grid is considered as 25-kV distribution feeder and 120 kV equivalent transmission system.

III. Simulation Results and Discussions

The system model is simulated under two conditions: (i) normal condition and (ii) three phase to ground fault condition for the two MPPT algorithms. The performance of the InC and P&O MPPT algorithms are evaluated under these two conditions.

A. Normal Condition

For normal condition, the Simulink model is simulated and the voltages, currents and power at different stages are recorded. In Fig. 6, the output power of PV array, output voltages of boost converter and PV array and output voltage and current of the inverter of first few milliseconds are presented for the case of InC MPPT algorithm. From the curve of output power of PV array, it is observed that the system becomes steady after few milliseconds. From the output voltage of the boost converter, it is clear that low PV array voltage is boosted up by the MPPT algorithm and boost converter. The output voltage and current of the inverter are found to be sinusoidal after the transient period. In the fluctuating period

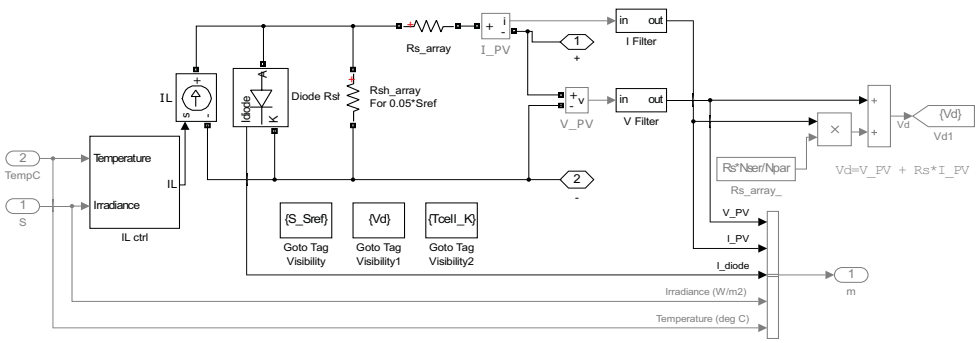


Fig. 3. Simulink model of PV array.

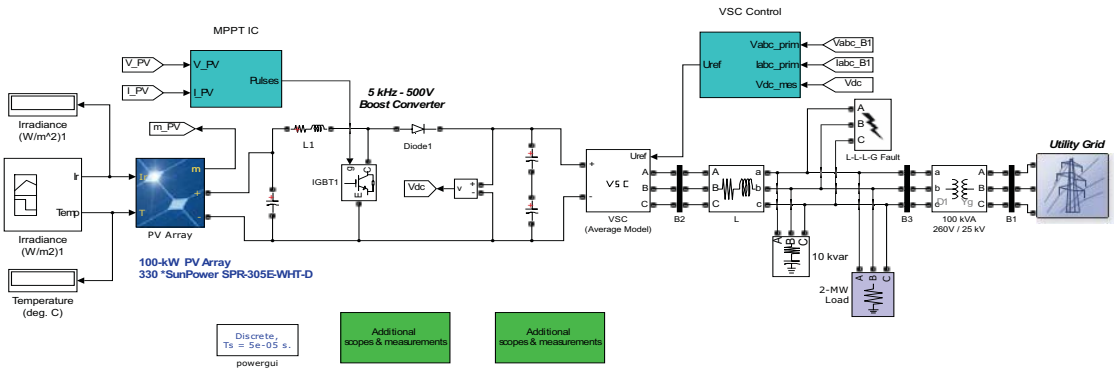
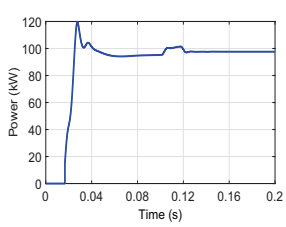
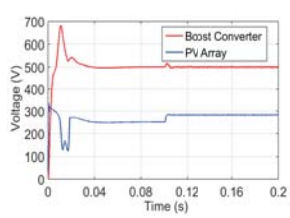


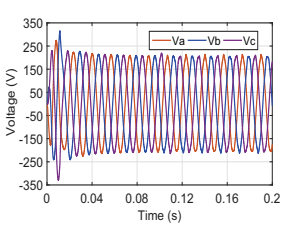
Fig. 5. Simulink Model of grid connected PV array with InC MPPT algorithm.



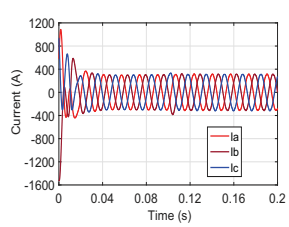
(a) Output power of PV array



(b) Output voltage of boost converter and PV array



(c) Output voltage of inverter



(d) Output current of inverter

Fig. 7. Voltage, current and power at different stages of the system under normal condition using P&O algorithm.

of the curves, the MPPT algorithm takes time to get a steady duty cycle and the PV cell becomes saturated during this time. Similar results are generated under the P&O MPPT algorithm and presented in Fig. 7. The trend of the curves are the same for the both MPPT algorithms. However, performance of the algorithms are quite different. A comparison of the various performance indicating indexes of the two MPPT algorithms is

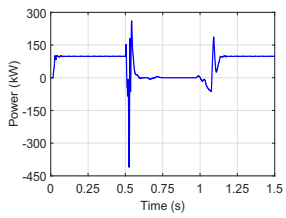
TABLE I
COMPARISON BETWEEN P&O AND INC ALGORITHMS UNDER NORMAL CONDITION

MPPT	Output Current (A)	Output Voltage (V)	Output Power (kW)	Response Time (s) (s)	Efficiency
InC	348.55	282.62	99.16	0.09	More
P&O	341.03	285.10	98.31	0.23	Less

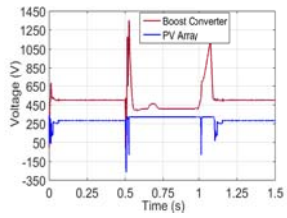
presented in Table I. The results show that the response time is less in case of the InC algorithm since this algorithm tracks the maximum power point more rapidly than the P&O algorithm. Further, the InC MPPT algorithm provides more steady state output power compared to the P&O MPPT algorithm. Thus, the efficiency of the InC algorithm is better than that of the P&O algorithm. In overall, the InC algorithm is better than the P&O algorithm under normal condition.

B. Transient Condition

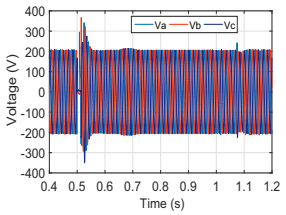
To study and compare the transient characteristics of the grid connected PV array under the two MPPT algorithms,three phase to ground fault is applied in the system. An L-L-L-G fault is applied in bus B_3 for 1 ms and the voltage, current and power of the system at the different stages are observed for the two algorithms. The output power of the PV array, the output voltage of the boost converter and the PV array and the output voltage and current of the inverter are presented in Fig. 8 and Fig. 9 for the InC and P&O MPPT algorithms, respectively. In all the curves, distortion is observed around 1 ms as three phase to ground fault occurs at this time. The distortion in



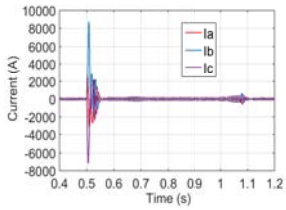
(a) Output power of PV array



(b) Output voltage of boost converter and PV array

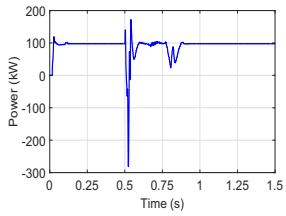


(c) Output voltage of inverter

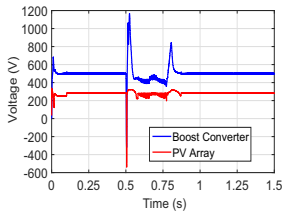


(d) Output current of inverter

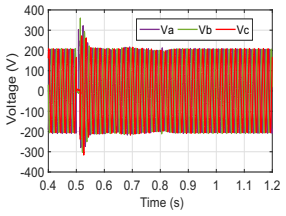
Fig. 8. Voltage, current and power at different stages of the system under transient condition using InC algorithm.



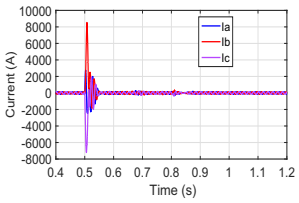
(a) Output power of PV array



(b) Output voltage of boost converter and PV array



(c) Output voltage of inverter



(d) Output current of inverter

Fig. 9. Voltage, current and power at different stages of the system under transient condition using P&O algorithm.

voltage, current and power continues for some time and then they become steady. Both P&O and InC algorithms show the similar behaviour. However, the recovery time is very large in case of the InC algorithm compared to the recovery time of the P&O algorithm. It is observed that the InC algorithm takes too much time to become steady again after the recovery time. A comparison of the performance of the InC and P&O algorithms is presented in Table II. The results show that under L-L-L-G fault the P&O algorithm provides less peak power, peak voltage, peak current and recovery time than the InC algorithm. So the performance of the P&O MPPT algorithm is better than the InC algorithm under L-L-L-G fault. Similar simulations are carried out with L-L and single L-G faults and obtained the similar insights.

TABLE II
COMPARISON BETWEEN P&O AND INC ALGORITHMS UNDER L-L-L-G FAULT

MPPT	Peak PV Array Power (kW)	Peak Voltage (V)	Peak Current (A)	Recovery Time (ms)	Efficiency
InC	409.6	367.8	8762.0	680	Less
P&O	279.8	360.0	8541.5	397	More

IV. Conclusion

The main focus of this paper remains under modelling and simulation of the grid connected PV system to study the stability of the system. A 100 kW grid connected PV array is studied under normal and transient conditions utilizing the P&O and InC MPPT algorithms. The InC algorithm offers faster response and expected output compared to the P&O algorithm under normal condition. On the other hand, the P&O algorithm shows better performance for any kind of transient condition. Since none of the two algorithms is efficient under the both normal and transient conditions, the future work of this work could be a design of a hybrid MPPT algorithm taking the features of the both algorithms to provide better performance under the both conditions.

References

[1] T.N.E.T.L. (NETL), "Understanding the benefits of smart grid", 2010.

[2] X. Fang, S. Misra, G. Xue and D. Yang, "Smart grid-The new and improved power grid: A survey," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, pp. 944-980, Dec. 2011.

[3] J.A. Lopes, N. Hatziargyriou, J. Mutaleand and N. Jenkins, "Integrating distributed generation into electric power system: A review of drivers, challenges and opportunities," *Electrical Power Systems Research*, vol. 77, pp. 1189-1203, 2007.

[4] A. Duragadevi, S. Aruselvi and S. P. Natarajan,"Study and implementation of maximum power point tracking (MPPT) algorithm for photovoltaic systems," *Proc. 1st International Conference On Electrical Energy Systems (ICEES)*, pp. 240-245, 2011.

[5] J. S. Katz, "Educating the Smart Grid," *Proc. IEEE Energy 2030 Conference*, pp. 1-3, Nov. 2008.

[6] T. Esram and P. L. Chapman,"Comparison of photovoltaic array maximum power point tracking techniques," *IEEE Transactions On Energy Conversion,* vol. 22, pp. 439-449, 2007.

[7] S. Tao, L. Liu, X. Xiao, X. Zhen and Y. Yang, "Photovoltaic MPPT algorithm based on hysteresis comparison and optimal gradient," *Proc. IEEE 15th International Conference on Harmonics and Quality of Power (ICHQP),* pp. 733-738, 2012.

[8] M. A. G. de Brito, L.P. Sampaio, L.G. Jr., G.A. e Melo, and C.A. Canesin "Comparative analysis of MPPT techniques for PV Applications", *Proc. International Conference on Clean Electrical Power (ICCEP),* 2011.

[9] M.G. Villalva, J.R. Gazoli and E.R. Filho, "Comprehensive approach to modeling and simulation of photovoltaic arrays," *IEEE Transactions on Power Electronics,* vol. 24, no. 5, pp. 1198-1208, 2009.

[10] H. M. Nordin and A. M. Omar, "Modeling and simulation of photovoltaic (PV) array and maximum power point tracker (MPPT) for grid-connected PV system," *Proc. the 3rd International Symposium and Exhibition in Sustainable Energy and Environment (ISESEE '11)*, pp. 114-119, June 2011.

[11] Photvoltaic Cell I-V Characterization Theory [Online] Available: http://www.ni.com/white-paper/7230/en/

[12] A. Safari and S. Mekhilef, "Simulation and hardware implementation of incremental conductance MPPT with direct control method using cuk converter," *IEEE Transactions On Industrial Electronics*, vol. 58, no. 4, pp. 1154-1161, April 2011.

[13] R. Faranda, S. Leva and V. Maugeri, "MPPT techniques for PV Systems:energetic and cost comparison," *Power and Energy Society, IEEE General Meeting-Conversion and Delivery of Electrical in the 21st Century*, 2008.

[14] M.G. Villalva and F.E. Ruppert, "Analysis and simulation of the P&O MPPT algorithm using a linearized PV array model," *Proc. 35th Annual Conference of IEEE Industrial Electronics,* pp. 231-236, 2009.

Performance Evaluation of Saturating Class – C Driver Circuit for Inductive Wireless Power Transfer

Md Khurram Monir Rabby[1], Shamim Ara Shawkat[1, 2] and Mohammad Shah Alam[1]

[1]IICT, Bangladesh University of Engineering and Technology (BUET), Bangladesh
[2]EECS, University of Tennessee, Knoxville, USA

Abstract—**This paper presents performance evaluation of saturating class – C driver circuit for inductive wireless power transfer. Exact link is designed following link optimization theory to select parameters of the objective circuit. By exploring specific parameters of drain to source voltage effect on overall circuit, single MOSFET is used in the proposed driver circuit design instead of two parallel MOSFETs used in conventional driver circuit. Simulation results show that the proposed driver circuit provides a significant improvement in the driver efficiency (90.34%) as well as in the overall efficiency compared to the existing driver circuit. Furthermore, from the performance comparison, it is observed that the transferred output power (14.94 mW) has also been improved considerably.**

Index Terms—**Wireless power transfer (WPT), ripple factor (RP), form factor (FF) and inductive power transfer (IPT).**

I. Introduction

Wireless power transfer (WPT) technology has been around for over 100 years, and applied in the automotive industry as well as biomedical systems. It is used in wireless charging to electric and hybrid machineries [1]. Power and efficiency of wireless power transfer via coupled magnetic resonance vary with the impacts of the distance, orientation, and centre deviation [2–3]. Although at long distance weak magnetic coupling for series resonance has negligence effect on sender due to the receiver, the overall efficiency is less. Therefore, for the desired frequency of operation, parallel secondary resonance is used [4–8]. In this case, to overcome the limitations of conventional WPT, a novel method using DC/DC converter is proposed for the reduction of reflected power by changing duty cycle of switching devices and impedance matching circuit to improve the efficiency of the overall circuit [9-12]. Another challenge of WPT technology is to design the driver circuit. The overall efficiency depends on this part also. Saturating class – C is more preferable due to its less system loss and more efficient capacity compared to all other driver circuit [1]. The success of WPT technology is applied in scientific and non-scientific field of interest. The main applicable area is biomedical sector. It is used in endoscope for communication, power transfer and motion-free capsule endoscopy inspection in a specifically designed jacket. In scientific purpose, it is used in powered train system, rotating spindle's diagnostic sensor, utilization of solar energy as well as the wireless transmission of the generated power to increase the usage of renewable energy resources in order to reduce the carbon-mono-oxide (CO) emission, implanted in tissue specific absorption rate (SAR) heating in the two resonant modes for electric vehicle (EV) charging application and on high-frequency DC - AC and AC - DC wireless power transfer stage where the inversion stage is current-fed topology [13-17].

In this paper, we propose an efficient saturating class – C driver circuit for inductive wireless power transfer circuit in biomedical field as well as small range power required area with a specific voltage and current rating. A step-by-step design procedure is carried out for driven optimum inductive link that operates as close as possible to their theoretical efficiency limit. Using this design technique, the circuit parameters are selected for the driven inductive link, saturating class – C driver circuit and other components of the overall circuit. In the proposed technique, depending on drain to source voltage single MOSFET is used instead of two parallel MOSFETs in the driver circuit. A test system is considered for simulation purpose where significant improvement is observed in overall system efficiency using proposed technique compared to existing one. For parametric selection of equation of existing design technique and circuit simulation purpose, MATLAB and SPICE simulator are used respectively.

The rest of the paper is organized as follows. Optimum link and primary coil driver of saturating class – C are discussed in section II. In section III, a step by step design process is presented. Simulation results as well as necessary discussion of circuit analysis are done in section IV. Finally, concluding comments are given in section V.

II. LINK & PRIMARY COIL DRIVER

A. Exact Link

An inductive links with a non-resonant primary and a parallel-resonant secondary are discussed in Fig. 1. This figure is followed by the optimization of the system efficiency and the displacement tolerance by critical coupling. The equations for the link components (i.e., primary coil, L_{s1} and secondary coil, L_{s2}) and performance are calculated by the equivalent impedance, Z_{eq} of the secondary reflected to the primary. The coupling factor, k and load, R_{load2}, have a minimal effect on the tuning of primary circuits. Here, C_2 is the stabilizing capacitor. General expressions for the link efficiency and gain are followed by the derivation of the operation modes for maximal link efficiency or critical coupling [1]. To represent the losses of primary and secondary resistors R_{s1} and R_{s2} are used respectively.

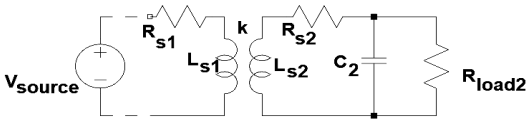


Fig. 1 Exact inductive link for non-resonant primary and parallel-resonant secondary [1].

B. Primary Coil Driver: Saturating Class C

The saturating-class-C amplifier consists of a parallel L-C (L_{eq} and C_1) tank having the resistive load, R_{s1} plus an additional coil resistance, R_{eq} (ref. Fig. 2). Tank coil, L_{eq}

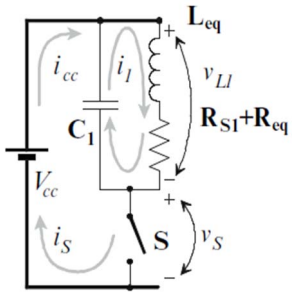


Fig. 2 Circuit diagram and operation of saturating class C primary coil driver [1].

indicates the tuned primary coil coupled to the secondary in an inductive link. The parasitic coil loss and the secondary consumption are represented by inseparable tank resistors, $R_{s1}+R_{eq}$. The driver has freely resonating primary tank for power supply and the driver transistor to compensate for the tank losses. This circuit is operated by switching, S on a power supply. The driver tank is operated at the tank-resonance frequency ω_{tank} as tank current, i_1. Switch voltage oscillation, v_s is found around the supply V_{cc}. These are operated as switches controlled by clock-like input signals converting from DC to AC [1].

III. Proposed Circuit Design Technique

The optimization procedure is followed to design the overall circuit [1]. According to this procedure, it is divided into two parts: the magnetic and the electronic link design. Furthermore, the electronic design is divided into three sub-parts which brings the number of major design steps to four:

Step 1: Geometry requirement and power specification.
Step 2: Magnetic design.
Step 3: Electronic design for primary and secondary circuit.
Step 4: Control simulation.

A. Magnetic Design

The steps of magnetic circuit design technique specified in [1] are followed to identify the parameters. The parameters are selected considering the maximization of coil coupling and minimization of the magnetic losses below:

1) Primary and secondary coil

The coil dimensions, the placement and the shape, the cores and the packaging selection are specified in steps [1].

2) Coupling factor, k

The most important parameter determined by the ratio of the two areas circumscribed by the coils, and by the distance between both coils [1].

3) Magnetic losses and frequency

A trade–off is maintained to select these two parameters. Higher frequency value matches the secondary to the load. However, the losses and radiation increase due to the generation of eddy current. Here the negligible coil coupling, skin and proximity effects are not considered [1].

4) Self and mutual inductance

They are determined by the universal formula using finite element flux distribution model [1].

B. Electronic Design

The optimization procedure is followed for electronic design. The electronic optimization design procedure

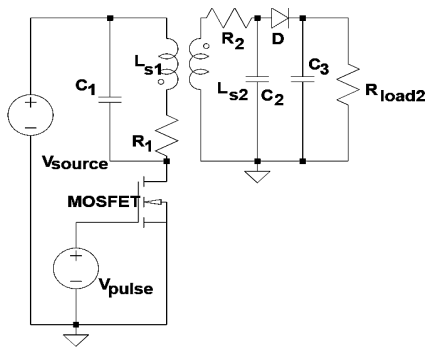


Fig. 3 Saturating class – C driver circuit design for inductive link with a parallel-resonant secondary.

described in [1] is used to specify parametric value of each component of the overall circuit. Electronic design specific parameters are selected within range of magnetic specified range.

The overall circuit using the proposed driver circuit is mentioned in Fig. 3. In this design procedure, the following parameters are specified below:

Here, R_{load2}: ohmic value of load from design specification, C_3: voltage stabilizing capacitor for load, D: diode used as rectifier, R_2: loss representation in secondary circuit, C_2: filtering capacitor, L_{s2}: secondary coil, k: coupling factor, L_{s1}: primary coil, R_1: loss representation in primary circuit, C_1: parallel tuned capacitor for primary coil, L_{s1} and V_{pulse}: for switching MOSFET.

C. MOSFETs Effect in the Efficiency of Driver

Transistor selection plays an important role in the performance of the saturating-class – C driver. Furthermore, primary coil and tank capacitors also affect the performance. There are five parameters: the on resistance, voltage maximum, current handling capacity, transistor speed and gate driving source. MOSFET's are preferred due to driving easiness comparative to BJT. Although power MOSFET has very low on-resistance, their large gate capacitance requires a special buffer. Its average gate current is as large as the actual driver consumption.

In our proposed MOSFET topology as switching purpose, single MOSFET is used instead of two parallel MOSFETs in the driver circuit to enhance performance. In this case, the MOSFET which has higher drain to source voltage, v_{ds} is selected. However, higher v_{ds} specified MOSFET has higher on resistance. Therefore, a trade-off is done to overcome this phenomenon. To do this, considering low on resistance, MOSFET having higher v_{ds} is initially selected. Here, v_{ds} must be at least 10 times greater than design required value. The advantage of this technique is that it increases current handling capacity. In rush current cannot do any damage during start-up. Moreover, due to low on resistance with increased current handling capacity the overall circuit performance increases.

IV. Simulation Results and Discussion

In order to prove the efficiency of proposed design technique, test small scale wireless power transfer circuit is designed. For the circuit design MATLAB is used as a simulation tool to find out parametric value for the specific range of desired output. To test whether designed circuit meet the requirement of design specification, SPICE simulator has been used.

Proposed improved technique has been applied in a sample test case to make the comparison. In this design, a telemetry unit is considered in a non-conductive housing that accepts a 6 mm solenoid with 20 mm diameter consuming a maximum of 4 mA from a 4 V-regulated supply having maximum target distance is 70 mm.

The parameters obtained using the design technique is presented in Table I.

TABLE I
PARAMETRIC VALUE OF CIRCUIT COMPONENT IN FIG. 3

Parameter	Value	Unit
$R_{d.c}$	1125	Ω
R_{load2}	625	Ω
f	1	MHz
k	0.71	%
L_{s1}	1.73	µH
L_{s2}	1.73	µH
C_1	14.55	nF
C_2	14.65	nF

After selecting the parameter in the overall circuit design, it is needed to testify whether the designed circuit meets the requirement or not. In this case, losses are considered in the equipment parametric value selection. Therefore, SPICE simulator has been used to justify the practical design environment. In addition to that, performance parameters are used to evaluate the output wave shape for circuit analysis.

In the proposed technique, single MOSFET is used instead of two parallel MOSFETs in driver circuit. The proposed technique is used in driver where comparative low on resistance but high v_{ds} featured MOSFET (Type – Si4490DY) is used (shown in Table II).

TABLE II
SI4490DY TYPE MOSFET FEATURE

Properties Type	Value	Unit
Manufacturer	Siliconix	-
Polarity	N-chan	-
Vds	200	V
Rds (on)	0.08	Ω
Qgate Cap	34	nF

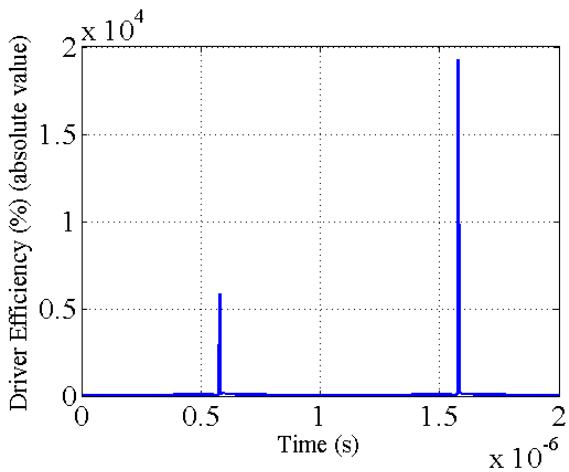


Fig. 4 Driver efficiency for proposed driver circuit design using single MOSFET.

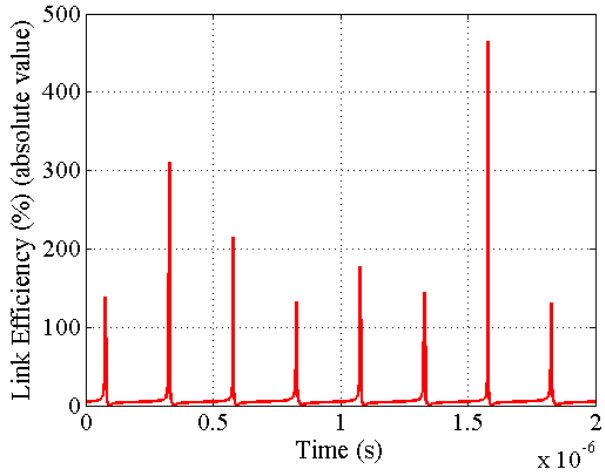


Fig. 5 Link efficiency for proposed driver circuit design using single MOSFET.

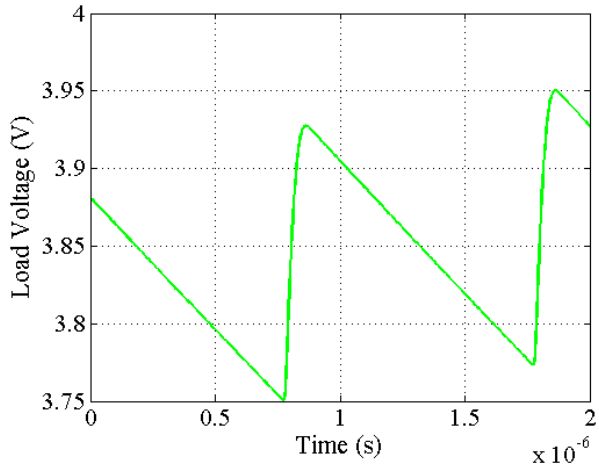


Fig. 6 Output voltage across load for proposed driver circuit design using single MOSFET.

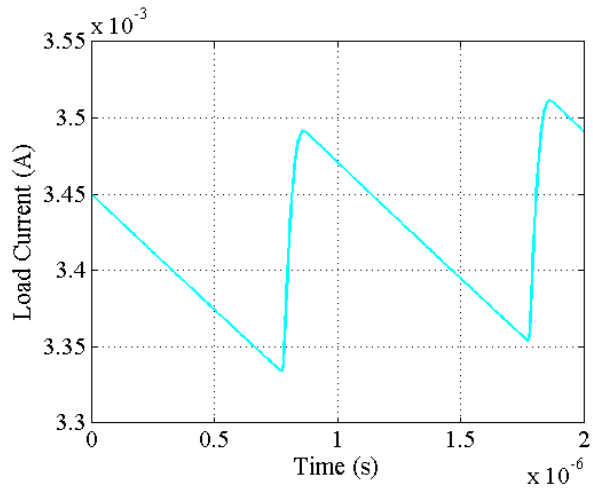


Fig. 7 Output current through load for proposed driver circuit design using single MOSFET.

This n-channel Si4490DY type MOSFET input source voltage is 21.4V. For DC source, DC current is found as 12.329 mA. Here, average source power, P_{source} is found as 263.8406 mW and power across driver, P_{driver} is calculated as 263.592882 mW. Therefore, driver efficiency, η_{driver} is calculated as 90.339%. Using power wave shapes of source and driver, absolute driver efficiency wave shape has been shown in Fig. 4

Similarly, for the link efficiency, primary side, RMS voltage and current as well as DC voltage and current are 15.749 V, 1.4489 A, 431.09 mV and 11.594 mA respectively.

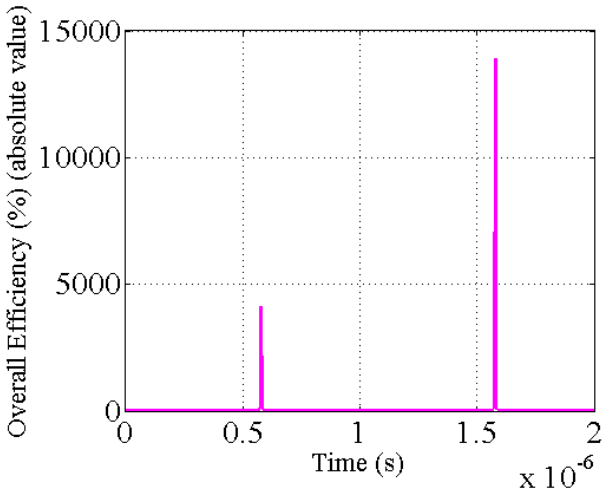


Fig. 8 Overall efficiency for proposed driver circuit design using single MOSFET.

Therefore, DC and AC primary power is 4.99806 µW and 22.8187261 W respectively. Across the inductor in the secondary side, RMS voltage and current have been found as 3.3655V and 310.34 mA respectively whereas DC voltage and current are 3.2813 mV and 4.0396 mA respectively. DC and AC secondary power are calculated as 13.2551 µW and 1.04444927 W respectively. Link efficiency is calculated as 8.5%. Like driver, absolute overall link efficiency has been shown in Fig. 5.

In load side RMS voltage and current are found as 3.8454 V and 3.4182 mA respectively. Here, DC values are almost same to the RMS value. Therefore, DC voltage and current are 3.8451 V and 3.8771 mA respectively (shown in Fig. 6 and Fig. 7). Some ripples are observed. However, ripple factor (RF) is calculated as 1.249% which is negligible. For overall efficiency calculation, DC and AC output power are calculated as 14.90784 mW and 13.1443 mW respectively. Overall efficiency is calculated 5.65%. Using source and load wave shapes of voltage and current, absolute overall efficiency wave shape has been presented in Fig. 8. To check for accuracy of the respective output wave shape form factor (FF) is calculated as 100%.

From performance parameter analysis, it is proved that using the proposed technique it is possible to transfer better quality and stable power through WPT system to a remote unit.

V. Conclusions

In this paper, we presented an efficient saturating class – C driver to improve the efficiency for inductive wireless power transfer circuit in which single MOSFET has been used instead of two parallel MOSFETs in the driver circuit. Magnetic link, electronic link and the saturating class – C driver have been designed considering optimum design strategy. Simulation results show that the proposed driver circuit provides a significant improvement in the driver efficiency as well as in the overall efficiency. Furthermore, from the performance comparison, it is observed that the transferred output power has also been improved considerably. Although single MOSFET application may decrease its redundancy, improved overall efficiency makes it better applicable to biomedical system and middle range wireless power transfer transmitter and energy harvesting area. Since the transferred overall efficiency is always an important issue for the above mentioned applications, simulation results indicate promise as a driver circuit for an inductive wireless power transfer.

Acknowledgment

The authors are thankful to Dr Koenraad Van Schuylenbergh and Dr Robert Puers, Department of Electrical Engineering, Katholieke University Leuven, Belgium for their valuable advices in performing in SPICE.

References

[1] Koenraad Van Schuylenbergh, Robert Puers, Inductive powering basic theory and application to biomedical systems, Springer, Belgium, 2009.

[2] Yurtin, J., "The Morton Antler lecture: So ... What's All This Wireless Power Transfer Stuff About Anyway??," in Electrical Contacts (Holm), 2014 IEEE 60th Holm Conference on, vol., no., pp.1-2, 12-15 Oct. 2014.

[3] Zhuo Yan; Yang Li; Chao Zhang; Qingxin Yang, "Influence Factors Analysis and Improvement Method on Efficiency of Wireless Power Transfer Via Coupled Magnetic Resonance," in Magnetics, IEEE Transactions on , vol.50, no.4, pp.1-4, April 2014.

[4] Mur-Miranda, J.O.; Fanti, G.; Yifei Feng; Omanakuttan, K.; Ongie, R.; Setjoadi, A.; Sharpe, N., "Wireless power transfer using weakly coupled magnetostatic resonators," in Energy Conversion Congress and Exposition (ECCE), 2010 IEEE, vol., no., pp.4179-4186, 12-16 Sept. 2010.

[5] Church, J.; Bagley, J.Q.; Bana, V.; Rockway, J.D., "Metamaterial wireless power transfer system (MWPT)," in Radio Science Meeting (Joint with AP-S Symposium), 2014 USNC-URSI, vol., no., pp.219-219, 6-11 July 2014.

[6] Jiejian Dai; Ludois, D.C., "A Survey of Wireless Power Transfer and a Critical Comparison of Inductive and Capacitive Coupling for Small Gap Applications," in Power Electronics, IEEE Transactions on, vol.30, no.11, pp.6017-6029, Nov. 2015.

[7] Young-Do Kim; Hee-Jin Lee; Jin-Young Bang; Chin-Wook Chung, "A new design of wireless power transfer system using hellical resonators applicapable to multi-channel power transmission," in Wireless Power Transfer Conference (WPTC), 2014 IEEE, vol., no., pp.277-279, 8-9 May 2014.

[8] Yiming Zhang; Zhengming Zhao; Kainan Chen, "Frequency Decrease Analysis of Resonant Wireless Power Transfer," in Power Electronics, IEEE Transactions on, vol.29, no.3, pp.1058-1063, March 2014.

[9] Moriwaki, Y.; Imura, T.; Hori, Y., "Basic study on reduction of reflected power using DC/DC converters in wireless power transfer system via magnetic resonant coupling," in Telecommunications Energy Conference (INTELEC), 2011 IEEE 33rd International, vol., no., pp.1-5, 9-13 Oct. 2011.

[10] Koh Kim Ean; Beh Teck Chuan; Imura, T.; Hori, Y., "Novel band-pass filter model for multi-receiver wireless power transfer via magnetic resonance coupling and power division," in Wireless and Microwave Technology Conference (WAMICON), 2012 IEEE 13th Annual, vol., no., pp.1-6, 15-17 April 2012.

[11] Lijuan Xiang; Yue Sun; Xin Dai; Yu Chen; Xiao Lv, "Route optimization for wireless power transfer network based on the CE method," in Electronics and Application Conference and Exposition (PEAC), 2014 International, vol., no., pp.630-634, 5-8 Nov. 2014.

[12] Xiu Zhang; Ho, S.L.; Fu, W.N., "Quantitative Design and Analysis of Relay Resonators in Wireless Power Transfer System," in Magnetics, IEEE Transactions on, vol.48, no.11, pp.4026-4029, Nov. 2012.

[13] Guozheng Yan; Dongdong Ye; Peng Zan; Kundong Wang; Guanying Ma, "Micro-Robot for Endoscope Based on Wireless Power Transfer," in Mechatronics and Automation, 2007. ICMA 2007. International Conference on, vol., no., pp.3577-3581, 5-8 Aug. 2007.

[14] Tianjia Sun; Xiang Xie; Guolin Li; Yingke Gu; Yangdong Deng; Zhihua Wang, "A Two-Hop Wireless Power Transfer System With an Efficiency-Enhanced Power Receiver for Motion-Free Capsule Endoscopy Inspection," in Biomedical Engineering, IEEE Transactions on, vol.59, no.11, pp.3247-3254, Nov. 2012.

[15] GiBum Lee; HoCheol Gwak; Young-Sik Kim; Wee-Sang Park, "Wireless power transfer system for diagnostic sensor on rotating spindle," in Wireless Power Transfer (WPT), 2013 IEEE, vol., no., pp.100-102, 15-16 May 2013.

[16] Suja, S.; Sathish Kumar, T., "Solar based wireless power transfer system," in Computation of Power, Energy, Information and Communication (ICCPEIC), 2013 International Conference on, vol., no., pp.93-99, 17-18 April 2013.

[17] Xingyi Shi; Waters, B.H.; Smith, J.R., "SAR distribution for a strongly coupled resonant wireless power transfer system," in Wireless Power Transfer Conference (WPTC), 2015 IEEE, vol., no., pp.1-4, 13-15 May 2015.

A Compact Spectacles Shaped Patch Antenna for UWB Applications

M. Tarikul Islam[1], M. Samsuzzaman[1], M. Z. Mahmud[2],M.T. Islam[2]
[1]Dept. of Computer and Communication Engineering
Patuakhali Science and Technology University, Patuakhai, Bangladesh
[2]Dept. of Electrical Electronic and Systems Engineering,
Universiti Kebangsaan Malaysia,, Malaysia
* tarikulislam1994@gmail.com

Abstract— In this paper, a new compact spectacles shaped patch UWB antenna using a microstrip fed line is proposed. The new design is consist of a spectacles shaped patch along with tapered slot ground plane.The proposed antenna has a compact electrical dimension of 0.24 λ × 0.21 λ × 0.016 λ and fed with a 50-ohm microstrip transmission line. Computer Simulation Technology (CST) based on finite-difference time-domain method and High Frequency Structural Simulator (HFSS) based on finite element method is implemented in this research. In simulation, a wider frequency bandwidth of 8.5 GHz (3.0 - 11.5 GHz) with excellent impedance matching which covers entire UWB frequency, stable radiation pattern and constant gain is achieved. The proposed antenna is low profile, inexpensive and has excellent characteristics of UWB.

Keywords—UWB; spectacles patch;low cost;microstrip feed line.

I. Introduction

UWB is one of the most interested wireless communication systems which can be used for monitoring, positioning, security, microwave imaging and various communication application. Due to its high gain, omnidirectional radiation pattern, high data resolution, low complexity, inexpensive properties, it is becoming more attractive research phenomena for students and wireless communication. The antennas with ultra wideband frequency have been broadly researched and developed after the declaration of the unlicensed bandwidth of 3.1 to 10.6 GHz as UWB by Federal Communication Commission (FCC) [1]. Printed slot type antennas are mostly accepted for UWB application. It is still a challenge to antenna designers to design a compact, cost effective, high gain ultra wideband antennas.

In recent years, due to its various number of benefits including stable radiation pattern, high gain, low profile and inexpensive fabrication the printed microstrip slot antennas were significantly researched. For UWB applications numerous antennas were designed. Among them, one of the antenna requires a large ground plane that rises dimension. As a result, that is not included in microwave integration [2]. Various line feeding and waveguide feeding antennas were offered for UWB applications. For achieving the characteristics of wide impedance bandwidth monopole architectures are commonly used, such as elliptical, pentagon, rectangular, square, hexagonal, annular ring and circular ring antennas [3-7]. A small elliptical ring antenna with a coplanar waveguide for UWB application was inspected by Ren et al. Wideband frequency has been attained by this antenna ranging 4.6 – 10.3 GHz which doesn't cover whole UWB. Again the full UWB was not attained by this antenna despite compacted size [8]. A printed rectangular with dual circular slot patch monopole antenna is proposed which acquires operating bandwidth with a range of 3 – 11.5 GHz which covers the entire UWB region. Although the size of the proposed antenna has been reduced and antenna performance has been enhanced significantly. The antenna is compact in size than antenna stated in ref [9] and inexpensive for fabrication for using low-cost FR4 substrate. The spectacles slot shape patch with microstrip line-fed and tapered slot ground make it perfect for UWB applications.

In this paper, a planar monopole microstrip patch with tapered slot ground plane organized high performance stable radiation pattern, high gain antenna with a compact dimension of 24 × 21 mm^2 is presented. Due to spectacles slot patch with microstrip line-fed and tapered slot ground plane the proposed antenna reached the wide bandwidth ranging 3-11.5 GHz which makes it significant for using ultra wideband application.

II. Antenna Geometry

The length and width of the microstrip monopole antenna are calculated by Eqs. (1) and Eqs. (2) where L indicates length and W indicate the width of the antenna, ε_r indicates the dielectric constant where we use FR4 substrate, f_0 is targeted center frequency and c represents the velocity of light [10].

$$L = \frac{c}{2 f_0 \sqrt{\varepsilon_r}} - 2\Delta l \quad (1)$$

$$W = \frac{c}{2 f_0} \sqrt{\frac{\varepsilon_r + 1}{2}} \quad (2)$$

The geometry of the proposed antenna is illustrated in Fig. 1. The antenna consists of a spectacles shape which is made up with two circulars including a rectangular slot patch and tapered slot ground plane with a dimension of 24 × 21 mm^2. The design of the antenna is implemented with resonance patch on the dielectric constant FR4 substrate, a ground plane and a microstrip fed line. The proposed antenna has a compact dimension having 1.6 mm of thickness with a relative permittivity of 4.4 and 0.02 of loss tangent. Good range of bandwidth, high gain and stable radiation patterns can be gained

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

by selecting properly shaped slot at patch and ground plane. The proposed antenna is one of the smallest and costs effective UWB planner antennas. The adjusted constraints of proposed antenna are: $W = 21$ mm, $L = 24$ mm, $l = 7.30$ mm, $Wn = 5.56$ mm, $Wf = 3$ mm, $L1 = 8.66$ mm, $L2 = 7.52$ mm, $L3 = 10.12$ mm, $w1 = 1.98$ mm, $r1 = r2 = 2.5$ mm, $h = 1.6$ mm.

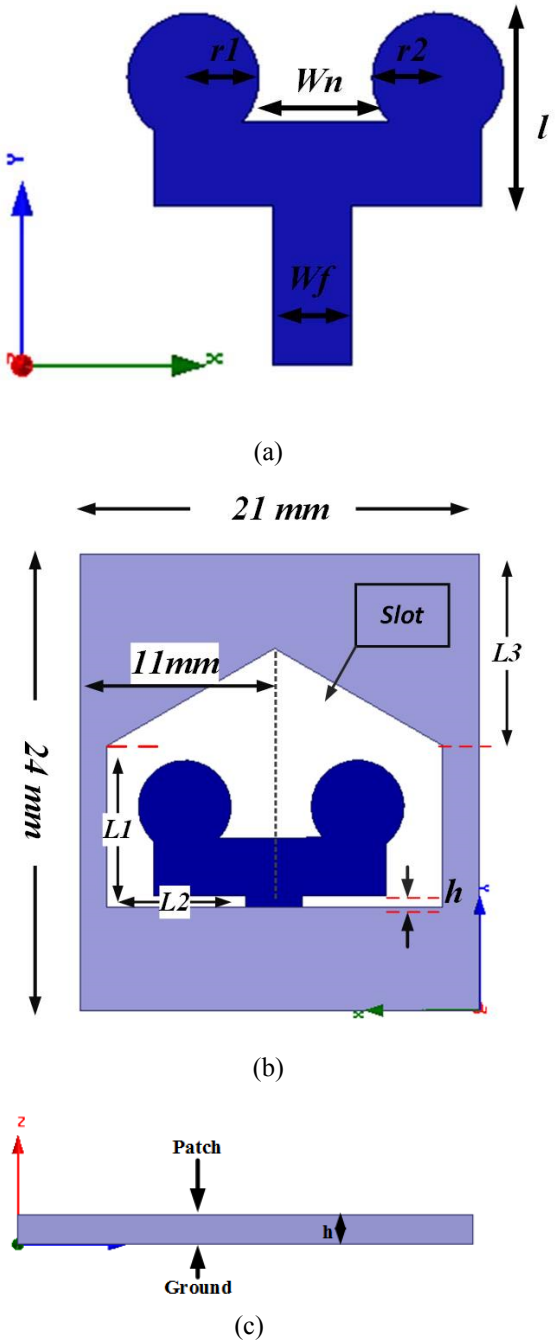


Fig. 1 The geometry of the proposed antenna. (a) Top view, (b) Bottom view, (c) Side view.

III. Results and Discussion

The performance of the proposed antenna is simulated and observed by CST and HFSS simulator. The antenna was fed by a 50-ohm transmission line and the dual circular with rectangular slot patch as microstrip transmission medium to get more wideband performance. High frequency working range can be obtained by using tapered slot ground structure. Simulated VSWR of the proposed antenna is demonstrated in Fig. 2, which indicates the frequency ranging from 3 to 11.5 GHz is the simulated operating bandwidth equal to or below 2. The result shows a little dispersion between two results of CST and HFSS due to their different working methods.

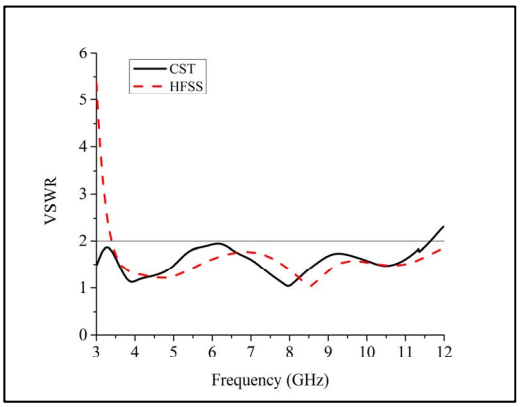


Fig. 2 Simulated VSWR in CST and HFSS.

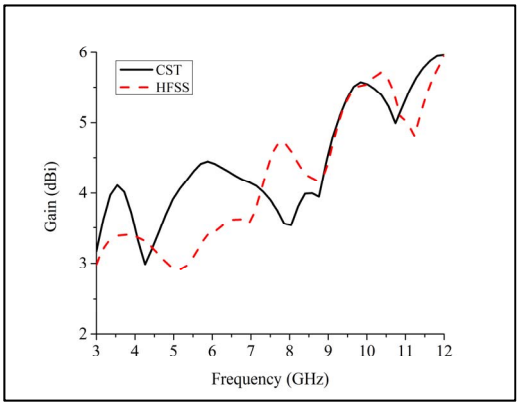


Fig. 3 Simulated gain with CST and HFSS.

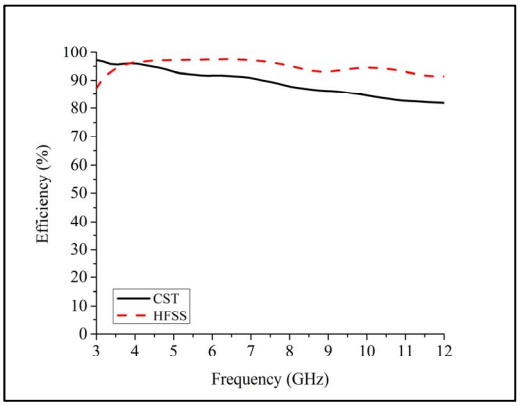
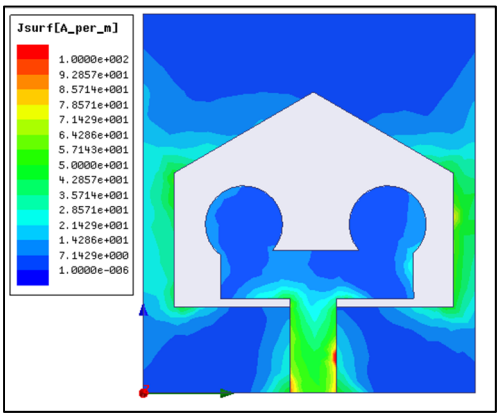
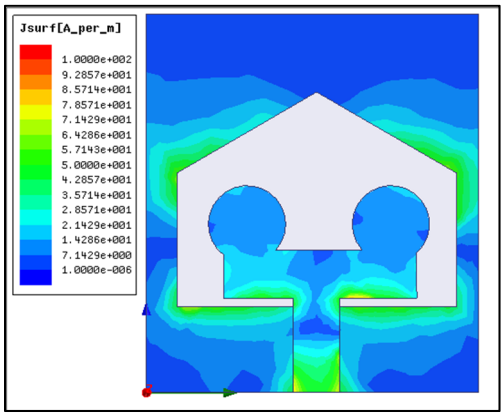


Fig. 4 Simulated Efficiency in CST and HFSS.

Despite being a small dimension, the proposed antenna can cover a full bandwidth range over 3.1-10.6 GHz that is allocated for UWB applications. The antenna peak gain is shown in Fig. 3 which indicates the maximum gain of 5.7 dBi at 10.2 GHz and under UWB and an average gain of 4.2 dBi. The observed radiation efficiency of the proposed antenna in CST and HFSS is indicated in Fig. 4. The efficiency range between lower band and the higher band is 81.84% and 97.20%. Both the efficiency and gain were affected for using low priced FR4 dielectric substrate materials. By using higher priced dielectric materials both the efficiency and gain can be improved. Fig. 5 shows the distribution of surface current of the proposed antenna for two different frequencies of 4.5 and 8.5 GHz. The observation says that a large amount of current is carried by feeding line at the point where the electric field is created. For the lower band, the current distribution is steadier comparing to upper band. Including upper and lower band of the total antenna portions, the excitation is strong enough.



(a)



(b)

Fig. 5 Surface current distribution at (a) 4.5 GHz and (b) 8.5 GHz.

The simulated phase variation of the antenna's input impedance is shown in Fig. 6. After observation, it is noticed that the phase variation across the whole operating bandwidth of the proposed antenna is almost linear. The linear phase variation regarding frequency indicates the same delay and pulse distortion for all of the frequency constituents. One of the fundamental properties of the antenna is directivity. It indicates the antenna directionality of radiation pattern. The equally radiated antenna in all directions has commendably zero directionality, which value would be 1 (or 0 dB).The variation of directivity is shown in Fig. 7 over a range of frequency. It is shown that the directivity of the proposed antenna is greater than 3 dBi around the entire frequency range. The 2D radiation pattern of the proposed antenna is shown in Fig. 8 The simulated radiation pattern of two significant plane –XZ and –YZ is represented for three different operating frequencies of 3.5, 6.5 and 9.5 GHz. From Fig. 8 it can be realized that the radiation pattern is almost omnidirectional along –YZ plane for both 3.5 and 6.5 GHz. At a higher frequency of 9.5 GHz, the radiation pattern is omnidirectional for both –XZ and –YZ plane simultaneously. The cross and co-polarization are defined by $E_{\emptyset}$ and E_{θ} respectively. The omnidirectional property of the proposed antenna increases along with the increment of frequency. The radiation patterns are significantly in stable condition throughout ultra-wideband.

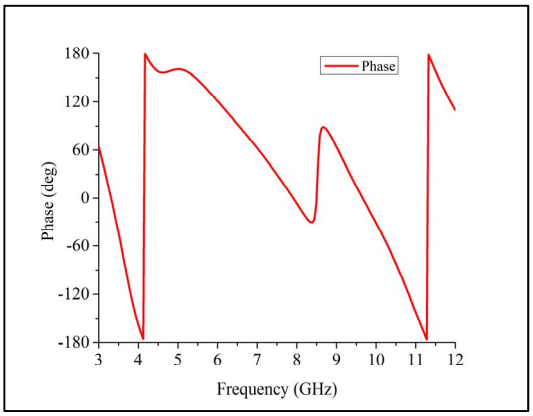


Fig. 6 Simulated phase of input impedance.

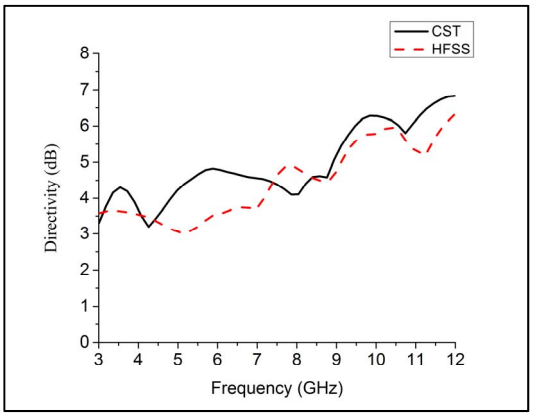


Fig. 7 Simulated Directivity of proposed antenna.

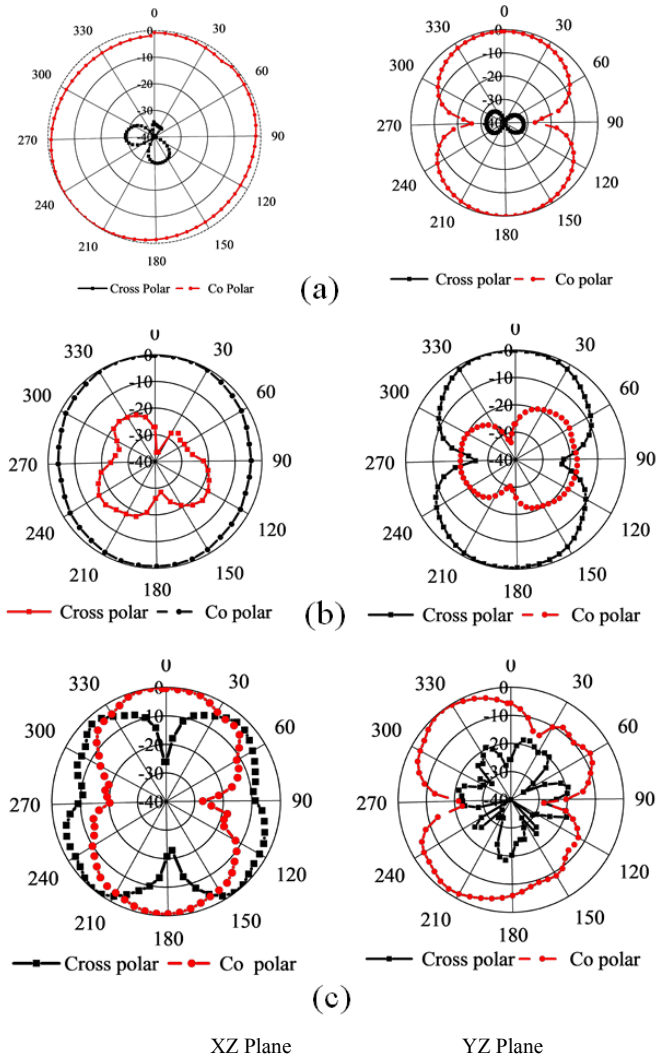


Fig. 8 Simulated 2D radiation patterns at different frequencies. (a) 3.5 GHz; (b) 6.5 GHz (c) 9.5 GHz.

IV. Conclusion

A compact better performed UWB monopole antenna with a dimension of 24 × 21 mm^2 and its measurements are presented in this paper. The proposed antenna was simulated on an inexpensive substrate with the thickness of 1.6 mm. The spectacles slot patch and tapered slot ground plane mark this antenna appropriate for attaining the impedance bandwidth of 117% (3-11.5 GHz) which is ideal for UWB applications. The proposed antenna is very simple in design, highly configured and will be easy for fabrication with a low manufacturing cost for using various purposes of UWB wireless applications.

References

[1] U. F. C. Commission, "FCC Revision of part 15 of the commission's rules regarding ultra-wideband transmission systems: First report and order," technical report, Feb2002.

[2] M. N. Shakib, M. T. Islam, and N. Misran, "Stacked patch antenna with folded patch feed for ultra-wideband application," *IET microwaves, antennas & propagation,* vol. 4, pp. 1456-1461, 2010.

[3] R. Azim, M. T. Islam, and N. Misran, "Microstrip line-fed printed planar monopole antenna for UWB applications," *Arabian Journal for Science and Engineering,* vol. 38, pp. 2415-2422, 2013.

[4] Y.-J. Ren and K. Chang, "An annual ring antenna for UWB communications," *IEEE Antennas and Wireless Propagation Letters,* vol. 5, pp. 274-276, 2006.

[5] S. K. Rajgopal and S. K. Sharma, "Investigations on ultrawideband pentagon shape microstrip slot antenna for wireless communications," *IEEE Transactions on Antennas and propagation,* vol. 57, pp. 1353-1359, 2009.

[6] J. Liang, C. C. Chiau, X. Chen, and C. G. Parini, "Study of a printed circular disc monopole antenna for UWB systems," *IEEE Transactions on antennas and propagation,* vol. 53, pp. 3500-3504, 2005.

[7] D. Li, S. Quan, and Z. Wang, "A printed ultra˜ wideband hexagon monopole antenna with WLAN band˜ notched designs for wireless communication," *Microwave and Optical Technology Letters,* vol. 51, pp. 1049-1052, 2009.

[8] M. Islam, M. T. Islam, and M. R. I. Faruque, "Design of an UWB patch antenna for dual frequency operations," *Research Journal of Applied Sciences, Engineering and Technology,* vol. 7, pp. 822-825, 2014.

[9] M. Hossain, M. R. I. Faruque, and M. T. Islam, "Design of a patch antenna for ultra wide band applications," *Microwave and Optical Technology Letters,* vol. 58, pp. 2152-2156, 2016.

[10] C. A. Balanis, *Antenna theory: analysis and design*: John Wiley & Sons, 2016.

Design and Performance Analysis of a Modified Outer Rotor Permanent Magnet Alternator for Low-Speed Wind Turbine

Md. Yeasin Arafat[1], Mahtab Murshed[2], Md. Mahamudul Hasan[3], M. Abdur Razzak[4]
Department of Electrical and Electronic Engineering
Independent University, Bangladesh
Plot-16, Block-B, Bashundhara, Dhaka-1229, Bangladesh

E-mail: nealmagik@gmail.com[1],Mahtab_murshed@hotmail.com[2], hemal.medio366@gmail.com[3], razzak@iub.edu.bd[4]

***Abstract*—In this paper, the design and analysis of a modified outer rotor PMA has been presented, which increases the efficiency and machine capacity by minimizing the cogging torque. The modified outer rotor type PMA will be suitable for extremely low-speed wind turbine based power generation system. The performances of the machine are analyzed by simulation using ANSOFT MAXWEL software, and the results are compared with the typical PMA model. All the performances are examined using finite element analysis (FEA) method to get higher accuracy in simulation stage.**

Keywords—Outer rotor PMA; cogging effect; farctional slotting; slot fill factor; finite elements analysis

I. Introduction

Recently the total annual global installed capacity of wind turbine based power generation system is more than 60000 MW,where more than 90% are installed in outside of south Asian region [1]. The main problem of wind turbine based power system in south Asian region is the low wind's speed around the year which behaves randomly [2, 3]. Wind power generation is suitable only in seashore or near seashore area because of adequate air flow [1-3], which restricts the propagation of standalone wind turbine based power system in Bangladesh [4]. Moreover,sometimeswind flow rate drastically increases in these areasespecially during the storm like cyclone, which makes most of the wind turbine system installed near the coastal areas out of working order [2] during stormy days.

Generally low-speed wind turbine blade with gear system are used to power up the high-speed PMA in low wind flow rate area, but the efficiency falls dramatically due to gear mechanism [5, 6]. When the gear system is removed to get higher efficiency of the system, direct coupling of alternator with turbine blade shaft is the only option, which requires eitherhigh-speed turbine blade or low-speed PMA. The first one is not suitable because of blade's material of the turbine [2, 3]. The second option is also quite difficult because of it generates higher cogging torque with lower machine efficiency [7], which makes the wind turbine system less suitablein low-speed wind flow areas[8].

In wind turbine, the outer rotor based permanent magnet alternator (PMA) is more popular than the typical inner rotor PMA due to its higher cross-sectional area where the outer rotor PMA can achieve higher efficiency than that of the innerrotor PMA. Also the permanent magnet pole position can be modified in outer rotor type PMAs because of available space in the rotor surface. But the outer rotor based alternator design faces problem with higher amount of cogging torque, which reduces the life time of the machine by increasing internal vibration, extra noise, and low quality output power, making PMAsunpopular [6]. That's why, in recent days, minimization of cogging effect in low-speed PMAspays great interest in research [4, 8, 9].

Lots of cogging torque minimization techniques are available now a day [8] although only one technique is not sufficient for minimizing the cogging torque. All is needed to combine a few methods by checking its characteristics and make some trial and error to find the maximum efficient point. However, combining different techniquesare not so easy to achieve. Currently available and mature techniques are teeth pairing of stator, optimal width of pole-arc, stator skewing, rotor tooth notching, PM structure and mounting modification, genetic algorithm, fractional slotting per pole, fill factor improvement etc. [4, 8-13]. In this research, only three methods are adopted for minimizing the generated cogging torque in the modified model as described in section-III. These techniques are cost effective, easy to implement and can reduce more than 95% cogging torque particularly inlow-speed PMA system [8].

II. PMA Modelling

The basic model and materials are discussed in this section. We are assuming that both models are used same outer rotor PMA.

A. Radial Flux Outer Rotor PMA

In PMA, usually radial flux permanent magnet alternator is used instead of axial flux permanent magnet alternator due to its robust and simple design and good performance characteristics at low-speed [10, 11]. In wind turbine system,air flow rate has a big impact on PMAfor the gearless system. So radial flux PMA is perfectly suitable for wind turbine systemalthough there is flexibility of using various typesof radial flux PMAs [7, 11-14]. Most of the time outer rotor PMA is the first choice over inner rotor PMA due to its higher cross-sectional area of the rotor [11]. Also the available extra space in rotor surface for magnet mounting can help to mount magnet pole perfectly with good ratio of embrace and exact magnet shape.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

Although, the typical model is available in the market but that is not suitable forlow-speed system. In general,verylow-speed PMA is not available now because of higher cogging torque generation. In thisresearch the typical model, as shown in Fig.1, is designed for 100 rpm system. The design parameters of the outer rotor PMA for both typical and modified models are listed in Table-I.

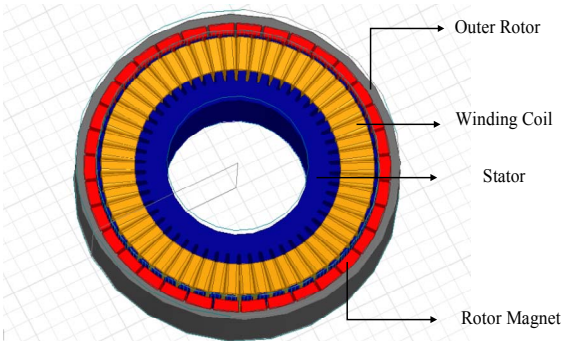


Fig. 1. Outer rotor based permanent magnet alternator

TABLE I. DESIGN PARAMETERS OF THE OUTER ROTOR PMA

Property	Typical	Modified
Rated power (Pout)	400 W	400 W
Reference speed (N)	100 rpm	100 rpm
Rated phase voltage (Va)	24 V	24 V
Thickness of PM (LPM)	7.0 mm	7.0 mm
Number of stator slots (Q)	18	54
Number of poles (P)	12	34
Length of the air-gap (g)	1.0 mm	1.0 mm
External diameter of the stator (Sout)	168 mm	168 mm
Internal diameter of the stator (Sin)	85 mm	85 mm
Length of the stator (L)	50 mm	50 mm
External diameter of the rotor (Rout)	195 mm	195 mm
Internal diameter of the rotor (Rin)	170 mm	170 mm

TABLE II. DATA TABLE FOR M19_26G SILICON STEEL SHEET

Property	Value
Material Grade	M19_26G or M350-50A
Thickness	0.50 mm
Minimum Magnetic Induction at 1 T, 50 Hz	1.45 W/kg
Minimum Lamination Factor	97%
Core loss at 1.0 T, 50 Hz	1.33 W/kg

TABLE III. MAIN PARAMETERS OF NDFEB

Property	Value
Residual flux density	1.15 T
Coercive force	915 kA/m
Maximum energy density	263.1 kJ/m3
Relative recoil permeability	1.03

B. Steel and Magnet Materials

Presently electrical steel sheet is cheaper and has low volume than any other type material is used in building the machine's body. Minimum thickness and maximum lamination factor make this material suitable for building rotor and stator. In this research, M19_26G grade electrical steel sheet is usedfor designing the rotor and stator's body. The magnetic properties of M19_26G are provided in Table II [15].

Generally rare earth magnet type NdFeB is used for PMA's rotor magnet for its higher flux density and energy density. Although, the price of NdFeB type of material is a bit high but comparing with other materialsit is the most efficient one[16, 17]. Table III shows the main parameters of NdFeB [17].

C. Impacts of Cogging Effect

Some difficulties are also arises along with these benefits in use of outer rotor PMA in such a low-speed system. Getting higher amount of output torque in a verylow-speed system can increase higher amount of magnetic interaction between stator slot and rotor magnetic pole [4, 9]. This higher interaction between these two elements can generate an extra torque which is known as cogging torque. Only use of permanent magnet with low air gap between rotor and stator can generate this torque.

Generally increase of internal vibration and noise, raise of internal temperature, low grade output power are main impacts of having high cogging torque in PMA where the lifetime of the alternator decreases significantly due to theseimpacts. The cogging torque equation stated [6]as

$$T_{cogg} = \sum_{m=1,2,3,\ldots\ \&\ i=\frac{N_s}{C}m}^{\infty} \left[T_{SSi} \left(\frac{N_s}{\cos\frac{\pi}{N_s} iP} \right) \times \sin\left(iP \times \theta - \frac{\pi}{N_s} iP \right) \right]$$

where$C = \frac{P \times N_s}{N_L}$ is the goodness factor, which is defined by the combination of magnet pole and stator slot number [18]. A small amount of changes in goodness factor minimizes the magnitude of cogging torque.

III. ADOPTING MODIFICATION

To minimize the cogging effects on the PMA some modifications, which are the combination of a number of techniques, are adopted in this research.

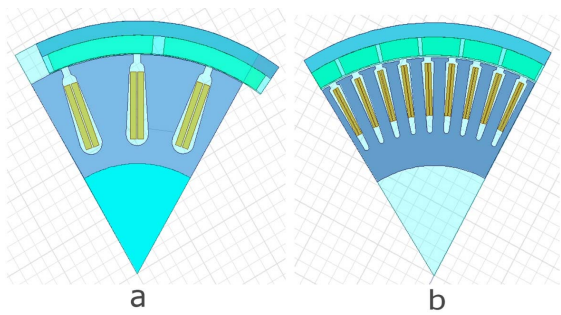


Fig. 2. Pole-slot combination model for (a) typical PMA, (b) modified PMA

A. Fractional Slot Technique

The fractional slotting per pole is applied because of minimizing the interaction between slot and pole of a PMA. The higher grade of fractional slotting means lower interaction between pole-slot hence lower cogging torque generations [6]. Moreover, fractional slot technique can decrease the harmonic

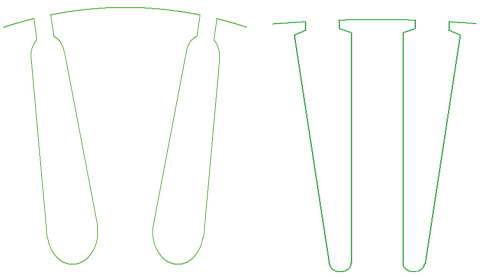
content to improve the electromotive force (EMF) significantly. After fractional slotting the modified model has higher pole-slot ratio of 1.5882 compared to the typical one which has 1.5 as shown in Fig. 2.

B. Magnet Mounting Technique

Although the same magnet mounting technique is applied to both model but this technique has several benefits which can also impacts on cogging effects. The surface mounting technique is themost useful and efficient way to mount magnet in the rotor surface [8]. Some other methods like surface embedded or vertical mounting technique is not much efficient or helpful in radial flux PMA. The surface embedded can be adopted in this particular design but the interaction between pole-slot will be much higher due to closer position of magnet surface, rotor body and stator slot.

C. Modified Slot Shape

In this proposed model the slot shape is also modified from the typical available model to increase the slot fill factor and reduce the interaction's effect of slot-pole on alternator's performance.The typical one is just follow the typical slot shape model which has lower slot fill factor capacity. This lower fill factor capacity of the typical one increases the internal temperature and maximize the ripple torque. The modified model has around 72% of slot fill factor whereas the typical one achieved around 63%. Higher slot fill factor need thicker conductor inside slot shape which reduces the winding resistance and minimizes the copper loss. This principle can improve the internal thermal condition of the rotor and stator. The stator slot shape diagrams for both models are depicted in Fig.3.



(a) (b)

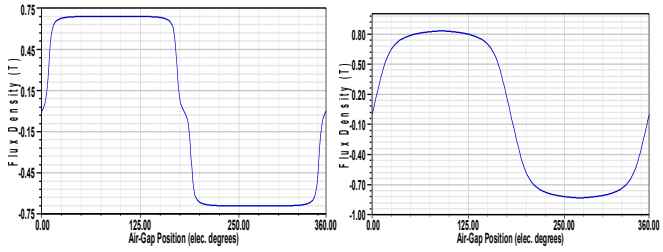
Fig. 3. Slot shape diagram (a) Typical(b) Modified

IV. Simulation Results and Comparison

The well-known machine design specially permanent magnet based machine design software called ANSOFT MAXWELL is employed to perform simulation based analysis where finite element method (FEA) is used for both models. With an interval of 0.0005 second the transient response of 0.1 second is traced.

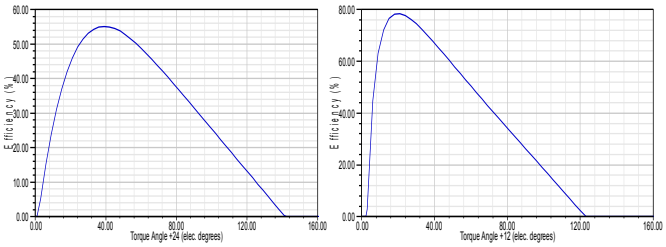
Fig.4 shows the air gap flux density where the peak of 0.69T and 0.83T are obtained for typical and modified model, respectively. It is also observed that the proposed model shape is much smoother than that of the typical one.

Fig.5 shows the efficiency plot of both the PMA models. The maximum efficiency of typical one is just over 55% whereas the efficiency traced around 50.66% at 100 rpm with a torque angle of 1.85 degrees. On the other hand the proposed model has maximum efficiency of 78% whereas the efficiency reached at 77.22% at 100 rpm with a torque angle of 13.22 degrees.



(a) (b)

Fig. 4. Air gap position vs. flux density (a) typical model (b) proposed model



(a) (b)

Fig. 5. Efficiency vs. torque angle (a) typical model (b) proposed model

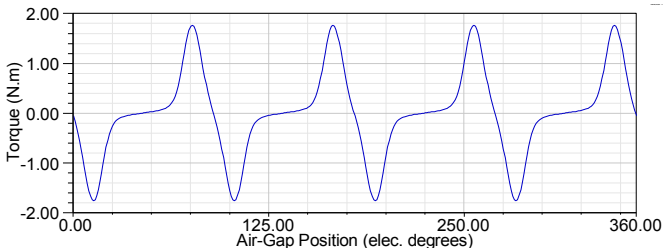


Fig. 6. Air gap position vs. cogging torque plot (typical model)

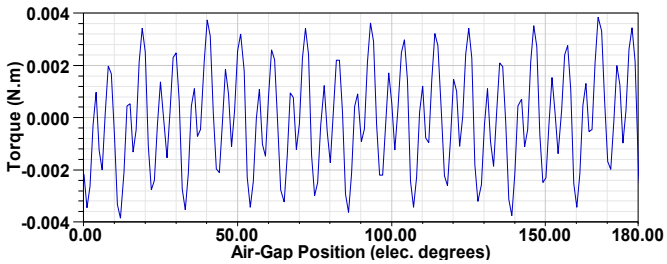
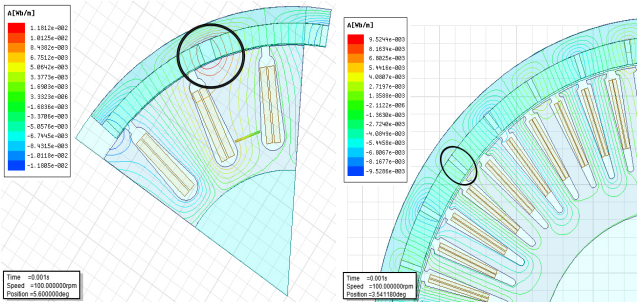


Fig. 7. Air gap position vs. cogging torque plot (proposed model)

The cogging torques for both typical and proposed models is graphically represented in Figs. 6 and 7, respectively. Typical model has maximum peak of 1.78 N-m whereas the proposed model has 0.0038 N-m.

The magnetic flux line analysis show the cogging torque generation visually which is shown in Fig.8 for both typical and proposed models. Generally, the higher magnetic flux is generated in the circled area due to higher interaction occurred. It is observed that the proposed modified model has lower magnitude of flux lines in the circled area than that of the typical one.



(a) (b)

Fig. 8. Magnetic flux lines analysis (a) Typical model (b) Modified model

The full load operational results are shown in Table IV. The armature copper loss, which is directly related to higher efficiency and lower cogging torque generation, is found to be three time lower in the modified model. Due to higher slot fill factor the total armature thermal load is also found to be three times lower than the typical model, although the maximum output power generation of typical model is higher.

TABLE IV. FULL LOAD OPERATION RESULT

Parameter	Typical Model	ProposedModel
Peak line induced voltage (V)	17.8368	27.114
RMS line current (A)	26.5118	12.6349
Armature thermal load (A^2/mm^3)	435.349	149.2
Frictional and winding loss (W)	12	12
Iron – core loss (W)	1.4831	6.06047
Armature copper loss (W)	376.351	100.013
Total loss (W)	389.834	118.074
Output power (W)	400.287	400.324
Input power (W)	790.121	518.398
Efficiency (%)	50.6615	77.2233
Synchronous speed (rpm)	100	100
Rated torque (N-m)	38.2246	38.2281
Torque angle (deg)	1.85494	13.2208
Maximum output power (W)	691.93	669.877

V. CONCLUSION

The improvement of the sudden fall off efficiency in low-speedwind turbine system by minimizing the cogging torque has been presented in this paper. Changing of slot configuration, pole mounting and fill factor improvement make the PMA more efficient by generatingless cogging torque. The thermal condition of the modified model improves due to good fill factor which make the PMA more robust and small in size. Generally this type of PMA will be suitable for low-speed application which can influencethe use of low-speed wind turbine at standalone system at low air flow areas in near future.

REFERENCES

[1] Global Wind Energy Council “Global Wind Report 2015”, April 2016

[2] Murshed, M.; Arafat, Y.; Razzak, M.A.; "Design of blades for a low-speed 400W wind turbine suitable for coastal area of Bangladesh," 3rd International Conference on Green Energy and Technology (ICGET), vol., no., pp.1-6, 11-11 Sept. 2015

[3] Murshed, M.; Arafat, Y.; Razzak, M.A.; "Analysis of Air Foils and Design of Blades for a Low-Speed 250W Horizontal Axis Wind Turbine Suitable for Coastal Areas of Bangladesh," 4th International Conference on the Developments in Renewable Energy Technology (ICDRET), vol., no., pp.1-6, 07-09 Jan. 2016

[4] Arafat, Y.; Murshed, M.; Razzak, M.A.; “Design and analysis of an outer rotor permanent magnet alternator for low-speed wind turbine," 3rd International Conference on Green Energy and Technology (ICGET), vol., no., pp.1-7, 11-11 Sept. 2015

[5] A. P. Ferreira, A. M. Silva, A. F. Costa, “Prototype of an Axial Flux Permanent Magnet Generator for Wind Energy Systems Applications”, Power Electronics and Applications, 2007 European Conference, pp. 1–9 (2007)

[6] Arafat, Y.; Murshed, M.; Hasan, M. M;Razzak, M.A.; “Impacts of Cogging Torque and Its Reduction for an External Rotor Permanent Magnet Alternator," 5th IEEE International Conference on Informatics, Electronics and Vision (ICIEV), pp.1-6, 13-14 May 2016

[7] Arafat, Y.; Murshed, M.; Razzak, M.A.; “Design and Analysis of an In-Runner Permanent Magnet Alternator for Low-Speed Wind Turbine” 4th International Conference on the Developments in Renewable Energy Technology (ICDRET), vol., no., pp.1-5, 07-09 Jan. 2016

[8] Sergeant, P.; Van den Bossche, A.P.M., "Influence of the Amount of Permanent-Magnet Material in Fractional-Slot Permanent-Magnet Synchronous Machines," IEEE Transactions on Industrial Electronics, vol.61, no.9, pp.4979-4989, Sept. 2014

[9] Wanjiku, J., et al. "Influence of Slot Openings and Tooth Profile on Cogging Torque in Axial-Flux PM Machines." IEEE Transactions on Industrial Electronics, 62.12 (2015): pp. 7578-7589

[10] Islam, M.S.; Mir, S.; Sebastian, T.; "Issues in reducing the cogging torque of mass-produced permanent magnet brushless DC motor," 38th IAS Annual Meeting. Conference Record of the Industry Applications Conference, vol.1, no., pp.393-400, 12-16 Oct. 2003

[11] Gyeong-Chan Lee; Su-Min Kang; Tae-Uk Jung, "Permanent magnet structure design of outer rotor radial flux permanent magnet generator for reduction cogging torque with design of experiment," International Conference on Electrical Machines and Systems (ICEMS), vol., no., pp.315-319, 26-29 Oct. 2013

[12] Daohan Wang; Xiuhe Wang; Dongwei Qiao; Ying Pei; Sang-Yong Jung, "Reducing Cogging Torque in Surface-Mounted Permanent-Magnet Motors by Non uniformly Distributed Teeth Method," IEEE Transactions on Magnetics, vol.47, no.9, pp.2231-2239, Sept. 2011

[13] Fei, W.; Luk, P., "A New Technique of Cogging Torque Suppression in Direct-Drive Permanent-Magnet Brushless Machines," IEEE Transactions on Industry Applications, vol.46, no.4, pp.1332-1340, July-Aug. 2010

[14] Arafat, Y.; Murshed, M.; Hasan, M. M; Razzak, M.A.; “Design Aspects and Performance Analysis of Inner and Outer Rotor Permanent Magnet Alternator for Direct Driven Low-Speed Wind Turbine," 2nd IEEE International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), pp.1-6, 27-28 Feb. 2016

[15] Typical data for M19_26G Silicon Steel Sheet, http://www.aperam.com/brazil/ing/arquivos/atalogelectricalsteeleng.pdf, [26 October 2015]

[16] Neumann T.W., Tompkins R.E., “Line start motor designed with NdFeB permanent magnet”, 8th International Workshop on Rare Earth Magnets and their Application, Dayton, Ohio, pp. 77 – 89 (1985)

[17] Ting Liu; Shoudao Huang; Qiuling Deng; Qingyun Pu; Keyuan Huang, "Effect of the number of slots per pole on performance of permanent magnet generator direct-driven by wind turbine," International Conference on Electrical Machines and Systems (ICEMS), vol., no., pp.1-4, 20-23 Aug. 2011

[18] Zhu, Z. Q., S. Ruangsinchaiwanich, and David Howe. "Synthesis of cogging-torque waveform from analysis of a single stator slot." IEEE Transactions on Industry Applications, 42.3 (2006): pp.650-657.

A Modified PWM Technique to Improve Total Harmonic Distortion of Multilevel Inverter

Shuvankar Podder,[1,*] Md Multan Biswas[2] and Md. Ziaur Rahman Khan[1]

[1]Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology
Dhaka-1205, Bangladesh

[2]Department of Electrical Engineering, University of South Carolina
Columbia, SC 29208, USA

Email: *podder.shuvankar@gmail.com

Abstract—**In this study, a modified carrier is proposed to generate PWM gating signal. Proposed PWM method reduces total harmonic distortion (THD) of the output voltage of inverter up to 5.8% than the traditional triangular carrier based PWM. Major harmonic components in output voltage are less in magnitude when proposed PWM technique is applied to inverter. Fundamental component of output voltage is also higher in proposed PWM method for all modulation index and frequency ratio. Performance of proposed modified carrier based PWM is evaluated by MATLAB simulation for single phase 4-cells 9-levels cascaded H-bridge MLI.**

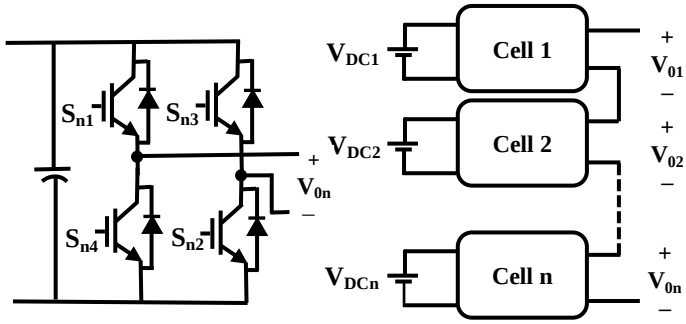
***Index Terms*— PWM, Multilevel inverter, THD.**

I. Introduction

Multilevel Inverter (MLI) is popular in medium and high voltage application due to its low rating of switch, low dv/dt, low rating of output filter. Typical structures of MLI are diode clamped MLI, capacitor clamp MLI and cascaded H-bridge MLI (CHBMLI) [1]. Most of the applications of CHBMLI involve reactive power compensation, PV power conversion, UPS, MRI etc. Due to more switches, the control of MLI is sophisticated. Usually carrier based PWM and space vector PWM techniques (SVPWM) are used to control MLI.

For CHBMLI, multi-carrier based sine- PWM (SPWM) is very popular. In SPWM, a set of triangular carrier signals are compared with a reference signal. For m-level inverter, (m-1) carrier signals are needed. Multicarrier PWM technique can be categorized broadly in Level- shifted PWM (LS PWM) and Phase-shifted PWM (PS PWM) depending on the arrangement of carrier signals. In conventional multicarrier LS-PWM, a set of triangular carriers with same amplitude A_C and frequency f_C are shifted along vertically. These set of carriers are disposed in such a way so that the band they occupy are contiguous. The reference waveform is centred in the middle of the carrier set when over-modulation is not considered. According to different methods of disposing the carriers set, LS-PWM is of three types. These are In Phase Disposition (IPD), Alternative Phase Opposition Disposition (AOPD) and Opposite Phase Disposition (OPD). In PS PWM all the triangular carriers have the same frequency and amplitude but carriers are phase shifted [2].

Carrier based MLI are discussed more in [2-4]. In studies [5-6] SVPWM for MLI are discussed. Babaei et el [7-8] introduces a novel MLI where look-up table is used to control MLI by FPGA. In this study, a new carrier based PWM technique is proposed for CHBMLI. The proposed carrier based PWM reduces total harmonic distortion (THD) and major harmonic components of the output voltage significantly. MATLAB simulation is used to investigate the performance of proposed modified carrier based PWM.



(a) (b)

Fig. 1 (a) H-bridge structure in a single cell. (b) General topology of n-cells CHBMLI

II. Cascaded H-bridge Multilevel Inverter

Figure 1(a) shows the H-bridge structure. Multiple H-bridges are connected in series in CHBMLI. Figure 1(b) shows the general topology of n-cell CHBMLI. The output voltage of the n-cell MLI is given by

$$V_0(t) = V_{DC1}+V_{DC2}+V_{DC3}+\ldots\ldots\ldots + V_{DCn} \quad (1)$$

If all the DC voltage sources in CHBMLI is V_{DC}, then by controlling the switching devices of H-bridge cell, the output voltage obtained are $+nV_{DC}$, $+(n-1)V_{DC}$, $+(n-2)V_{DC}$,…… $+2V_{DC}$, $+V_{DC}$, 0, $-V_{DC}$, $-2V_{DC}$, ….., $-(n-1)V_{DC}$, $-nV_{DC}$. So for n-cell CHBMLI, maximum m-level of output voltage can be obtained where

$$m = 2n+1 \quad (2)$$

In this study, 4 cells 9 levels CHBMLI is considered to compare the performance of proposed PWM technique with the conventional technique.

III. Traditional LS-PWM

Traditional triangular carriers based In Phase Disposed Level Shifted PWM (IPD LS-PWM) is shown in figure 2. For 4 cells 9 levels CHBMLI, 8 triangular carriers are compared with the sinusoidal reference signal to generate sine PWM signals. When the reference signal is greater than the carrier signal then the active device corresponding to that carrier signal is switched on. In m-levels inverter, for constant frequency carrier based LS-PWM, the amplitude modulation M_a and frequency ratio M_f are defined as

$$M_a = \frac{A_m}{(m-1)A_C} \quad (3)$$

$$M_f = \frac{f_C}{f_r} \quad (4)$$

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

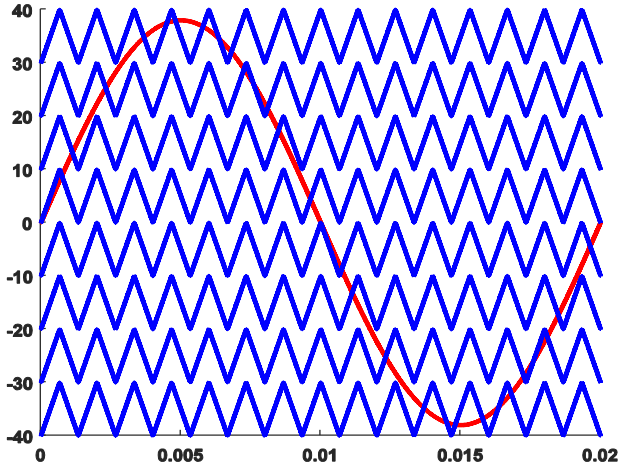


Fig 2. Sinusoidal reference signal and triangular carrier set (carriers are in IPD mode) for 9-levels CHBMLI (M_f=15 and M_a=0.9).

Where, A_m, A_C are the amplitude of reference signal and carrier signal respectively and f_m, f_C are the frequency of reference signal and carrier signal respectively.

IV. Proposed PWM Technique

In CHBMLI, the DC voltages of individual cells are added up to yield a desired output voltage. The objective of MLI is to generate a output voltage that is similar to sine wave. The more the output voltage is similar to sine wave, the less will be the calculated THD. Switching of the H-bridge gives a time sampled DC voltage at the output of each cell. For 4-cells MLI, the output voltage $V_0(t)$ can be written as

$$V_0(t)=[V_{DC1}(t)T_1 + V_{DC2}(t)\,T_2 + V_{DC3}(t)T_3+V_{DC4}(t)T_4]\,/T_{SW} \quad (5)$$

Where, T_{SW} is the considered time, V_{DC1} (t)-V_{DC4} (t) are DC input voltage of individual cell and T_1-T_4 are output time of four cells within that considered time. In this study input voltage is V_{DC} for each cell. These $V_{DCi}\,T_i$ terms are called volt-sec output.

In conventional triangular carrier based SPWM, for a particular carrier frequency, the volt-sec output from a H-bridge cell is fixed. If the duration of this volt-sec output can be made smaller, then the output of CHBMLI becomes closer to the target AC voltage. For conventional triangular PWM, the volt-sec can be made smaller by increasing carrier frequency.

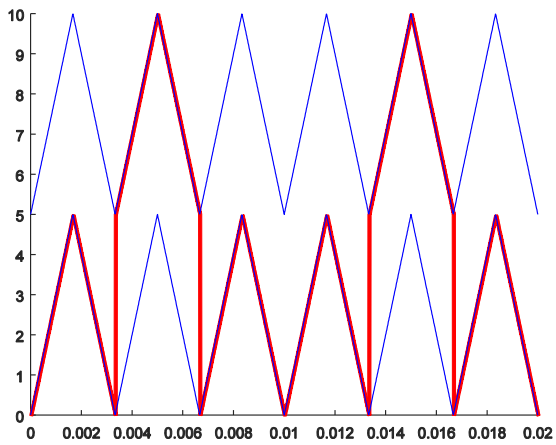


Fig. 3 Proposed carrier signal (red colour). Two sets of level shifted triangular carrier (blue colour) are used to generate proposed carrier signal. Two cycles of proposed modified carrier is displayed.

But it will be seen in next section that, increasing M_f does not improve THD after a certain level. To solve the problem a modification in carrier is proposed.

The original triangular carrier signal of amplitude A_C is divided into three equal parts. The modified carrier is made symmetrical to minimize the harmonic [9]. Figure 3 shows the proposed carrier signal.

Figure 4 illustrates how proposed modified carriers improve output voltage. In T_{C1}, desired output voltage is sinusoidal and increasing (green colour). In the time T_{C1}, average output voltage is zero in figure 4(b) and this is a finite value in figure 4(a). In figure 4(a), as $+V_{DC}$ volt-sec output portion exists in the output voltage during T_{C1}, the average output is finite here. So, in time T_{C1}, the waveform deviation of output voltage obtained from traditional PWM is larger from the output voltage obtained from the proposed PWM. In time T_{C2}, output volt-sec is $+2V_{DC}$ one time and the rest is $+V_{DC}$ in figure 4(b) while output volt-sec is $+2V_{DC}$ two times in figure 4(a). So, similarly, it can be shown that the waveform deviation of output voltage obtained from proposed PWM is less than that obtained from traditional PWM. In time T_{C3}, although the output voltage is almost similar in figure 4(a) and 4(b), the output voltage from proposed PWM becomes more close to desired AC voltage for higher carrier frequency. Thus the proposed PWM generates less deviated output voltage from the target AC sinusoidal voltage and therefore reduce THD.

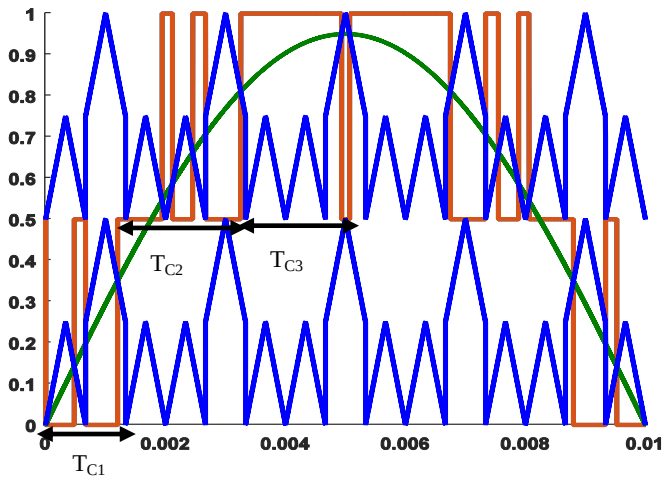


Fig. 4 (a) Sinusoidal reference (green colour), proposed carrier set (blue colour) and output voltage (red colour) for two-cells CHBMLI (M_f=10 and M_a=0.9).

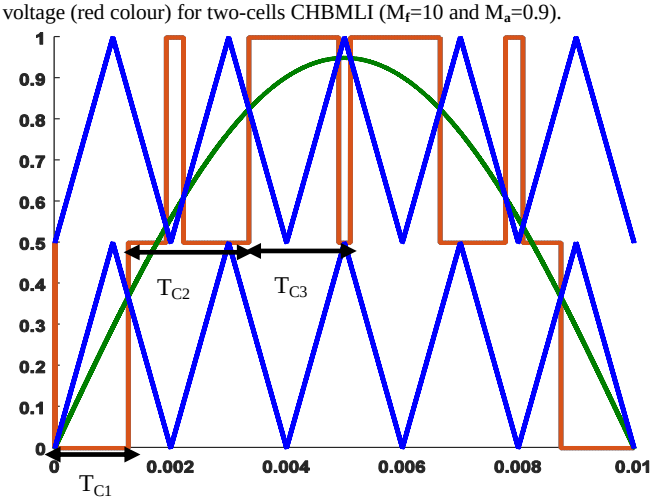


Fig. 4 (b) Sinusoidal reference (green colour), conventional triangular carrier set (blue colour) and output voltage (red colour) for two-cells CHBMLI (M_f=10 and M_a=0.9)

V. Simulation Result

Performance of proposed multicarrier PWM technique is compared with the conventional triangular carrier PWM technique for a 4-cells 9 levels single phase CHBMLI. Performance is compared in terms of

(1) Output voltage waveform
(2)THD and
(3) Harmonic spectrum.

Figure 5(a) and 5(b) show the output voltage comparison for conventional and proposed PWM. IPD LSPWM is considered here. In both cases, amplitude modulation M_a = 0.9 and frequency ratio M_f = 21. In figure 5(a), at proposed PWM, the volt-sec output at a particular time is divided into smaller part than the volt-sec output in conventional PWM as shown in figure 5(b). Smaller volt-sec output is desired to reduce THD.

Total harmonic distortion is the measurement of harmonics present in the particular signal. It is defined as the ratio of sum of powers of all harmonic components to the power of the fundamental frequency. THD is expressed as

$$\mathrm{THD} = \sqrt{\frac{V_2^2 + V_3^2 + V_4^2 + \ldots\ldots + V_n^2}{V_1^2}} \tag{6}$$

Here V_n is the root mean square (RMS) magnitude of the nth order harmonic component. Power contained in a harmonic component is proportional to the square of the RMS magnitude of that component.

Figure 6(a) & 6(b) show the THD comparison for proposed and conventional carrier based PWM where carriers set are in IPD mode. Here two amplitude modulation are considered; M_a = 0.85 and M_a = 1. For these two M_a , frequency modulation M_f is varied from 3 to 51 with triple harmonic. From figure 6, it is evident that proposed modified carrier based PWM gives better THD performance than conventional PWM for all M_f . THD of output voltage in both PWM techniques is lowest around M_f = 9 in both cases of M_a = 0.85 and M_a = 1. It is found from the figures that at high M_f, THD does not vary noticeably which is mentioned previously. THD can be reduced by increasing M_a; though the maximum limit of M_a is 1. In MLI, THD can be reduced significantly by increasing the number of level of output voltage. But increasing the number of level requires more complex control of inverter.

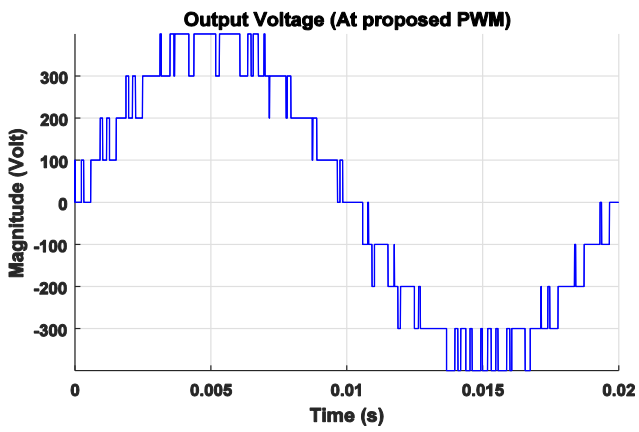


Fig. 5(a) Output voltage of 4-cells 9-levels CHBMLI at proposed PWM (Mf=21 and Ma =0.9)

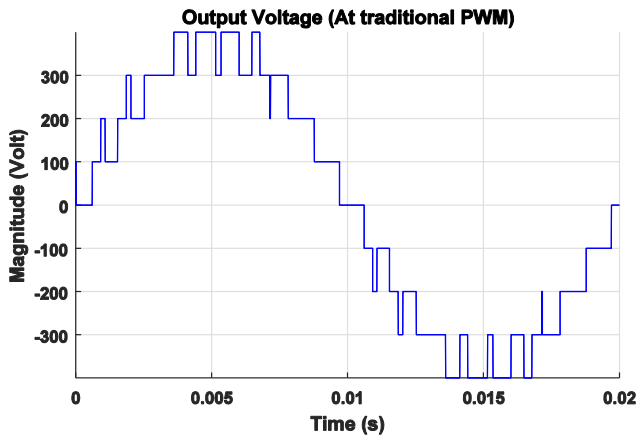


Fig. 5(b) Output voltage of 4-cells 9-levels CHBMLI at traditional triangular carrier based PWM (Mf=21 and Ma =0.9)

So there exists a trade-off between the number of levels and desired THD of output voltage. IEEE standard for the THD of output current from inverter is less than 5%. To meet this requirement, usually LCL second order filter is used for grid connected inverter. In [10-11], design of LCL filter for grid connected inverter is discussed in details.

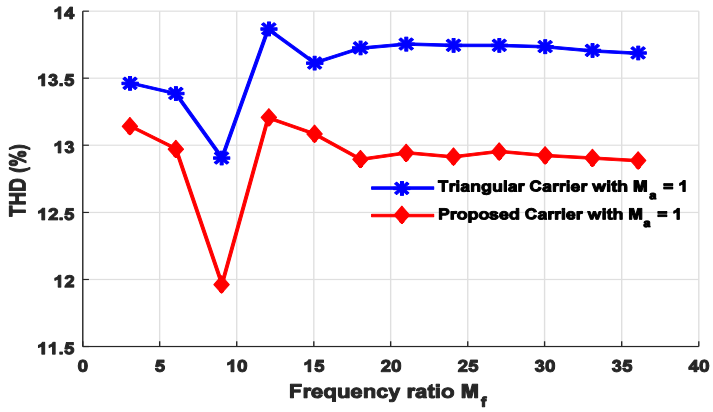


Fig. 6(a) Comparison of THD of output voltage from a 4-cells CHBMLI at traditional PWM (blue colour) and proposed PWM (red colour) with varying Mf (Ma =1)

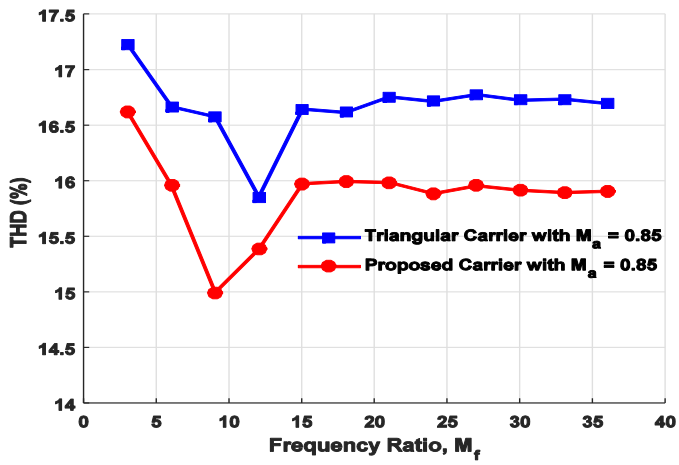


Fig. 6(b) Comparison of THD of output voltage from a 4-cells CHBMLI at traditional PWM (blue colour) and proposed PWM (red colour) with varying Mf (Ma =0.85)

Figure 7(a) and 7(b) show the harmonic spectra of output voltage for proposed and conventional PWM techniques. For both case M_a = 0.95 and M_f = 21. Input voltage of each cell is considered 100 V. Major harmonic components of output voltage obtained from traditional triangular carrier based modulation and proposed modified carrier based modulation is shown in table 1. From the table, it is evident that at proposed modified carrier based PWM; all major harmonics are less in magnitude. Again magnitude of fundamental component is 1.4% higher in proposed technique. Harmonic component at carrier frequency has the maximum amplitude. So position of largest harmonic can be made further from fundamental component by increasing the carrier frequency. Here single phase CHBMLI is considered. In case of three phase inverter, harmonic contents in output voltage become further less in magnitude. Magnitude of the fundamental component can be increased if over-modulation is adapted [12-13]. In three phase inverter, fundamental component can be increased up to 15% without over-modulation by adding a third harmonic voltage to the reference sinusoidal signal [13]. Reduction of magnitude of all major harmonic components can lead to reduce the rating of output LCL filter up to 10% [10].

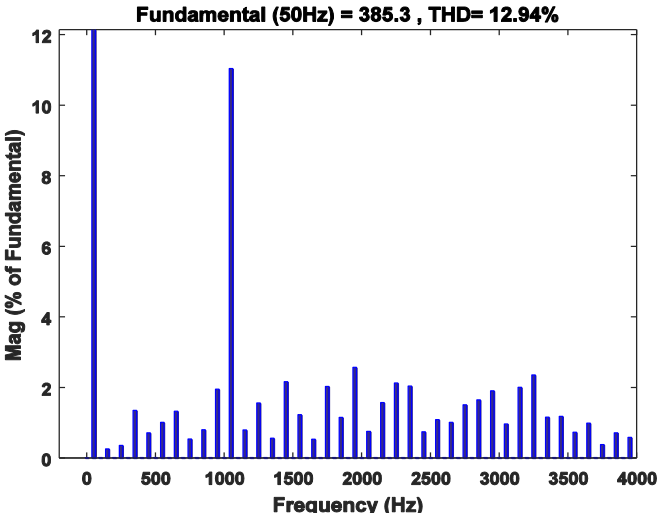


Fig.7(a) Harmonic spectra of output voltage of 4-cells CHBMLI at proposed PWM (Mf=21, Ma=0.95, magnitude of fundamental=385.3V THD=12.94%)

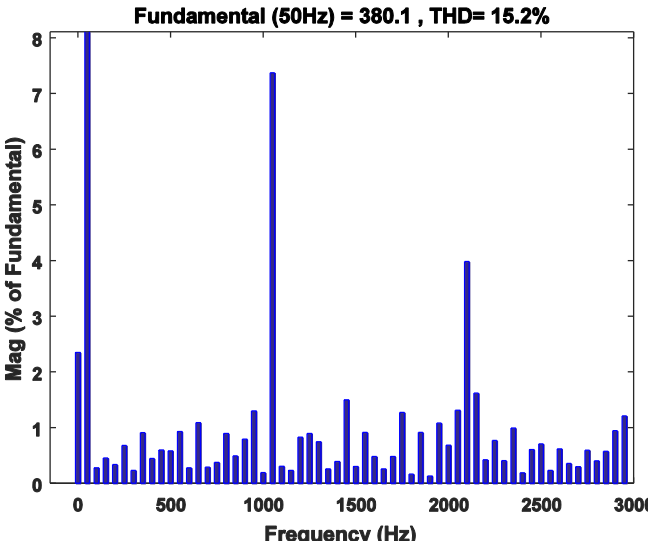


Fig. 7(b) Harmonic spectra of output voltage of 4-cells CHBMLI at triangular carrier based PWM (Mf=21, Ma=0.95, magnitude of fundamental=380.1V THD=15.2%)

TABLE I
COMPARISON OF MAJOR HARMONIC COMPONENTS

Harmonics	Traditional	Proposed
Fundamental	380.1	385.3
3rd (% of Fundamental)	0.452	0.295
5th (% of Fundamental)	0.767	0.400
7th (% of Fundamental)	1.289	1.381
9th (% of Fundamental)	0.75	0.68

VI. CONCLUSIONS

A modified carrier signal is proposed to generate gating signal for CHBMLI. The proposed modification in carrier signal results in less THD in output voltage than conventional method. Reduction in THD makes output voltage near sinusoidal and so the size of output filter reduces. One limitation of proposed PWM technique is that this method increases the switching frequency. The number of switching can be reduced by varying the frequency of the carriers in different level of LS PWM carrier arrangement instead of using constant frequency throughout the carrier band. But for constant frequency carrier set, this proposed modulation technique provides superior performance in harmonic elimination task which is the prime concern in this study.

REFERENCES

[1] Rodriguez, J., Lai, J.-S., Zheng Peng, F.: 'Multilevel inverters; a survey of topologies, controls, and applications', *IEEE Trans. Ind. Electronics*. vol.49, no. 4, pp.724-738, November 2002.

[2] L. M. Tolbert, and T. G. Habetler, "Novel multilevel inverter carrier-based PWM method," *IEEE Trans. Ind. Application.*, vol. 35, no. 5, pp.1098–1107,September/October 1999.

[3] B. P. McGrath, D. G. Holmes, "Multicarrier PWM strategies for multilevel inverters," *IEEE Trans. Ind. Electronics.*, vol. 49, no. 4, pp.858-867, August 2002.

[4] P. Palanivel S.S. Dash, "Analysis of THD and output voltage performance for cascaded multilevel inverter using carrier pulse width modulation technique," *IET Power Electron.*, vol. 4, Iss. 8, pp. 951–958, 2011.

[5] D. Zhao, V.S.S.P.K. Hari, G. Narayanan, and R. Ayyanar," Space-Vector-Based hybrid pulsewidth modulation techniques for reduced harmonic distortion and switching loss," *IEEE Tran.on Power Electronics*, vol. 25, no. 3, pp. 760-774, March 2010.

[6] G. Narayanan, D. Zhao, H. K. Krishnamurthy, R. Ayyanar, and V. T. Ranganathan," Space vector based hybrid PWM techniques for reduced current ripple," *IEEE Trans .on Industrial Electronics*, Vol. 55, no. 4, pp. 1614-1627, April 2008.

[7] E. Babaei, S. Alilu , S. L. A. Karnik, "A New General Topology for Cascaded Multilevel Inverters With Reduced Number of Components Based on Developed H-Bridge," *IEEE Trans. on Industrial Electronics*, vol.61 , no. 8, pp. 3932-3939, October, 2013.

[8] E. Babaei, "A cascade multilevel converter topology with reduced number of switches," *IEEE Trans. Power Electron*, vol. 23, no. 6, pp. 2657-2664, November 2008

[9] Holmes D.G., Lipo T.A. (2003). "Pulse Width Modulation for Power Converters Principles and Practice." IEEE Series on Power Engineering, John Wiley & Sons, Inc.

[10] Renzhong X.; Lie X.; Junjun Z.; Jie D., "Design and research on the LCL Filter in three-phase PV grid-connected inverters," *International Journal of Computer and Electrical Engineering*, vol. 5, No. 3, June 2013

[11] F Liu, X Zha, Y Zhou, S Duan, "Design and research on parameter of LCL filter in three-phase grid-connected inverter" *Power Electronics and Motion Control Conference*, 2009. IPEMC '09. IEEE 6th International

[12] Ramón Cáceres, Ivo Barbi, "A boost DC–AC converter: analysis, design, and experimentation" *IEEE Trans. Power Electron*, vol.14, no. 1, pp. 134-141,January 1999.

[13] A. Elserougi, A.S. Abdel-Khalik, A. Massoud, S. Ahmed, "Studying the effect of over-modulation on the output voltage of three-phase single-stage grid-connected boost inverter" *Alexandria Engineering Journal*, vol. 52, no. 3, pp. 347–358, September 2013.

Study of Blackout Prevention of a Power System

Jubair Yusuf[#1], Fuad Sakib[*2], Q. Ahsan[#3]

[#]Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology
Dhaka 1000,Bangladesh
[1]jyusuf177@gmail.com,[2]fuadsakib@gmail.com
[3]qahsan@eee.buet.ac.bd

***Abstract*— The blackout of a power system creates severe inconvenience to its consumers and manyfold losses to the utility. Therefore, the prevention of the blackout is a serious concern of the utility. Auto and adaptive load shedding and islanding are common operational techniques to keep the system stable at a severe perturbation due to faults or loss of generation or switching at bulk loads. This paper investigates the effectiveness and conditions of the use of different blackout prevention techniques with a view to help engineers develop operational strategies of a power system.**

***Keywords*— Load-shedding, islanding, rate of change of frequency, critical clearing time, blackout.**

I. Introduction

In a power system, a large-scale blackout rarely occurs and if it does, it causes enormous economical and social damages, including the security of fragile people (old persons, hospitals, etc.). For example, large scale blackouts happened in the Northeast of the United States and in the Canada on August 14, 2003, and in Italy on September 28, 2003. The first one affected approximately 50 million people of eight U.S. states and two Canadian provinces. The economic losses were about $7–10 billion [1,2]. The second one affected about 57 millions people in Italy and 180 GW of load was interrupted [3–5]. This blackout in Italy is the largest in Europe.

During abnormal conditions of the system ,the frequency fall and the rate of change of frequency (ROCOF) are major concerns. To keep the change of frequency within an allowable range the load-shedding is an established practice. The objective of load shedding is to balance load and generation. Since the amount of overload is not known at the instant of disturbance, the load shedding continues till the frequency stabilizes.

Sometimes ,the blackout can be prevented in real time through controlled disintegration of a system into a number of islands together with generation tripping and or load shedding [6-9].The Disintegration of a grid into a number of islands is, usually, considered as a last resort [8].The basis for islanding is never unique and depends upon the utility in particular.

A large number of papers on the application of different load shedding techniques and islanding approaches for the stability and prevention of blackout of a power system is reported in the literature [6],[8],[10-12], particularly in the last two decades.

This paper presents the investigation results of the application strategies of different blackout prevention techniques. The simulation uses a ten bus small but representative system. The investigation shows that the instability scenarios of a power system are different even for the same type of fault at different buses. It also reveals that the application of a particular technique with a particular level or scheme depends on the magnitude of frequency deviation and ROCOF to retain the stability of a system under abnormal condition. The study results recommend that the strategic planning for blackout prevention should be different for off-line and on-line applications.

In this paper, section I presents an introduction along with literature review whereas section II describes the blackout prevention technology along with the flow chart. The system under study and simulation results are given in Section III and the concluding remark is presented in section IV.

II. Blackout Prevention Technique

Abnormal condition in a power system generally leads to a fall in system frequency. The usual solution to rescue the system from this sort of state is to clear the fault within it's critical fault clearing time .If the fault can't be cleared within it's critical fault clearing time , then load shedding is effective to make the system stable. However, in some cases the load shedding is not required as the system makes itself stable by providing additional input from its stored kinetic energy or from the spinning reserve or by lowering the system frequency within acceptable limit. In some cases, the magnitude of load shedding may be inappropriate, that is more or less than required, to make the system stable by maintaining the system frequency within acceptable range. In some cases, only load shedding cannot rescue the system from total collapse. In that case, the faulty part of the system needs to be disintegrated , creating two or more islands when islanding alone is not enough to make the system of an island stable, islanding with load or generation shedding, depending on the load-generation balance, provides an effective solution. Islanding within an island may also be required. However, that is an exceptional case.

The technique used in this paper considers all these issues and develops a comprehensive solution to the instability of a power system. The technique is based on three schemes: 1) traditional load shedding, 2) ROCOF based load shedding, and 3) disintegration of the grid. The traditional load shedding scheme is implemented through UF (under frequency) relays. The scheme is activated only when the system frequency, falls below a certain threshold value f_{th}. The implementation of other schemes requires FD (frequency droop) relays.

A. Sequential Steps

The technique is implemented in steps, sequentially. Different steps of the technique are as follows.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

Step 1 Both frequency and the rate of change of frequency of the power system are continuously measured .
Step 2 When frequency is below a predetermined threshold value f_{th1} , then it is assumed that an abnormal condition has occurred.
Step 3 If $f<f_{th1}$ and the ROCOF is lower than a predetermined value f_{th2} , then only load shedding is enough to make the system stable. But if the ROCOF is greater, then islanding is the solution.
Step 4 After islanding , the frequency and ROCOF of each island are continuously measured to check the system stability.

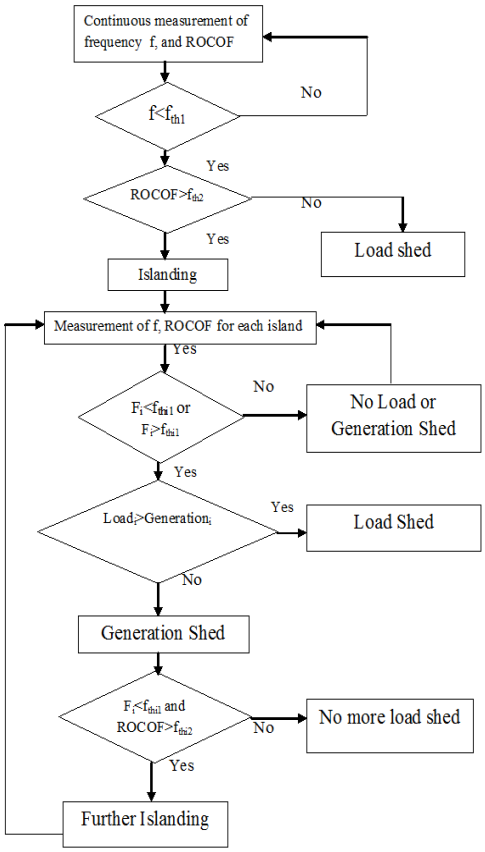


Fig. 1 Flow Chart of the Used Technique

Step 5 If the frequency of the ith island is lower or greater than f_{thi1} and ROCOF is less than a predetermined value f_{thi2} , then the technique requires the verification of load-generation balance, that is whether the island is load rich or generation rich. If the island is load rich , then the load is shedded and if the island is generation rich, then the generation is shedded.
Step 6 After that if $f_i < f_{thi1}$ and $ROCOF>f_{thi2}$ then further islanding is effective to save the system from collapse.

III. Simulation Results

A. System Description

The proposed blackout prevention technique is applied to a 10 bus power system. Out of 10 buses,3 are generator buses. In the system all the transformers are fixed tap transformers and it has a total generation of 49.87 MW. The single line diagram of the power system is shown in figure 2

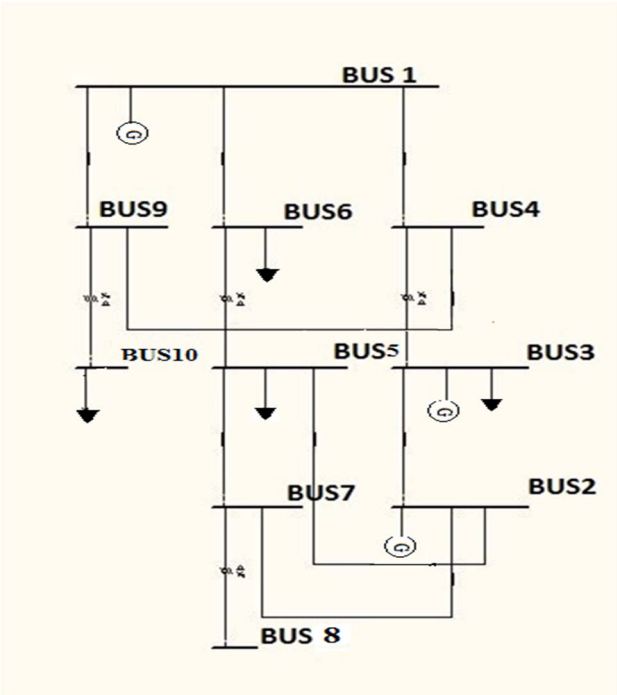


Fig. 2 Single Line Diagram of the ten Bus System

B. Stability Indices

Initially, the response to fault at different buses in terms of stability indices are evaluated. Single Line to Ground fault is considered at different buses and critical fault clearing time, frequency deviation and ROCOF are evaluated.

TABLE I
Critical Clearing Time for Faults at Different Buses

Fault Location (Bus Number)	Critical Clearing Time (second)
1	0.068
2	0.028
3	0.048
4	0.010
5	9.48
6	0.248
10	0.428

Critical clearing times for faults at different buses are presented in Table I. The table clearly shows that although the fault type is same but the fault critical clearing time at different buses is different. That is, the system response towards stability is significantly different for even the same type of fault at different buses.

The frequency falls of the system for a fault at bus 1 with different fault clearing time, starting from 5 cycles to 25 cycles at an interval of 5 cycles, are depicted in figure 3. Note that the critical fault clearing time of bus 1 is 3.4 cycles. The figure shows that the frequency deviation increases with the increase of fault clearing time, as expected.

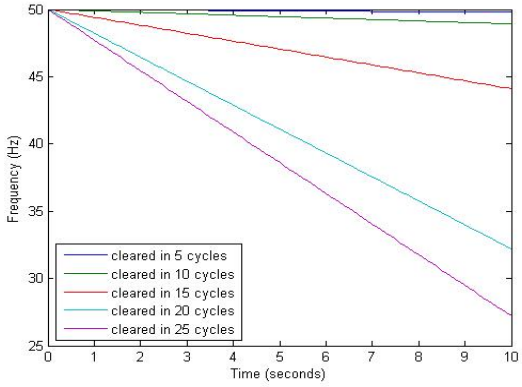


Fig. 3 Frequency Fall for a Single Line to Ground Fault at Bus 1 with Different Fault Clearing Time.

TABLE II
Comparison of System Frequencies and ROCOFs of Different Fault Locations for Different Fault Clearing Time

Fault Conditions		Frequency		ROCOF (Hz/sec)
Location	Fault Clearing Time (Cycles)	After 1 Sec	After 2 Sec	
Bus 1	5	49.98	49.96	0.021
	10	49.89	49.79	0.107
	15	49.41	48.82	0.59
	20	48.22	46.44	1.78
	25	47.72	45.44	2.28
Bus 2	5	49.998	49.996	0.002
	10	49.992	49.98	0.008
	15	49.98	49.95	0.025
	20	49.94	49.88	0.058
	25	49.88	49.77	0.117

The system frequencies and ROCOFs are compared of two different fault locations, buses 1 and 2 in Table II. It is observed from Table II the ROCOF and frequency deviation are higher for the fault location of higher critical clearing time.

C. Blackout Prevention Strategies

1) Auto-Load Shedding Scheme

If a fault occurs in a power system and it is cleared within critical clearing time , note that to clear the fault within the critical clearing time a circuit breaker with appropriate speed should be selected so that the fault is cleared within the time.

However, it is observed that if faults are cleared after critical clearing time, then the system becomes unstable. To regain the stability, load shedding is needed. Fig. 4 depicts the system frequency variation for faults without clearing the fault within critical clearing time at buses 1 and 2 and load shedding of 31% and 7% respectively.

In both the cases, the faults are cleared at 10 cycle, the figure shows that the frequency drops about 1.8% within 0.8 second with a ROCOF of 0.1. But with this implementation of auto load shedding scheme the system becomes almost stable within 2.5 second. It also shows that the frequency responses are different for the faults at two different locations, as expected.

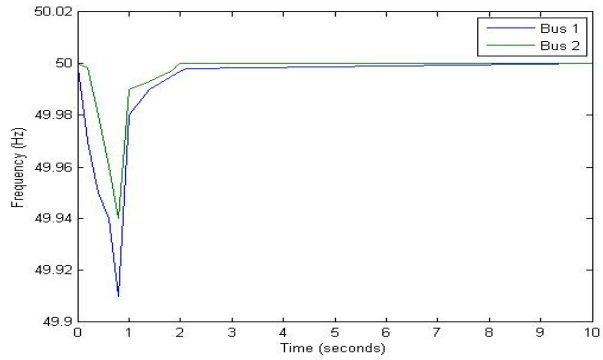


Fig. 4 Blackout Prevention with Auto-Load Shedding Scheme

2) Islanding

Under certain conditions, only load shedding is not enough to make the system stable. The fault at bus 4 is an example of it. In this case the proposed technique recommends islanding.

Scenarios of islanding are exemplified in two different cases.

Case I : A fault at bus 1 is considered and it is also considered that the fault is not cleared up to 20 cycles when the system ROCOF is 1.78 second. It is observed that the system becomes unstable even with 80% load shedding. Then the two islands are formed as shown in fig. 5. In this case no load shedding is implemented. The simulation result shows that the system remains unstable. However, it is observed that both the islands become stable if a total load shed 47.5% is implemented.

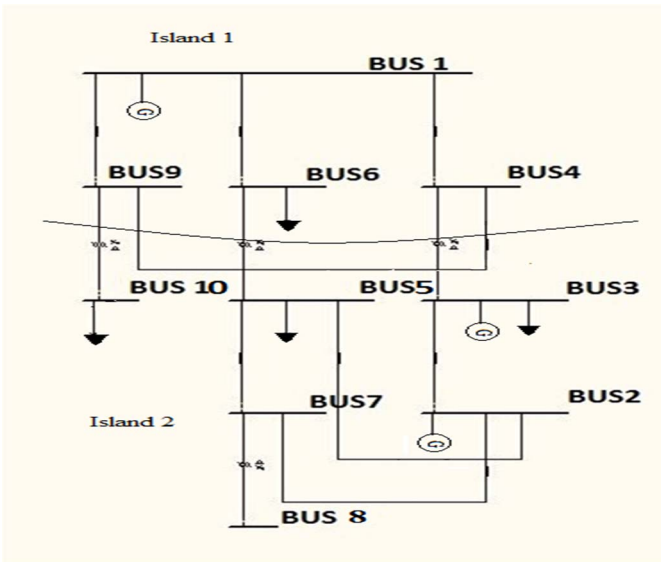


Fig. 5 System is Disintegrated to Form Two Islands

Case II : If a fault at bus 4 is simulated and it is not cleared up to 0.5 cycle, then the system becomes unstable. With the islanding of two as in case I along with the load shedding the system remains unstable. If the number of islands is increased to three as shown in fig. 6, the system is not recovered from the state of instability. However, when four islands are formed as shown in fig. 7, only one island, island 3 becomes stable. Other three islands can not be made stable even with the maximum amount of load shedding.

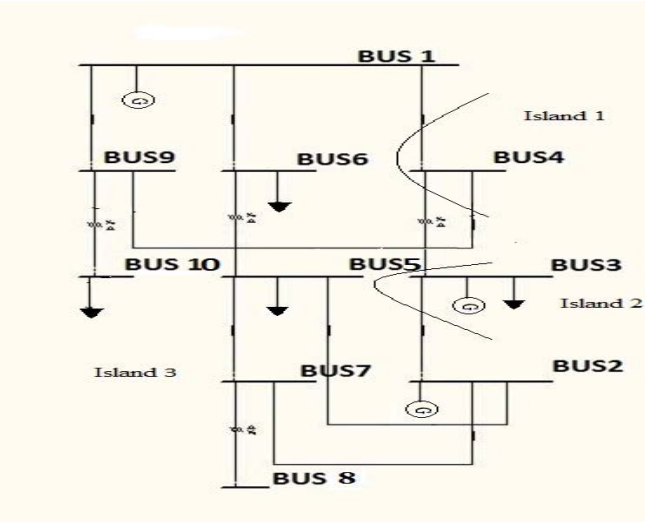


Fig. 6 Three Islands to Recover The System from The State of Instability for a Fault at Bus 4

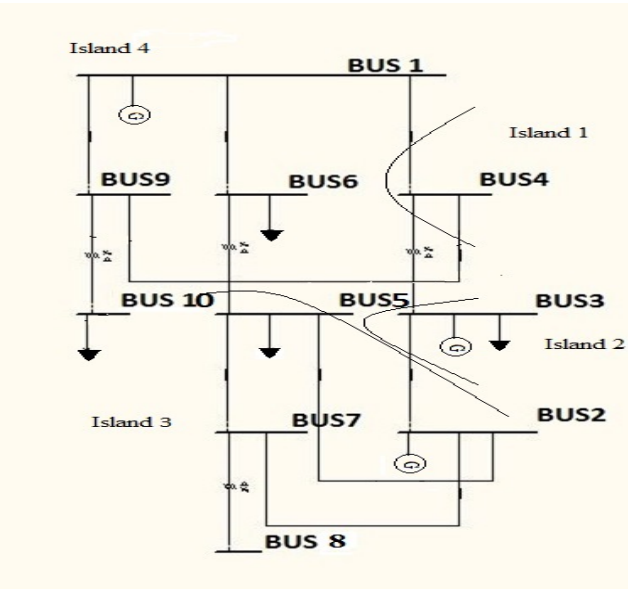


Fig. 7 Four Islands out of which Only Island 3 is Stable.

The simulation results reveal that islanding scheme is condition dependent. Therefore, for the off line planning of islanding scheme requires an exhaustive number of event simulations, Using the simulation results an islanding strategy may be developed. However, for on line planning, islanding scheme should be dependent of the frequency deviation and ROCOF.

IV. Conclusion

The blackout of a power system is an extreme scenario of power interruptions. It generates severe inconvenience among the consumers and a huge financial loss to the utility. There are both offline and online application for blackout prevention. Offline application includes the predetermined preventive measures like the correct placement of circuit breakers. Online application relies on the technique based on frequency and ROCOF. From the simulation results and the sequential steps that have been taken to run the simulation, it is clear that this technique is capable of handling abnormal conditions to the power system. When auto load shedding scheme is incapable of handling abnormalities, then islanding can make the system stable. Different sizes of islands can be simulated and the results can be compared with a view to develop an optimum solution to maintain the stability of faulted network.

References

[1] *"NERC Steering Group Technical analysis of the August 14, 2003, blackout: what happened ,why, and what did we learn?"* Report to the NERC Board of Trustees, July 13, 2004.

[2] *"U.S. – Canada Power System Outage Task Force, Final report on the August 14, 2003 blackout in the United States and Canada: causes and recommendations,"* U.S. – Canada Power System Outage Task Force, April, 2004.

[3] Corsi, S., Sabelli, C. *"General blackout in Italy September 28, 2003"*, IEEE PES General Meeting, June 6–10, 2004.

[4] Berizzi, A. *"The Italian 2003 blackout*", IEEE PES General Meeting, June 6–10, 2004.

[5] UCTE, *"Interim report of the investigation committee on the 28 September 2003 blackout in Italy, UCTE"*, October 27, 2003.

[6] H. You, V. Vittal, and Z. Yang, *"Self-healing in power systems: An approach using islanding and rate of frequency decline-based load shedding," IEEE Trans. Power Syst.*, vol. 18, no. 1, pp. 174–181, Feb. 2003.

[7] P. Kundur, *"Power System Stability and Control"*. New York: McGraw-Hill, 1994.

[8] S. S. Ahmed, N. C. Sarker, A. B. Khairuddin, M. R. B. A. Ghani, and H. Ahmad, "*A scheme for controlled islanding to prevent subsequent blackout," IEEE Trans. Power Syst.*, vol. 18, no. 1, pp. 136–143, Feb.2003.

[9] M. V. Venkatasubramanian, M. Kezunovic, and V. Vittal, "*Detection, prevention and mitigation of cascading events," Power Systems Engineering Research Centre publication no. 08-18*, Sep. 2008.

[10] N. Grudinin and I. Roytelman, "*Heading off emergencies in large electric grids," IEEE Spectrum.*, vol. 34, pp. 43–47, Apr. 1997.

[11] M. Thaden, *"Analysis of a major load island outage on the Potomac electric power company system," IEEE Trans. Power Syst.*, vol. 14, no.1, pp. 306–311, Feb. 1999.

[12] Y. Liu and Y. Liu, *"Aspects on power system islanding for preventing widespread blackout,"* in *Proc. IEEE Int. Conf. Networking, Sensing and Control*, Ft. Lauderdale, FL, Apr. 23–25, 2006, pp. 1090–1095.

[13] P. Mahat, Z. Chen, and B. Bak-Jensen, *"Under frequency load shedding for an islanded distribution system with distribution generators*," *IEEE Trans. Power Del.*, vol. 25, no. 2, pp. 911–917, Apr. 2010.

[14] Q. Ahsan, A. H. Chowdhury, S. S. Ahmed, I. H. Bhuyan, A. Haque, and H. Rahman, "*Technique to develop auto load shedding and islanding scheme to prevent power system blackout"* in *IEEE Trans. Power Syst.*, vol. 27, no.1, pp. 198-205, Feb. 2012.

[15] Jun Yan*, Student Member, IEEE*, Yufei Tang*, Student Member, IEEE*, Haibo He*, Senior Member, IEEE*, and Yan Sun*, Member, " IEEE Cascading Failure Analysis With DC Power Flow Model and Transient Stability Analysis"* in IEEE Transactions on Power Systems (Volume:30 , Issue: 1) Page(s):285 - 297 Jan. 2015

[16] M. Vaiman, K. Bell, Y. Chen, B. Chowdhury, I. Dobson, P. Hines, M. Papic, S. Miller, and P. Zhang, "*Risk assessment of cascading outages: Methodologies and challenges,*" *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 631–641, May 2012.

[17] Benjamin A. Carreras, David E. Newman, and Ian Dobson*, Fellow, "IEEE North American Blackout Time Series Statistics and Implications for Blackout Risk"* IEEE Transactions on Power Systems (Volume: PP , Issue: 99) Page(s):1 – 9. January 2016

[18] M. Shafiullah, H. Rahman, M.Q. Ahsan, *"Impact study on Islanding operation and frequency based auto load shedding to improve service reliability using CYME PSAF",* Proceedings of the 7th International Conference on Electrical and Computer Engineering, ICECE 2012, Dhaka, Bangladesh, pp. 583-586, Dec. 20-22, 2012.

[19] K. C. Tham, "*Power System Stability*" University of Queensland, October 2003.

[20] P.M. Anderson, A.A. Fouad, *"Power System Control and Stability",* Second Edition, IEEE press, NY, 2003.

Transaction Authorization from Know Your Customer (KYC) Information in Online Banking

Prakash Chandra Mondal,[1,*] Rupam Deb,[2] and Mohammad Nurul Huda[1]

[1]Department of Computer Science and Engineering, United International University
Dhaka, Bangladesh

[2]School of Information and Communication Technology, Griffith University
Brisbane, Australia

[*]prakash.cse2009@gmail.com

Abstract— **Online banking is getting popularity due to location independence, 24/7 services and responsiveness. Financial services through the internet are running under various threats like phishing, pharming (cyber attack intended to redirect a website's traffic to another fake site), malware, Man-In-The-Middle (MITM) attack and the evolving sophistication of compromise techniques. One time password (OTP) in online banking system alleviate the risk and make it secure. In various methods of OTP and Mobile Transaction Authentication Number (mTAN), device can be lost or stolen, delivery in delay, etc. Compliance with Anti-Money Laundering (AML), Know Your Customer (KYC) and sanctions requirements continues to be a key focus area for financial institution (FI) management, and firms must ensure they are following appropriate compliance procedures to meet the increasing regulatory demands [1, 2]. Addressing existing limitation of OTP, this paper proposes Challenge Question (CQ) from dynamic KYC database for transaction authorization before committing any financial transaction from online banking application. Analysis and simulation results show that the proposed method provides equal control as existing OTP/mTAN.**

Index Terms—**Two factor authentication (2FA); Know your customer (KYC); online banking authentication method; internet banking; online banking; Transaction authorization;**

I. Introduction

Online banking is an alluring way of doing business as well as doing all banking activity independent/irrespective of location. This awesome facility also incurred a highest point of risk as it is using public network over the world. The main endeavour of the financial institutions (FIs) is to provide a consistent and secure process of authentication to their users minimizing potential avenues of attack, especially attack vectors beyond the control of either the user or the FIs.

However, authentication method for internet based financial system is still pursuing to ensure highest level of security. Existing strong authentication method is sometimes a nuisance to the real users who are operating their regular business using internet based financial system. Current activity of any of the user has a linear relationship with users many past usage patterns such as geo-location, nature of regular transaction, range of regular transaction, time of regular transaction, hardware to perform transaction, frequency of transaction, purpose of regular transaction, bio-metric behaviour and many other non bio-metric behaviours.

In this research paper, we introduced KYC information verification in place of one time password (OTP) while performing transaction in order to verify the user more intensively. In that case KYC must be privatized with widespread dynamic user input. The application affluent KYC from the dynamic update from the user interaction through the application; on the other hand user can add more confidential information or random question with answer to the KYC database to make the authentication process much stronger and secure. We will also consider ranking on the KYC information for CQ which will be asked to the user during a transaction. The CQ will be assigned to ask the user based on the risk factors assessment result. High rated CQ will be asked to the user when the risk assessment result is higher and low rated CQ will be asked when the risk assessment result is lower.

The rest of the paper is organized as follows. Section II overviews the related work. Description of our proposed model is introduced in Section III. Section IV discusses the result and compares with other method of OTP. Finally, Section V concludes the paper and presents future work.

II. Related Work

In recent couple of decades, many authentication methods have been developed. There are a variety of technologies and methodologies FIs can use to authenticate transaction of the customers. These methods include the use of customer passwords, personal identification numbers (PINs), digital certificates using a public key infrastructure (PKI), physical devices such as smart cards, one-time passwords (OTPs) [4][5][8], USB plug-ins or other types of "tokens", transaction profile scripts, biometric identification, and others. The level of risk protection afforded by each of these techniques varies. The selection and use of authentication technologies and methods should depend upon the results of the financial institution's risk assessment process [3][10]. Mohamed Hamdy Eldefrawy, Khaled Alghathbar and Muhammad Khurram Khan, 2011 [5] presents a novel two-factor authentication scheme whereby a user's device produces multiples OTPs from an initial seed using the proposed production scheme. The initial seed is produced by the communications partners' unique parameters. Applying the many from one function to a certain seed removes the requirement of sending SMS-based OTPs to users, and reduces the restrictions caused by the SMS system. Security risks associated with current Mobile Transaction Authentication Number (mTAN) System are discussed and some modifications in existing system are proposed for using mTAN in a more secure way by Waqar Ahmad Khan, Ghalib Asadullah Shah, Yasir Saleem and Yasir Saleem, 2016 [6]. Alain Hiltgen, Thorsten Kramp and Thomas Weigold, 2005 [7] presents two challenge/response Internet banking authentication solutions, one based on short-time passwords and one certificate-based, and relate them to the taxonomy of credential stealing and channel breaking attacks. They further outline how these solutions can be easily extended for non-repudiation (i.e., transaction signing), should more

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

sophisticated content manipulation attacks become a real problem. Providing foolproof security for financial applications is a rigorous activity. Security architect needs to consider various design considerations to make the applications bullet proof. We further outline how these solutions can be easily extended for non-repudiation (i.e., transaction signing), should more sophisticated content manipulation attacks become a real problem.

Some of the factors [4] considered by the Business Signatures solution are:

Credential Risk: User does not have the correct password and other personal identification data expected.

Transaction Risk: User is trying to make a payment over a threshold amount or change passwords or other personal security information.

Location Risk: The user is not coming from an approved or previously authenticated Internet location with an expected Browser Profile and Computer Profile.

Behavioural Risk: The user coming at an unusual time of day, performing a transaction involving an unusual payee, or unusual amount is given and does not match with users' past behaviour.

III. Proposed Model

The main endeavour of the model is to identify the real online banking user from his/her KYC database information. In this model of financial transaction through online we proposed single login using user ID and password with CQ for transaction authorization from customer KYC information database before committing a financial transaction. The purpose of CQ is to authorize a transaction in place of OTP. The CQ's are selected to authorize a transaction based on the risk assessment of the user who has initiated the transaction. Those CQ are carried higher grade which contains private information of the customer. In this method CQ are choosing from a collection of CQ where the level of the CQ is defined by the risk factor calculation results.

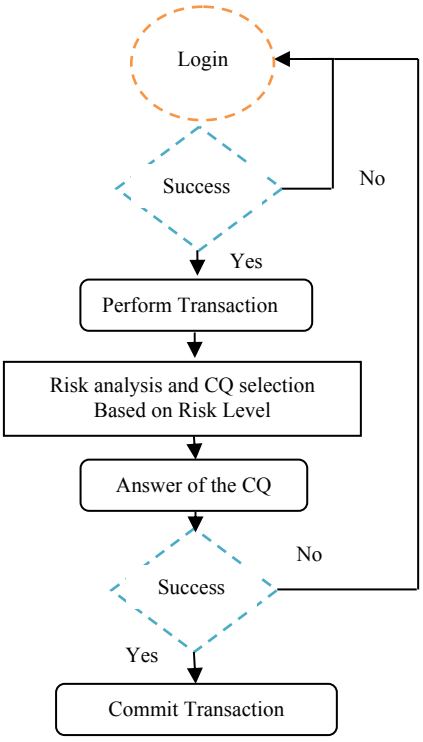


Fig. 1 Transaction authorization model using CQ from KYC database

Risk factors are calculated considering Credential Risk, Behavioral Risk, Transaction Risk and Location Risk. The process works in two stages first, Risk factor calculation and selects CQ for transaction authorization and next and final stage is sending the selected CQ to the customer through mail or SMS.

The brief idea of the model is depicted in the Fig. 1. In the figure the initial step is single factor authentication called login using user ID and password. Analyze transaction performers risk level and select CQ. In the final state verify the transaction performer based on the CQ response. This CQ is the replacement of OTP/nTAN or some other existing 2FA authentication models.

A. Risk Assessment

Risk and threat in online transaction depends on various factors. The prime goal of the FIs is to permit right user for access to financial network and restrict all. In this context, we analyse previous activities of the user and calculate similarity between previous and present introducing multiple techniques and algorithms.

1) Behavioural Risk:

User behaviour is assessed in this scope by computing the probability of the current access time from the previous access times historical log. Activities of the user in the financial application can be collected from the recording different kind of user activities during his/her uses. For this simulation we have used live activities log from a client web based application [12] where hourly access time for a particular user depicts in Fig. 2. Time dependency of the user access has been partitioned hourly basis for each hour in 24 hours.

Number of access of a particular user is C_h, where

$C_h = \sum_{t=h-1}^{t=h+1} a^{t\pm\nabla t}$ and h represents current hour

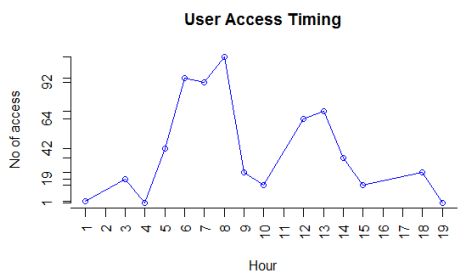


Fig. 2 Historical access time for a user plotted using R studio

From the above behavioural access time data, it has been determined the probability of uses the application in the current time. The probability result determines the selection of CQ which will be assigned to ask for transaction authorization before finally commit a financial transaction. The probability of the current time from the historical timing of a particular user is calculated using Hidden Markov Model (HMM) (1).

$$P(x, \pi) = P(x_1, \ldots, x_N, \pi_1, \ldots\ldots, \pi_N) = P(x_N| \pi_N)\, P(\pi_N \mid \pi_{N-1}) \ldots\ldots P(x_2 \mid \pi_2)\, P(\pi_2 \mid \pi_1)\, P(x_1 \mid \pi_1)\, P(\pi_1) = P(\pi_1)\, P(x_1 \mid \pi_1) \prod_{N=2}^{N} P(X_N|\pi_N)\, P(\pi_N|\pi_{N-1}) \quad (1)$$

CQ level will be assigned based on the derived probability result. Less probability indicate suspicious or fake user on the other hand high probability means trusted and real user. Suspicious user is checked by high rated CQ and vice versa.

In this assessment process we measure least time difference between two consecutive transactions from a dataset [12] of a

live website. We get user activities time records from the user activities log and assessed over 36 least intervals from the collection, which is figured out in the Fig. 4 bar chart. Fraudsters always try to commit transaction within a short time or using some automated script. Using this evaluation we considered short time intervals between the transactions which will prevent to run script and quick transactions. On the other consideration we assessed user's regular short time intervals between the transactions which will match with user's regular behaviour of transaction to indentify real user.

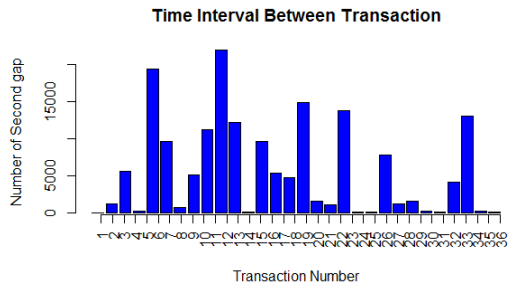


Fig. 4 Time least interval among consecutive quickest transactions plotted by R studio

Any unnatural time space between the transactions will be treated as high risk transaction. If two transactions referred within little time space (30 seconds or less based on the application response time) that will temporarily blocked the transaction and notify the user as it may generated by fraudster or malicious scripting. In other case if any transaction occurs in less than minimum time interval of users previous activities will result to face a high rated CQ during transaction.

2) Credential Risk

In case of credential risk this paper proposes a strong authentication method as all kinds of credential risk are vital and seldom used *i.e* personal information changing, failed to provide correct details. This criteria match between provided input and KYC database existing data. User failed to provide correct answer to the KYC information will be treated as suspicious and high rated CQ will be assigned for further verification.

3) Transactional Risk

Transactional risk assessed based on the KYC profile data limit and regular transaction domain. When user exceeds regular transaction maximum value or is inferior to regular transaction minimum value then it will be treated as suspicious transaction and high rated CQ will be applied during the transaction. In case of user exceeding the transaction profile (TP) upper limit the transaction will block the account temporarily without asking for CQ and any other MFA. From the World Bank dataset [18] to detecting suspicious transaction we calculate standard deviation STD=$\sqrt{\frac{\sum(X-\bar{X})^2}{n}}$, maximums and minimums of the last 30 to 100 transactional data.

Hence, MIN-STD > Suspicious transaction amount > MAX+STD.

4) Location Risk

Location risk has been determined by translating the ip address to geo-location. User will be asked different time to keep record of his/her upcoming movement into the system. The system will compare with the given information and current location prior to giving access to the system or making a transaction. In case of hardware address used by the user will be compared with the current hardware (desktop, laptop, mobile, tab etc.). If the hardware used for the transaction is different from all previous hardware/devices the CQ will be harder based on the non similarity.

B. Assigning the CQ from the risk analysis

More personalized information will be high grading and more public questions will be graded as low value. Based on the risk factor taxonomy the challenge question will be carried to the user for further verification during login or prior to the transaction. The TABLE I describes some of the samples of grading method of the CQ. Frequently ask question (FAQ) for the user will pop-up to the user during his work though the application and will update the answer accordingly.

TABLE I. SAMPLE CQ WITH GRADING

Sample Challenge Question(CQ)	***Rate(1-5)***
One of your childhood friends picture with name?	5
Which color you like best?	5
How often you go to abroad?	1
How many devices you use for online transaction?	5
Usually when do you use the online transaction?	3
What is upper limit of most your transactions?	3
What is the purpose of online transaction?	1
What is your highest educational degree?	3
What is your profession?	3
Which fruit you like best?	5
Which meal you enjoy more?	5
Your Passpor No/Ration Card/ Aadhar Card/Driving License	5
Your Permanent Account Number (PAN) Card	5
National ID card/ TIN/VAT ID Number	5
Your Marriage date?	5
Your spouse born in?	3
Your first job in?	1
User question and answer entry	5
Your family name?	1
Birthday of your spouse?	3
The Drink you like best?(FAQ)	5
Are you going abroad in coming month?(FAQ)	5
What is your first/last educational institute name?	5

The outcome of the risk analysis select the grade of the CQ as described in the TABLE II.

TABLE II. CQ SELECTION PROCEDURE

Challenge Question	***CQ Level 1***	***CQ Level 2***	***CQ Level 3***
Geo-location change	NO	NO	YES
Transaction device Change	NO	YES	NO
Changing nature of transaction	YES	NO	NO
Changing regular transaction purpose	YES	NO	NO
Exceeding regular transaction frequency	NO	YES	NO
Exceeding regular transaction limit	NO	NO	YES
Exceeding profile transaction limit	Temporarily Blocked and notify		
New user	YES	NO	NO
Exceeding general transaction timing	NO	YES	YES
Transaction differs from CQ No.31	Temporarily Blocked and notify		
First failed login attempt	NO	NO	YES
Second failed login attempt	NO	YES	YES
Third failed login attempt	YES	YES	YES
Forth failed login attempt	Temporarily Blocked and notify		

CQ sends to the customer's email or SMS to the mobile shows in the Fig. 5.

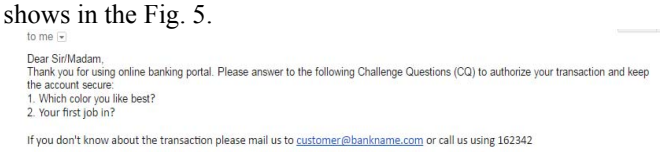


Fig. 5 Transaction authorization CQ to the customer email

Answer to the CQ sends via SMS or email should enter into the online banking portal for authorization a transaction.

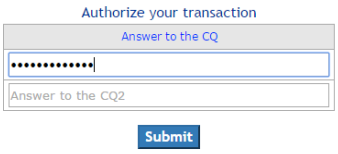


Fig. 6 Answer to the CQ from email of SMS

IV. Result and Comparative Analysis

This paper proposes an alternative method of authorize/signing for financial services through the internet. It is a method to minimize financial fraud forgery on online financial network. The main challenge to avoid fraudulent activity in the financial network is keep the system away from unauthorized person. The proposed method presented here includes sensitive personal information called KYC information to verify the actual owner of the account for online financial activity. It considers all the known and upcoming possible way to theft information and unauthorized entry into the online financial system. The dimension to the risk of the hacking and information stealing is unlimited and tendency for these illegal operations evolving time to time; the method proposed a way to extend the KYC database as required by risk assessment in a certain intervals. As the FIs already preserved customers KYC so, it is effective and fruitful to continue reuse of KYC data for authentication purpose rather than using additional cost involved mechanism.

By this approach the evolving security threat will be minimized in a significant stage moreover the model performance is preserved and improved eventually in each user activities. Authentication measures are a dynamically assigned procedure to restrict transaction from the unauthorized user. In this process it is mostly recommended to user CQ in lieu of OPT/nTAN described in TABLE 1. CQ is graded in various level of standard from 1 to 5 with an independent mechanism based on the heuristic result of the risk.

In this paper we used users' own attributes to indentify a real user and a verfication method by using his/her own attributes from KYC database information. KYC is most secrete and private information to identify a particular account holder in the FIs moreover simulation shows mitigated all possible ways of compromizing the system. We can state that the proposed system will ensure equal security as existing OTP/nTAN.

TABLE III. Comparison Security Properties Traditional Scheme and Our Proposed Model [4]

Attacks	OTP/nTAN	KYC based CQ
Evesdropping attack	✓	✓
Replay attack	✓	✓
Dictionary attack	✓	✓
Brute force attack	⌂	✓
Man-in-the-middle attack	X	✓
User impersonation attack	✓	✓

Notation: √ Satisfied ⌂ Partially Satisfied × Not Satisfied

KYC based CQ and OTP returns a recommended model as it has multiple advantages with no chance of getting stolen or lost, CQ does not view in the portal and answer receives in a encrypted form to prevent possible phishing attack for fraudulent activity, no chance of HASH change, no chance to be misguide by man-in-the-middle (MITM) attack.

V. Conclusions

In this paper, we propose dynamic KYC based transaction authorization method to ensure secure and flawless financial access to the actual account holder of the online bank. Analysis and simulation results show that the proposed method provides equal control as existing OTP authorization minimizing some dynamic risk of being stolen and delay delivery of SMS. Our proposed method is costless and does not incur any hurdle to carry an additional hardware. As there is no chance of key theft this method can be used anywhere from private or public computer. Our method will ensure dynamic security by tackling most of the vulnerabilities in internet based financial transactions.

References

[1] Know your customer (KYC) laws by country. [Wikipedia] Available: https://en.wikipedia.org/wiki/Know_your_customer

[2] FATF, GAFI. "Guidance on the risk-based approach to combating money laundering and terrorist financing," June 2007.

[3] FFIEC Guidance: Authentication in an internet banking environment, federal financial institutions examination council (FFIEC), Retrieved February 4, 2006.

[4] Wen-Bin Hsieh, Jenq-Shiou Leu, "Design of a time and location based one-time password authentication scheme," 2011 7th International Wireless Communications and Mobile Computing Conference (IWCMC), pp 201-206, 4-8 July 2011.

[5] Mohamed Hamdy Eldefrawy, Khaled Alghathbar, Khurram Khan, "OTP-based two-factor authentication using mobile phones," Eighth International Conference on Information Technology: New Generations (ITNG), pp. 327-331, 2011.

[6] Waqar Ahmad Khan, Ghalib Asadullah Shah, Yasir Saleem and Yasir Saleem , "Modified mobile transaction authentication number system for 2-layer security," Intelligent Systems Engineering (ICISE) International Conference, 15-17 Jan 2016.

[7] Alain Hiltgen, Thorsten Kramp, Thomas Weigold, "Secure internet banking authentication," IEEE Security & Privacy, vol.4, no. 2, pp. 21-29, March/April 2006.

[8] Shailesh Kumar Shivakumar, Babu Krishnamurthy, "Advanced security design for financial applications, External Document," White paper of Infosys 2016.

[9] Syeda Farha Shazmeen, Shyam Prasad "A practical approach for secure internet banking based on cryptography," International Journal of Scientific and Research Publications, Volume 2, Issue 12, December 2012.

[10] Lior Golan, Amir Orad, Naftali Bennett, "System and method for risk based authentication," U.S. Patent US 20050097320 A1, May 5, 2005.

[11] World Bank dataset Available:https://finances.worldbank.org/Financial-Intermediary-Funds/Financial-Intermediary-Funds-Cash-Transfers/h4s8-nwev

[12] Data Set from clients live website Available: http://reporting4results.com/actiontracker/login.php

Solid State Transformer Based on Cascade Multilevel Converter for Distribution Network

Muhammad Masud Rana,[1,*] Rafiur Rahman,[2] and Md. Monibor Rahman[1]

[1] Dept. of Electrical and Electronics Engineering
Patuakhali Science and Technology University, Patuakhali, Bangladesh
[2] Dept. of Electrical and Electronics Engineering,
Khulna University of Engineering and Technology, Khulna, Bangladesh
*masud_eee10@yahoo.com

Abstract— **Solid State Transformer (SST), a power electronics based transformer, tends to replace conventional bulk size transformer for distribution of electric power. In this paper a SST has been designed in which cascading has been done for interfacing high voltage power using low power switching devices. Pulse Width Modulation (PWM) has been used for switching technique. The configuration has been simulated in MATLAB Simulink software and its performance at different stage has been shown. Results shows that voltage and current stress is equal to each bridge.**

Index Terms— **Solid state transformer (SST); Electric power; Pulse width modulation (PWM); MATLAB Simulink software; Voltage stress; Current stress.**

I. Introduction

Now a day's power electronic device playing a significant role to increase reliability and controllability in transmission and distribution system. Various FACT devices such as static var compensator (SVC), static synchronous compensator (STATCOM), unified power flow controller (UPFC) have been using since many years to increase the power handling capacity of transmission line [1].

Recently a concept of power electronics based transformer termed as solid state transformer (SST) has been developed which along with stepping up or down voltage level provides some extra benefits such as voltage regulation, power factor correction, var compensation, active power control fault isolation, reduction of size, availability of LVDC link and so on [2]. Again in recent years there is a trend of establishing smart grid in distribution system. SST will be a key component to achieve this. Besides use of renewable energy source is increasing day by day. It will help to integrate the renewable source and associated storage device to grid [3].

Basic structure of Solid State Transformer is shown in Fig.1 [2]. It is basically a high frequency transformer associated with power electronics based converter. Here incoming 50Hz AC voltage is rectified into DC voltage. This DC voltage is converted into high frequency (several kHz) ac voltage and stepped down through a high frequency transformer. Then this high frequency ac voltage is rectified into DC voltage. Finally This DC voltage is converted into useable AC voltage [2].

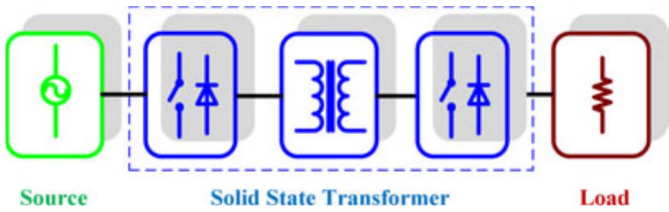


Fig.1 Configuration of solid-state transformer.

In this paper a 3 level cascade SST has been designed. The designed system is simulated in Simulink in MATLAB software and performance has been analysed.

II. Topologies

There exits different SST topologies. Some of them have bidirectional power conversion characteristic, some don't. There are four SST configurations as shown in Fig. 2 that cover all possible SST Topologies [4]. They are

a) single-stage with no DC link
b) two-stage with low voltage DC (LVDC) link
c) two-stage with high voltage DC (HVDC) link and
d) three-stage with both HVDC and LVDC links.

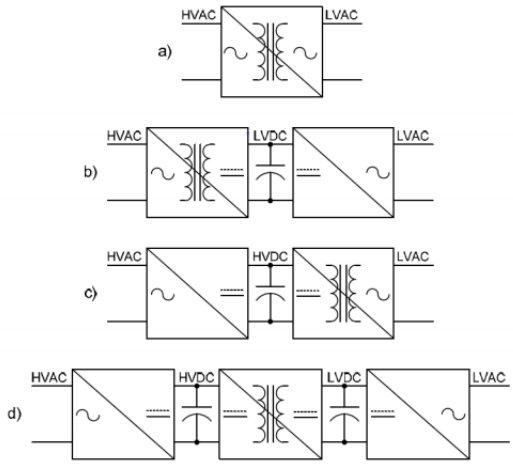


Fig.2 SST configurations: a) single-stage, b) two-stage with LVDC link, two-stage with HVDC link, and d) three-stage.

The single stage topology is the simplest of all in configuration and control. It does not have some benefits provided by the DC link voltage e.g. power factor correction [5]. In this case the input ac voltage is directly converted in to ac voltage a then step downed, rectified back into power frequency ac voltage.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

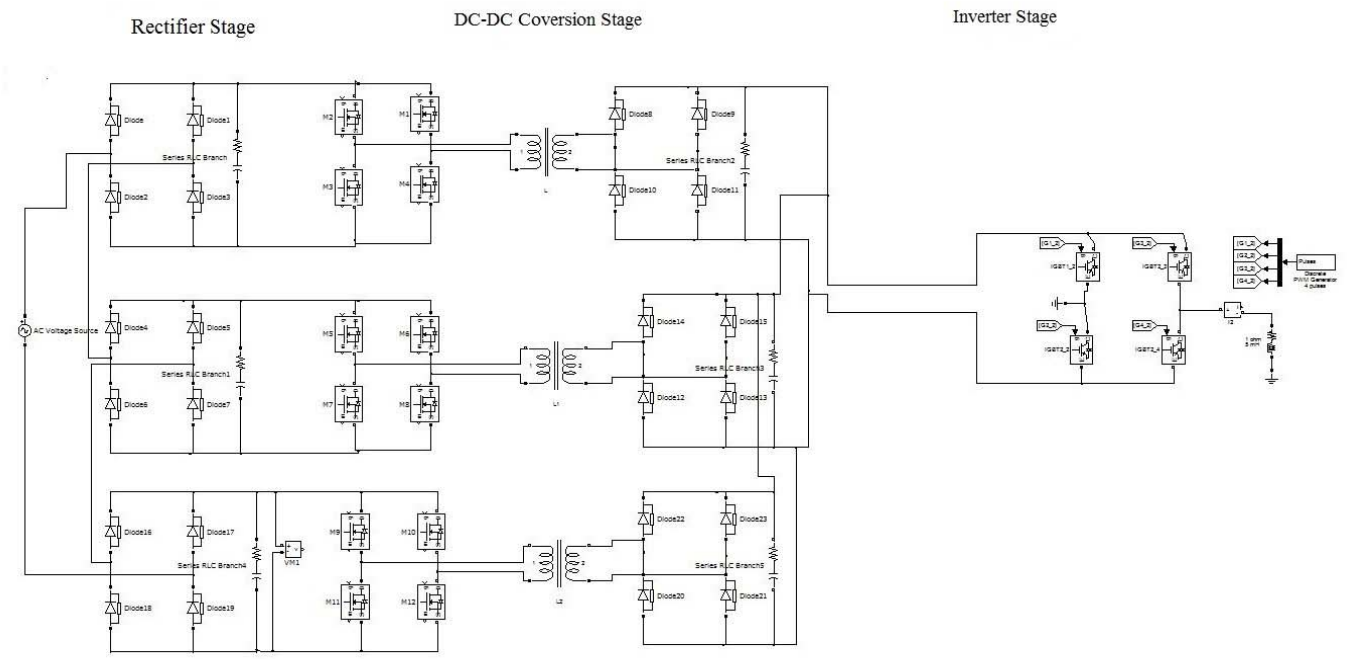


Fig.3 Cascsde SST Configuration.

In two stages SST there will be a dc link either in MV or LV section. It allows having separate dc bus and integration of discrete energy source. Three stage SST is the most important of all. It contains a rectifier in front end, a dc-dc converter, an inverter at the end to interface user. The thing what made it more significant is the presence of a dual active bridge converter which helps to achieve active power flow control, voltage regulation, fault isolation etc.

III. Designing of Cascade Multistage SST

A cascade three stage Solid Stage Transformer Configuration has designed in Simulink which is shown in Fig. 3. Here in high voltage side the converters are connected in series. So the voltage stress is equally divided among them. In low voltage side converters are in parallel connection which permits an equal lower current shearing. Cascading allows use of switing devices of lower rating than system voltage.The main challenge of this system is to ensure equal voltage shearing in each converter and equal power flow throgh all dual active bridge converter.To over come this challenge voltage and power balance control method is used in this type of configuration [6].

IV. Simulation Result and Discussion

The designed model of SST shown in Fig. 3 has been simulated in MATLAB Simulink software.

TABLE I
Specification of Designed SST

Power Rating	300 KVA	Input Voltage	7.2 KV
Output Ac Voltage	220 V	Output Dc Voltage	222 V
Input Current	46 A	Output Ac Current	1060 A

Primary distribution voltage is characterized by 33kv and secondary distribution voltage by 11kv. Since a single phase model is designed, input voltage taken is 7.2kv. MOSFET is selected as switching devise and switching frequency selected is 1 kHz. Results obtained in simulation is shown in Fig. 4-10

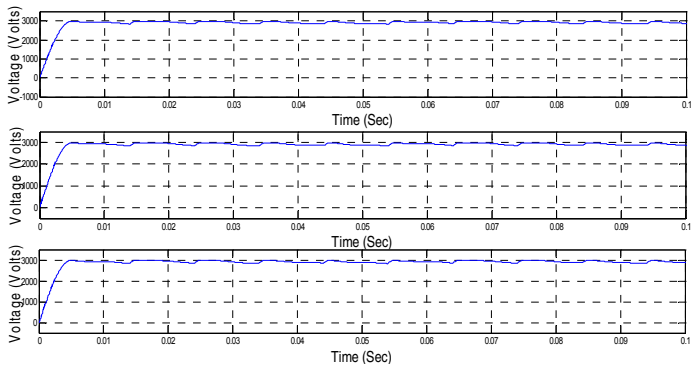


Fig.4 Voltage across each bridge.

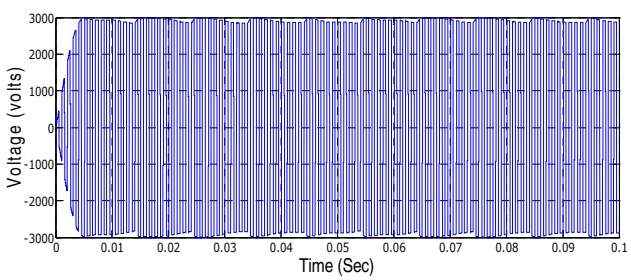


Fig.5 High frequency voltage in dc-dc converter.

The equal voltage sharing off the bridge is shown in Fig. 4. It can be seen that voltage across each bridge is 3000 volts. High frequency voltage input to the transformer is shown in Fig. 5. This high frequency voltage will help to decrease the size of the transformer.

The output dc bus voltage is shown in Fig. 6. Renewable energy source and distributed energy storage devices can be added to this dc bus voltage. Also this dc voltage is converted into ac 50Hz voltage to supply user.

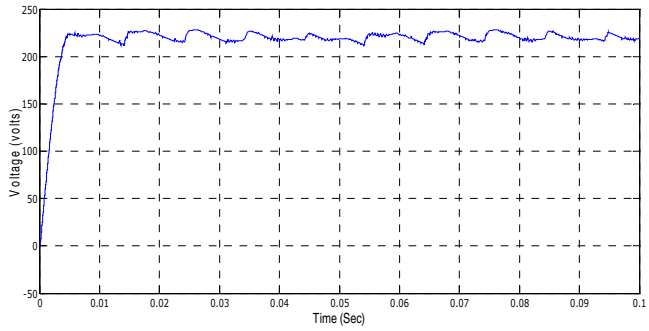


Fig.6 Output dc voltage

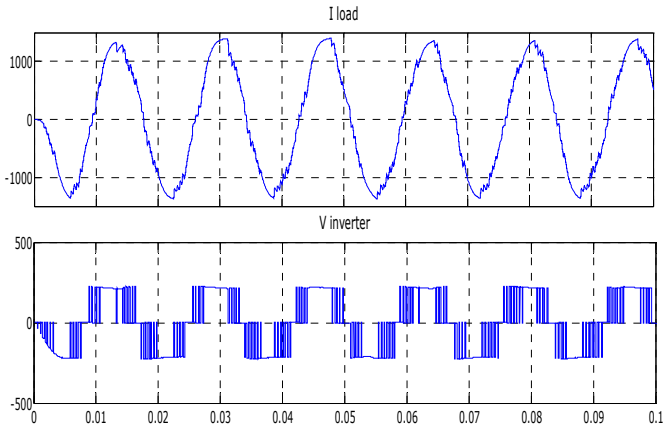


Fig.7 Output AC voltage and current.

Output ac voltage and current profile is shown in Fig. 7. A sine wave inverter is used in this stage. Fig. 8-9 shows input and output real and reactive power of SST. It is seen that input reactive power is less than output. That means SST helps in improving power factor.

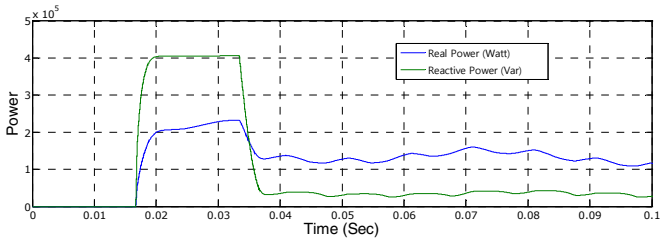


Fig.8 Input real and reactive power

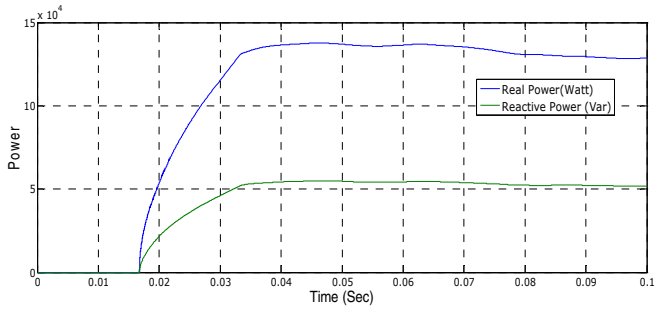


Fig.9 Output real and reactive power

V. Conclusions

Solid State Transformer is a useful technology which facilitates distribution of electrical power in many ways. In this paper a cascade three stage SST has been designed and simulated which helps using semiconductor devices of low rating. Main challenge of this configuration is to ensure equal voltage and power shearing between each level. Again due to its additional level control becomes complicated and less reliable. Voltage and power balance control method con solve this problem. Use of SiC switching devices instead of Si can reduces the number of level which may also leads to less complicated control system

References

[1] Xu She, Member, IEEE, Alex Q. Huang, Fellow, IEEE, and Rolando Burgos, Member, IEEE, “Review of Solid-State Transformer Technologies and Their Application in Power Distribution Systems” IEEE Journal of Emerging and Selected Topics in Power Electronics, vol. 1, no. 3, Sept 2013.

[2] Xu She, Member, IEEE, Xunwei Yu, Student Member, IEEE, FeiWang, Student Member, IEEE and Alex Q. Huang, Fellow, IEEE “Design and Demonstration of a 3.6-kV–120-V/10-kVA Solid-State Transformer for Smart Grid Application. IEEE Transactions on Power Electronics, vol. 29, no. 8, Aug 2014.

[3] Alex Huang, Xu She, Xunwei Yu, Fei Wang, and Gangyao Wang, “Next Generation Power Distribution System Architecture: The Future Renewable Electric Energy Delivery and Management (FREEDM) System” ENERGY 2013 : The Third International Conference on Smart Grids, Green Communications and IT Energy-aware Technologies

[4] Sixifo Falcones, Student Member, IEEE, Xiaolin Mao, Member, IEEE, and Raja Ayyanar, Senior Member, IEEE, “Topology Comparison for Solid State Transformer Implementation”

[5] Kang, M.; Enjeti, P.N.; Pitel, I.J., "Analysis and design of electronic transformers for electric power distribution system," Industry Applications Conference, 1997. Thirty-Second IAS Annual Meeting, IAS'97., Conference Record of the 1997 IEEE , vol.2, no., pp.1689-1694 vol.2, 5-9 Oct 1997.

[6] T. F. Zhao, G. Y. Wang, J. Zeng, S. Dutta, S. Bhattacharya, and A. Q. Huang, "Voltage and power balance control for a cascaded multilevel solid-state transformer," in Proc. IEEE APEC, 2010, pp.761-767.

Improvement in the exciton and EHP generation rate of P3HT:PC_{60}BM based BHJ OSCs with the incorporation of SWCNTs

S.M.Khan[1*], A.A.Hasan[1], A.Saha[1], S.M.Mominuzzamamn[2]

[1]Department of Electrical and Electronic Engineering, BRAC University, Bangladesh

[2]Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology, Bangladesh

*saadmohdkhan665@gmail.com

Abstract— **In this paper the impact of SWCNTs content on the optical performance and EHP generation rate of P3HT:PC_{60}BM based BHJ OSC is presented using the Optical transfer matrix theory and Onsager-Braun formulism. The results from our optical simulation show that the presence of SWCNTs in the active layer composite of P3HT:PC_{60}BM based BHJ OSC improves the exciton generation rate by broadening the absorption spectra towards the NIR region. The exciton generation rate and intensity of absorption was seen to improve as the added content of SWCNT increased. Furthermore the active layer thickness dependent exciton dissociation probability and EHP generation rate was found to be optimal when the added content of SWCNT was at 1%.**

***Index Terms*—Organic Solar Cells, Bulk-heterojunction, Single Walled Carbon Nanotubes, Electron-Hole pair**

I. INTRODUCTION

Organic solar cells (OSCs) have emerged to be the most promising alternative to silicon based solar cell devices due to its low material cost, ease of fabrication, thin film weight and high mechanical flexibility[1,2]. OSCs incorporating bulk heterojunction (BHJ) instead of planar heterojunction (PHJ) have resulted in better efficiencies [3]. The enhancement in the performance of BHJ OSC can be attributed to the increase in the number of donor-acceptor (D-A) interfaces available for the dissociation of the photo generated excitons[3]. At present the most extensively studied BHJ OSC involves the use of an active layer consisting of the conjugated polymer (P3HT) as the donor and fullerene derivative (PC_{60}BM) as the acceptor [1,3].Despite initially showing great promise, the power conversion efficiency (PCE) of P3HT:PC_{60}BM based BHJ OSCs has remained relatively low when compared to its inorganic counterpart [2]. The low PCE results primarily due to the organic semiconductors low carrier mobility(retarded due to the organic semiconductor's intermolecular hopping transport mechanism) and low visible and Near-infrared(NIR) region absorptivity [2].The aforementioned draw backs directly hampers the two key processes governing the power conversion efficiency (PCE) of a BHJ OSC: 1) the generation of excitons in the active layer upon photon absorption (photon absorption/exciton generation rate) and 2) the dissociation of the excitons into free charge carriers at D-A interface of the active layer (EHP generation rate)[1]. One dimensional nano structured materials such as SWCNTs are known to have high carrier mobility due its ballistic property and can also absorb photon wavelengths in the NIR region [2,4]. This fascinating optical and electrical property of SWCNTs makes them a promising candidate as an additive to the P3HT:PC_{60}BM based active layer composite. In this paper the optical performance and EHP generation rate of P3HT:PC_{60}BM based BHJ OSCs of various SWCNT content is investigated. The simulations presented in this paper are based on the Optical transfer matrix theory and Onsager-Braun formulism:

II. SIMULATION SCHEME:

In a BHJ OSC, the active layer consists of a random complex blend of the acceptor and donor and hence both the components are equally involved in the absorption of light and generation of photo excitons[5]. For simulation and modeling purposes the BHJ device is usually described using the metal-insulator metal model, where the whole active layer is considered to be a single homogenous semiconductor device with an effective hole mobility(donor's hole mobility is considered), effective electron mobility(acceptor's electron mobility) and effective band gap [6]. The simulation based study performed in this paper consists of two parts: 1) optical simulation and 2) EHP generation rate determination .From the optical simulation we generate the:

1) Wavelength dependent Absorption coefficient profile
2) Absorption spectra
3) Active layer thickness dependent exciton generation rate

The exciton generation rate determined from the optical sim--ulation is then used to calculate electron-hole pair (EHP) generation rate.

III. OPTICAL MODELING:

The optical transfer matrix theory was used to simulate the optical performance of P3HT:SWCNT:PC_{60}BM based BHJ OSC [5]. The model characterizes the whole organic solar cell as a multilayered device, where every layer is defined by its particular sets of optical constants and thickness. The model considers that in a thin film device, light is at normal incidence to the substrate. As shown in Fig.1 (a), the model considers that a plane wave was incident from left at an m layer structure (layer 1 to layer m) between a semi-infinite transparent ambient and semi-infinite substrate. The resultant optical electric field in the thin film device due to the interference effect at any point x in layer j consists of: 1) positive complex component $E_j^+(x)$ and 2) negative complex component $E_j^-(x)$. The optical properties of each layer (j) can be characterized by its complex index of refractions consisting of the extinction coefficient k_j and refractive index n_j , [$\widetilde{n_j} = n_j + ik_j$]and thickness d_j.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

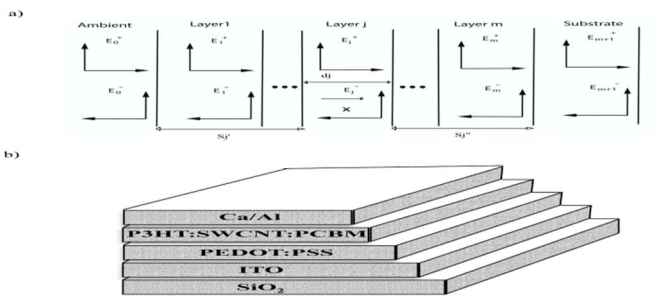


Fig .1(a) Multilayered structure with m layers (b) Device structure of BHJ OSC.

The light intensity at a distance x in a particular layer j can be expressed as:

$$I_j(x,\boldsymbol{\lambda}) = T_j I_0(\lambda)\begin{bmatrix} e^{-\alpha_j x} + \rho_j''^2 . e^{-\alpha_j(2d_j - x)} \\ +2\rho_j''. e^{-\alpha_j d_j}. \cos\left(\frac{4\pi n_j}{\lambda}(d_j - x) + \delta_j''\right) \end{bmatrix} \quad (1)$$

Where $I_0(\lambda)$represents the intensity of the incident light, T_j is the internal intensity transmittance, $\alpha_j = \frac{4\pi k_j}{\lambda}$ is the absorption coefficient , p_j'' and d_j'' are the absolute and the argument of the complex reflection coefficient. The first term in equation 1 is due to the optical electric field propagating in the positive x direction $E_j^+(x)$, the second term is as result of the field propagating in the negative x direction $E_j^-(x)$ and the third term is due to the optical interferences of the first two terms. The rate of energy dissipated per unit volume Q can then be determined as:

$$Q(x,\lambda) = \alpha(\lambda) I(x,\lambda) \quad (2)$$

Where $\alpha(\lambda)$ is the absorption coefficient. The exciton generation rate $G(x)$ is assumed to be equal to the photon absorption rate and is given by:

$$G(x) = \int \frac{\lambda}{hc} Q(x,\lambda) d\lambda \quad (3)$$

Where h is the plank constant and c is the speed of light in vacuum and λ is the wavelength.

IV. EHP GENERATION RATE DETERMINATION:

The Onsager-Braun Model was used to determine the active thickness dependent EHP generation rate [6]. According to the model the excitons generated at the active layer upon photon absorption dissociate into bound EHPs at the D-A interface via ultrafast electron transfer from the donor to the acceptor .The bound EHP can either decay to the ground state or dissociate into free charge carriers. According to Braun, the dissociation probability of a bound EHP is a temperature and field dependent process and is given by:

$$P(x,T,F) = \frac{k_{diss}(x,T,F)}{k_{diss}(x,T,F) + k_f(T)} \quad (4)$$

Where k_{diss} is the dissociation rate, k_f is the decay rate and is assumed to be constant ($1.5\times10^6 s^{-1}$)[6], the dissociation rate is expresses as:

$$k_{diss}(x,T,F) = \frac{3\gamma}{4\pi a^3} e^{-\frac{E_B}{k_B T}}\left(1 + b + \frac{b^2}{3} + \cdots\right) \quad (5)$$

γ represents recombination strength, a is the length of the carbon-carbon bond , E_B represents the EHP pair binding energy($b = q^3/(8\pi\varepsilon k_B^2 T^2)$), F is the electrical field strength.

The recombination strength is given by langavin as:

$$\gamma = \frac{q}{<\varepsilon>} < \mu > \quad (6)$$

Where $< \mu >$and $< \varepsilon >$ are the spatially averaged sum of electron and hole mobilies and dielectric constant. The electric field is assumed to be uniforme inside the active layer and is given by:

$$F = \frac{V_a - V_b}{L} \quad (7)$$

Where V_a is the applied voltage, V_b is the built in voltage (difference between the work function of the electrodes) and L is the active layer thickness.
The EHP generation rate can then be assumed to be:

$$U = P \times G \quad (8)$$

Where P is the exciton dissociation probability and G is the exciton generation rate and U is the net EHP generation rate.

V. RESULT AND ANALYSIS:

The configuration of the BHJ OSC device shown in Fig.1(b) is simulated as follows:
[SiO_2/ITO(100nm)/PEDOT:PSS(25nm)/P3HT::SWCNT:PCBM(L)/Ca(7nm)/Al(100nm)]. Where SiO_2 is the substrate, ITO is the cathode, PEDOT:PSS acts as the hole transport layer, P3HT:PCBM is the composition of the active layer with L thickness and Al is the anode. The optical constant of all the layers use in the simulation were extracted from literature and is shown in Fig.2 (a) [7,8]. The optical constants of the SWCNT incorporated P3HT:PC_{60}BM based active layer composite were modified in accordance to SWCNTs content and is shown in Fig.2 (b). The resultant absorption coefficient profile used for the optical simulation after the modification of the optical constants is shown in Fig.2(c).

Fig. 3 shows the impact of SWCNT content on absorption spectra of the P3HT:PC_{60}BM based BHJ OSCs at an active layer thickness of 70nm simulated using the optical transfer matrix theory. From Fig. 3 it can be seen that the reference device (P3HT:PC_{60}BM) displayed two major peaks and one shoulder peak situated at 340nm, 510nm and 630 nm. In general, polymers (P3HT) are known to exhibit strong absorption in the visible region covering the 450-750nm range, while fullerene derivatives (PC_{60}BM) display stronger absorption in the region ranging from 300-450nm as can be seen in Fig.2 (a) and hence the peaks observed at 510 and 630 nm can be attributed to the presence of P3HT in the active

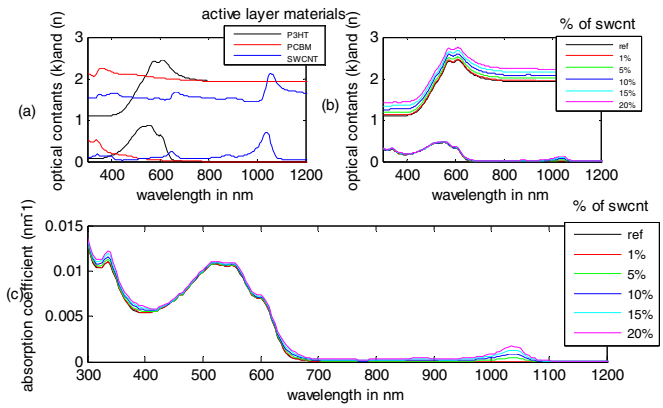


Fig. 2 (a) Optical constants of the active layer materials (b)Modified Optical constants of SWCNT content dependent P3HT:PC_{60}BM composite (c) Absorption coefficient of SWCNT content dependent P3HT:PC_{60}BM composite

layer composite, while the peak appearing at 310 nm corresponds to $PC_{60}BM$. The reference device however did not show any absorbance in the NIR region. From Fig. 3 it can be deduced that the incorporation SWCNTs widened the absorption spectra toward towards the NIR region by introducing an additional peak at 1035nm and although the position of the peaks in the visible and NIR spectrum does not change with the incorporation of SWCNTs, the intensity of all the peaks achieved in the reference device does increase as the added content of SWCNT in the active layer composite increases. The improvement in the absorption intensity of the shoulder (630nm) and major peaks (340,510nm) can be ascribed to the second van Hove singularities of s-SWCNTs and the peak (1035 nm) observed at the NIR region for SWCNT based devices can be attributed to the first van Hove transition of s-SWCNTs [8].

Fig. 4 shows the impact of SWCNT content on the exciton generation rate of the P3HT:$PC_{60}BM$ based BHJ OSCs. From Fig. 4 we can deduce that the exciton generation rate of BHJ OSCs of any active layer composite gradually increases with its active layer thickness but shows apparent oscillation (peaks and roughs).The oscillation's observed are due to the to the optical interface effect [5].Our simulation result from Fig. 4 also show that the exciton generation rate of P3HT:$PC_{60}BM$ based BHJ OSC's at any particular active layer thickness increases as the content of SWCNT incorporated in the active layer composite increases. The increase in the exciton generation rate of SWCNT incorporated P3HT:$PC_{60}BM$ can be correlated to the absorption spectra in Fig 3. Since the absorption spectra of P3HT:$PC_{60}BM$ based devices incorpoarted with SWCNTs is stronger and broader than P3HT:$PC_{60}BM$ device, it will lead to greater photon harvestation and eventually higher exciton generation rate

Fig. 5 shows change in the effective hole mobility of a P3HT:$PC_{60}BM$ composite due to the incorporation of SWCNTs within its active layer. The effective hole mobility data of P3HT:$PC_{60}BM$ based composites of various SWCNT content were taken from literature [9, 10]. From Fig. 5 we can deduce that the active layer composite with SWCNT content of 1% displayed the highest effective hole mobility. However incorporation of SWCNTs to the active layer beyond 1% resulted in mobilities lower than the reference device. The improvement in the effective hole mobility of the P3HT:$PC_{60}BM$ composites for SWCNTs content less than 1% can be attributed to SWCNTs ballistic properties, which

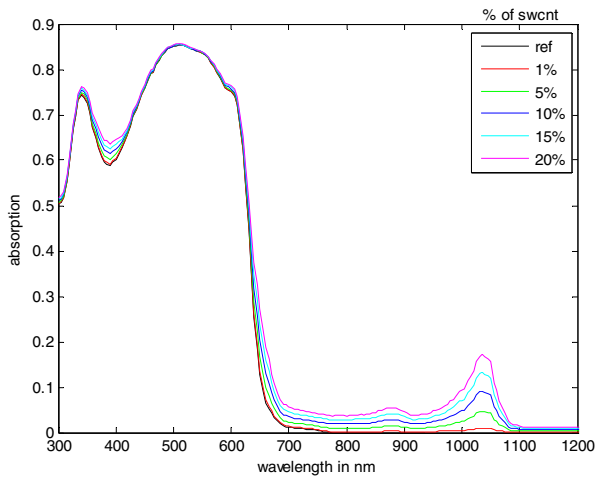


Fig. 3 Absorption spectra of (SWCNT content dependent) P3HT:SWCNT:$PC_{60}BM$ based BHJ OSCs

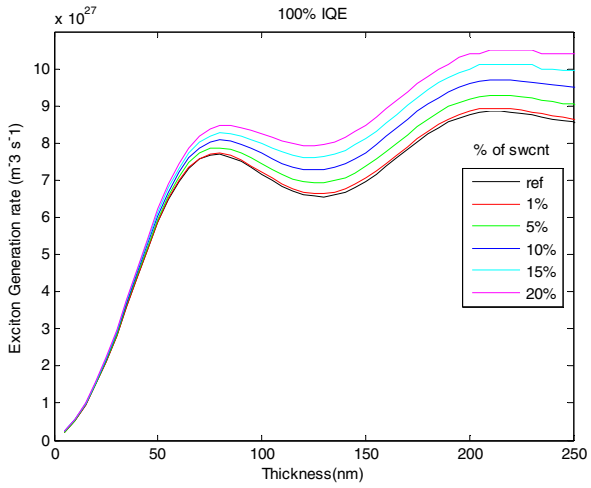


Fig. 4 Active layer thickness dependent exciton generation rate of (SWCNT content dependent) P3HT:SWCNT:$PC_{60}BM$ based BHJ OSCs

leads to the formation of percolating pathways in the P3HT matrix, the presence of percolating pathways via nanotube leads to greater electrical conductivity and hence greater effective hole mobility [11,12]. The drastic decrease in the effective hole mobility of P3HT:$PC_{60}BM$ composite for SWCNT content higher that 1% can be attributed to 1) the presence of metallic SWCNTs, which suppresses the effective hole mobility by introducing recombination sites and 2) to SWCNTs natural tendency to form bundles of um scale, which reduces the amount of interface available for the dissociation of excitons and thus furthering suppressing the effective hole mobility [3,13]. Fig. 6 shows the SWCNT content dependent exciton dissociation probability of P3HT:$PC_{60}BM$ based composite plotted as a function of the active layer thickness. In our work we have assumed that the addition of SWCNTs do not affect the: 1) overall dielectric constant of the active layer composite and thus the excition binding energy remains constant for any SWCNT based active layer composite and 2)electron mobility of $PC_{60}BM$, since SWCNTs are genearlly p type materials[12].Thus the major parameters affecting the exciton dissociation probability of the P3HT:$PC_{60}BM$ composite in our analysis is its effective hole mobility.From Fig. 6 it is evidient that the exciton dissociation probability of BHJ OSCs of any active layer composite gradually decreases with its thickness. The composite with SWCNT content of 1% displayed the highest exciton probability at any active layer thickness. Composites with

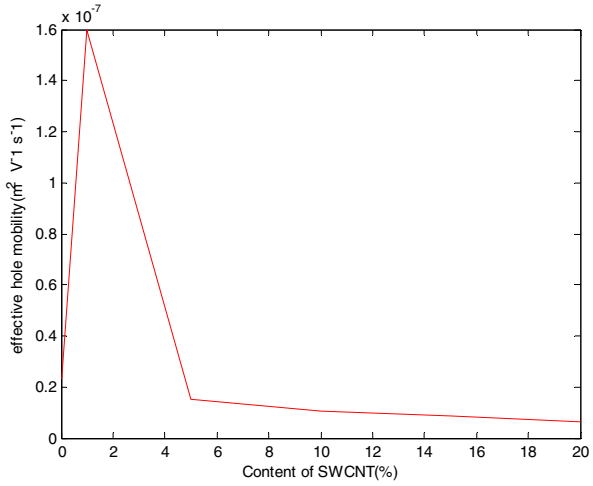


Fig. 5 SWCNT content dependent effective hole mobility of P3HT:SWCNT:$PC_{60}BM$ based BHJ OSC

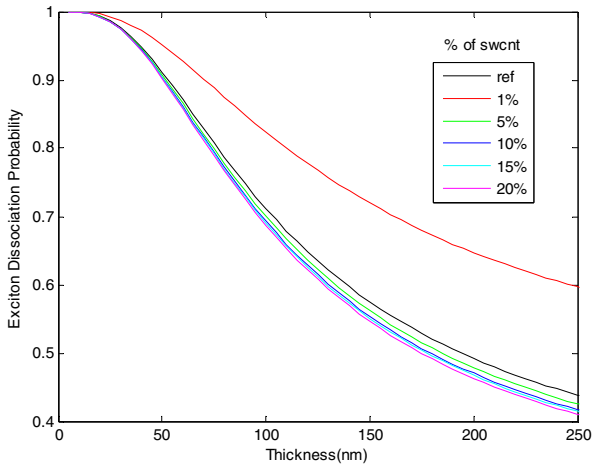


Fig. 6 Active layer thickness dependent exciton dissociation probability of (SWCNT content dependent) P3HT:SWCNT:$PC_{60}BM$ based BHJ OSCs

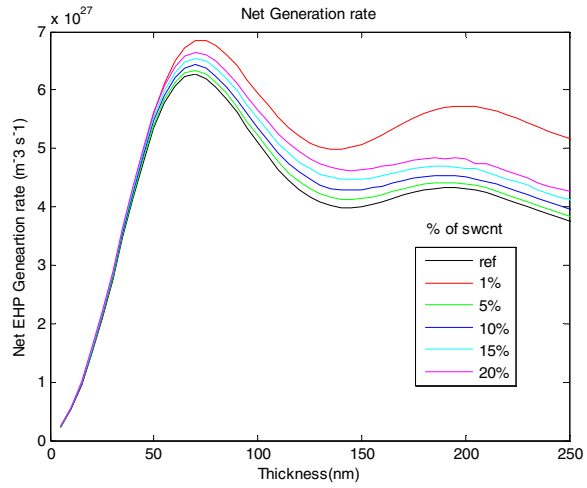


Fig. 7 Active layer thickness dependent EHP generation rate of (SWCNT content dependent) P3HT:SWCNT:$PC_{60}BM$ based BHJ OSCs

higher content of SWCNT (>1%) registered exciton dissociaion probabilites lower than the reference device.

Fig. 7 shows EHP genearation rate of the P3HT:$PC_{60}BM$ based composites at different SWCNT content.The two major parameters affecting the EHP gneneration rate in our analysis : 1) exciton generation rate and 2) exciton dissociation probability. From Fig. 7 we can deduce that the highest EHP generation rate of any active layer composite was achieved at an optimized thickness (70-80nm). Since P3HT:$PC_{60}BM$ is taken as a reference device, its optimized thickness (70nm) is used as a reference for comparison purpose. The highest thickness optimized EHP generation rate was exhibited when the SWCNT content was at 1% ($6.84\times10^{27}m^{-3}s^{-1}$), which was effectively 9.24% greater than that of the reference device ($6.26\times10^{27}m^{-3}s^{-1}$), making it the optimal device. The enhancement in thickness optimized EHP generation rate at low content (1%) of SWCNT is mainly due to greater exciton dissociation probability and exciton generation rate of the composite in comparison to the reference device. The thickness optimized EHP generation rate for SWCNT content higher than 1% was found to be lower than that of the optimal device but remained higher than the reference device despite having exciton dissociation probabilities lower than the reference device. The greater thickness optimized EHP generation rate fort SWCNT content higher than 5% with respect to the reference device thus could be attributed to the greater excition generation rate of the composite in comparison to the reference device.

CONCLUSION:

In conclusion the impact of SWCNTs content on the spectral performance (Absorption spectra, Exciton generation rate) and the EHP generation rate of P3HT:$PC_{60}BM$ based BHJ OSC were simulated and analysed using the Optical transfer matrix and Onsager-Braun formulism. The results obtained from the optical simulation show that the incorporation of SWCNTs broadens the absorption spectra by enhancing the absorptivity in the visible and NIR region and although the position of the absorption peaks in the absorption spectra do not get shifted with the addition of SWCNT, the intensity of the absorption peaks does increase as the added content of SWCNT increases. It has also been observed that the thickness optimized EHP generation rate of P3HT:$PC_{60}BM$ based BHJ OSCs was at its optimal when the added content of SWCNT was 1%, further additon of SWCNTs resulted in EHP generation rates greater than the reference device but lower than the optimal device. The performance of the devices with higher content of SWCNT could be enhanced by improving the processes used to produce pure semiconducting SWCNTs and by homogenously dispersing the SWCNTS within the the active layer of the BHJ OSC.

REFERENCES:

[1] Y.-H. Lin *et al.*, "Comparative study of spectral and morphological properties of blends of P3HT with PCBM and ICBA," *Organic Electronics*, vol. 13, no. 11, pp. 2333–2341, Nov. 2012.

[2]L. Lu, T. Xu, W. Chen, J. Lee, Z. Luo, I. Jung, H. Park, S. Kim and L. Yu, "The Role of N-Doped Multiwall Carbon Nanotubes in Achieving Highly Efficient Polymer Bulk Heterojunction Solar Cells", *Nano Letters*, vol. 13, no. 6, pp. 2365-2369, 2013.

[3] G. Cakmak, H. Y. Guney, S. Aydın Yuksel, and S. Gunes, "The effect of functionalized single walled carbon nanotube with octadecylamine on efficiency of poly-(3-hexylthiophene): [(6, 6)] phenyl C61 butyric acid methyl ester organic solar cells," *Physica B: Condensed Matter*, vol. 461, pp. 85–91, Mar. 2015.

[4] M. S. Konsta-Gdoutos, Z. S. Metaxa, and S. P. Shah, "Highly dispersed carbon nanotube reinforced cement based materials," *Cement and Concrete Research*, vol. 40, no. 7, pp. 1052–1059, Jul. 2010.

[5]L. Liu and G. Li, "Thickness optimization of organic solar cells by optical transfer matrix", *2011 11th IEEE International Conference on Nanotechnology*, 2011.

[6]L. Koster, E. Smits, V. Mihailetchi and P. Blom, "Device model for the operation of polymer/fullerene bulk heterojunction solar cells", *Physical Review B*, vol. 72, no. 8, p. 085205, 2005.

[7]"McGehee Group - Stanford Materials Science and Engineering", *Web.stanford.edu*, 2016. [Online]. Available: http://web.stanford.edu/group/mcgehee/transfermatrix/. [Accessed: 28- Aug-2016].

[8] S. L. Guillot *et al.*, "Precision printing and optical modeling of ultrathin SWCNT/C60 heterojunction solar cells," *Nanoscale*, vol. 7, no. 15, pp. 6556–6566, Mar. 2015.

[9] J. M. Lee, J. S. Park, S. H. Lee, H. Kim, S. Yoo, and S. O. Kim, "Selective electron- or hole-transport enhancement in Bulk-Heterojunction organic solar cells with N- or b-doped carbon Nanotubes," *Advanced Materials*, vol. 23, no. 5, pp. 629–633, Nov. 2010.

[10] A. E. Rudenko, P. P. Khlyabich, and B. C. Thompson, "Random Poly(3-hexylthiophene- co -3-cyanothiophene) copolymers via direct Arylation polymerization (DArP) for organic solar cells with high open-circuit voltage," *ACS Macro Letters*, vol. 3, no. 4, pp. 387–392, Apr. 2014

[11] E. Kymakis, N. Kornilios, and E. Koudoumas, "Carbon nanotube doping of P3HT: PCBM photovoltaic devices," *Journal of Physics D: Applied Physics*, vol. 41, no. 16, p. 165110, Jul. 2008.

[12]S. Cataldo, P. Salice, E. Menna, and B. Pignataro, "Carbon nanotubes and organic solar cells,"*Energy Environ. Sci.*, vol. 5, no. 3, pp. 5919–5940, 2012.

[13] S. Ren, M. Bernardi, R. R. Lunt, V. Bulovic, J. C. Grossman, and S. Gradečak, "Toward efficient carbon Nanotube/P3HT solar cells: Active layer morphology, electrical, and optical properties," *Nano Letters*, vol. 11, no. 12, pp. 5316–5321, Dec. 2011

Microstrip Planar Antenna for Wideband Wireless Communication with Defected Ground Structure

Rezaul Azim[1,*] and Mohammad Tariqul Islam[2]

[1]Department of Physics, University of Chittagong, Chittagong 4331, Bangladesh

[2]Department of Electrical, Electronic and Systems Engineering, Universiti Kebangsaan Malaysia, 43600 Bangi, Malaysia
razim71@gmail.com

***Abstract*— In this article a microstrip antenna is reported for wideband wireless communication application. The main radiator of this design is printed in the front side of a piece of FR4 dielectric substrate while a finite ground plane is etched on the rear side. To achieve wide operating band, the ground plane is reduced to partial one and defected by inserting slots on its upper edge. It is found that this modification established a proper coupling between ground structure and the radiator, and the investigated antenna attained a measured fractional bandwidth of 130%. Furthermore, the proposed antenna demonstrates omnidirectional radiation characteristics with good gain and efficiency.**

***Index Terms*—Antenna; defected groud plane; microstrip patch; wireless communication; slot.**

I. Introduction

As an essential component for transmitting and receiving of signals, antenna plays a vital role in modern day wireless communication applications. The IEEE defines the antenna as "a means for radiating or receiving radio waves" [1]. A receiving antenna turns an electromagnetic wave into a signal or voltage at its terminals that can afterward be processed by the receiver. To satisfy the system's requirement, in telecommunication an antenna can be of different forms. Therefore, an antenna can be a part of conducting wire, a patch, a reflector, an aperture, a lens or an array of elements. A suitable antenna has to satisfy the system requirements and increase complete system performance. Different forms of antennas are reported in open literature for different purposes. Some of the reported designs are large, non-planar and physically obtrusive and difficult to be integrated in RF circuits and are rarely be used in handy communication gadgets.

Recently design of wideband antennas is getting interest in industry and academia. When compared to other kinds of antennas, the planar microstrip design etched on a slice of printed circuit boards (PCB) is the most suitable for portable devices. Usually design of planar type of antenna comprises of a patch and a ground structure printed in two sides of dielectric substrate of the PCBs. They are relatively easy to fabricate due their planar compact structure, light weight and can easily be integrated in microwave and RF devices. The microstrip planar antennas can be applied to good number of applications such as telecommunication, GPS, WiMAX, WiFi, WLAN, radio and TV broadcasting, RFID, satellite communication, radar, microwave imaging etc. All these applications are dedicated to well-being of the human society.

Despite of numerous attractive features, the use of basic microstrip patch antennas is limited in numbers of modern day wireless applications due to their narrow operating bandwidth (usually 2- 5%). Different methods have already been implemented to widen the operating band of basic planar antenna. The simplest method is to use substrate material with higher thickness. But, the radiated power as well as the radiation efficiency of the antenna is decreases when thick substrate material is used to fabricate the antenna [2]. Stacked patches with multiple layers have also been used to enhance the antenna's operating band [3]. However, manufacture of stacked patch antenna is a challenging job. Introduction of different types of slot/s in the patch and change of patch geometry have also been reported for the increment of planar microstrip antenna [4]. To increase the bandwidth, use of metamaterial also has been observed. However usually metamaterial is unavailable in nature and be engineered [5]. Another efficient technique to increase the bandwidth is to defect the ground plane [6-10]. Defected ground plane not only widen the impedance bandwidth but also reduces the cross-polarized radiations of the patch antenna. For example in[7], to enhance the operating band, the top side of the antenna's ground structure was changed to make a saw-tooth shape top edge and this modification enhances the bandwidth by 43.6%. To widen the operating band, in [8] a rectangular slot of dimension 6 × 0.9mm^2 was inserted on the upper side of the ground structure. In [9], two complementary were introduced on the ground plane to attain an impedance bandwidth of 3.1GHz. Most the above mentioned designs exhibits broad impedance bandwidth but some of them possesses large size which restrict their use in practical applications while others have complex structure results in fabrication difficulties and elevated cost.

In this article, a planar microstrip antenna with defected ground structure is proposed. To widen the impedance band of the basic antenna, numbers of rectangular slots have been inserted in the upper side of the ground plane. It is found that the introduction of these slots helps to elevate the matching between the ground structure and radiator and thus the investigated antenna attained an impedance band of more than 11GHz.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

II. FOOTRPRINT OF THE PROPOSED ANTENNA

The layout of the proposed planar design is depicted in Fig. 1. A planar radiator of size 13.5 × 14.5mm^2 is etched on one side of a 1.6mm thick PCB substrate of dielectric constant 4.6 and tanδ = 0.02 and a conducting ground plane is etched on the rear side of the radiator. The radiator is fed by a 50Ω microstrip feedline of length 5.5mm and width 2.76mm and is connected to an SMA for signal transmission and reception. The radiator is 0.5mm away from the upper edge of the ground plane and the overall size of this design is 29 × 20.5mm^2.

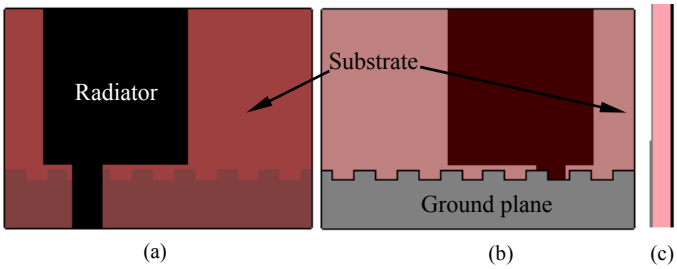


Fig. 1. Design layout(a) Front view, (b) rear view (c) side view.

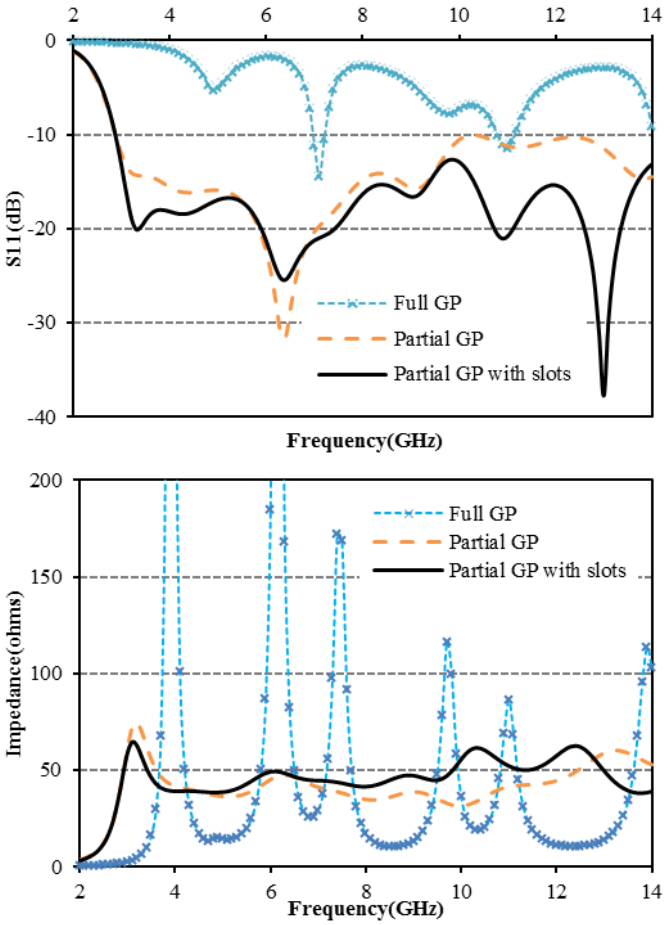


Fig. 2. S_{11} (top) and input impedance (bottom) for different ground plane sizes.

In the design of patch antenna, the ground structure shape and size can be suitably selected to attain a wider impedance bandwidth. Since the electric current in any patch antenna is distributed both on the ground structure and radiator, it is impossible to avoid the release of EM wave from the ground plane. The shape and size of the ground plane are thus significantly control the performance of the patch antenna [11]. In Fig. 2 the S_{11} parameter of the investigated antenna with different ground plane shapes is depicted. It can clearly be commented that, the investigated antenna with full ground plane possesses smallest operating band and worst performance as compared to the antennas with partial ground plane and partial ground plane with multiple slots. The S-parameter is significantly improved when the ground plane is reduced to partial one and its value drops to -31dB. Finally insertion of slots on the partial ground plane helps the design antenna to attain an operating band ranging from 2.86GHz to more than 14GHz with very good S-parameter values. When compared to the antenna with full and partial ground plane, the input impedance of the antenna with multi slotted ground plane is very much close to the 50Ω line which demonstrates a proper matching. It is also evident from plot in Fig. 2 that the designed antenna exhibits multiple resonances at 3.3GHz, 6.35GHz, 9GHz, 10.9GHz,13GHz and combining of all these resonances leads to the characterization of very wide operating band.

To understand the operating principle i.e. wide band characteristics of the designed antenna, the average and vector currents at 3.3GHz, 6.35GHz, 9GHz and10.9GHz (that are much nearer to the resonance frequencies) are displayed in Fig. 3. It can be seen from Fig. 3 that most of the currents is accumulated near the sides of the ground plane and radiator while the currents at the center of the radiator are very weak. In the current distribution plot the resonances are seen on the edges of the ground plane and radiator.

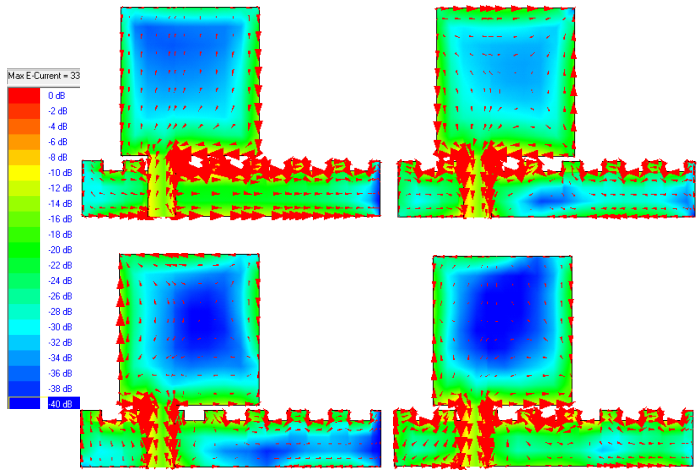


Fig. 3. Surface current distribution at (a) 3.3GHz, (b) 6.35GHz, (c) 9GHz and (d) 10.9GHz

III. RESULTS AND DISCUSSION

The measured and simulated S-parameter of the proposed antenna shown in inset of Fig. 4 is plotted in Fig. 4. The input impedance characteristic is measured using a PNA series vector network analyzer (E8362C). It can be realized from the measured S-parameter curve that the fabricated antenna is able to attain an impedance band from 2.95 to more 14GHz. This wideband ensured that the proposed antenna is a suitable

candidate to be used in WiFi, WiMAX, WLAN, UWB, C-band and X-band satellite communications. The measured results in the graph are slightly differ from simulated one which can be due to imperfection in prototyping, use of long feeding cable in the measurement and higher loss tangent of inexpensive FR4 substrate.

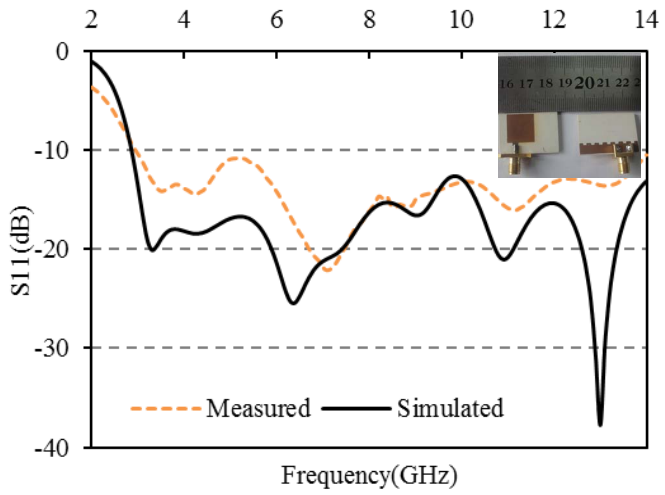


Fig. 4. Measured and simulated S_{11} curves of the proposed antenna.

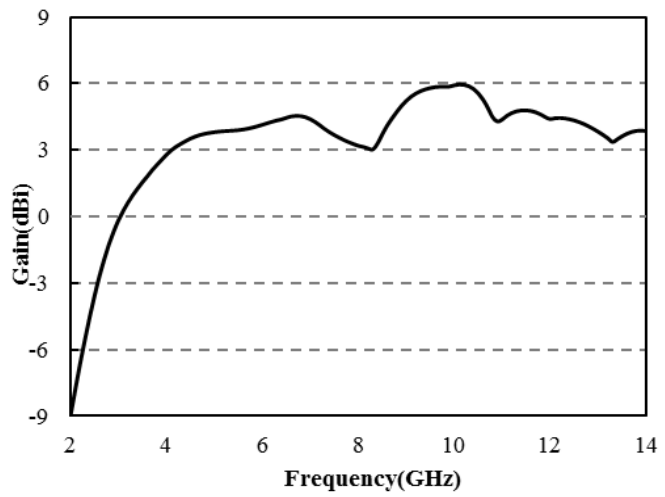


Fig. 5. Peak gain of the proposed antenna.

Fig. 5 displays the gain of the presented antenna. From the gain curve it seen that in 2-14GHz band the reported antenna is able to achieve a peak gain of 5.95Bi. The achievable average gain is 4.03dBi. The efficiency of the designed antenna is displayed in Fig. 6 from where it can be evident that the investigated antenna attained good efficiency. The average efficiency is more than 83% and with this efficiency the antenna can efficiently radiate the signal.

The radiation characteristics of the proporeported antenna is shown in Fig. 7. To explain the radiation characteristics of the proposed antenna, *E*-and *H*-plane patterns are plotted in two diemnsional polar plot. It can be commented from the plot that at lower operating band, the patterns of both the *E*- and *H*-plane are omidirectional. However, at higher frequencies some nulls have been introduces especially in the *E*-palne pattern which may be due to higher order current modes as well as asymetric patch. Despite this nulls, the designed antenna surely exhibits omnidirectional radiation patterns which is a prime requisite for many wireless applications.

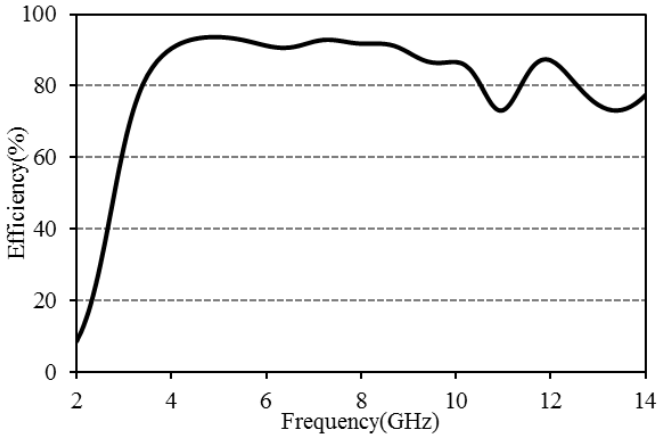


Fig. 6. Efficiency of the proposed design.

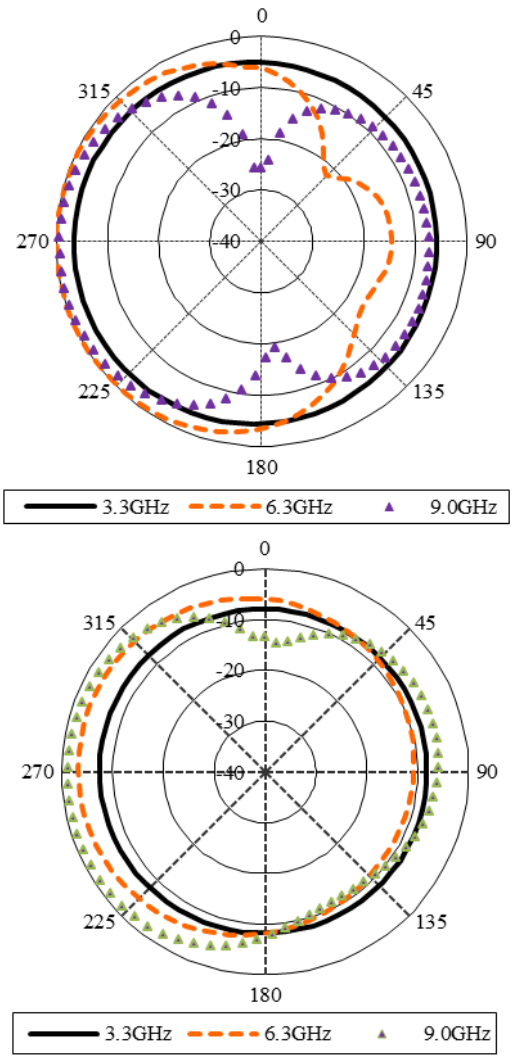


Fig. 7. *E*-(top) and *H*-bottom) field pattern at different frequencies.

IV. CONCLUSION

Planar microstrip antennas are widely used in modern day communication devices due to their numerous advantages. In this article a microstrip patch antenna that consists of a regular shape radiator and a partial defected ground structure is presented. The measured and simulated result shows that the investigated antenna is able to exhibits multiple resonances and combination of these resonances helps to attain a wideband operating band ranging from 2.95 to more than 14GHz. Moreover, it demonstrates symmetric omnidirectional radiation characteristics, good gain and efficiency which make it an appropriate candidate for numerous wireless communication applications.

REFERENCES

[1] IEEE Std 145-1993, " IEEE standard definitions of terms for antennas", The Institute of Electrical and Electronics Engineers, New York, 1993.

[2] D. M. Pozar, "Microstrip antennas," *IEEE Proc*., vol. 80, pp. 79 - 91, 1992.

[3] M. N. Shakib, M. T. Islam, and N. Misran, "Stacked patch antenna with folded patch feed for ultra-wideband application," *IET Microwave Antennas Propag.*, vol. 4, no. 10, pp. 1456 -1461, 2010.

[4] S. N. Ather and P. K. Singhal, "Truncated rectangular microstrip antenna with H and U slot for broadband," *Int J Eng Sci Technol.*, vol. 5, no. 1, pp. 114 - 118, 2013.

[5] L. M. Si, H. Sun, Y. Yuan, and X. Lv, "CPW-fed com-pact planar UWB antenna with circular disc and spiral split ring resonators," *in Proc. of the Progress in Electromagnetics Research Symposium*, Beijing, China, 2009, pp. 502 - 505, 23 – 27.

[6] A. Ghosh, D. Ghosh, S. Chattopadhyay, and L. Singh. "Rectangular microstrip antenna on slot type defected ground for reduced cross polarized radiation", *IEEE Antennas Wirel. Propag. Lett.*, vol. 14, pp. 321-324, 2015.

[7] R. Azim, M. T. Islam and N. Misran, "Design of a planar UWB antenna with new band enhancement technique," *Appl. Comput. Electrom.,* vol. 26, no. 10, pp. 856 - 862, 2011.

[8] R. Azim, M. T. Islam, N. Misran, and A. T. Mobashsher, "Compact UWB planar antenna for broadband aplications", *Inform Midem.*, vol. 41, no.1, pp. 37 - 40, 2011.

[9] C. K Ghosh, S. Mondal, and S. K. Parui, "A compact multiband microstrip antenna using complementary slots on the ground plane," *Microwave Opt Technol Lett*., vol. 58, no. 1, pp. 47-51, 2016.

[10] R. Azim, M. T. Islam, and N. Misran, "Printed circular disc compact planar antenna for UWB applications", *Telecommun Syst.*, vol. 52, no. 2, pp. 1171 - 1177, 2013.

[11] Y. Lu, Y. Huang, H. T. Chattha, and P. Cao, "Reducing ground-plane effects on UWB monopole antennas", *IEEE Antennas Wirel. Propag. Lett.*, vol. 10, pp. 147-150, 2011.

Short Channel Effects Suppression in a Dual-Gate Gate-All-Around Si Nanowire Junctionless nMOSFET

M W Rony,[1,*] Pankaj Bhowmik,[1] Harley R. Myler,[1] *Senior Member*, *IEEE* and Provakar Mondol[2]

[1]Electrical Engineering, Lamar University
Texas, USA

[2]Electrical & Electronic Engineering, KUET
Khulna, Bangladesh

[*]ronyeee44@gmail.com

***Abstract*— Advances in short channel transistor technology has allowed the emergence of these devices in modern chips. These transistors experience many different types of Short Channel Effects (SCE) that are addressed in this work. Using TCAD simulations, we report a new and effective technique to reduce SCE in a novel dual-gate silicon nanowire Gate-All-Around (GAA) junctionless transistor. A channel length of 10 nm and a diameter of 5 nm is employed to mitigate SCE in terms of Subthreshold Slope (SS), Leakage current (Ioff) and Drain Induced Barrier Lowering (DIBL). A 1 nm thick HfO2 gate is used as a dielectric material. Analysis from deep depletion to strong accumulation mode with 2e19 cm^{-3} n-type channel doping (Vds = 0.8V and T = 300 K) is reported. The dual-gate GAA junctionless nanowire transistor shows improved electrical characteristics such as high Ion/Ioff (>10^6), good subthreshold slope (~68 mV/decade) and low drain induced barrier lowering (~8mV/V) compared to conventional inversion mode transistors and single nanowire GAA devices that makes the device a candidate for energy efficiency improvements.**

Keywords*— *Dual-gate GAA, Junctionless Nanowire Transistor (JNT), Si nanowire, Sentaurus 3D TCAD Device

I. Introduction

With increased downscaling of CMOS technology, the formation of sharp gradient drain/source with shorter channel length becomes a barrier for high speed and low power devices. Multi-gate transistors can be a strong candidate for new architectures with low subthreshold slope, immunity against SCEs and optimized power consumption. To overcome the difficulties of abrupt source/drain junction formation, a highly doped architecture with a single type doping profile from source to drain without any junction is proposed that allows best electrostatic control using the GAA concept [1]. Unlike a conventional inversion mode transistor, the proposed dual-gate GAA device simplifies the fabrication process significantly with few limitations; e.g., ultra-abrupt source/drain junctions to fabricate shorter channel devices [2,3]. In these junctionless shorter channel devices, low leakage and high saturation current characteristics are possible through the formation of a silicon channel using a highly doped nanowire [2,4]. As the technology converges on smaller features, the SCEs are becoming prominent and several papers have reported different ways to effectively reduce SCEs. To suppress SCEs new device structures such as dual material gate (DMG) silicon-on-insulator (SOI) MOSFET [5], fully depleted silicon-on-diamond substrate [6] have been proposed. Further, Numata and Takagi [7] have designed a device for subthreshold slope and threshold voltage control and showed that the optimization of the thickness and the permittivity of buried oxides in a fully depleted SOI have a significant impact on the control of the subthreshold slope under sub-100-nm regime. Moreover, Alexandru and Ionescu [8] worked on an analytical model to predict subthreshold slope improvement. Basak, et al. [9], have used a lightly doped source to reduce leakage current in junctionless TFET. However, the reported data are only for characteristics and SCE analysis of long channel junctionless nanowire transistors (JNT).

In this work, a novel short channel Si nanowire GAA junctionless device is simulated in TCAD with a dual-gate GAA architecture. Two different models of GAA Si nanowire nMOSFET with 10 nm gate length were designed; the first one is a well-known accumulation mode Si nanowire GAA junctionless transistor and the second one is a dual-gate Si nanowire GAA junctionless transistor. Short channel electrical behaviors of these two devices are compared. The simulated result of the dual gate GAA architecture shows improved overall performance than that of conventional devices in terms of Ion/Ioff ratio with high on and low off state leakage currents, subthreshold slope and DIBL. The proposed devices are simulated using quasistationary 3D TCAD simulation under the same physical conditions.

II. 3D Device Simulation in TCAD

The Synopsys Sentaurus 3D TCAD (ver. K-2015.06-SP1) device tool was used for a quasistationary numerical device simulation of the silicon nanowire transistors. A density

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

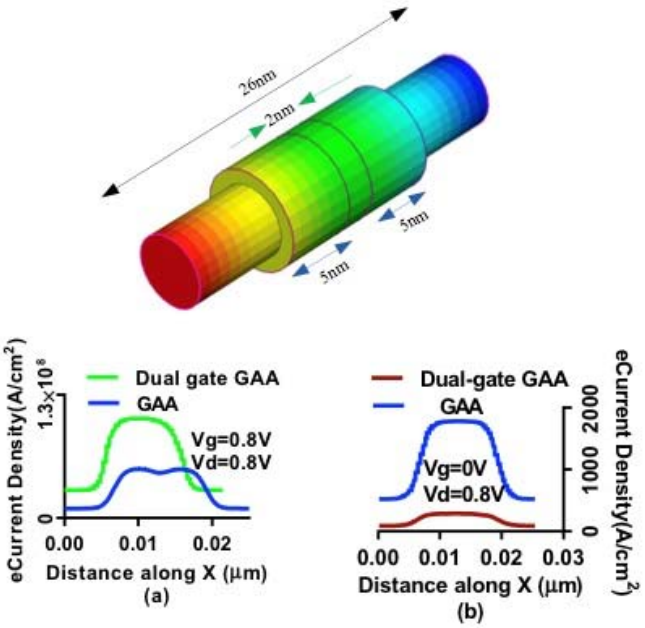


Fig. 1. (top) Dual-gate GAA Si nanowire junctionless transistor, (a) On state and (b) Off state eCurrent density.

gradient quantization model that is coupled to the 3D Poisson equation is employed to solve Ids for different gate and drain voltages in the TCAD tool. Quasi fermi potential and electrostatic potential equations can be used to calculate carrier current and carrier density in a 3D architecture of nanowire nMOS. A Fermi-Dirac statistic is used along with semi-classical Slotboom bandgap narrowing model for the heavily doped silicon region [10,11].

III. Device Simulation

Surface potential in the channel region that injects carrier from drain to source determines electrostatics and scaling criteria of multi gate transistors such as thin body SOI and GAA nanowire MOSFETS. As suggested by Young [12], the electrostatic potential Φ in the channel region can be calculated by using Poisson's equation in the form shown in (1).

$$\Phi\,(r, z) = c_0\,(z) + c_1\,(z)\,r + c_1\,(z)\,r^2 \qquad (1)$$

Where z and r are the direction of current and perpendicular to

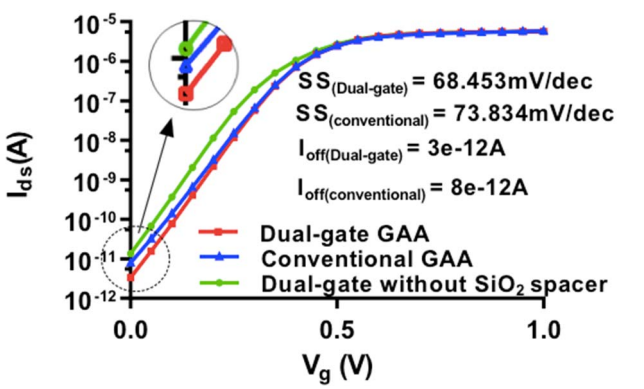


Fig. 2. I_d-V_g characteristics shows an improved subthreshold slope and reduced off state leakage current at Vd=1 and Vg=1 V.

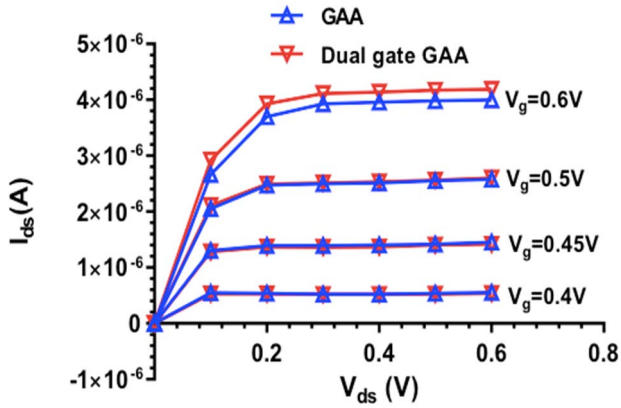


Fig. 3. The simulated output characteristics (I_d-V_d) curve.

the current direction respectively. The solution for the 1D Poison's equation after simplification has an exponential form given by

$$\Phi_S\,(z) \propto \exp\,(-z/\lambda) \qquad (2)$$

Where λ is the natural or screening length. A device will have minimal short channel effects if L_{eff} is approximately 6× longer than λ_N [1]. The analytical solution for the GAA structure is given by (3):

$$\lambda_{GAA-NW} = \sqrt{\frac{2\in_{Si} t_{Si}^2 \ln(1+\frac{2t_{ox}}{t_{Si}})+\in_{ox} t_{Si}^2}{16\in_{ox}}} \qquad (3)$$

Where t_{Si}, ε_{Si}, t_{ox} and ε_{ox} are the thickness and permittivity of Si and dielectric respectively [13]. Lastly, the scaling parameter α is defined as

$$\alpha = \frac{L_{eff}}{2\lambda} \qquad (4)$$

In this study a nanowire of 5 nm diameter is wrapped around with a gate oxide (HfO2) of 1 nm thick and 10 nm long but split into two equal halves that introduces a double gate GAA architecture. This is shown in Fig. 1. A 2 nm long and 1 nm thick SiO2 spacer is used between the two gates that effectively separates the two gate regions. A 7 nm long source/drain extension and 10 nm long channel are heavily doped (2e19 cm^{-3}) with the same kind of n-type material. A mid-gap work function is used here to obtain a desirable threshold voltage that can be changed by gate electrode engineering.

TABLE I
EXTRACTED PARAMETERS FOR DUAL GATE GAA AND GAA NMOSFET

Transistor Parameters	Dual gate GAA	GAA
SS (mV/dec)	68	73
DIBL (mV)	8	13
I_{on} (A/μm)	5.7e-06	5.7e-06
I_{off} (A/μm)	3.369e-12	7.98209e-12
G_{mmax} (S/μm)	1.96552e-05	1.83842e-05
R_{on} (Ω)	50.8772	54.4328
R_{out} (Ω)	4.10112e+06	2.07389e+06

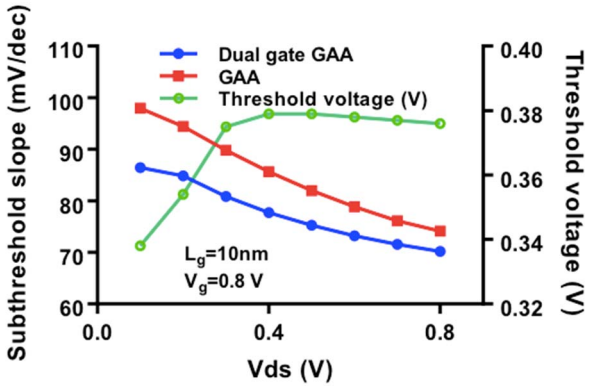


Fig. 4 Subthreshold slope (SS) and threshold voltage (V_{th}) variation with V_{ds}

IV. SCE Parameter Extraction

I_{ds}-V_g characteristics of the 10 nm long nanowire junctionless nMOSFET design is depicted in Fig. 2 and I_{ds}-V_{ds} characteristics for different gate to source voltages up to 0.8 V is shown in Fig. 3. In Table I, the numerical values of extracted SCE and key MOSFET parameters for the simulated devices, using standard methodologies defined in section IV, are reported.

A. *Off-state Leakage Current (I_{off})*

The junctionless dual-gate GAA transistor is characterized along with the regular GAA for various conditions. Figures 1 through 5 show various characteristics for comparison of the n-type GAA JNTs. Under identical conditions, the off state leakage current (I_{off}) of the dual- gate GAA is reduced by 0.005 nA with a high Ion/Ioff ratio (>10^6).

B. *Subthreshold Slope (SS)*

Subthreshold slope can be calculated by taking the inverse of sub threshold swing from the Ids-Vg curve. Fig. 5 describes the variation of subthreshold slope for various drain to source voltages. The junctionless transistor with 10 nm dual-gate exhibits improved performance compared to a regular junctionless transistor in terms of subthreshold slope ($SS = dV_G/d(log(I_D))$ =68 mV/dec, Vd=1V).

C. *Drain Induced Barrier Lowering (DIBL)*

Due to Drain Induced Barrier Lowering (DIBL), the threshold voltage gets lowered, which in turn increases the leakage current. In the dual-gate GAA architecture presented here, the threshold voltage variability due to a small DIBL compared to existing GAA devices reflects a reduction in leakage current. As can be seen in Fig. 4, the threshold voltage starts increasing with a gradual increase in drain to source voltage. This transition indicates that, at low drain voltage, the threshold voltage is lower than that of the drain voltage. As a result, the DIBL is negative and with a higher drain voltage, off state leakage current is low compared to the lower drain voltage. Therefore, at higher drain voltage it develops higher current when the device is on and lower leakage current when the device is off at lower drain voltage. At that point, DIBL (V_{th} (V_d = 0.3) -V_{th} (V_d =1)) is 8mV/V as measured. The steep subthreshold slope, which indicates a higher switching speed and lower threshold voltage variation due to small DIBL, are key factors for future high performance technology. They also contribute to a high I_{on}/I_{off} ratio as the leakage current is small.

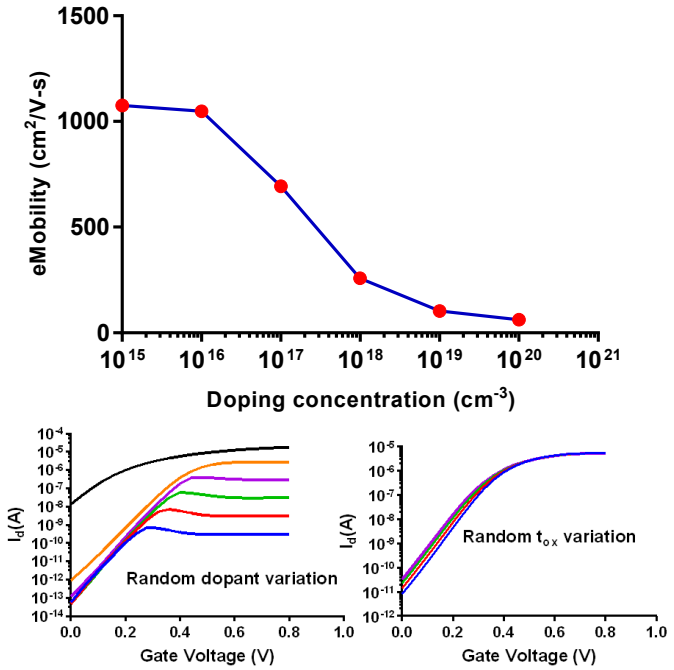


Fig. 5. Mobility degradation and process variability induced by random dopant and oxide thickness fluctuation of dual-gate GAA nanowire device.

V. Process Variation

The electron mobility degradation and the random process variability and is depicted in figure 5, where we can see the decreasing mobility with an increase in doping.

As the scaling of transistors continues the statistical variability of the device properties becomes important because self-averaging of device properties turn out to be less effective. A 3-D nonequilibrium Green's function simulation is performed to analyze the discrete doping profile behavior. It has been argued that due to higher self-averaging effects and lower impact on channel length a junctionless transistor will be less susceptible to doping profile randomization. The studied variability induced by random dopant shows the increasing threshold voltage as the dopant goes high in concentration. The randomization of process variability has been extracted by applying impedance field method in Sentaurus device modeling.

VI. Discussion

Our results show that the dual-gate GAA JNT has a strong SCE immunity compared to regular GAA. In an inversion mode transistor, a depletion region is formed near the interface of the source/drain and channel that reduces the effective channel length of the device. As the channel length is scaled down, the depletion region becomes a larger fraction of the total channel

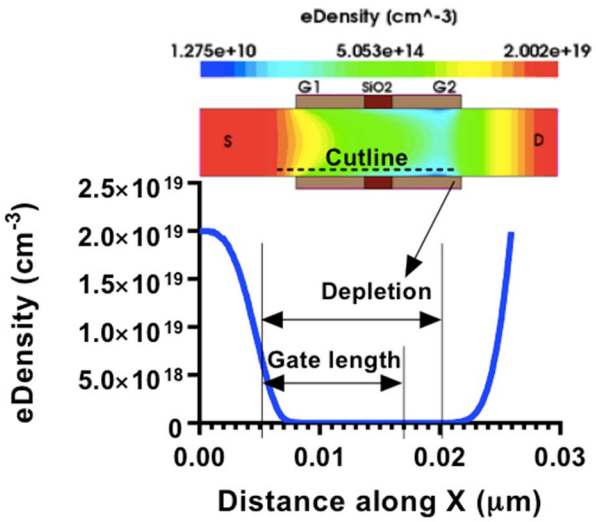


Fig. 6. Electron Density along the X direction shows the depletion region extension during off state of the dual-gate GAA device.

length, which results in degradation of SCEs [14]; however, in the case of a nanowire GAA, only a single doping profile is used throughout the drain-channel-source region with a wraparound gate that provides better electrostatic control. During the off state an electrostatic squeezing effect (ESE) occurs and most of the carriers are pushed away into the drain and source region that results in a depleted channel region. In wrap-around gate transistors, it has been shown that the depletion region results from the electrostatic squeezing effect (ESE) that extends the depletion region longer than the effective channel length [4]. In the case of the dual-gate GAA, we show that the ESE exerts higher electrostatic force from both sides of the SiO2 spacer that extends the depletion region even longer than that achieved in the single-gate design. Fig. 6 shows the simulated depletion contour of the dual-gate GAA JNT and the extended depletion region can be seen clearly.

The higher ESE improves the immunity to SCEs and contributes to the reduction of off state leakage current in short channel devices. On the other hand, our device has a low doping concentration (2e19 cm^{-3}) that leads to a high external resistance (Rex) resistance [15]. Nevertheless, during the on state more carriers are accumulated, resulting in high on current due to the dual-gate GAA design. This work can be further analyzed to reveal the detailed relationship between spacer gap and SCE immunity. The fabrication feasibility of dual-gate GAA JNTs with shorter channel length can also be studied.

VII. Conclusion

A novel dual-gate GAA n-channel JNT design is demonstrated that successfully reduces SCE and improves electrostatic control. SCE parameter extraction assessment was done by following standard methodologies. A calibrated TCAD device simulation was developed and conducted to produce data and curves that show reasonable compatibility to expected results. Moreover, the device shows excellent gate controllability due to a newly introduced dual-gate GAA architecture in short channel devices. Also the mobility degradation and random process variability analysis for dopant fluctuation come up with suppressed SCEs. The randomization induced by dopant fluctuation also refer the excellent behavior for high speedy and low power devices. All experimental data in this research work suggest that dual-gate GAA JNT is a viable candidate for future scaled technology.

References

[1] K. Kuhn. (2012, July) "Considerations for Ultimate CMOS Scaling," in *IEEE Transactions on Electron Devices*, vol. 59, no. 7, pp. 1813-1828. DOI: 10.1109/TED.2012.2193129

[2] J. P. Colinge, C. W. Lee, A. Afzalian, N.D. Akhavan, R. Yan, I. Ferain, & A. M. Kelleher. (2010). Nanowire transistors without junctions. *Nature nanotechnology*, *5*(3), 225-229. DOI:10.1038/nnano.2010.15

[3] J. P. Colinge, D. Lederer, A. Afzalian, R. Yan, C.W. Lee, N. D. Akhavan, & W. Xiong. (2009). Properties of accumulation-mode multi-gate field-effect transistors. *Japanese Journal of Applied Physics*, *48*(3R), 034502. DOI:10.1143/JJAP.48.034502

[4] C. W. Lee, A. Afzalian, N.D. Akhavan, R. Yan, I. Ferain, & J. P. Colinge. (2009). Junctionless multigate field-effect transistor. *Applied Physics Letters*, *94*(5), 053511. DOI:10.1063/1.3079411

[5] A. Chaudhry and M. J. Kumar. (2004, March). "Controlling short-channel effects in deep-submicron SOI MOSFETs for improved reliability: a review," in *IEEE Transactions on Device and Materials Reliability*, vol. 4, no. 1, pp. 99-109. DOI:10.1109/TDMR.2004.824359

[6] A. Daghighi. (2013). "A novel structure to improve DIBL in fully-depleted silicon-on-diamond substrate." *Diamond and Related Materials* 40: 51-55.DOI:10.1016/j.diamond.2013.10.010

[7] T. Numata and S. Takagi. (2004, Dec.). "Device design for subthreshold slope and threshold voltage control in sub-100-nm fully depleted SOI MOSFETs," in *IEEE Transactions on Electron Devices*, vol. 51, no. 12, pp. 2161-2167,. DOI:10.1109/TED.2004.839760

[8] A. Rusu and A. M. Ionescu. (*2012*). "Analytical model for predicting subthreshold slope improvement versus negative swing of S-shape polarization in a ferroelectric FET," *Mixed Design of Integrated Circuits and Systems (MIXDES), Proceedings of the 19th International Conference*, Warsaw, 2012, pp. 55-59.

[9] S. Basak, S., Asthana, P. K., Goswami, Y., & Ghosh, B. (2015). Leakage current reduction in junctionless tunnel FET using a lightly doped source. *Applied Physics A*, *118*(4), 1527-1533. DOI: 10.1007/s00339-014-8935-9

[10] Ancona, M. G., & Iafrate, G. J. (1989). Quantum correction to the equation of state of an electron gas in a semiconductor. Physical Review B, 39(13), 9536. DOI:10.1103/PhysRevB.39.9536

[11] Sentaurus Device User Guide Version K-2015.06-SP1. Synopsys.

[12] K. K. Young, "Short-channel effect in fully depleted SOI MOSFETs," in *IEEE Transactions on Electron Devices*, vol. 36, no. 2, pp. 399-402, Feb 1989.

[13] C. P. Auth and J. D. Plummer, "Scaling theory for cylindrical, fully-depleted, surrounding-gate MOSFET's," in *IEEE Electron Device Letters*, vol. 18, no. 2, pp. 74-76, Feb. 1997.

[14] T. Skotnicki, G. Merckel and T. Pedron. (1988, march) "The voltage-doping transformation: a new approach to the modeling of MOSFET short-channel effects," in *IEEE Electron Device Letters*, vol. 9, no. 3, pp. 109-112. DOI:10.1109/55.2058

[15] C. H. Park, M. D. Ko, K.H Kim, R.H. Baek, C. W. Sohn, C. K. Baek & J. S. Lee (2012). Electrical characteristics of 20-nm junctionless Si nanowire transistors. *Solid-State Electronics*, *73*, 7-10. DOI: 10.1016/j.sse.2011.11.032

A Tunable Energy Efficient UWB Transmitter Design for Medical Applications

Ishmam Mahbub,[1,*] Nouroz Rahman Amon ,[1]and Pran Kanai Saha[1]

[1] Department of Electrical and Electronics , Bangladesh University of Engineering and Technology (BUET)
Dhaka-1205, Bangladesh

[*]Ishmam.Mahbub@yahoo.com

Abstract— **This paper describes two configurable ultra-wideband (UWB) pulse generators for imaging and communication specially targeted for medical applications. Transmitted pulse is controllable in amplitude and duration, enabling flexible shaping of the Power Spectral Density (PSD) of the signal. Both circuits use IR-UWB based on edge combining methods. Pulses of 220ps-380ps (center frequency 5GHz-8GHz) duration and 140mV-320mV peak to peak amplitude is achievable. Both the circuits draw energy of around 4.4pJ/pulse. A maximum data rate of 3.3Gbps is achievable for communication.**

Keywords—ultra-wideband(UWB), medical imaging, high data rate communication.

I. INTRODUCTION

After the revision of part 15 in the FCC regulations for UWB systems in 2002, usage of UWB systems for data communication, ground-penetrating radar (GPR) and imaging became a growing research interest [1]. Specially, in recent times, UWB imaging techniques have shown potential for non-invasive early breast cancer detection [2], cardiology imaging [3], obstetrics imaging [4]. Because of the difference in the electrical properties of different tissues in the UWB range, microwave imaging with UWB techniques are considered a future alternative for traditional imaging methods like X-ray or MRI [5].

In this paper, we have developed an all digital impulse radio (IR) based 5th derivative Gaussian pulse (5th DGP) UWB transmitter. Now, two of the main challenges in transmitter designing are: firstly, whether or not these pulses comply with FCC regulations [1] and secondly, whether or not the whole system is low-power, low-cost. In this regard, the popularly used 1st and 2nd derivative of Gaussian pulse does not satisfy FCC regulations [6]-[7]. It is reported that the 5th derivative of the Gaussian pulse has the most effective spectrum under FCC limitation floor and can be used without filtering [8].

Now, for a battery powered device, the issue of power is very crucial. On this issue, impulse radio (IR) signaling has been chosen for its low complexity, low power implementation [9]-[11]. While simple and small in size, it can reach high data rates. Also, this avoids any need for an up conversion mixer, whose CMOS implementation is known to be extremely power hungry [12]-[13]. Finally, a digital implementation is adopted, because an analog transmitter circuit consumes large amount of static power and also is more complex to implement as compared to a digital circuit.

Existing works on 5th DGP UWB transmitters can be found in [14]-[16]. But these circuits do not provide any pulse shaping facilities. Contrasting these works, the proposed circuit has the ability to control both pulse duration and amplitude. Firstly, these calibration features in the proposed circuit can overcome the problem of variations in the circuit performance due to parasitic capacitances and process variations. Now, a tunable pulse generator can be used for optimizing the absorbed and radiated power of an isolated target. Power optimization is needed for enhancing discrimination and evaluation of electrical properties associated with a target [17]. Tunable generators can be used to achieve various radiation intensities, varying penetrating depth and different range resolutions by changing the shape and center frequency of the pulse spectrum, due to different frequency dependent electrical characteristic of targets.

II. UWB TRANSMITTER CIRCUIT DESIGN

We have presented two topologies for low-power implantable transmitter. Both circuits are impulse radio based and use edge combining approach. The first circuit can be used for imaging purposes and the second circuit for communication. Both circuits are clock independent as used in [18] and for modulation, a combination of Bi-Phase modulation and On-Off Keying is used [18].

Figure 1(a) shows the transmitter circuit for medical imaging. The circuit consists of two identical Gaussian mono-pulse (GMP) generators and control voltage (ctrl_5) is used to control the delay for the overlapping of the two GMP pulses to create 5th derivative pulse. All the delays in the circuit are tuned by current starved inverters as shown in Figure 1(c) and Figure 1(d).

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

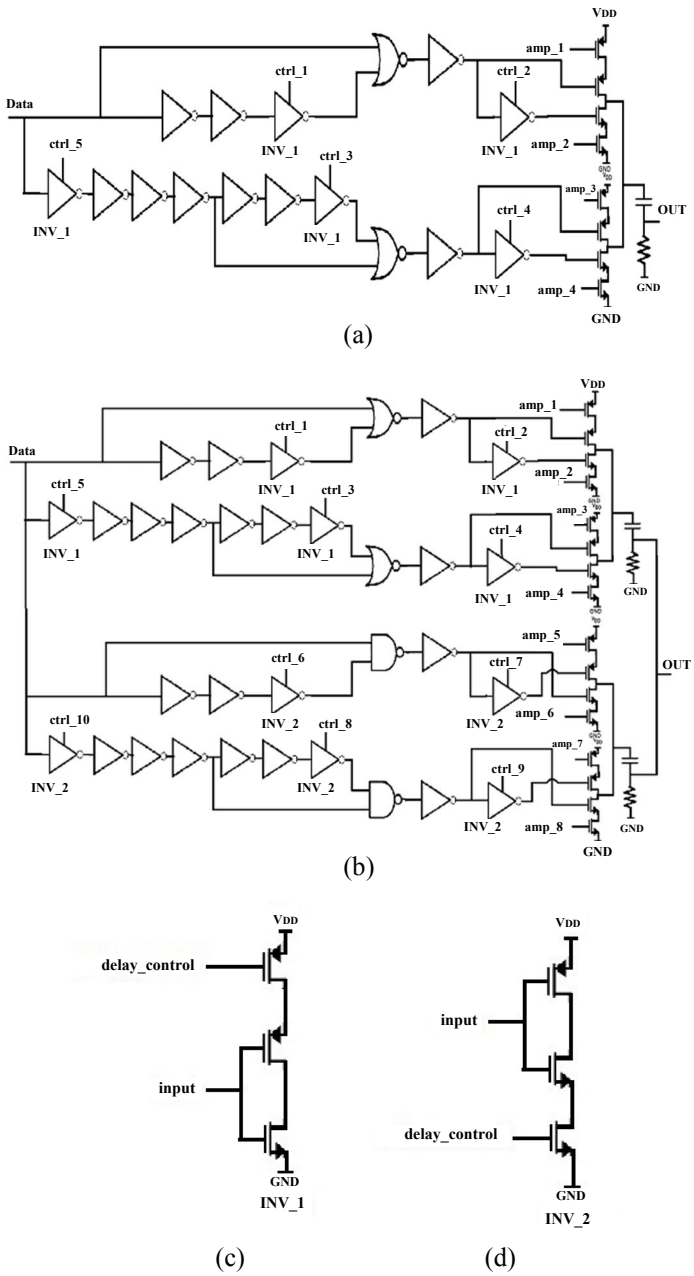


Figure 1: (a) Transmitter for medical imaging (b) Transmitter for data communication. Circuit of (c) INV_1 and (d) INV_2

Now, the circuit does not use any external clock and uses the edges of the input data to create the pulses. The input is delayed, inverted and fed to NOR gate with the input itself to produce a triangular pulse. This delay is governed by control voltage ctrl_1 and ctrl_3. The width of the triangular pulse depends on the delay provided, which thereafter determines the period of the final pulse. So, ctrl_1 and ctrl_3 determines the period of the pulses. The generated triangular pulse and its delayed inverted version (controlled by ctrl_2 and ctrl_4) are fed to NMOS and PMOS transistors to generate two separate GMP. Also as we vary the time period of the final pulse, we need to vary the delay between the triangular pulse and the inverted triangular pulse.

Finally, four current sources, controlled by amp_1, amp_2, amp_3 and amp_4, are provided to control current through the transistors fed by triangular pulses. Thus, the amplitude of each of the four segments of the 5th derivative Gaussian pulse can be dictated. By changing amp_3 to VDD and amp_4 to 0V, this circuit can also work as a GMP generator.

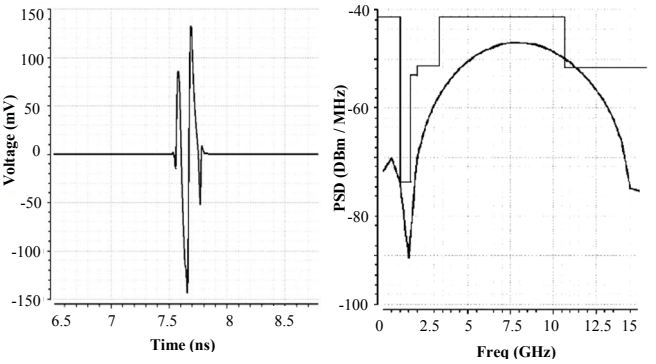
Figure 1(b) shows the transmitter for data communications. For communication, bi-phase signals are required. This circuit is principally same as the one described before, except to create a pulse of the opposite phase, NAND gates are used as opposed to NOR gates. The circuit, to indicate data status, creates two opposite phased pulses for two opposite transitions (0 to 1/ 1 to 0).

III. SIMULATION RESULTS

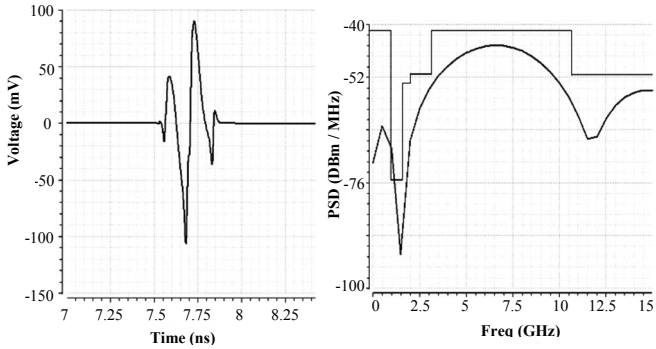
For simulation and circuit design Cadence was used. The transmitter circuits were designed using IBM 90nm Standard CMOS process. For the imaging circuit, the area of the layout was 56µm×28µm. After the simulation, 140mV-320mV peak to peak amplitude was achieved at any center frequency. The time period was adjustable from 220ps-380ps (center frequency 5GHz-8GHz). Figure 2 shows three pulses of 240ps, 315ps and 360ps. The 240ps pulse (Figure 2(a)) takes 884µW total power and energy of 4.42pJ/pulse, the 315ps pulse (Figure 2(b)) takes total power of 911µW and 4.55pJ/pulse energy and 360ps pulse (Figure 2(c)) takes 840µW of power and 4.31pJ/pulse of energy. All power and energy measurements are at a pulse repetition frequency (PRF) of 200MHz. Total power depends on pulse amplitude and data rates. The pulses have -10dB bandwidth covering the entire UWB range.

A 100µm×47µm layout was done for the circuit proposed for communication. Figure 3 shows simulated results of pulse trains at PRF of 1GHz which consumes 4.51pJ/pulse energy and 9mW power. As can be seen from this figure, a positive and a negative edge produces two opposite phased pulses. Two or more consecutive 0s' or 1s' will not produce any pulse. So, the modulation is a combination of Bi-phase modulation and on-off keying. A maximum PRF of 3.3GHz is achievable by the circuit as shown in Figure 4. At PRF of 3.3GHz, total power of 13.25mW is consumed while an energy of 4.41pJ/pulse is needed.

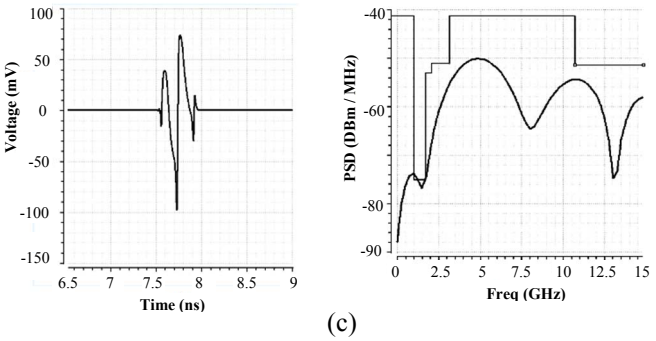
Finally, Table 1 compares this work with other existing works. Comparing these works, it can be remarked that while GMP generators are faster and more energy efficient than the proposed circuit, they have the disadvantage of not complying with the FCC regulations. The proposed circuit has a lower per bit energy consumption and faster data rates than any existing 5th DGP transmitter and provides additional pulse shaping.



(a)



(b)



(c)

Figure 2: (a) 240ps (b) 315ps (c) 360ps UWB pulse and its Power Spectral Density (PSD)

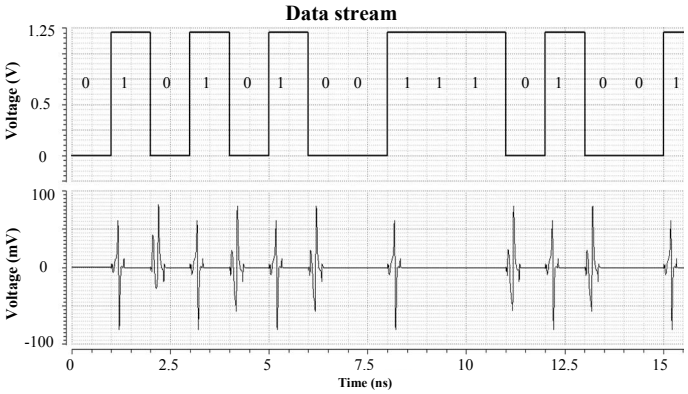


Figure 3: Pulse train at PRF of 1GHz

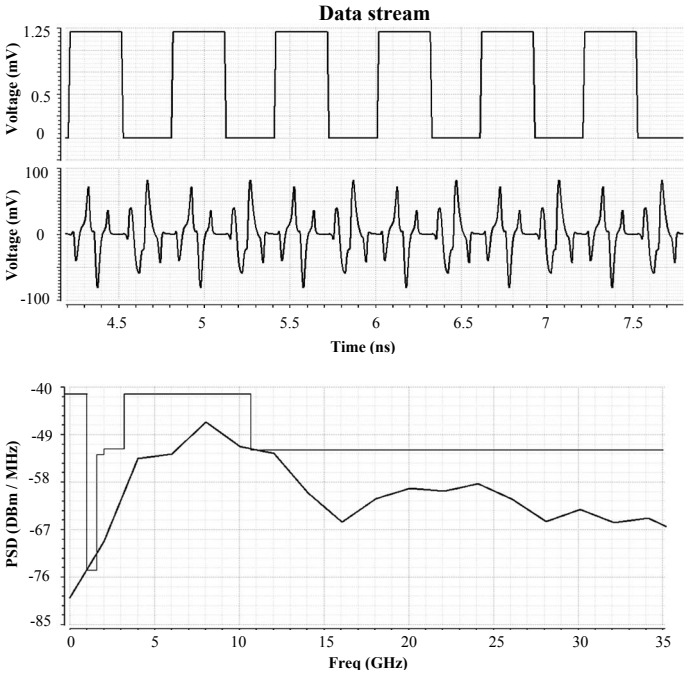


Figure 4: Pulse train at PRF of 3.3GHz and PSD of the pulses.

TABLE I: PERFORMANCE COMPARISON

Ref	Technology (CMOS)	Pulse shape	Modulation	Energy Consumption pJ/pulse	P-P pulse Amplitude (mV)	Pulse Width (ps)	Max data rate (Gbps)	-10dB Bandwidth (GHz)	Power Supply (V)
[19]	90nm CMOS	GMP	BPSK	1.13	500	142	7	3-12	1.2
[18]	130nm CMOS	GMP	BPSK + OOK	2.23	390	183	5	5-10	1.2
[20]	0.18µm CMOS	5th DGP	OOK	8.9	211	500	0.2	3-10	1.8
[21]	0.18µm CMOS	5th DGP	BPSK,OOK	7	140	800	0.500	3.1-7	1.2
[22]	0.5µm CMOS	5th DGP	BPSK	(1.3mW)	289-337	3360-4560	0.01	0-0.96	5
This work	90nm CMOS	5th DGP	BPSK+OOK	4.41	140-320	220-380	3.3	3-12	1.2

IV. Conclusion

The proposed circuits were designed using IBM 90nm standard CMOS process and simulations were done. Pulses of 140-320mv peak to peak and 220-380ps duration are successfully achievable by the circuit. For communication, a maximum of 3.3Gbps data rate is possible with a combination of OOK and BPSK modulation. Energy consumption for both circuits is around 4.4 pJ/pulse.

V. Acknowledgement

We would like to thank the EEE department of Bangladesh University of Engineering and Technology (BUET) for providing constant support throughout the research work.

VI. REFERENCES

[1] Federal Communications Commision, Government Printing Office, Washington.: 'First Report and order: Revision of part 15 of the commision's rules regarding ultra-wideband transmission systems' ET Docket 98-153, April 2002.

[2] Y.Chen E. Gunawan Y.Kim, K. Low C. Soh and L. Thi "The UWB Microwave Breast Cancer Detection: Generalized Models and Performance Prediction" in Proceedings of the 28th IEEE EMBS Annual International Conference New York City, USA, Aug 30-Spet3,2006.

[3] T. E. McEwan, "Body monitoring and imaging apparatus and method", US Patent No. 5,573,012, issued on Nov 12, 1996.

[4] E.M. Staderini, "UWB raders in medicine", IEEE AESS System Magazine, Jan 2002, pp. 13-18.

[5] C. Allen, T. Brown, K. Schwieger, E. Zimmermann, W.Q. Malik, D.J. Edwards, L. Ouvry, and I.Oppermann, "Ultra Wideband: applications, technology and fuure perspectives" in Proc. Int. Workshop Convergent Tech. Oulu, Finland, June 200s.

[6] Stoica, L., Tiuraniemi, S., Rabbachin, A., and Oppermann, I.: 'An ultra wideband TAG circuit transceiver architecture'. IEEE Conf. on Ultra Wideband Sys. and Tech., Kyoto, Japan, May 2004, pp. 258–262

[7] Kim, H., Park, D., and Joo, Y.: 'Design of Scholtz's monocycle pulse generator'. IEEE Conf. on Ultra Wideband Sys. and Tech., Reston, VA, USA, Nov. 2003, pp. 81–85.

[8] Sheng, H, Orlik, P., Haimovich, A.M., Cimini, L.J., and Zhang, J.: 'On the spectral and power requirements for ultra-wideband transmission'. IEEE Int. Conf. on Communications, Anchorage, AL, USA, March 2003, Vol. 1, pp. 738–742

[9] R. J. Fontana, "Recent system applications of short-pulse ultra-wideband (UWB) technology," IEEE Trans. Microw. Theory Tech., vol. 52, no. 9, pp. 2087-2104, September 2004. I.S. Jacobs and C.P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G.T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271-350

[10] A. Batra, J. Balakrishnan, G. R. Aiello, J. R. Foerster and A. Dabak, "Design of a multiband OFDM system for realistic UWB channel environments," IEEE Trans. Microw. Theory Tech., vol. 52, no. 9, pp. 2123-2138, September 2004..

[11] IEEE 802.15.TG4a page [Online]. Available: http://www.ieee802.org/15/ pub/TG4a.html M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[12] H. Miranda and T.H. Meng, "A programmable pulse UWB transmitter with 34% energy efficiency for multichannel neuro-recording systems" Custom Integrated Circuits Conference (CICC), 2010 IEEE, Vol.1, no.4, pp 19-22,2010.

[13] D.D. Wentzloff, "Pulse-based ultra-wideband transmitters for digital communication," Ph.D. Massachuessets Inst. Technol. (MIT), 2007

[14] S. B. T. Wang, A. M. Niknejad, and R. W. Brodersen, "A Sub-mW 960-MHz ultra-wideband CMOS LNA," in IEEE Radio Frequency Integ. Cir. (RFIC) Symp., Long Beach, CA, June 12-14, 2005, pp. 35-38.

[15] R. Blazquez, P. R. Newaskar, F. S. Lee, and A. P. Chandrakasan, "A baseband processor for pulsed ultra-wideband signals," in IEEE Custom Integ. Cir. Conf. (CICC), Orlando, FL, October 3-6, 2005, pp. 587-590.

[16] H. Kim, D. Park and Y. Joo, "All-digital low-power CMOS pulse generator for UWB system", Electronics Letters 25th November 2004, Vol. 40, No. 24.

[17] J. Taylor, "Ultra-wideband radar technology", Boca Raton: CRC press 2001.

[18] P.K. Saha, M.N. Karim, "A versatile pulse rate and low power CMOS IR UWB transmitter for medical imaging", Wireless Symposium (IWS), 2015 iEEE International.

[19] M. A. Arafat , A. B. M. Harun-ur-Rashid , "A novel 7 Gbps low-power CMOS ultra-wideband pulse generator", IET Circuits, Devices & Systems (Volume:6 , Issue: 6).

[20] L. Nagy, J. Radic, A. Djugova, M. Videnovic-Misic, "Ultra low power low complexity tunble 3-10GHz IR-UWB pulse generator", informacije MIDEM – Journal of Microelectronics, Electronic components and Materials, vol. 42, no. 3, pp.185-191, 2012.

[21] H. Bahrami, S.A. Mirbozorgi, L.A. Rusch, B. Gosselin, "Integrated UWB transmitter and antenna design for interfacing high-desnsity brain microprobes", 2015 IEEE International Conference on Ubiquitous Wireless Broadband (ICUWB)

[22] Y. Joo, H. Kim, S. Jung, "A tunable pulse generator for sub-GHz UWB systems",2009 52nd IEEE International Midwest Symposium on Circuits and System.

Low Loss Porous-Core Photonic Crystal Fiber for Long-Haul Broadband THz Transmission

Shekh Farhad Uddin Ahmed*, Sk. Nayemuzzaman, and Mohammad Faisal
Department of Electrical and Electronics Engineering,
Bangladesh University of Engineering and Technology, Dhaka-1205, Bangladesh.
*sk.farhad.eee@gmail.com

***Abstract*—This paper presents a novel porous-core photonic crystal fiber (PCF) having extremely low material absorption loss at terahertz window suitable for THz transmission. Finite element method (FEM) is used to investigate the properties of the proposed PCF with a porous core surrounded by a compact cladding structure having hexagonal air holes. Simulation results reveal that effective material loss (EML) of ~ 0.028 cm^{-1} can be achieved at operating frequency of 1 THz which is only 12.5% of the bulk material loss of Topas. Additionally, the confinement loss of ~ 0.0005 cm^{-1} is obtained at 1 THz which is relatively low with a nearly zero flat dispersion profile of 0.196 ± 0.183 ps/THz/cm which are significant features for long distance broadband THz transmission.**

***Index Terms*—Effective material loss, FEM, Photonic crystal fibers (PCFs), terahertz waveguides.**

I. Introduction

Terahertz (THz) waveguides have drawn huge research interest in recent times. Electromagnetic waves having frequencies from 0.1 to 10 THz are categorized as THz waves. There are many attractive applications of THz waves in the fields of biological imaging [19], spectroscopy, sensing, medical diagnostics, security [20] and modern communication systems etc. Huge progress has been achieved in designing THz waveguides in recent years but most of the present THz systems are inconveniently large and heavily depend on free space radiation. There is still a lack of low-loss transmission waveguides in the THz region since free-space wave propagation suffers many inconveniences such as, misalignment of transmitter and receiver, free-space absorption loss caused by the surroundings atmospheric condition etc. That's why efficient long-haul transmission of broadband THz waves is still a challenge. For overcoming these difficulties, guided transmission was proposed by researchers in the place of unguided transmission [3]. It was observed that THz waves could propagate through all kinds of conventional metallic waveguides [4]. Next, plastic sub-wavelength fibers [5] entered the scene because of their low material losses. Later, researchers came up with photonic crystal fiber (PCF) designs [6], [7], hollow-core fibers [8] and polymer porous-core fibers [9] to reduce the losses of THz waveguides to a much lower level. Among these alternatives, the model of sub-wavelength porous-core fiber demonstrates some advantages, such as, extremely low material absorption loss, flexibility in design and fabrication, and small core diameters [9]. Additionally, the control of dispersion profile is more convenient in those fibers. Because of the sub-wavelength diameter of the core region, the fundamental guided mode remains well confined within smaller air holes in the core area surrounded by the large air cladding and hence low absorption loss is obtained. To date, various porous-core PCF designs with remarkably low material loss properties have been demonstrated. Research trends of THz communication systems is being headed towards designing more efficient porous-core PCFs because of its low loss and high mode power confinement in the air holes of the core region [7], [8]. A novel honeycomb porous-core photonic band gap fiber (PC-PBGF) was proposed by Nielsen et al. [10]. It showed four times lower EML than that of the conventional PBGF. Later, Bao et al. [11] has demonstrated the first fabrication of the honeycomb PC-PBGF. M. Uthman et al. proposed Teflon based porous-core and porous-cladding PCF [12] for THz wave guidance. Recently, Kaijage et al. [13] reported a porous core fiber having low material loss of 0.076 cm^{-1} but some crucial design issues for THz guidance such as dispersion and power fraction were not taken into account. More recently, octagonal porous-core PCF [14] with an EML of 0.054 cm^{-1} and a Kagome PCF having a porous-core [15] with an EML of 0.035 cm^{-1} are also reported along with improvements in the dispersion profile.

In this paper, we propose a novel PCF having a TOPAS-air microstructure for THz wave guiding applications. It contains large hexagonal air holes in the cladding region and a porous core including smaller circular air holes. Important optical waveguide properties of the proposed PCF such as chromatic dispersion, confinement loss, material absorption loss, and power fraction are numerically investigated. The PCF demonstrates an EML of 0.028 cm^{-1} which is lower than any previous designs so far reported [10-15] along with low confinement loss and nearly zero flat dispersion profile. These features suggest that the designed porous PCF could be an attractive candidate for long-distance THz transmission applications.

II. Fiber Construction

Fig. 1 shows the proposed PCF structure with detailed design parameters. The proposed fiber not only shows ultra-low material loss profile but also ensures low confinement

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

loss. The low-loss transmission through this fiber is based on the total internal-reflection so the effects of external environment can be neglected totally. Therefore, the fiber is expected to become a smart media for THz wave guidance.

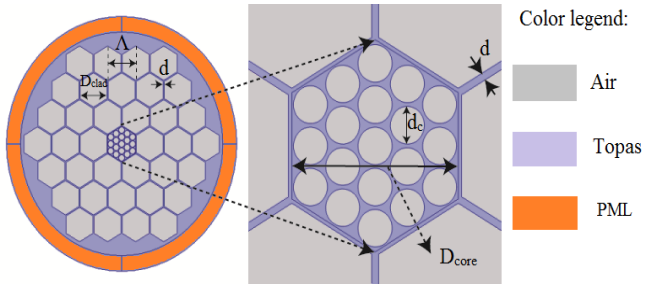


Fig. 1. Geometric cross section of the proposed PCF having hexagonal air holes in the compact cladding area surrounded by a PML and circular air holes in the core.

The background material is TOPAS with refractive index n = 1.53. The diameter of hexagonal air-holes in the cladding region is D_{clad} and the center to center distance (also known as the pitch) of the adjacent air holes is Λ. The spacing between adjacent air-holes in the cladding is termed as the strut length d and it is uniform in the whole cladding structure because of the hexagonal shape of the air holes. This is advantageous as only 3 layers are sufficient for obtaining desired level of guiding property of the proposed PCF. The core is made porous by introducing a hexagonal microstructure having circular air-holes with diameter d_c. The diameter of the hexagonal central core microstructure is D_{core}.

The relation between D_{clad} and D_{core} is, $D_{clad} = D_{core} - d$, and also $D_{core} = \Lambda$. The relation between d_c and D_{core} is, $d_c = K\times(D_{core}\times\cos(\pi/6))/(L\times2\times\cos(\pi/6)+1)$ where, L is equal to the number of layers of circular air holes in the core which is 2 in this case. K is the air filling factor of the core. By varying K, we can change the porosity of the core. The range of possible values of K is from 0 to 1. For K = 0, the core will be solid and for K = 1, the air holes of the core region will touch each other at the boundary. A perfectly matched layer (PML) around the structure is used to absorb the back-scattering wave. The thickness of the PML region used is taken about 10% of the fiber radius.

III. Theoretical Background

The full vector finite element method (FEM) based software COMSOL, version 5.0 is used for simulation. During simulation, the core diameter, D_{core} is varied while keeping d constant, and also d is varied for the same range of D_{core}. The losses of this kind of fiber consist of mainly two components namely the absorption loss and confinement loss. The absorption loss also called Effective Mode Loss (EML) can be calculated by [15]

$$\alpha_{eff} = \frac{(\epsilon_0/\mu_0)^{1/2}\int_{A_{mat}} n\alpha_{mat}\,|E|^2\,dA}{2\int_{All} S_z\,dA} = \alpha_{mat}\eta \qquad (1)$$

where, ε_0 and μ_0 are the permittivity and the permeability respectively in vacuum, n is the refractive index of the material and α_{mat} is the absorption loss of bulk material. For TOPAS $\alpha_{mat} = 0.2$. The integration in the numerator of Eq. (1) is only performed over the solid material region (A_{mat}), while that in the denominator is performed over the entire region. This is due to the fact that air is transparent in the THz regime. In the equation, S_z is the Poynting vector in the direction of propagation which is expressed as $S_z = \mathrm{Re}(E\times H^*)z$, where E is the electric field component and H is the magnetic field component. Also, η indicates the absorption loss ratio of proposed fiber to the bulk material.

The confinement loss of the proposed compact cladding porous core PCF is calculated by using the following expression [10]

$$\alpha_{CL} = \left(\frac{4\pi f}{c}\right)\mathrm{Im}(n_{eff}) \qquad (2)$$

where f is the frequency of the light, c is the speed of light in vacuum and $\mathrm{Im}(n_{eff})$ is the imaginary part of effective index of the guided mode.

Mode power fraction is used to quantify the amount of power propagation through the different regions like air-holes in the core, background material and air-regions in cladding and it is expressed by [15]

$$Power\ fraction\ in\ X\ region = \frac{\int_X S_z\,dA}{\int_{All} S_z\,dA} \qquad (3)$$

where, X in the numerator is identifying all air-holes in that area and the denominator is identifying total background material and air-holes regions at cladding as well as core. It is desirable to pass most of mode power through the air holes of core as air is assumed to be loss free for THz radiation.

IV. Results and Discussions

We have carried out several simulations for studying the nature of loss with the variations of several design parameters. At first, we have varied the strut length along with the core diameter and observed the effect of it on the EML and confinement losses. Figure 2 depicts the trend of both losses. Both types of loss can be reduced by decreasing d as observed in Fig. 2.

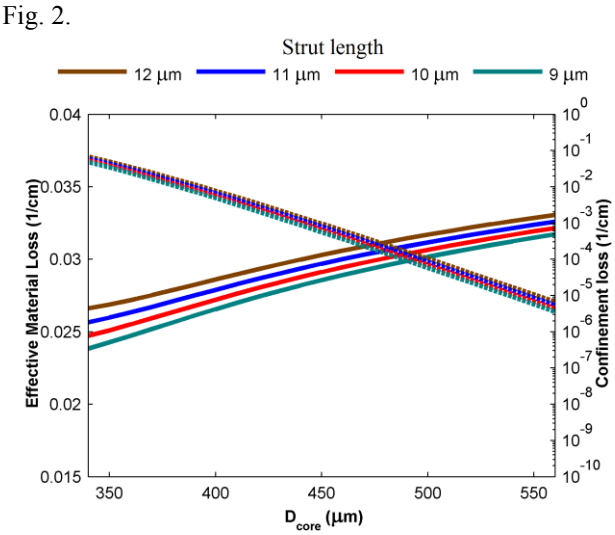


Fig. 2. Variation of effective material loss (solid lines) along with confinement loss (dotted lines) as a function of core diameter for different strut length with a porosity of 78.04% at 1 THz.

We choose the lowest value of d that is 9 μm for later simulations. However, we can not reduce d as we wish as the dimension of strut length is governed by the mechanical strength of the structure.

Now the core diameter D_{core} is varied along with core porosity while keeping strut length, d constant; and the losses are recorded. Figure 3 presents the variation of EML and confinement loss as a function of D_{core} for different core porosities at constant frequency of 1.0 THz. It is seen from the figure that for the same core diameter there is a significant reduction of EML with the increment of core porosity. Increasing core porosity means filling more air into the core and a corresponding decrease in core material which is responsible for loss. Dry air has the lowest material loss in the THz frequency range. Therefore by tuning the air-filling fraction in the core, material or spectral losses can be controlled. However, the increase in core porosity increases the confinement loss of the fiber as the difference between the average refractive indices of core and cladding region decreases with the increase of porosity.

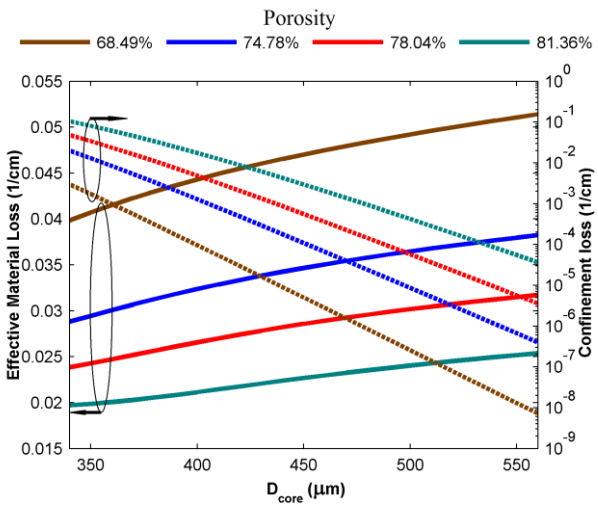


Fig. 3.Variation of effective material loss (solid lines) along with confinement loss (dotted lines) as a function of core diameter for different core porosity.

Mode power fraction is used to quantify the amount of power propagation through the different regions like air-holes in the core, background material and air-region in cladding. It is an important parameter with lots of practical interest when we are to choose the optimal values of the design parameters, such as, core diameter and core porosity for fabrication of the fiber. It's not practical to just use the values giving the lowest value of EML and confinement loss. We need to consider the respective values of power fraction also. In Fig. 4, the variation of fraction of power as a function of D_{core} and core porosity is depicted for 1 THz. We find that with the increase in porosity, power fraction in air-core decreases. This is because with the increase in porosity, material portion of core decreases which reduces the index difference between core and cladding and thereby weakens the light confinement in core. We have to consider the values of EML, confinement loss and the fraction of power in the air holes of core region in a combined way to achieve the optimal design of the PCF. The design that provides the lowest EML causes less than 50% power fraction in the air-core. It won't be much suitable for long distance transmission. From Fig. 3, we can see that at 81.36% porosity, the values of EML are the lowest. But from Fig. 4, it is evident that the fraction of power in air-core is low at this porosity level. Additionally, very high porosity at the core may cause much complication in the fabrication process in addition to increase in confinement loss. So we choose the second highest porosity level which is 78.04%. The optimal value for D_{core} is taken as 450 μm where the fraction of power in air-holes of core gets above 50% which is satisfactory for long haul transmission of THz waves [13], [14]. At the same time, the confinement loss for this porosity and core diameter is 0.0005 cm^{-1} at 1 THz.

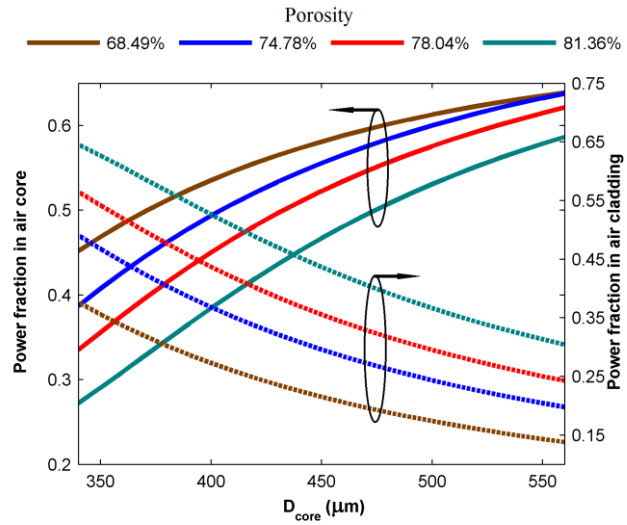


Fig. 4. Variation of power fraction in the air-core (%) along with power fraction in cladding (%) as a function of core diameter for different porosity.

In Fig. 5, the value of total loss consisting of EML and confinement loss is plotted against core diameter with a constant value of strut length of 9 μm. We find that lower total loss could be attainable for higher porosity, however, we consider 78.04% porosity which gives slightly higher loss than that of 81.36% porosity. It should be noted here that the contribution of EML to total loss is much higher than that of confinement loss.

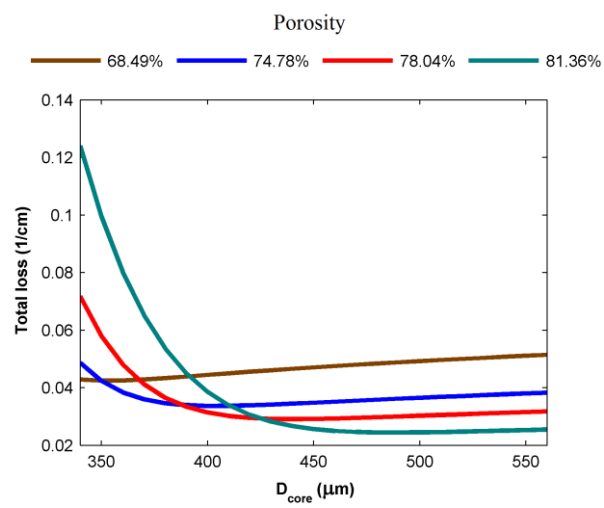


Fig. 5. Variation of total field power loss as a function of core radius and core porosity where the strut length is 9 μm.

Next we examine the dispersion characteristics of the proposed PCF. The dispersion parameter of the PCF can be determined using the equation below [14]

$$\beta_2 = \frac{2}{c}\frac{dn_{eff}}{d\omega} + \frac{\omega}{c}\frac{d^2 n_{eff}}{d\omega} \quad (4)$$

where the central angular frequency, $\omega = 2\pi f$ and f is the frequency of the wave. Figure 6 presents the dispersion profile of the PCF at optimized design parameters. It can be observed that it has a nearly zero flat dispersion over 850 GHz band in THz window. The variation of dispersion parameter β_2 is only ±0.1833 ps/THz/cm around the value of 0.1957 ps/THz/cm in this band which is a significant improvement over the values obtained in [14], [15] and very attractive for transmission applications.

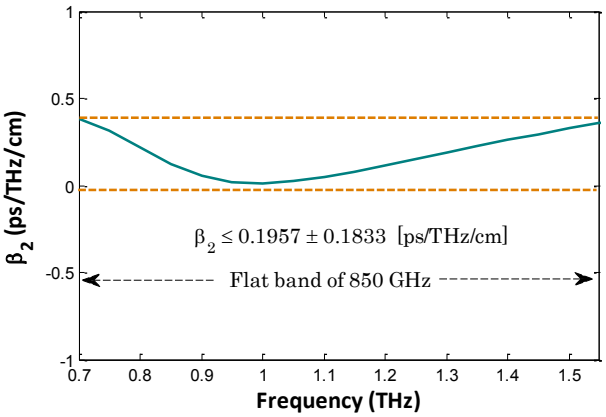


Fig. 6. Dispersion profile of the PCF as a function of frequency with core diameter of 450 µm, core porosity of 78.04% and the strut length is 9 µm.

Thus, we can see that simulation results demonstrate an EML of 0.028 cm^{-1} at the operating frequency of 1 THz with optimized design parameters. This is the lowest value of EML for porous-core PCFs reported so far in the literature. Table I gives a brief comparison with recent works related to porous-core PCF in THz regime.

TABLE I
MOST RECENT PROGRESS IN THZ PCF WAVEGUIDES

Authors	Frequency [THz]	EML [cm^{-1}]
Kaijage et al.[10]	1	0.070
Imran Hasan et al.[14]	1	0.056
Ali et al. [21]	1	0.037
Hasanuzzaman et al.[15]	1	0.035
Our work	1	0.028

The proposed PCF can be easily fabricated using extrusion fabrication process [17] or stack and draw process [16] since the construction is very simple. The only complication may arise due to the hexagonal shape of the cladding air holes but with the advent of latest technologies [18] in the fabrication world, this design can become feasible.

V. CONCLUSIONS

A novel porous-core PCF with compact cladding having hexagonal airholes is designed and analyzed for extremely low loss THz wave guidance. The optimum fiber design parameters confirm effective material loss as low as ~ 0.028 cm^{-1} which is only 12.5% of material absorption loss of bulk Topas, with a very low confinement loss of ~ 0.0005 cm^{-1}at 1 THz. Also a neatly zero flat dispersion profile of 0.196 ± 0.183 ps/THz/cm is demonstrated by this PCF. We therefore believe that the proposed novel PCF could offer significant advantages and improvements over current designs, promising excellent performance for various applications in THz regime.

REFERENCES

[1] M. Tonouchi, "Cutting-edge terahertz technology," *Nature Photon.*, vol. 1, no. 2, pp. 97–105, 2007.

[2] N. F. Ren, M.-Y. Chen, B. Sun, and L. Ge, "Design of subwavelength optical fibre for low-loss terahertz transmission," *J. Eur. Opt. Soc.-Rapid Publications,* vol. 8, pp. 13068-1–13068-4, Oct. 2013.

[3] B. Ung, A. Mazhorova, A. Dupuis, M. Rozé, and M. Skorobogatiy, "Polymer microstructured optical fibers for terahertz wave guiding," *Opt. Exp.,* vol. 19, no. 26, pp. B848–B861, Dec. 2011.

[4] [K. L. Wang and D. M. Mittleman, "Metal wires for terahertz wave guiding," *Nature,* vol. 432, no. 7015, pp. 376–379, Nov. 2004.

[5] L. J. Chen, H. W. Chen, T. F. Kao, J. Y. Lu, and C. K. Sun, "Low loss subwavelength plastic fiber for terahertz waveguiding," *Opt. Lett.,* vol. 31, no. 3, pp. 308–310, Feb. 2006.

[6] A. H. Ghasemi and H. Latifi, "Localized modes in a defectless photonic crystal waveguide at terahertz frequencies," *Opt. Lett.,* vol. 37, no. 13, pp. 2727–2729, Jul. 2012.

[7] K. Nielsen, H. K. Rasmussen, A. J. L. Adam, P. C.M. Planken, O. Bang, and P. U. Jepsen, "Bendable, low-loss Topas fibers for the terahertz frequency range," *Opt. Exp.,* vol. 17, no. 10, pp. 8592–8601, May 2009.

[8] M. Skorobogatiy and A. Dupuis, "Ferroelectric all-polymer hollow Bragg fibers for terahertz guidance," *Appl. Phys. Lett.,* vol. 90, no. 11, pp. 113514-1–113514-3, Mar. 2007.

[9] S. Atakaramians, S. A. Vahid, B. M. Fischer, D. Abbott, and T. M. Monro, "Porous fibers: A novel approach to low loss THz waveguides," *Opt. Exp.*, vol. 16, no. 12, pp. 8845–8854, Jun. 2008.

[10] K. Nielsen, H. K. Rasmussen, P. U. Jepsen, and O. Bang, "Porous-core honeycomb bandgap THz fiber," *Opt. Lett*., vol. 36, no. 5, pp. 666–668, Mar. 2011.

[11] H. Bao et al., "Fabrication and characterization of porous-core honeycomb bandgap THz fibers*," Opt. Exp.*, vol. 20, no. 28, pp. 29507-29517, 2012.

[12] M. Uthman et al., "Design and characterization of low-loss porous-core photonic crystal fiber," *IEEE Photon. J.,* vol. 4, no. 6, pp. 2315-2325, 2012.

[13] S. F. Kaijage, Z. Ouyang, and X. Jim, "Porous-core photonic crystal fiber for low loss terahertz wave guiding," *IEEE Photon. Technol.Lett.,* vol. 25, no. 15, pp. 1454–1457, Aug. 1, 2013.

[14] Md. Imran Hasan, S. M. AbdurRazzak, G. K. M. Hasanuzzaman, Md. Samiul Habib, "Ultra-Low Material Loss and Dispersion Flattened Fiber for THz Transmission" *IEEE Photon. Technol. Lett.*, vol. 38, pp. 546–547, June 2002.

[15] G. K. M. Hasanuzzaman, Md. Selim Habib, S. M. AbdurRazzak, Md. Anwar Hossain, "Low Loss Single-Mode Porous-Core Kagome Photonic Crystal Fiber for THz Wave Guidance," *IEEE Journal of Lightwave Tech.*, vol. 33, pp. 4027–4031, Oct. 2015.

[16] J. Lægsgaard and A. Bjarklev, "Microstructured optical fibers – fundamentals and applications," *Journal of the American Ceramic Society*, vol. 89, pp. 1–12, Jan. 2006.

[17] K. M. Kiang, K. Frampton, T. M. Monro, R. Moore, J. Tucknott, D. W. Hewak, D. J. Richardson, and H. N. Rutt, "Extruded singlemodenonsilica glass holey optical fibres," Electronics Letters, vol. 38, pp. 546–547, June 2002.

[18] R. T. Bisen and D. J. Trevor, "Solgel-derived microstructured fibers: Fabrication and characterization," in Proc. Opt. Fiber Commun. Conf. (OFC), Mar. 2005, p. OWL6.

[19] R. M. Woodward*et al*., "Terahertz pulsed imaging of skin cancer in the time and frequency domain," *J. Biol. Phys*., vol. 29, no. 2/3, pp. 257–259, 2003.

[20] D. J. Cook, B. K. Decker, and M. G. Allen, "Quantitative THz spectroscopyof explosive materials," in Proc. OSA Conf., Opt. Terahertz Sci. Technol., pp. 1–4, Mar. 2005.

[21] S. Ali, N. Ahmed, S. Aljunid, B. Ahmad, "Ultra-flat low material loss porous core THz waveguide with near zero flat dispersion," *Electron.Lett.*, vol. 52, pp. 863 - 865, May 2016.

Highly GeO_2 Doped Polarization Maintaining PCF for Dispersion Compensation

Sk. Nayemuzzaman[1], Shekh Farhad Uddin Ahmed, Md. Imrul Hasan and Mohammad Faisal
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology, Dhaka-1205, Bangladesh
[1]nayem.buet@gmail.com

***Abstract*—In this paper a highly GeO_2 doped single mode equiangular spiral photonic crystal fiber (HDES-PCF) is proposed to achieve high birefringence with extremely high negative dispersion, higher nonlinearity with low confinement loss for E+S+C+L+U band of interest. From the simulation results, large negative dispersion of −602.7803 ps/nm/km at 1.33 µm and −903.5316 ps/nm/km at 1.55 µm is achieved. High birefringence 5.8×10^{-2} at 1.33 µm, 5.3×10^{-2} at 1.55 µm with very low confinement loss is also achieved. Our proposed structure is also suitable for nonlinear applications because of nonlinearity ~133.9298 $W^{-1}km^{-1}$ is achieved at 1.33 µm and ~79.1229 $W^{-1}km^{-1}$ is achieved at 1.55 µm wavelength. The proposed PCF can be used in dispersion compensation, optical sensor, higher bit rate coherent communication, and polarization maintaining fiber etc.**

***Index Terms*— Birefringence, Photonic crystal fiber, Dispersion compensation, Finite element method**

I. Introduction

Photonic crystal fiber is a new technology that paves the way in optical communication due to its unique properties. In comparison to conventional optical fiber, PCF shows basic properties like birefringence, nonlinearity and chromatic dispersion that can be tailored to achieve extraordinary outputs. For example, birefringence can be produced by adopting asymmetric structures in the PCF design. Many methods can be applied to increase birefringence. The fiber core can be designed to be asymmetrical (e.g., double or triple defects in the fiber core). High birefringence up to the order of 10^{-2} is found in recent research works. Birefringence of 0.01544 is reported in [1], 0.0315 is reported in [2] and 0.022 is reported in [3]. Higher value of birefringence is a key concern for coherent communication and sensing application etc.

Nonlinearity is also an important property of PCF that was also discussed in the reports mentioned above. Nonlinear coefficient of 58 $W^{-1}km^{-1}$ and 68 $W^{-1}km^{-1}$ is reported in [2], [3] respectively. However higher nonlinearity is achieved using GeO_2 doped core. Spiral design with higher nonlinearity of is reported for visible range [4]. The ES-PCF is much compact and simple as compared to conventional PCF. So increasing nonlinearity will let us use PCF for nonlinear applications like four-wave mixing, supercontinuum generation, and second harmonic generation. In addition, it is also desired to decrease the complexity of PCF design.

Standard single mode fiber (SSMF) used in the long distance transmission channel can face pulse broadening due to its higher value of positive dispersion. It limits the transmission capacity. To mitigate this problem dispersion compensating fiber (DCF) in the transmission link can be used as an effective way to compensate the positive dispersion in the system. This is necessary to neutralize the dispersion as it creates great trouble in higher bit rate transmission. The main purpose of DCF should have large negative with flattened dispersion [5]. Also higher birefringence is necessary to maintain the polarization state which is useful in first order polarization mode dispersion (PMD) compensation. Birefringent PCFs also have potential applications in signal processing systems, electro-optical modulation systems and optical sensor designs [6]. Researchers around the world worked past years on birefringence and negative dispersion. Silva *et al.* in [7] proposed a Ge doped core PCF having low negative dispersion of −212 ps/nm/km that demands long fiber requirements. The DCFs proposed by Islam *et al.* based on equiangular spiral photonic crystal fibers (ES-PCFs) in [8] and [9] provide average negative dispersion of −227 and −393 ps/nm/km respectively. Another DCF reported by Tee *et al.* [10] that exhibits high average negative dispersion of −457.40 ps/nm/km. Moreover, polarization properties were not considered here. Recently, an octagonal PCF has been proposed in [11] to achieve large negative dispersion of −558.96 ps/nm/km. This structure has different sized elliptical air-holes that make the fabrication process too complex. So having large negative dispersion is a challenging goal for dispersion compensating applications. In this paper, an equiangular spiral PCF is designed to achieve suitable dispersion property along with highly GeO_2 doped core for higher nonlinearity. The elliptical core offers higher birefringence.

II. Design Methodology and Fabrication

Figure 1 shows the cross-section of proposed Highly Doped-Equiangular Spiral PCF (HDES-PCF) with concentrated GeO_2 doped elliptical core in the center. The cladding structure has gradually increasing air-holes C1, C2, C3, C4, C5, C6, C7 and C8 as in Fig. 1. The host material is silica. Two ellipses E0, E1 and two extra air-holes A0, A1 are added to introduce asymmetry in the design. This PCF has 6 arms having 8 rings per arm with equiangular spiral design. The position of the center of each circular air-hole is adopted from the following equations.

$$r_{spiral} = r_0 . e^{\theta cot\alpha} \quad (1)$$

$$\alpha = \frac{(n-2)\pi}{2n} \quad (2)$$

Where α is the constant angle between a radius of the equiangular spiral and the tangent at the end point of the radius and r_{spiral} is the distance at any point along the spiral arm from the centre of the structure. Each air-hole is separated

by an angular displacement of θ with respect to the previous successive air-hole in the same arm and $\boldsymbol{n}$ is the number of arms.

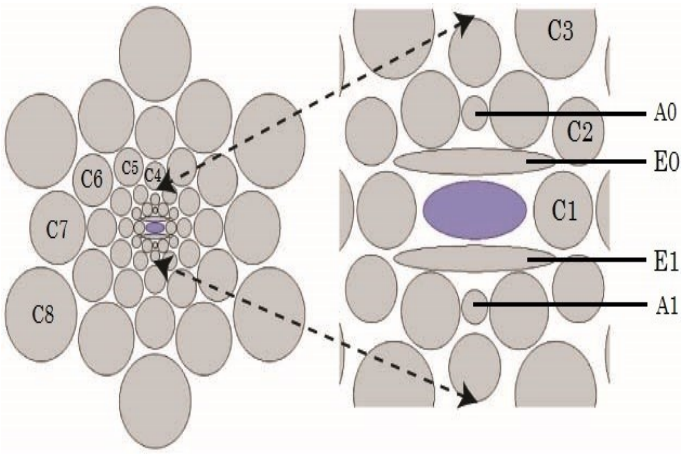


Fig. 1: Geometry of our proposed PCF

The total number of air-holes in each arm represent the number of rings in the HDES-PCF and the total number of air-holes is thus simply the product of the number of arms with the number of rings. The *x, y* coordinates of the centre of each air-hole can be found from Equation (1). The HDES-PCF structure shown in Fig. 1 has 6 arms and 8 rings per arm with $\theta = 30^{\circ}$. In our spiral design r_0 is 0.9 µm. The major axis of the elliptical core is 0.525 µm and minor axis is 0.225 µm. So its ellipticity ratio is 2.33 and the cladding air-hole radius C1, C2, C3, C4, C5, C6, C7 and C8 are 0.3, 0.263, 0.413, 0.6, 0.825, 1.125, 1.575 and 2.03 µm respectively. Radius of the two extra air-holes A0, A1 in the design is 0.13 µm for each. A0 is 0.75 µm from origin in the +y axis and A1 is −0.75µm is in the −y axis. Two extra added ellipses E0 and E1 are used to increase the amount of asymmetry to achieve higher birefringence. The major axis of both ellipses is 0.825 µm and minor axis is 0.105 µm. In Cartesian co-ordinate system the position of the center of E0 is (0, 0.375) and E1 is (0, −0.375). Ellipticity ratio of both ellipses is 7.86. In order to reduce the backscattering and absorb the leakage at the boundary, a circular perfectly matched layer (PML) of 1.5 µm thickness is provided around the structure.

Fabrication of PCF is a key process. Different fabrication techniques have been proposed so far. Bisen et al. [12] developed the sol-gel casting technique to fabricate PCFs with any structure and the air-hole size, shape and spacing may all be adjusted independently. So this method can provide additional design flexibility that will be necessary for our proposed PCF. Doping of high concentration of GeO_2 in silica glass can be done by modified chemical vapour deposition [13].

III. Results and Discussion

Simulations have been carried out using COMSOL Multiphysics 5.0 using FEM to obtain the optical properties of the PCF. Dispersion was calculated by the equation as reported in [14]. Wavelength dependent refractive index of SiO_2 can be found using Sellmeier's equation [15]. Refractive index of GeO_2 doped core can be found from [16]. We have investigated our simulation using 50% mole GeO_2, where the nonlinear refractive index, n_2 for GeO_2 doped SiO_2 can be calculated by [17]. Fiber birefringence is found from $B = |n_x - n_y|$, where n_x, n_y are the real part of effective refractive index for the x-polarized mode and y-polarized mode respectively.

The fundamental optical field distribution for x- and y-polarized modes at the operating wavelength of 1550 nm is shown in Fig. 2. According to simulation, it is seen that x- and y-polarized modes are strongly tight in the high-index center core region.

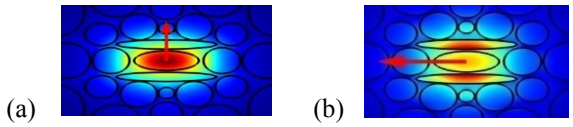


Fig. 2: Field distributions of fundamental modes for (a) y and (b) x polarization at 1550 nm.

The asymmetric design is created by adding 2 extra ellipses and 2 extra air-holes that significantly increase linear birefringence. Figure 3 shows the wavelength dependence of birefringence of the proposed structure for optimum parameters for different air-hole diameter. Birefringence of 5.8×10^{-2} at 1.33 µm and 5.3×10^{-2} at 1.55 µm are obtained. These higher values are suitable for polarization maintaining (PM) applications. Equiangular spiral design has no constant pitch Λ as it's pitch exponentially increases. So in our design only first 4 dominating rings were considered and their average pitch is calculated Λ = 0.746 µm. We successfully investigated the effect of changing radius. A tolerance of ±2% in air-hole diameter is considered. From Fig. 3, we observed that increasing air-hole radius increases birefringence. Per 1% radius changes 0.001 birefringence changes because the birefringence rapidly increases due to index difference between two orthogonal polarization axes. Tightly constructed air-holes create much refractive index difference between orthogonal axes. As the structure is quite congested, so optimum value is chosen keeping fabrication limitation in mind. The elliptical core is 50% GeO_2 doped silica. The core shape has huge effect on birefringence. It is clear that elliptical doped core shows higher birefringence than circular doped core. The extra added ellipses and circular air-holes have larger effect on birefringence.

In this paper we have also investigated the dependence of birefringence over doping concentration of core. Increasing doping concentration of GeO_2 increases the refractive index of the core. So the effective refractive index increases that creates large index difference between two orthogonal polarized modes.

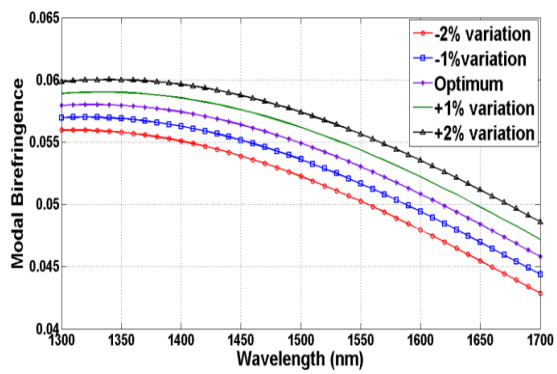


Fig. 3: Birefringence as a function of wavelength for different cladding air-hole radius for y-polarized mode.

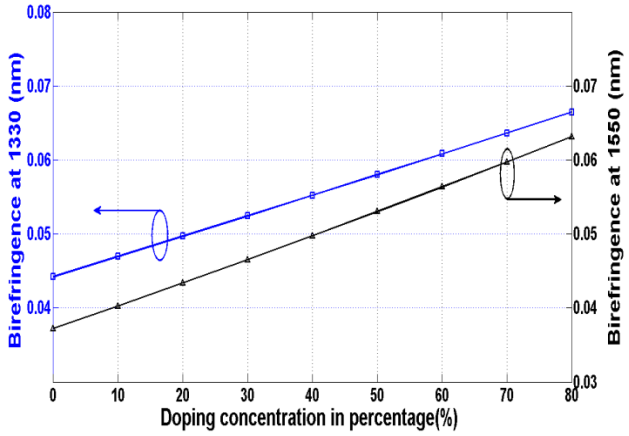


Fig. 4: Birefringence at 1330 nm and 1550 nm vs doping concentration in percentage.

Figure 4 explains the dependence of birefringence with respect to doping concentration. From Fig. 4 it is clear that for both operating wavelength birefringence is linearly increasing. It is observed that per 10% doping increases, average birefringence of 0.003 is increased. In our design 50% doping is kept as optimum because doping is a difficult fabrication process. Furthermore, very high doping increases scattering loss, intrinsic loss and heating problem as well. In order to achieve large negative dispersion, the diameter of the air-hole in the first ring is chosen higher than the second ring because it directly affects the dispersion characteristics. Figure 5 shows wavelength dependence of dispersion of the proposed HDES-PCF for y-polarized mode with optimum design parameters. y-polarized mode is chosen as the operating mode, because x-polarized mode is lossy and also exhibits low nonlinearity. In our study, we have assumed the average pitch, Λ as 0.746 µm, radius of the first ring 0.3 µm, second ring is 0.263 µm. The rest rings are increased gradually. According to the simulation results, the proposed structure shows a negative dispersion coefficient of about −602.7803 ps/nm/km at 1.33 µm and −903.5316 ps/nm/km at 1.55 µm. Lowering the spiral radius r_0 increases the negative value of dispersion [18] but lowering r_0 lowers birefringence. So to cope with opposite condition optimum r_0 is chosen as 0.9 µm. In our simulation we have demonstrated the radius variation of all circular air-holes. From Fig. 5 it is observed that increasing radius of air rings decreases negative dispersion and decreasing air hole radius increases negative dispersion.

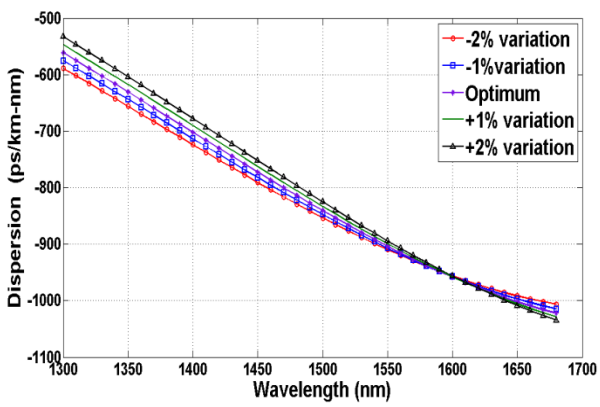


Fig. 5: Dispersion as a function of wavelength for different cladding air-hole radius for y-polarized mode.

After observing the dispersion curve from Fig. 5 we then examined the dispersion accuracy of the proposed design. In PCFs, ±1% variations in fiber global diameter may occur [19] during the fabrication process. So it is possible that the global diameter variation of± 2% may require ensuring dispersion flatness [20]. From the observation it is clear that variation of radius does not change dispersion abruptly. So our design is fabrication imperfection protected. Confinement loss of our designed fiber is very low in comparison to the number of air-holes. Large air-hole in the seventh and eight rings ensures low confinement loss in the fiber. Also GeO_2 has higher refractive index than silica. So it can confine light better than silica glass. At 1550 nm confinement loss is ~0.00024 dB/m. If we increase the number of air-holes in the cladding region, we can reduce confinement loss. From Fig. 6 +1% radius variation decreases loss to ~0.000059 dB/m and -1% variation increases loss up to 0.0011 dB/m.

Fiber nonlinearity is an interesting topic that is hugely investigated for its numerous applications like optical parametric amplification and supercontiuum generation. Lowering the core size has huge effect on nonlinearity. Spiral design gives the flexibility to control fiber nonlinearity [21]. Using higher index material with large nonlinear refractive index we can achieve large non linearity with small effective area. That is why GeO_2 is used to increase the nonlinearity. The radius of the air ring near the core is kept higher to achieve large nonlinearity. Figure 6 demonstrates the air-hole radius variation and its effect on nonlinearity.

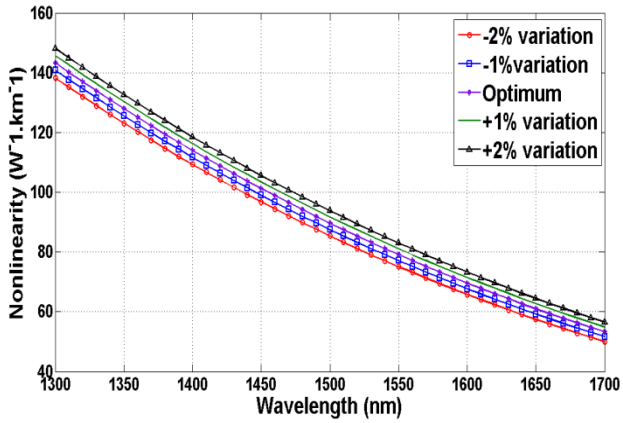


Fig. 6: Nonlinearity as a function of wavelength for different cladding air-hole radius variation for y-polarized mode.

Changing radius has very small effect on nonlinearity. The single modeness of the proposed fiber can be verified with the help of V parameter. V_{eff} parameter of the proposed fiber can be calculated by using the following formula [22].

$$V_{eff} = \frac{2\pi\Lambda}{\lambda}\sqrt{n_{core}^2 - n_{eff}^2} \quad (3)$$

Here n_{core} is refractive index of the doped core and n_{eff} is the effective refractive index that is found from FEM. From Fig. 7 it is ensured that in the whole band of operation our proposed fiber is single mode, that is V_{eff} is always less than π. As our proposed fiber is an efficient candidate for DCF, it has to be attached to SMF. The splice loss occurs due to due to mismatch of mode field diameters (MFD) of SMF and DCF. So splice loss (*Ls*) has to be considered. Splice loss is calculated using the formula [23],

$$Ls = 20\,Log\left(\frac{1}{2}\frac{MFD1}{MFD2} + \frac{1}{2}\frac{MFD2}{MFD1}\right) \quad (4)$$

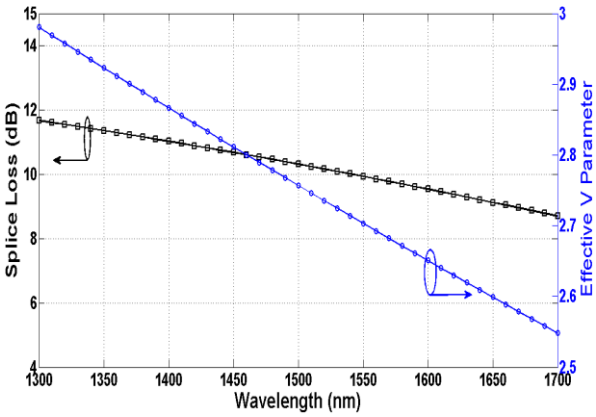


Fig. 7: Effective V parameter and Splice loss (dB) vs. wavelength.

Where, MFD1 and MFD2 are the mode field diameters of the proposed PCF and SMF, respectively. Overall diameter of a SMF is likely 8 µm to 10 µm. In our investigation MFD2 is taken 8µm. MFD1 is found from effective area. Figure 7 explains the splice loss with respect to wavelength. Increasing wavelength increases effective area and thus decreases loss. At 1550 nm splice loss is 10.0106 dB. This loss can be reduced by using splice-free interconnection technique [24].

Finally, a comparison table has been made between guiding properties of the proposed HDES-PCF and some other fibres designed for dispersion compensation at 1550 nm wavelength. Table. 1 shows the comparison taking into account birefringence (B), chromatic dispersion (D) and nonlinear coefficient (γ).

TABLE 1

Reference	D (ps/nm/km)	B × 10^{-2}	γ ($W^{-1}km^{-1}$)
[1]	-	1.544	48
[2]	−25	3.15	58
[3]	−110	2.20	68
[18]	−327	1.79	41
Our work	--903.532	5.3	79.12

IV. Conclusion

The proposed HDES-PCF exhibits ultra-high negative dispersion of −903.5316 ps/nm/km and birefringence of 5.3×10^{-2} at operating wavelength of 1.55 µm; and dispersion of −602.7803 ps/nm/km with birefringence of 5.8×10^{-2} at 1.33 µm. This design provides very low confinement loss and higher nonlinearity. So our proposed HDES-PCF may be a potential candidate for broadband dispersion compensation in high-bit-rate transmission network, coherent communication, and polarimetric sensing applications etc.

.

References

[1] H. Rui, "Highly birefringent photonic crystal fiber with a squeezed core and small modal area," *Optik - International Journal for Light and Electron Optics*, Volume 127, Issue 13, July 2016.

[2] Md. Rabiul Hasan , Md. Shamim Anower and Md. Imran Hasan "Polarization maintaining highly nonlinear photonic crystal fiber with closely lying two zero dispersion wavelengths", Opt. Eng., vol.55, no. 5, 056107, May 2016.

[3] T. Yang, E. Wang, H. Jiang, Z. Hu, and K. Xie, "High birefringence photonic crystal fiber with high nonlinearity and low confinement loss," *Opt. Express,* vol. 23, pp. 8329-8337, 2015.

[4] M. N. Hossain, M. S. Alam, D. M. N. Hasan, and K. M. Mohsin, "A highly nonlinear spiral photonic crystal fiber for tailoring two zero dispersion wavelengths in the visible region," *Photonics Letters of Poland*, vol. 2, no. 3, pp. 143–145, 2010.

[5] M. S. Habib, R. Ahmad, S. Habib, and M. I. Hasan, "Residual dispersion compensation over the S + C + L + U wavelength bands using highly birefringent octagonal photonic crystal fiber," *Appl. Opt.*, vol. 53, no. 14, pp. 3057-3062, 2014.

[6] M. R. Hasan, M. S. Anower, and M. I. Hasan, "Polarization maintaining highly nonlinear photonic crystal fiber with closely lying two zero dispersion wavelengths," *Optical Engineering*, vol. 55, no. 5, pp. 056107, May. 2016.

[7] J. P. Silva et al., "Ge-doped defect-core microstructured fiber design by genetic algorithm for residual dispersion compensation," *IEEE Photon. Technol. Lett.*, vol. 22, no. 18, pp. 1337–1339, Sep. 2010.

[8] M. A. Islam et al., "Design of a polarization-maintaining equiangular spiral photonic crystal fiber for residual dispersion compensation over E+S+C+L+U wavelength bands," *IEEE Photon. Technol. Lett.*, vol. 24, no. 11, pp. 930–932, Jun. 2012.

[9] D. C. Tee et al., "Photonic crystal fiber in photonic crystal fiber for residual dispersion compensation over E+S+C+L+U wavelength bands," *IEEE Photon. J.*, vol. 5, no. 3, art. ID 7200607, Jun. 2013.

[10] M. S. Habib, and E. Khandker "Highly birefringent photonic crystal fiber with ultra-flattened negative dispersion over S + C + L + U bands," *App. Opt.*, vol. 54, no. 10, pp. 2786-2789, April 2015.

[11] A. Agrawal, J. Kejalakshmy, M. Uthman, B. M. Rahman, and K. T. Grattan, "Ultra Low Bending Loss Spiral Photonic Crystal Fibers in Terahertz Regime," in CLEO: 2011 - Laser Applications to Photonic Applications, OSA Technical Digest (CD) (Optical Society of America, 2011), paper CThN1.R.T.

[12] Bisen, D.J. Trevor, "Sol–gel-derived microstructured fibers: fabrication and characterization", in: *Optical Fiber Communication Conference (OFC) (Optical Society of America, 2005),* 2005, Paper OWL6.

[13] R. Sidharthan, S. Yoo, D. Ho, L. Zhang, W. Qi, M. S. Yue, L. Zhu, X. Dong, and S. C. Tjin, "Fabrication of 74 mol% GeO_2-Doped Fibers and Mid-IR Supercontinuum Generation," in *Conference on Lasers and Electro-Optics*, OSA Technical Digest (2016) (Optical Society of America, 2016), paper SM2R.4.

[14] K. Saitoh and M. Koshiba, "Ultra-flattened chromatic dispersion controllability using a defected core photonic crystal fiber with low confinement loss." OPTICS EXPRESS 8365, Vol. 13, No. 21.(2005).

[15] J. W. Fleming, "Dispersion in GeO_2-SiO_2 glasses," Appl. Opt., vol. 23, no. 24, pp. 4486–4493, Dec. 1984.

[16] A. Boskovic, L. Gruner-Nielsen, O. A. Levring, S. V. Chernikov, and J. R. Taylor, "Direct continuous-wave measurement of n_2 in various types of telecommunication fiber at 1.55 µm," *Opt. Lett.,* vol. 21, no. 24, pp. 1966–1968, Dec. 1996.

[17] Y. Yatsenko and A. Mavritsky, "D-scan measurement of nonlinear refractive index in fibers heavily doped with GeO_2,"*Opt. Lett.,* vol. 32, no. 22, pp. 3257–3259, Nov. 2007.

[18] M Samiul Habib, K M Nasim, M Selim Habib, M I Hasan, Redwan Ahmad (2013),"Relative dispersion slope matched dispersion compensating highly birefringent spiral microstructure optical fibers using defected core," *Optical Engineering,* 52: 9. 096110-5. September.

[19] S. M. A. Razzak et al., "Design of a decagonal photonic crystal fiber with ultra-flattened chromatic dispersion,"*IEICE Trans. Electron.*E90-C(11), 2141–2145, 2007.

[20] F. Poletti et al., "Inverse design and fabrication tolerances of ultraflattened dispersion holey fibers," *Opt. Exp.* 13(10), pp.3728–3736 ,2005.

[21] S. Revathi, Srinivasa Rao Inbathini, and Rizwan Ali Saifudeen, "Highly Nonlinear and Birefringent Spiral Photonic Crystal Fiber," *Advances in OptoElectronics*, vol. 2014, Article ID 464391, 6 pages, 2014.

[22] N.A. Mortensen, J.R. Folkenberg, M.D. Nielsen, and K.P. Hansen, "Modal cut-off and the V-parameter in photonic crystal fibers," *Opt. Lett.* vol. 28, no. 20, pp. 1879–1881, 2003.

[23] M. A. Islam et al., "Design optimization of equiangular spiral photonic crystal fiber for large negative flat dispersion and high birefringence," *IEEE J. Lightw. Technol.,* vol. 30, no. 22, pp. 3545–3551, 2012.

[24] S. G. Leon-Saval et al., "Splice-free interfacing of photonic crystal fibers," *Opt. Lett.,* vol. 30, no. 13, pp. 1629–1631, 2005

Design and Implementation of Home Security System using PZT

Sadeka Sultana, Tawhida Akand and Mrinmoy Dey*
Department of Electrical and Electronic Engineering
Chittagong University of Engineering and Technology,
Chittagong-4349, Bangladesh
sadekazinia@yahoo.com, tawhidacuet@yahoo.com, mrinmoycuet@gmail.com*

***Abstract*—Security is a major concern in our day-to-day life since it can prevent intentional or unintentional interference in human life with the proper operation of security system. In this project based research work, two set of piezoelectric transducer (PZT) are used.When anybody stands on PZT(I), the buzzer indicates the problem detection and LED light shows the problem occurring place. The buzzer and LED are centralized in a map so that one can monitor a large area with this system design. Microcontroller is used for continuous supply and PZT(II) is used to charge the battery. Piezoelectric transducer of high sensitivity must be selected in the system as security system detectors should have high sensitivity to sense the failure of security within the shortest period of time. This system is best suitable for corporate offices, ATMs and home security.**

Keywords—Piezoelectric transducer, microcontroller, buzzer, LED.

I. Introduction

Home security has changed a lot from the last century and will be changing in coming years [1]. Security is an important aspect or feature in the smart home application [2]. System for detecting and quantifying pedestrian traffic has a wide range of application from data collection for municipal planning to perimeter security systems for homeland security [3]. Sometimes only a guard is not capable to keep a good eye on a large area. If the security system is centralized one can keep a good eye on the whole area, which reduces the number of labour and labour cost.It is a nonconventional energy source. For a remote place where power transmission line is impossible to instalment this self-sustained system is useful [4].

Different types of sensors are used in security systems. V. Karri and J. S. Daniel Lim, used more sensors set up in a remote array, a PIC micro-controller and a GSM moduleto develop a security device that is capable of notifying a person or several persons using a SMS message initiated by one touch dialing [2]. Ms. Gunjan Jewani et al. worked on knock based security system (KBSS) system using piezo sensor and ARDUINO [5]. M. Sathis Kumar et al. established smart surveillance system based on the embedded system along withPIRsensor network and GSM [6]. A research ofmicrocontroller based home security system with remote monitoring was done by Nikhil Agarwal and G. Subarmanaya Nayak in which IR Laser sensors are used to detect any obstacle while monitoring the windows and doors [7]. Passive Infrared Sensor was used to develop a Security System [8]. Rpindarsingh Brar showed ARDUINO Based Industrial Security System using Piezoelectric sensor [9]. A paper was published on “Electronic Eye for security System” by Saurabh Vinayak Lawate and M.S. Ali using video camera for security purpose [10]. Aayush Agarwal and R. C. Joshi proposed WSN and GSM based remote home security system [11].Ando Juniors et al. proposed a micro generator piezoelectric for portable device from the energy harvesting [12].

II. Methodology

Different types of sensors are used in security systems. Forced-plates that are suitable for traffic monitoring and security detection also have significant applications in other fields such as bio-mechanics research on both human and animal subjects [3].

Table I shows various sensors for home security and it also represents different types of sensor used for different objects.

TABLE 1: VARIOUS SENSORS FOR HOME SECURITY [12]

Sensor type	Sensor generally used
Motion	Passive infrared
Vibration & Force	PZT
Temperature	Bandgap
Humidity	Capacitive
Light	LDR
Smoke	Ionization

III. System Design

The system has two set of piezoelectric transducer array. One set is used as a switch to microcontroller and another set is used as a power source.If system is secured, green LED indicator on the map indicates it. If anybody stands on the PZT set (II), the buzzer is switched on and alarmed security guard for a short while. Then green LED is off and red LED lit on to indicate the problem until the program is restarted.If both LED radiates their light that means problem was occurred there, but now the area is safe.The whole system is represented by a block diagram shown in Fig. 1.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

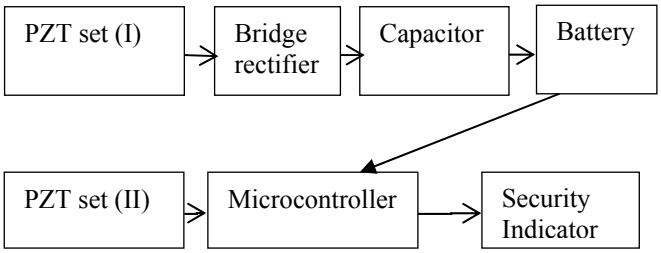


Fig. 1 Block diagram

IV. System Construction & Operation

The following equipments are used in this project:

- PIC16F877A
- Red LED
- Green LED
- Buzzer
- Oscillator (20Mhz)
- Schotky diode
- Resistor (100 ohm)
- PZT
- Capacitor
- Battery
- Frame
- Insulator for PZT
- Protection for PZT
- Spring

A. *Charging circuit*

Piezoelectricity is defined as the property of some dielectric materials that have developed a polarization that are subjected to mechanical strain deformation. This mechanical tension produced by the biasing force realignment of atoms thereby creating loads of polarization and thus an electric field.However, the energy generated through the piezoelectric effect is not sufficient for directly powering most electronic devices [13]. The circuit consists of a full wave rectifier, capacitor and the battery intended to be charged [13].

PZT set (I) is installed there. PZT set produces energy for every pressure on it. The rectifier converts PZT signal(AC)to DC signal. Then using a capacitor this is stored in a battery. Here schottky diode is used. Schottky diode has a threshold voltage which is smaller than that of a p-n junction diode [14].

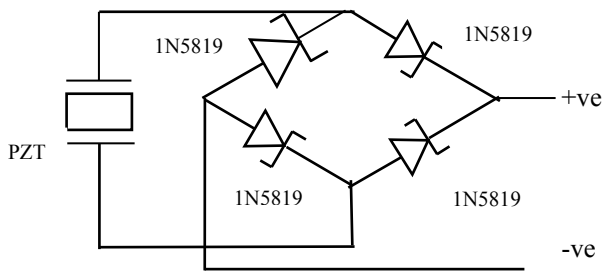


Fig. 2 Circuit diagram for power supply PZT set (I)

The microcontroller is in digital I/O mode powered by PZT set (I). Buzzer is directly connected to PZT. Oscillator provides the clock pulse for the microcontroller. Here 20 MHz Oscillator is used.The capacitors are used here for stability of the clock pulse.MCLR pin is used to reset the microcontroller.

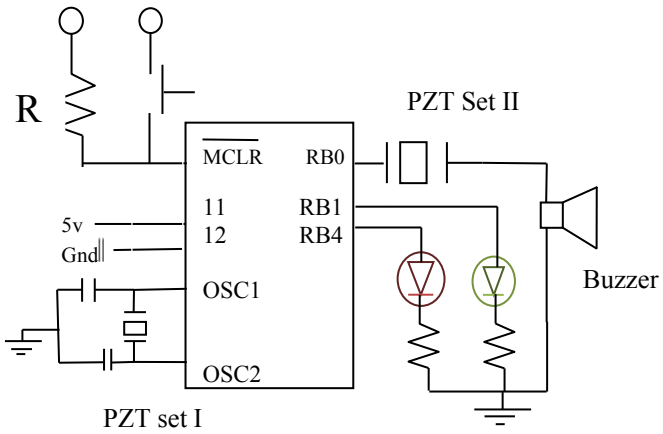


Fig. 3 Microcontroller circuit

Fig. 4 Structure of PZT Plate

V. Flow Chart

Fig. 5 shows flow chart for self-sustaining microcontroller based security system

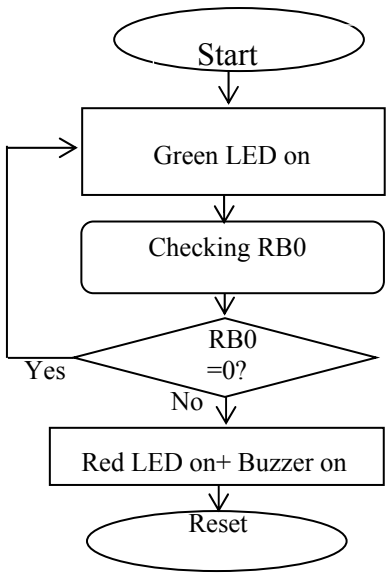


Fig. 5 Flow chart for self-sustaining microcontroller based security system

VI. Design of Frame

The frame is made of steel. A spring is placed in the frame to bring the PZT at its original place. The spring is called hard spring due to its high spring constant.

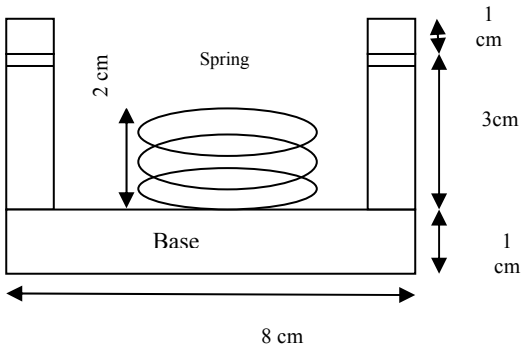


Fig. 6 Design of frame

VII. Overall Setup

The PZT is covered and insulated with wool and placed in the frame. The whole area is covered with PZT, which are divided into two set. PZT set (I)and PZT set (II). There are clips too lock the PZT with the frame. Fig. 7 shows overall mechanical set up.

(a)

(b)

Fig.7 Mechanical Setup (a) Top view (b) Front view

VIII. System Implementation

This system was implemented in Chittagong Collegiate School (CCS), Chittagong-4000 situated in Ice Factory road on 3 July 2015.Fig. 8 shows the implementation of the system.

Fig. 8 Implementation of the project

IX. Experimental Analysis

The output voltage increases with the increasing number of plate and output current increases with the increasing number of plates.Table II shows the relation between voltage, current and number of plates.

Table II Output Voltage & Current

No. of Plates	Voltage (V)	Current (mA)
6	21.9	0.22
8	27.4	0.25
10	29.5	0.28

Fig.9 shows the relationship between voltage and number of plates.Generated voltage is proportional to the number of PZT plates.

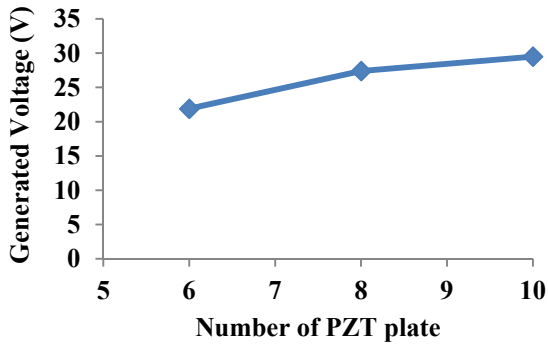


Fig. 9 Number of PZT plate vs. Generated Voltage

Fig. 10 shows the relationship between current and number of plates. Current is also proportional to the number of PZT plates.

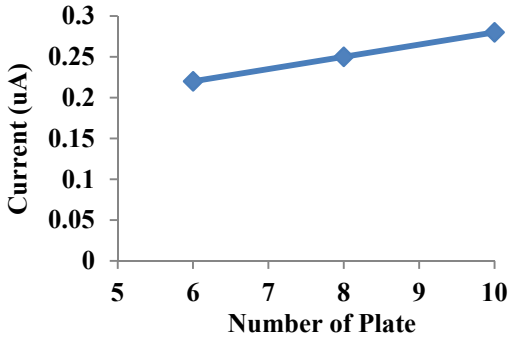


Fig. 10 Number of Plate vs. Current

X. Simulation Output

Green led indicates everything is secured.Push button and power supply indicates the PZT. When push button is not pressed means PZT is original mode. When push button is pressed means PZT is pressurized. PZT displacement is proportional to the force and voltage is proportional to displacement. The total security system is represented in Fig. 11, Fig. 12 and Fig. 13.

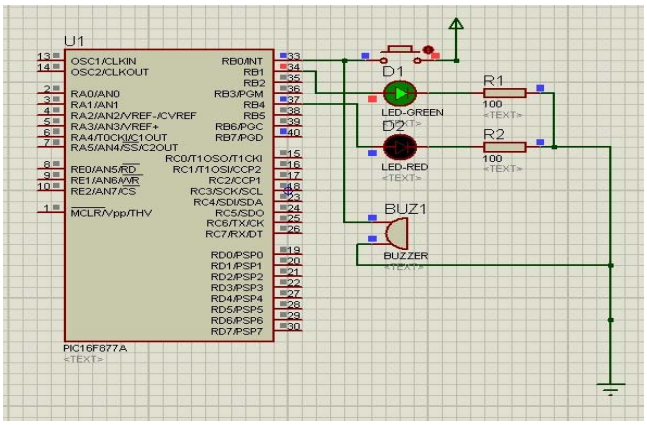


Fig. 11 When everything is secured

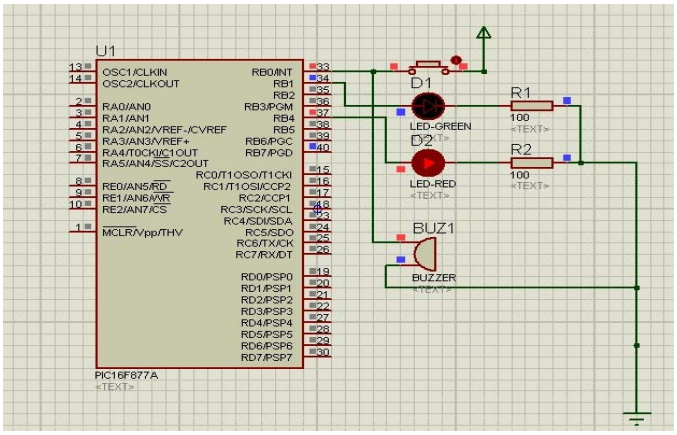


Fig. 12 When a thief stands on PZT

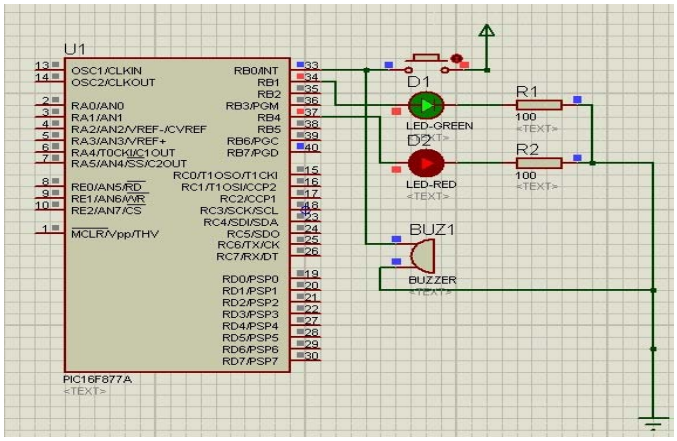


Fig. 13 When PZT is again original mode

XI. Advantages

This security system has following advantages:

- Simplicity
- Cost effective
- Easy to maintain
- Easy to control
- Reduces labor cost
- Reduces number of labor
- Centralized
- Low power consumption

XII. Conclusions

In this project based research work, the designed security system having the facilities of self-powered and centralized to detect the abnormal conditions. The implemented security system is very simple because of the convenient way to control and maintain. The system is cost effective and no manual work is necessary to control it.

Acknowledgment

The authors would like to thank the Department of Electrical and Electronic Engineering (EEE) of Chittagong University of Engineering and Technology (CUET), Bangladesh.

References

[1] The-History-of-Home-Security 4th July 2010 [Online]. Available: http://ezinearticles.com.

[2] V. Karri and J. S. Daniel Lim, "*Method and Device to Communicate via SMS After a Security Intrusion,"1st International Conference on Sensing Technology*, Palmerston North, New Zealand, November 21- 23, 2005.

[3] B. Yan, R. Blacow, S.E. Prasad and D. Waechter, "Perimeter Security Detection System Based on Piezoelectric Strain Measurement,".

[4] The History of Home Security [Online]. Available: http://www.alarm.org.

[5] Ms. Gunjan Jewani, Prof Shubangiborkar,"Review On knock Based Security System,"*International Journal on Recent and Innovation Trends in Computing and Communication,* vol. 3, no. 2.

[6] M. Sathishkumar and S.Rajini,"Smart Surveillance System Using PIR Sensor Network and GSM,"*International Journal of Advanced Research in Computer Engineering & Technology (IJARCET),*vol. 4, no. 1, January 2015.

[7] Nikhil Agarwal and G. SubarmanayaNayak, "Microcontroller Based Home Security System with Remote Monitoring,"*International Conference on Electronic Design and Signal Processing (ICEDSP)*, 2012.

[8] PemaChodon, Devi Maya Adhikari, Gopal Chandra Nepal, RajenBiswas, SangayGylshen, Chencho,"Passive Infrared Sensor Based Security System,"*International journal of Electrical, Electronic and Computer System.*

[9] RpindarsinghBrar,"*ARDUINO Based Industrial Security System using Piezoelectric sensor,".*

[10] SaurabhVinayakLawate and M.S. Ali, "Electronic Eye for security System," *International Journal of Electronic and Electrical Engineering*, vol. 7, no. 9, 2014.

[11] AayushAgarwal and R. C. Joshi, "WSN and GSM based Remote Home Security Syestem," *International Conference on Recent Advances and Future Trends in Information Technology (IRAFIT),* 2012.

[12] O. H. Ando Junior, M. A. J. Coelho, C. F. Malfatti and V. J. Brusamarello, "Proposal of a Micro Generator Piezoelectric for Portable Devices from the EnergyHarvesting,"*International Conference on Renewable Energies and Power Quality (ICREPQ-2014)*, Cordoba, Spain, 8-10 April, 2014.

[13] Sheikh Izzal Azid andSushil Kumar, "Analysis and Performance of a Low Cost SMS Based Home Security System," *International Journal of Smart Home,*vol. 5, no. 3, July 2011.

[14] Henry A. Sodano, Daniel J. Inman and Gyuhae Park,"*Generation and Storage of Electricity from Power Harvesting Devices,"*

COASYM-MAC: A Cooperative Asymmetric MAC Protocol for Wireless Sensor Network

Md. Mahedee Hasan[1,*], Amit Karmaker[1], Shafika Showkat Moni[2] and Mohammad Shah Alam[1]

[1]IICT, Bangladesh University of Engineering and Technology, Dhaka-1205, Bangladesh
[2]CSE, Rajshahi University of Engineering and Technology
[*]mahedee.hasan@gmail.com

Abstract— **Link asymmetry, arisen from varying coverage distance between two adjacent nodes, poses a real challenge for designing efficient receiver initiated MAC protocol for extremely low duty-cycled wireless sensor networks (WSN). One of the viable solution to overcome this limitation is to exploit the benefit of cooperative communication. In this paper, we propose a cooperative MAC protocol for WSN that effectively reduces asymmetry by taking advantage of alternative paths between a pair of nodes with link asymmetry. Simulation results show that our proposed scheme achieves significant improvement in performance compared with Asym-MAC, a state-of-the-art MAC protocol for WSN with asymmetric links. This performance improvement is effectively demonstrated in terms of energy consumption, average delay per packet and packet reception ratio.**

Index Terms—**Wireless Sensor Network, Asymmetry, MAC.**

I. Introduction

Wireless Sensor Network is the most promising technology in the last decade. As wireless sensor nodes operates very low battery power, researchers pay a great attention to increase the lifetime of sensor nodes. Most of them consider the link between two nodes is fully symmetric. But when the link is asymmetric then there is a lot of possibilities to resend same data that creates extra energy loss. A substantial percentage of wireless links are asymmetric for low power WSN [1]. Fig. 1 shows some situations when we can say the links are asymmetric. In 1(a), when sender can sends data but receiver can't. 1(b), When receiver can sends beacon but sender can't and 1(c), neither sender nor receiver can send data. Wang et.al [2] describes some common causes of link asymmetry: Limited transmission range, power limitation and interference.

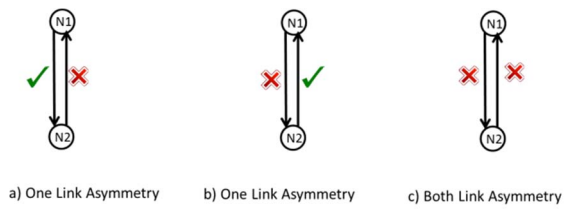


Fig. 1 Asymmetric link situation.

Sang et.al. [1] analyse the asymmetric link situation by transmitting 100 packets in different power level. According to Table I, most of the links are symmetric at higher power level but when the power level decreases then the possibility of link asymmetry increases. In this paper we propose a better scheme that more strictly handle the asymmetric link situation at low power WSN. The remaining part of this paper is organized as follows: Section II discusses related works. In section III, problem statements are discussed. Section IV presents the proposed COASYM-MAC protocol, section V describes the experimental result. Finally, section VI concludes the paper.

TABLE I
LINK ASYMMETRY IN DIFFERENT POWER LEVEL

Tx Power	Symmetric (<10%)	Asymmetric (10-90%)	Unidirectional (>90%)
-19dBm	50%	43%	07%
-14dBm	65%	22%	13%
-5dBm	88%	06%	06%

II. Related Work

Recent years have witnessed a lot of MAC protocol had been proposed to reduce the energy consumption of sensor nodes. Duty cycling [3][4] is most prominent technique to turnoff transceiver periodically to save energy. Later sender initiated protocol has been introduced [5] where sender takes initiative to seek attention of receiver by sending a long preamble. It has advantages over synchronous MAC because in sender initiated protocol sensor nodes don't need to create any schedule by using clock synchronization. But long preamble causes extra overhead and high probability of collision. Later these long preamble has packetized and one of the effective packetized preamble base MAC protocol is X-MAC [6]. In X-MAC receiver can send early acknowledge between two short preambles. But these short preamble can't fully overcome extra overhead. Then receiver initiated protocol came into focus [7]. Unlike sender initiated protocol, here receiver takes initiative by sending beacon to seek sender attention. Receiver initiated protocol reduces extra overhead over sender initiated protocol. Later ASYM-MAC [8], has been proposed where they work over RI-MAC. ASYM-MAC combined sender and receiver initiated protocol. By default, system operates on receiver initiated mode. When sender can't get beacon from receiver it goes to sender initiated mode and sends its data followed by long preamble. But when sender gets beacon from receiver but sender can't send data due to link asymmetry then this protocol can't work. In this paper we try to solve this problem by using relay node communication when links are asymmetric. By using relay node we also avoid transition from receiver initiated to sender initiated. Nacef et.al. [9] proposed CL-MAC where they used cooperative communication in low power sensor network. In CL-MAC sender sends chunk of short preamble packet to all its relay and destination nodes. These short preamble informs all relay and receiver to wake up at a common rendezvous (RDV) point. Time synchronization is the main drawback of these protocol. Chunk of short preamble also creates extra overhead. ARQ-CRI [10] work over CL-MAC and try to ignore clock

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

synchronization. When a receiver sends a beacon then sender transmit a coop message that informs receiver wait some moments for the relay selection. When relay selection takes place then sender transmit its data both receiver and relay node. It avoids the clock synchronization but it doesn't give any particular solution when sender fails to transmit coop message.

III. Problem Statement

As stated earlier, in ASYM-MAC all nodes are receiver initiated mode, when sender fails to get beacon from intended receiver due to asymmetric link it goes to sender initiated mode and send preamble to receiver and then data transmission takes place. We find some drawbacks of these protocol.

A. Preamble and Data Collision

Preamble and Data can collide when a node suddenly goes to sender initiated mode. When Receiver 1 sends beacon to Sender 1 and Sender 2. Sender 2 finds the beacon but Sender 1 can't find due to asymmetric link. Then, sender 1 goes to sender initiated mode and start sending preamble. This preamble and data of Sender 2 can collide at Receiver end as shown in Fig. 2.

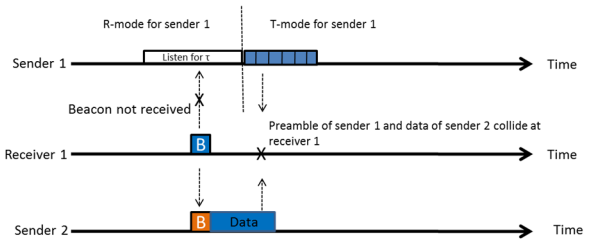


Fig. 2 Preamble and data collision.

B. No Mechanism in Data Loss Situation

We already know that link asymmetry can happen from both ends so it is possible to find asymmetric link when sender 1 wants to send data to the receiver. ASYM-MAC doesn't describe any mechanism how a sender can sends data in this scenario.

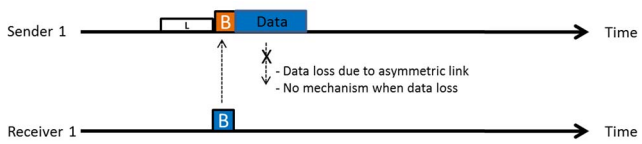


Fig. 3 Data loss due to link asymmetry.

C. Long Preamble

When a node goes to sender initiated mode it sends long preamble which creates extra overhead.

IV. Proposed Method

Our main contribution of this protocol is to reduce energy consumption and collision. As we already know that link asymmetry may or may not be persistent. Here we consider three cases that can happen in a low power wireless sensor network:

A. In Case of Link Symmetric

When a receiver awakes it sends a beacon to all of its neighbor. Sender, which has data to send takes a short random backoff and perform Clear Channel Assessment (CCA) . If it finds channel is free then it immediately sends data. Otherwise it will wait for next beacon. In Fig. 4, all three sender has data to send and listen the channel. Receiver sends beacon after wakeup. After receiving beacon all senders takes random backoff to avoid collision. Sender 3 takes the least random backoff and win the channel. After sending data it waits for the acknowledgement.

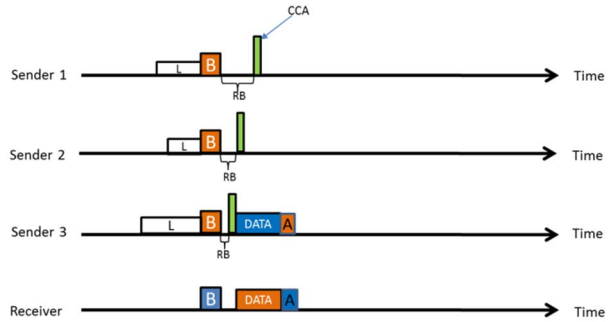


Fig. 4 Symmetric link situation.

Algorithm 1 present the pseudo-code that summarize the operation when link is symmetric.

Algorithm 1 Link Symmetric

```
/*Sender s sending data to receiver r */
if a beacon packet received from r then
        sender takes random backoff and check CCA
        if win the channel
            send data to r and go to sleep mode
        else wait for next beacon
        end if
end if
/* receiver r received data from sender s*/
 if a data packet received from s then
send ACK to sender
end if
```

B. In Case of Beacon Loss Due to Link Asymmetry

When a sender (s) has data to send it waits for the beacon from intended receiver (r). Sender waits for T second which is twice of the wakeup time of a node. If sender cannot get beacon from intended receiver due to link asymmetry within T second then it uses a cooperative node which is common neighbour of both sender and receiver. In dense network, it is very difficult to keep all neighbour node information due to memory constraints [1]. That's why we don't use neighbour information for selecting relay node. After T second, sender waits for the beacon of another neighbour. If sender gets a beacon from another neighbour (n), it takes a very Short Random Backoff (SRB) which is shorter than the backoff used in Algorithm 1. The reason behind very short backoff is that, when a node use relay node for transmit its data it needs to get high priority. When it gets the channel it sends a Claim Relay (CR) message which contains intended receiver ID that informs neighbor node (n) it has a data to send with other node. Neighbor node then waits another T second. If it cannot get a beacon from the intended receiver (r) it finds intended receiver is not its neighbor and it goes to sleep mode. If

neighbor node gets a beacon from intended receiver it takes a very Short Random Backoff (SRB) so that it can win the channel. Then two way handshake takes place. Neighbor node sends a Relay Request (RR) message. Intended receiver confirms it through Relay Acknowledge (R-ACK) message. Two way handshaking is needed because link from neighbor node (n) to intended receiver (r) can also be asymmetric. When neighbor node gets R-ACK message it immediately send a Claim Relay Acknowledge (CR-ACK) message to sender. These CR-ACK not only allow the sender to sends its data but also it informs other nodes in its range that it is in transmission mode. Other nodes immediately go to sleep. Once relay has been selected (n), sender (s) sends its data. If receiver (r) correctly received it, receiver (r) sends an acknowledgement(ACK). Otherwise relay (n) detects the absence of ACK and retransmit the data. Fig.5, shows the timing diagram in case of beacon loss due to link asymmetry.

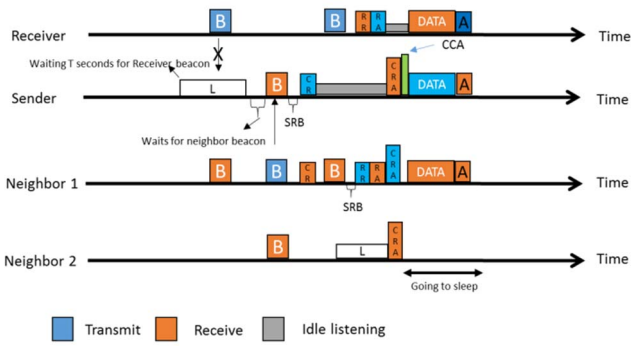


Fig. 5 Beacon loss situation.

C. *In Case of Data Loss Due to Link Asymmetry*

Sometimes link asymmetry also happens from sender to receiver. Sender (s) has data to send and waits for receiver (r) beacon. After successful channel win through random backoff sender (s) sends its data. But receiver (r) can't receive data due to asymmetric link. Sender assumes the link is asymmetric because it waits t time (which is half of beacon interval) but can't receive ACK from receiver. Then sender (s) chooses a cooperative node which is common neighbour of both sender and receiver. In Fig.6, shows the timing diagram in case of data loss due to link asymmetry.

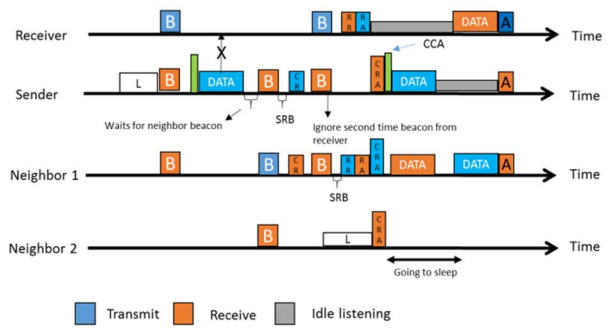


Fig. 6 Data loss situation.

Algorithm 2 present the pseudo-code that summarize the operation of relay selection when beacon loss (receiver to sender asymmetry) and data loss (sender to receiver asymmetry).

Algorithm 2 Relay Selection

```
/*Sender s sending data to receiver r */
if beacon packet not received from r within T
else if ACK not received from r within t then
    sender waits for the beacon from another neighbour
end if
if sender receives a beacon from another neighbour then
    it takes SRB and send Claim Relay (CR) message
else wait for beacon from another neighbour
end if
/* Neighbour n received CR from Sender s */
if beacon received from receiver r within T then
  it takes SRB and sends Relay Request (RR) message
else go to sleep
end if
if receiver r acknowledge by R-ACK then
    neighbour n sends CR-ACK to sender s
    sender s sends data to receiver r and neighbour n
        if neighbour successfully detect ACK then
            go to sleep
        else neighbour n retransmit it
        end if
else go to sleep
end if
```

V. Simulation

A. *Simulation Environment*

The performance of our protocol is evaluated and compared with known asymmetry base MAC protocol ASYM-MAC using OMNeT++ simulator. We calculate the average energy consumption, average delay per packet and packet reception ratio with link asymmetry. The simulation network composed of a sink node and 100 nodes deploy into 10 X 10 grid. Both protocol run in same parameter. Node wake up time is uniformly distributed between 0.01 to 0.99 second with fixed beacon interval 0.5 second. Transition time from receiver initiated to sender initiated is 1 second for ASYM-MAC asymmetric.

B. *Simulation Results*

Average Energy Consumption: Fig. 7 shows the average energy consumption of a node defined as total energy consumed (data packet, control packet, idle listening) by all nodes divided by number of nodes. In this case we vary the percentage of asymmetric links. Figure shows the relationship between average energy consumption and percentage of asymmetric links. We observe that when the number of asymmetric links increase then the power consumption by a node is also increase. In ASYM-MAC, we identify that when a node suddenly goes to sender initiated mode and send preamble then other nodes in its range have no idea about these transition. Then a lot of collision occur between preamble and beacon which creates extra energy consumption.

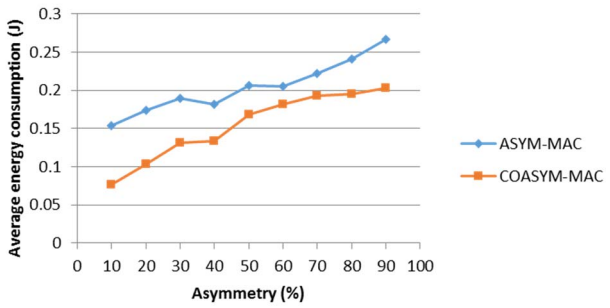


Fig. 7 Average energy consumption per asymmetry

Packet Reception Ratio: Fig. 8 compares packet reception ratio (PRR) with number of link asymmetry. Packet reception ratio is defined as the ratio of number of received packet and number of transmitted packet. Figure shows when the number of asymmetric link is below 30% then PRR of both protocol remain same. But when number of asymmetric link increase then the PRR curve of ASYM-MAC degrades quickly as compared with our COASYM-MAC. We also noticed that if link asymmetry occur from sender to receiver then PRR of both protocol around same but if the link asymmetry occur either sender to receiver or both end then COASYM-MAC gives great PRR over ASYM-MAC. In our simulation we consider only 1 node full asymmetric per 10 nodes. But if we consider more node full asymmetric then the PRR curve of ASYM-MAC degrades higher rate.

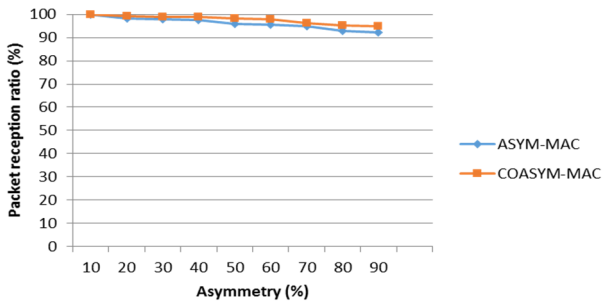


Fig. 8 Packet reception ratio per asymmetry

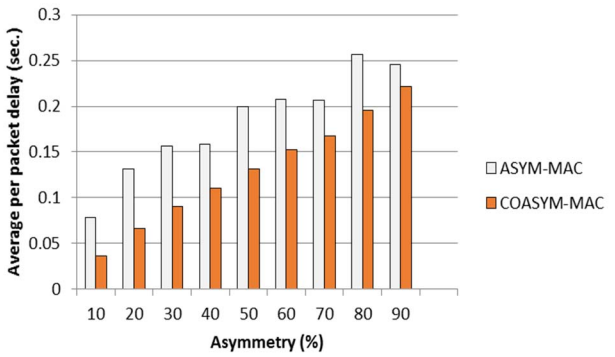


Fig. 9 Average delay (per packet) per asymmetry

Average delay per packet: Average delay per packet defined as the total period of time need to transmit a data packet to intended receiver divided by total number of packets. Fig. 9 shows the relationship between average delay per packet and number of asymmetric link. Within 50% of link asymmetry, average delay per packet of COASYM-MAC is almost half compared with delay per packet of ASYM-MAC. When link asymmetry is more than 80% then most of the nodes need to select relay node to transmit data to their intended receiver. But compare with ASYM-MAC our protocol gives better performance.

VI. Conclusions

A significant number of MAC protocols have been proposed that offer cooperative communication in wireless sensor network. Unfortunately most of the protocols don't consider the asymmetric link situation. In this paper we consider asymmetric link situation from both ends and choose relay node on demand. We evaluated our proposed COASYM-MAC in OMNeT++ simulator. Compared to ASYM-MAC, COASYM-MAC achieves low energy consumption, higher packet reception ratio and low average delay per packet.

References

[1] L. Sang, A. Arora and H. Zhang, "On exploring wireless link via one-way estimation," *In Proc MobiHoc'07*, p. 11-21, Sep. 2007.

[2] G. Wang, D. Turgut, L. Boloni, Y. Ji and D. Marinescu, "A MAC protocol for wireless sensor netwoks with asymmetric link," Journal of Ad-hoc Networks, Volume 6, Issue 3, pp. 424-440, May. 2008.

[3] W. Ye, J. Heidemann and D. Estrin, 'An energy-efficient MAC protocol for wireless sensor networks," *In Proc. INFOCOM'02*, p. 1567-1576, Jun. 2002.

[4] T. Dam and K. Langendoen, "An adaptive energy-efficient MAC protocol for wireless sensor networks," *In Proc. SenSys'03*, p. 171-180, Nov. 2003.

[5] J. Polastre, J. Hill., D. Culler. "Versatile low power media access for wireless sensor networks," *In Proc. SenSys'04*, p. 95-107, Nov. 2004.

[6] M. Buettner, G.Yee, E. Anderson and R. Han. "X-MAC: a short preamble MAC protocol for duty-cycled wireless sensor networks," *In Proc. SenSys'06*, p. 307-320, Nov. 2006

[7] Y. Sun, O. Gurewitz and D. Johnson, "RI-MAC: a receiver-initiated asynchronous duty cycle MAC protocol for dynamic traffic loads in wireless sensor networks," *In Proc. SenSys'08*, p. 1-14, Nov. 2008

[8] M. Won, T. Park and S. Son, "Asym-MAC: A MAC Protocol for Low-Power Duty-Cycled Wireless Sensor Networks with Asymmetric Links," IEEE communications letters, Volume 18, Issue 5, pp. 809-812 May. 2014.

[9] A. Nacef, S. Senouci, Y. Ghamri-Doudane and A. Beylot, "A cooperative low power MAC protocol for wireless sensor network", *In Proc. ICC'11* , May. 2011.

[10] D. Nguyen, L. Tran, O. Berder and O. Sentieys, "A low-latency and energy efficient MAC protocol for cooperative wireless sensor networks," *In Proc. GLOBECOM'13*, p. 3826-3831, Dec. 2013.

A Deep Domain Adaption Approach for Object Recognition Using Multiple Model Consistency Analysis

Biprodip Pal,* Boshir Ahmed

Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh

*biprodip.cse@gmail.com, boshir_bd@yahoo.com

Abstract—**Domain adaption tends to transfer knowledge across domains following dissimilar distribution and where target domain has inadequate labelled samples. When knowledge is transferred from abundantly irrelevant sources negative transfer may occur resulting in poor classification of test samples. Deep learning research illustrates the semantic clustering as well as transferability of deep convolutional features for numerous tasks including domain adaption. Traditional clustering based domain adaption approaches are practical to handle negative transfer scenario. This paper presents a scheme that uses graph based consistency analysis of one supervised and another unsupervised model to effectively transfer knowledge using deep features. This approach uses local neighbourhood analysis to classify hard samples that are identified using consistency analysis of models. This method yields encouraging experimental results on benchmark domain adaption dataset compared to a single deep feature based supervised support vector machine classifier, demonstrating effective use of target domain data.**

Index Terms— **Domain Adaption, Clustering, Deep Learning, SVM, Object Detection.**

I. Introduction

Usually, the training and test data are drawn from the same feature space and the same distribution for most of the machine learning models. Change into this frequent assumption asks for rebuilt of statistical models from scratch using newly collected training data. Knowledge transfer or transfer learning between task domains would be enviable in such cases. In the transductive transfer learning or domain adaption setting, the source and target tasks are the similar but dissimilarity exists between the source and target domains [1]. Numerous machine learning applications effectively using transfer learning include image classification [2] [3], software defect classification [4], text sentiment classification [5] and multi-language article classification [6].

Due to the recent domination of deep learning techniques in most of the machine learning challenges made it a superior candidate for domain adaption research. Deep learning algorithms are particularly focused to learn multiple levels of representation. The principle objective is to learn good expressive representations [7]. Domain adaption works include Convolutional Neural Network (CNN) for different event detection domains that automatically learns effective feature representations from pre-trained word embedding, position embedding as well as entity type embedding and reduces the error propagation [8].Moreover, Chen et al. used two different CNN based path to model the two different domains for describing people based on clothing attributes [9]. Apart from these techniques, approaches considering the clustering or manifold processing for domain adaption are also interesting. Supervised manifold assumption check smooth & consistent predictions through assigning high weights to relevant manifold [10]. Gao et al. used multiple model local structure mapping for knowledge transfer [11]. When the source domain task and data contribute to the performance reduction of learning in the target domain, negative transfer occurs. Local manifold based approaches also manage the negative impact of knowledge transfer well.

This paper harness the synergy between local manifold processing and traditional supervised classification on deep convolutional activation features (DeCAF). It has been shown, manifold processing leverage information from deep features to find hard samples from target domain and local neighbourhood consideration for these samples improves performance of classification while in domain adaption scenario.

II. Transductive Transfer Learning

In many real-world applications, training and future data come from different feature space and have the dissimilar distribution. Thus traditional machine learning and data mining algorithms suffer from the problem of reduced performance. While classifying in one domain, but available sufficient training data is from another domain of interest, knowledge transfer or transfer learning, if done successfully, would greatly improve the performance. Usually the latter domain is in different feature space or a different data distribution is followed [1] Transfer learning reduces the cost of building systems on each different domain of interest.

We symbolize the source domain data as $D_S = \{(x_{S1}; y_{S1}),\ldots,(x_{Sn} ; y_{Sn})\}$ where $x_{Si} \in X_S$ is the data instance and $y_{Si} \in Y_S$ is the consequent class label. Similarly, the target-domain data is denoted as D_T $\{(x_{T1}; y_{T1}),\ldots,(x_{Tn} ; y_{Tn})\}$, where the input $x_{Ti} \in X_T$ and $y_{Ti} \in Y_T$ is the corresponding output. In most cases, $0< n_T << n_S$. Thus transfer learning can be defined as in definition 1.

Definition 1 (Transfer Learning). *Given a learning task T_S from source domain D_S and learning task T_T at a target domain D_T , transfer learning aims to develop the learning of the objective predictive function $f_T(.)$in D_T using the knowledge in D_S and T_S, where $D_S \neq D_T$, or $T_S \neq T_T$.*

When the test or target domain are different, it requires that part of the unlabeled target data be available during training phase to obtain the marginal probability for the target data.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

Definition 2 (Transductive Transfer Learning/Domain Adaption). *Given a learning task T_S from source domain D_S and learning task T_T at a target domain D_T ,transductive transfer learning or domain adaption aims to develop the learning of the objective predictive function $f_T(.)$in D_T using the knowledge in D_S and T_S, where $D_S \neq D_T$, or $T_S=T_T$. In addition, a few unlabeled target-domain data are required to be available at training time* [1].

When source and target domains only have a few overlapping features, transferring knowledge of feature representations is a general approach to deal with the new domain. When source and target domains have a lot of overlapping features, transferring the knowledge of instances is the common approach for domain adaption.

If knowledge transfer has a negative impact on target domain learning it is called negative transfer. For two too divergent tasks brute-force transfer may cause negative transfer. Clustering techniques are found useful to analyze relatedness among tasks [12] [1], this paper also uses a clustering based approach to find relevant hard samples in target domain using clustering analysis.

III. Deep Features

Building every system from scratch for dissimilar domains is much time consuming and costly. Thus researches are focused towards reuse of common knowledge discovered from existing systems for target domain. Transferable knowledge is extracted from source domain by finding expressive features. Appropriate representation learning algorithm is therefore an important part.

A deep learning algorithm is a particular kind of representation learning procedure that discovers several levels of representation where more abstract aspects of the data is presented with higher-level features. Automated discovery of abstraction is the central concept behind all deep learning tactics. It believes that more abstract representations of data tend to be more useful and they represent the semantic content of the data unlike those of the low-level features of raw data.

Functions that can be densely represented by a depth d design, might require comparatively more elements (often exponential) to be represented by a depth d-1 architecture. Moreover, the number of computational elements one can afford depends on the available number of tuning examples. Consequently, using an insufficiently deep architecture may lead to poor generalization to represent some functions [13]. Jeff et al. illustrated that using features extracted from the activation of a deep CNN that is trained in a fully supervised fashion, a fixed set of object recognition tasks can be repurposed to novel generic tasks. The state of the art implementation of the deep CNN design proposed by Krizhevsky et al. produces activation features at different levels [15].

Deep architecture design hypothesizes that the activations of the neurons in its late hidden layers might serve as very strong features for several tasks of object recognition.

The nth hidden layer activations of the deep CNN is denoted as a feature $DeCAF_n$. Thus $DeCAF_7$ denotes features from the 7th hidden layer – i.e., just before propagating through the final dense layer to produce the label predictions. Empirical analysis reveals the outstanding performance of semantic clustering of DeCAF with respect to a diverse tasks, like domain adaptation, fine-grained recognition challenges as well as scene recognition. DeCAF not only cluster same category instances across domains but also offer better within category grouping **[14][15].**

IV. Proposed Approach

Several classification models are available from any particular training domain. Because of the inductive bias of the specific learning technique and the distributional differences among the training domains, different models are often effective at different regions or structures to gather variety of knowledge. Therefore, it's hard for a single model to perform well in new and different domain until the features are elegant for the new domain. But, more base model requires more time to be trained and during decision phase more data needs to be processed. Thus unlike traditional approaches, our approach consists of single base model S, trained in a supervised manner from the source domain using DeCAF.

In domain adaption, P(y|x) the true conditional probability of class label y given a test example x, is not known a priori and x is from a different distribution. So, unsupervised technique to cluster the test samples is used. This approach uses an unsupervised model U built from the deep features in the test domain. Due to the expressive quality of the DeCAF and relevant empirical success in several works a simple way of clustering the samples using k-means algorithm is considered for the test domain. Unlike other papers who did not considered deep features, this work avoids spectral clustering. For a set of n observations $(x_1,x_2,\ldots x_n)$, k-means clustering aims to partition the observations into k ($\leq n$) clusters $C=\{c_k, k= 1,2,\ldots K\}$ with the following objectives defined in Eq. (1)[16].

$$J(C)=\arg\min \sum_{k=1}^{K} \sum_{x_i \in c_k} |x_i - \mu_k|^2 \qquad (1)$$

For this task, the traditional model averaging are intended in global weight assignment for models or fixed by favoring certain model (e.g., in single-classifier selection). But in domain adaption setting, different test examples may favor predictions from different base models or set of features. So, these global weighting schemes may not perform well. Consequently, a graph-based approach to approximate the optimal model weights is used. When the base model and target domain clustering carry conflicting concepts for some examples, those examples are defined as hard examples in our approach. Those samples are predicted by voting among its neighbors inside the test set.

A. Graph-Based Weight Estimation

The desired model should produce more accurate label prediction for x when x is from target domain. To put together similarity between the unknown actual target function and model prediction, P(y|x) is inferred through modelling of underlying P(x) modelled from the unlabelled test set. However, precise estimates of the probability density is hard, high-dimensional data instances x. Thereby like the approach of Jing et al. this approach uses clustering of the test data since low density areas between the clusters are represented by the cluster boundaries [11].

For a test examples X and base model S, graph $Gs = (V;E_s)$ is built where the vertex set V contains all the test examples and there is an edge connecting two test examples if and only if the examples are classified into the same class by S. For the clustering approach in the target domain another graph $G_U = (V;E_u)$ is used where V is same as Gs. The test samples are

clustered into *c clusters*. Edges in E_u connect two samples if and only if the two examples are in the same cluster.

To approximate the level of inconsistency each test example x is assigned a weight value between the local structures around x in *Gs* and *Gu*. Under the clustering assumption, if two examples are in the same class they probably will belong to the same cluster in *Gu* . So total common neighbors of x found in *Gs* and *Gu* approximates the inconsistency level for each sample x.

Suppose, *Vs* and *Vu* represent the sets of neighbors for *x* in *Gs* and *Gu* respectively. The model weight w_x is proportional to the consistency of the local structures between *Gs* and *Gu* at x:

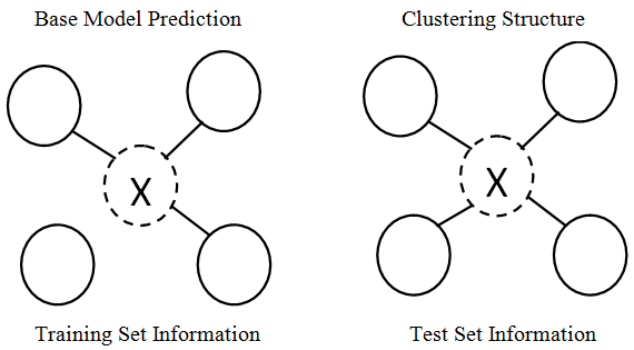


Fig. 2 Consistency measure trough neighbourhood graph.

$$w_x = \frac{\sum_{u \in V_S} \sum_{v \in V_T} 1\{u=v\}}{|V_S + V_U|} \qquad (2)$$

Thus degree of similarity in test samples labeling is reflected by w_x. For instance, figure 2 shows the G_s and G_T of a test example x constructed from the clustering algorithm and a base model S on the test set. According to eq. (2), the similarity between the clustering structure and S is 0.75 for x. The clustering assumption requiring the manifold structure of the data is related to the conditional probability P(y|x), needs to be hold for the weight approximation. Since deep convolutional features ensures reliable clustering of images, clustering techniques can verify the consistency of S in the target domain. The approach is summarized below in Algorithm 1. For ease of discussion this approach is named Multiple Model Consistency Based Model (MMCBM).

Algorithm 1: MMCBM

Input: (1) A training set X_s from source domain
(2) A test set X_T from a different domain
(3) A threshold ∂ and cluster number *k*.

Output: The set of predicted labels *L* for examples in X_T.

Steps:

1. Find the deep transferable features set *F* for each sample x_i using CNN activations as described in section 3.
2. Train base classifier model *S* from train domain samples X_s using corresponding features from *F*.
3. Find classification labels $L_S = \{l_{s1}, l_{s2}, \ldots l_{sn}\}$ for test samples X_T.
4. Perform base clustering on the test domain samples X_T by k-means to minimize eq. 1.
5. Find clustering labels $L_U = \{l_{u1}, l_{u2}, \ldots l_{un}\}$ for test samples.
6. Form the graph G_S and G_U from L_S and L_U .
7. FOR each $x_i \in X_T$, and defined threshold ∂
   - Compute consistency between G_S and G_U through weight w_{x_i}
   - if $w_{x_i} > \partial$
   - Label of x_i is $l_i = l_{si}$
   - else
   - put x_i into the list of hard samples *H*.
8. Apply kNN neighborhood analysis to find the labels l_i of the hard samples $x_i \in H$.

V. Experiments

As a measure of the classification performance accuracy, precision and recall is used. Formula are given in for a single class in (3),(4) & (5). In case of multiclass classification the result is the average for all classes.

$$\text{Accuracy} = \frac{\sum True\ Positive + \sum True\ Negative}{\sum Total\ Population} \qquad (3)$$

$$\text{Precision} = \frac{\sum True\ Positive}{\sum True\ positive + \sum False\ Positive} \qquad (4)$$

$$\text{Recall} = \frac{\sum True\ Positive}{\sum True\ positive + \sum False\ Negative} \qquad (5)$$

For our experiments we use the benchmark Office dataset[17].The dataset contains three domains: Amazon(A), which consists of product images taken from amazon.com; and Webcam(W) and Dslr(D), which consists of images taken in an office environment using a webcam or digital SLR camera, respectively.

The provided DeCAF features for office dataset are $DeCAF_6$ and $DeCAF_7$.Empirical results showed that $DeCAF_6$ is superior for better clustering as well as domain adaption settings[18]. As discussed in section 3, $DeCAF_6$ is the activations of the layer 6 of the mentioned deep CNN. Thus for this approach, $DeCAF_6$ features are considered.

Table 1 presents multi-class accuracy averaged for three different train and test splitting of the dataset for the domain shifts from each of the A,W and D domain to other. Table 1 uses notion A-D to denote domain shift from Amazon to Dslr. A total of 5 categories has been considered, for a random number of train sample variation for each class. Since support vector machine (SVM) is one of the superior candidate with best results in most of the classification challenges, experimental result is compared between a SVM only approach and MMCBM approach. MMCBM also considers SVM base model *S*. Both approach uses linear kernel for SVM and uses $DeCAF_6$ features.

Table 1 depicts that the accuracy of the MMCBM is better compared to that of the LSVM only approach. While shifting from D-W the performance of both are identical since the domain variation are not so significant thus negative information transfer is not abundant. Among all of the 18 cases at 18 cases the MMCBM comes out with better accuracy and all other cases they are similar. Table 2 illustrates the precision and recall analysis of the classification result. It can be clearly seen that for most of the cases, MMCBM is better in terms of precision and recall. While transferring from D-A, despite the accuracy is better, LSVM based approach classifies test samples more precisely. Performance is considerably better for MMCBM in terms of accuracy and recall at all scenario.

VI. Conclusion & Future Works

In this work, we analyze the single classifier and graph based consistency analysis model called MMCBM for the domain adaption setup. Given a target domain, how to identify the harmony between the domain and previous domains, and transfer knowledge from the previous domains to the target one is the aim of domain adaption task. Our approach uses graph based local neighborhood consideration to solve the problem of negative knowledge transfer in target domain.

Deep representations are expressive and reusable**.** Though deep features are famous for domain adaption tasks with support vector classifiers, performance comparison with traditional state of the art SVM algorithm with deep features shows encouraging performance of MMCBM approach. Future consideration for this work includes more precise classification through precision improvement. The reliability of the approach also needs to be empirically analyzed using different clustering techniques and supervised algorithm as candidate for replacement of the used k-means and SVM algorithm.

TABLE I. ACCURACY COMPARISON FOR DIFFERENT DOMAIN CHANGE SCENARIO

Total Training Samples/Class	**Methods**	**Domain Change**					
		A - D	**A - W**	**D - A**	**D - W**	**W - A**	**W - D**
10	Decaf+LSVM	.79	.81	.82	.99	.84	.98
	Decaf+MMCBM(ours)	.88	.88	.85	.99	.85	1
8	Decaf+LSVM	.86	.81	.81	1	.80	1
	Decaf+MMCBM(ours)	.92	.93	.84	1	.82	1
5	Decaf+LSVM	.77	.84	.83	.99	.85	.96
	Decaf+MMCBM(ours)	.80	.89	.84	.99	.87	.99

TABLE II. PRECISION AND RECALL VALUES FOR BOTH APPROACHES

Total Training Samples/Class	**Methods**	**Domain Change**											
		A - D		**A - W**		**D - A**		**D - W**		**W - A**		**W - D**	
		P	R	P	R	P	R	P	R	P	R	P	R
10	Decaf+LSVM	.88	.79	.91	.81	.88	.82	.99	.99	.87	.84	.98	.98
	Decaf+MMCBM(ours)	.88	.92	.88	.94	.85	.90	.99	.99	.85	.88	1	1
8	Decaf+LSVM	.92	.86	.89	.81	.87	.81	1	1	.85	.80	1	1
	Decaf+MMCBM(ours)	.92	.95	.93	.94	.84	.89	1	1	.82	.87	1	1
5	Decaf+LSVM	.87	.77	.89	.84	.88	.83	.99	.99	.87	.85	.97	.96
	Decaf+MMCBM(ours)	.80	.89	.89	.92	.84	.88	.99	.99	.87	.91	.99	.99

REFERENCES

[1] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," Journal of Big Data, vol. 3, no. 1, pp. 1-40, 2016.

[2] W. Li, L. Duan, D. Xu, and I. W. Tsang, "Learning with augmented features for supervised and semi-supervised heterogeneous domain adaptation," IEEE transactions on pattern analysis and machine intelligence, vol. 36, no. 6, pp. 1134-1148, 2014.

[3] B. Kulis, K. Saenko, and T. Darrell, "What you saw is not what you get: Domain adaptation using asymmetric kernel transforms," in Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, pp. 1785-1792, IEEE, 2011.

[4] J. Nam and S. Kim, "Heterogeneous defect prediction," in Proceedings of the 2015 10th joint meeting on foundations of software engineering, pp. 508-519, ACM, 2015.

[5] C. Wang and S. Mahadevan, "Heterogeneous domain adaptation using manifold alignment," in IJCAI Proceedings-International Joint Conference on Artificial Intelligence, vol. 22, p. 1541, 2011.

[6] P. Prettenhofer and B. Stein, "Cross-language text classification using structural correspondence learning," in Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, pp. 1118-1127,Association for Computational Linguistics, 2010.

[7] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," IEEE transactions on pattern analysis and machine intelligence, vol. 35, no. 8, pp. 1798-1828, 2013.

[8] T. H. Nguyen and R. Grishman, "Event detection and domain adaptation with convolutional neural networks," in Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, vol. 2, pp. 365-371, 2015.

[9] Q. Chen, J. Huang, R. Feris, L. M. Brown, J. Dong, and S. Yan, "Deep domain adaptation for describing people based on fine-grained clothing attributes," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5315-5324, 2015.

[10] L. Ge, J. Gao, H. Ngo, K. Li, and A. Zhang, "On handling negative transfer and imbalanced distributions in multiple source transfer learning," Statistical Analysis and Data Mining, vol. 7, no. 4, pp. 254-271, 2014.

[11] J. Gao, W. Fan, J. Jiang, and J. Han, "Knowledge transfer via multiple model local structure mapping," in Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 283-291, ACM, 2008.

[12] S. Ben-David and R. Schuller, "Exploiting task relatedness for multiple task learning," in Learning Theory and Kernel Machines, pp. 567-580, Springer, 2003.

[13] Y. Bengio, "Learning deep architectures for ai," Foundations and trends in Machine Learning, vol. 2, no. 1, pp. 1-127, 2009.

[14] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, "Decaf: A deep convolutional activation feature for generic visual recognition.," in ICML, pp. 647-655, 2014.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, pp. 1097-1105, 2012.

[16] A. K. Jain, "Data clustering: 50 years beyond k-means," Pattern recognition letters, vol. 31, no. 8, pp. 651-666, 2010.

[17] K. Saenko, B. Kulis, M. Fritz, and T. Darrell, "Adapting visual category models to new domains," in European conference on computer vision, pp. 213-226, Springer, 2010.

[18] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?," in Advances in neural information processing systems, pp. 3320-3328, 2014.

Performance Limitations due to Combined Influence of ASE and Raman Amplifier Induced Crosstalk in a WDM System with Direct Detection Receiver

F.H. Tithi,[1,*] and S. P. Majumder[2]

[1]Ph.D. Student, EEE Department, Bangladesh University of Engineering and Technology
Dhaka-1205, Bangladesh

[2]Professor, EEE Department, Bangladesh University of Engineering and Technology
Dhaka-1205, Bangladesh

[*]tithi_18@yahoo.co.uk

Abstract—**An analytical approach is presented to evaluate the bit error rate performance of a wavelength division multiplexing (WDM) fiber-optic transmission link with distributed Raman amplifier taking into account the combined effect of Amplified Spontaneous Emission (ASE) and Raman scattering induced crosstalk for direct detection receiver. The analysis is developed to find the photodetector output due to Raman scattering and ASE noise. The signal to noise plus crosstalk ratio at the receiver output is derived. The bit error rate (BER) performance results are numerically evaluated for 10 Gbps and different fiber length and WDM channels. It is found that at a BER of 10^{-9}, the allowable transmission distance is limited by ASE and crosstalk due to Raman scattering. The maximum allowable pump power is 1 mW, 3 mW, 4 mW , 6 mW for channel bandwidth 100 GHz, 50 GHz, 25 GHz and 12.5 GHz respectively for fiber length of 50 Km at BER 10^{-9}. The effect of ASE is more prominent at higher pump power**

Index Terms— **crosstalk, wavelength division multiplexing (WDM), Raman amplifier, bit error rate (BER), amplified spontaneous noise (ASE).**

I. INTRODUCTION

Distributed Raman amplifier is an important issue in Wavelength Division System (WDM) in an optical fiber transmission system [1]. Now a day's significant woks have been carried out to analyze the performance of distributed Raman amplifier due to Stimulated Raman Scattering (SRS) and Amplified Spontaneous Noise (ASE) with direct detection receiver [2-7].

The analytical formula of ASE noise power is presented in ref [8] for different pumping configuration as a parameter of fiber length and shown that the effect of ASE power is lower, higher lower and medium lower for forward pumping, backward pumping and bidirectional pumping. A new method is invented to evaluate the bit error rate performance for distributed Raman amplifier considering ASE noise double Rayleigh scattering (DRB) noise in optical transmission system [9].

J. Bromage *et all* experimentally and theoretically show that if the peak power of the pump pulses is too high then amplified spontaneous noise (ASE) will be high and results degrading noise figure [10].

Recently, the Gamma distribution is used to observe the effect of gain saturation, power penalty and noise distribution on BER in ASE seeded WDM system [11].More recently, we presented an analytical approach to evaluate the effect of Raman amplifier induced crosstalk, optical signal to crosstalk ratio, bit error rate, power penalty for a WDM transmission system with direct detection receiver [12] and also considering optical on-off keying (OOK) with heterodyne detection [13].

In this paper, we present an analytical model to evaluate the effect of Raman induced crosstalk and ASE power on the BER performance of a WDM link with distributed Raman fiber amplifier using direct detection. Results are evaluated in terms of ASE power, Raman scattering induced crosstalk and BER for different WDM system parameters.

II. SYSTEM MODEL

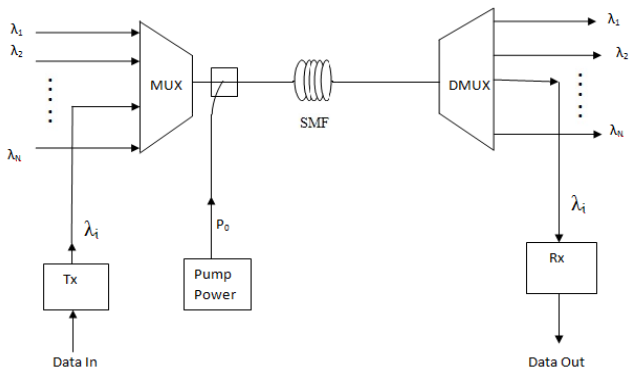


Fig. 1. Block diagram of a single span WDM transmission system with Raman amplifier with forward pumping configuration.

The WDM transmission system with Raman amplification is shown in Fig.1. The signal channels are wavelength multiplexed using a WDM multiplexer and coupled to the transmission fiber along with the pump signal. The WDM signals propagate through the SMF along with the pump signal and are separated at WDM node using an Optical Add Drop Multiplexer (OADM). The dropped WDM channel is received by a direct detection receiver and data detection is carried out.

III. SYSTEM ANALYSIS

The transmitted WDM signal of N number of channel can be represented as

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

$$S_{WDM} = \sum_{k=0}^{\infty}\sum_{n=1}^{N}\sqrt{2P_{s,n}(0)}.a_k^n.Cos\left[\frac{2\pi c}{\lambda_n}t\right] \quad (1)$$

where $P_{s,n}(0)$ is the n-th channel input signal power, N is the number of WDM channel.

Pump power at z=0 distance is P_0 and the electric field due to pump can be expressed as

$$S_{pump}(t) = \sqrt{2P_0}Cos\left[\frac{2\pi c}{\lambda_p}t\right] \quad (2)$$

Electric field of the WDM signal at a distance z can be expressed as

$$e_{WDM(t,z)} = \sum_{n=1}^{N} e_{sn}(t,z)$$

$$= \sum_{n=1}^{N}\sqrt{2P_{s,n}(z)}Cos\left[\frac{2\pi c}{\lambda_n}(t) - \beta.z\right] \quad (3)$$

where $P_{s,n}(z)$ is the power of the n-th channel at a distance z for the transmitter which can be expressed as[2]

$$P_{s,n}(z) = P_{is}P_{ip}e^{-\alpha z}\exp\left[GP_{ip}(n-1)Z_e\right]\left[P_{ip}\sum_{m=1}^{N}e^{GP_{ip}(m-1)Z_e}\right]^{-1} \quad (4)$$

where

$$G = \frac{g^{/}\Delta f}{2A};$$

$$Z_e = \frac{1-e^{-\alpha z}}{\alpha}$$

The eqn (4) can be re-written as,

$$P_{s,n}(z) = \frac{P_{is}e^{-\alpha z}e^{GP_{ip}nZ_e}}{\sum_{m=1}^{N}e^{mGP_{ip}Z_e}} \quad (5)$$

Here $g^{/}$ is the slope of the Raman gain profile, Δf is the interchannel frequency spacing, P_{ip} is the input pump power, P_{is} is the input signal power, Z_e is the effective propagation distance and α is the fiber attenuation coefficient. The optical channel bandwidth B_0 is taken to be $B_0 = \Delta f + B_g$. where B_g is the guard band.

The optical crosstalk power of a N-channel WDM system at a distance, z is given by [2],

$$P_{ct}(z) = \left(1 - N\exp\left[-GNP_{ip}Z_e\frac{(N-1)}{2}\right].\left[\frac{\sinh\left(\frac{NP_{ip}GZ_e}{2}\right)}{\sinh\left(\frac{N^2P_{ip}GZ_e}{2}\right)}\right]\right)P_{ip} \quad (6)$$

The optical signal to crosstalk ratio, $OSCR$ at a distance z is given by, $OSCR = \frac{P_{s,n}(z)}{P_{ct}(z)}$

$$= \frac{\dfrac{P_{is}e^{-\alpha z}e^{GPi_pnZ_e}}{\sum_{m=1}^{N}e^{mGP_{ip}Z_e}}}{\left(1 - N\exp\left[-GNP_{ip}Z_e\frac{(N-1)}{2}\right].\left[\dfrac{\sinh\frac{NP_{ip}GZ_e}{2}}{\sinh\frac{N^2P_{ip}GZ_e}{2}}\right]\right)P_{ip}} \quad (7)$$

The electric field due to optical crosstalk corresponding to n-th channel can be represented as

$$e_{ct,n}(t,z) = \sqrt{2P_{CT}(z)}\cos\left[\frac{2\pi c}{\lambda_n}t\right] \quad (8)$$

ASE power is [7] ,

$$P_{ASE} = h\upsilon B_0\eta T\left[G0 - 1 + \frac{G_0\alpha}{GP_{ip}}\left(e^{\alpha z} - \frac{1}{G_0}\right)\right] \quad (9)$$

where, ηT is thermal equilibrium phonon number

Gain, $G_0 = \exp(GZ_eP_{ip})$

Photon energy $h\upsilon = 6.63\times10^{-34}\ j$

Optical bandwidth, B_0=2B_e=2B

If n-th channel is dropped by the OADM, the current at the output of photodiode is then given by

$$i(t) = R_d\left|e_{s,n}(t) + e_{ct}(t) + e_{ASE}(t)\right|^2$$

$$= R_d\left|\sqrt{2P_{s,n}}\cos\omega_s(z,t) + \sqrt{2P_{ct}}\cos\omega_c(z,t) + \sqrt{2P_{ASE}}\cos\omega_{ASE}(z,t)\right|^2 \quad (10)$$

Where R_d is the photodiode responsivity,

$$= R_d\left|\begin{array}{l} P_{s,n} + P_{s,n}\cos\omega_s(z,t) + 2\sqrt{P_{s,n}P_{ct}}\cos[\omega_s(z,t)+\omega_c(z,t)] + \\ 2\sqrt{P_{s,n}P_{ct}}\cos[\omega_s(z,t)-\omega_{ct}(z,t)] + P_{ct} + P_{ct}\cos2\omega_c(z,t) \\ +2\sqrt{P_{s,n}P_{ASE}}\cos[\omega_s(z,t)+\omega_{ASE}(z,t)] + 2\sqrt{P_{s,n}P_{ASE}}\cos[\omega_s(z,t)-\omega_{ASE}(z,t)] \\ +2\sqrt{P_{ct}P_{ASE}}\cos[\omega_c(z,t)+\omega_{ASE}(z,t)] + 2\sqrt{P_{ct}P_{ASE}}\cos[\omega_c(z,t)-\omega_{ASE}(z,t)] \\ +P_{ASE} + P_{ASE}\cos2\omega_{ASE}(z,t) \end{array}\right| \quad (11)$$

Now, $\omega_s = \omega_c$ and removing higher frequency term,

The photocurrent at the output of bandpass filter,

$$i_o(t) = R_d\left|\begin{array}{l} P_{s,n} + P_{ct} + P_{ASE} + 2\sqrt{P_{s,n}P_{ct}} + 2\sqrt{P_{s,n}P_{ASE}}\cos[\omega_s(z,t)-\omega_{ASE}(z,t)] \\ +2\sqrt{P_{ct}P_{ASE}}\cos[\omega_c(z,t)-\omega_{ASE}(z,t)] \end{array}\right| \quad (12)$$

The average signal current is given by

$$I_s = \overline{i_{s,j}}(t) = R_d.\overline{P_{s,n}}(z) \quad (13)$$

Signal power, $P_s = [R_d P_{sn}]^2$ (14)

The variance of crosstalk is given by

$$\sigma^2_{CT} = MS(i_{ct}(t)) = \overline{i^2_{ct}}(t) = [R_d P_{ct}(z)]^2 \quad (15)$$

The receivers shot noise

$$\sigma_n^2 = 2eB(i_s(t) + i_{ct}(t) + i_{ASE}(t)) = 2eBR_d[P_{s,n}(z) + P_{ct}(z) + P_{ASE}(z)] \quad (16)$$

The variance of thermal noise, $\sigma_{th}^2 = I_{th}^2 B$, where I_{th} is the thermal noise current spectral density in $A/\sqrt{Hz}$.

Signal –crosstalk beat noise,

$$\sigma_{s-ct}^2 = \left(2R_d\sqrt{P_{s,n}P_{ct}}\right)^2 = 4R_d^2 P_{s,n} P_{ct} \quad (17)$$

ASE-crosstalk beat noise,

$$\sigma_{ASE-ct}^2 = \left(\frac{2R_d\sqrt{P_{ASE}P_{ct}}}{2}\right)^2 = 2R_d^2 P_{ASE} P_{ct} \quad (18)$$

Signal-ASE beat noise,

$$\sigma_{s-ASE}^2 = \left(\frac{2R_d\sqrt{P_{s,n}P_{ASE}}}{2}\right)^2 = 2R_d^2 P_{s,n} P_{ASE} \quad (19)$$

Electrical signal to crosstalk plus noise ratio at the receiver LPF output is given by

$$SCNR = \frac{I_s^2}{\sigma_{CT}^2 + \sigma_{shot}^2 + \sigma_{th}^2 + \sigma_{s-ct}^2 + \sigma_{ASE-ct}^2 + \sigma_{s-ASE}^2} \quad (20)$$

The bit error rate (BER) can be evaluated as

$$BER = \frac{1}{2} erfc \frac{\sqrt{SCNR}}{2\sqrt{2}} \quad (21)$$

IV. RESULTS AND DISCUSSION

Following the analytical approach presented in sec.III, the performance of a WDM link with crosstalk and ASE noise power with distributed Raman amplifier is evaluated considering the following system parameters.

TABLE I
VALUES OF PARAMETERS USED FOR PERFORMANCE EVALUATION

Parameter	Value
Effective area of core, A_{eff}	50μm^2
Operating wavelength, λ_{eff}	1550 nm
Fiber attenuation constant	0.046 km^{-1}
Raman gain slope, g'	4.9×10^{-20} m/(W GHz)
Input signal power, P_0	2 mW
pump power, P_0	10~50 mW
Channel spacing, $\Delta f \approx B_0$	6.25~100 GHz
Receiver Bandwidth, B	5 -10 GHz
Guard Band, B_g	2 GHz

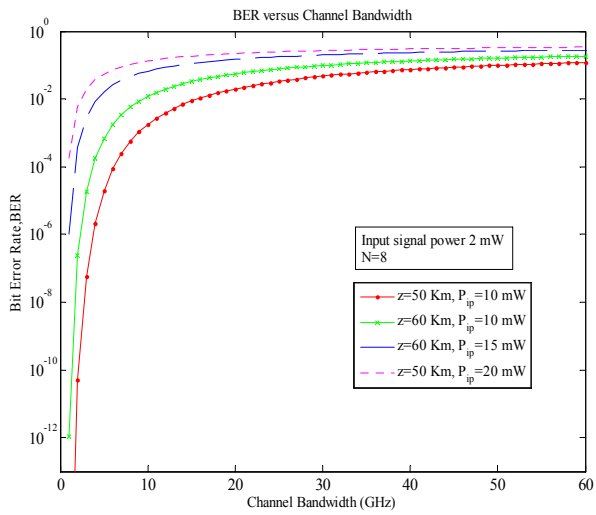


Fig.2. Plots of BER versus channel bandwidth of a WDM system with number of channel N=8 and input power of 2mW.

The plots of BER versus channel bandwidth are shown in Fig.2 for the different pump power and fiber length. It is found that BER increases with pump power and also fiber length for given input signal power 2 mW and number of WDM channel N= 8.

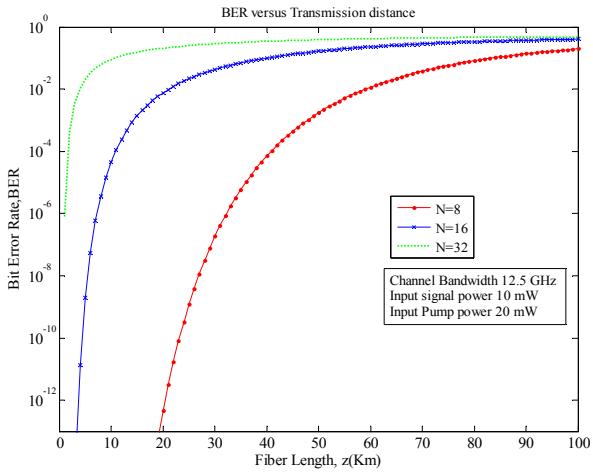


Fig.3. BER versus fiber length for direct detection receiver .

The effect of BER on fiber length is shown in Fig.3 for direct detection receiver considering Raman induced crosstalk and ASE noise power. From this figure it noticed that with increasing fiber length, BER also increases. It is also observed that BER increases with increasing total number of WDM channel. For example at 30 Km fiber length the BER is 10^{-7}, 10^{-2}, 10^{-1} for N=8, 16 and 32 respectively. Here we used N= 8,16 and 32 for channel bandwidth 12.5 GHz, input signal power 10 mW and input pump power 20 mW.

The variation of BER with pump power with a given transmission distance of z= 50 Km, input signal power is 2 mW and different channel bandwidth is shown in Fig.4 for direct detection receiver. It noticed that BER is affected by Raman amplifier induced crosstalk and ASE power due to pump power. The allowable transmission distance versus no. of WDM channel is also plotted in Fig.5 for given input signal of 10 mW, input pump power of 20 mW and different values

of channel bandwidth. It is clearly noticed that allowable transmission distance is reduced by Raman amplifier induced crosstalk and ASE power for higher number of WDM channel.

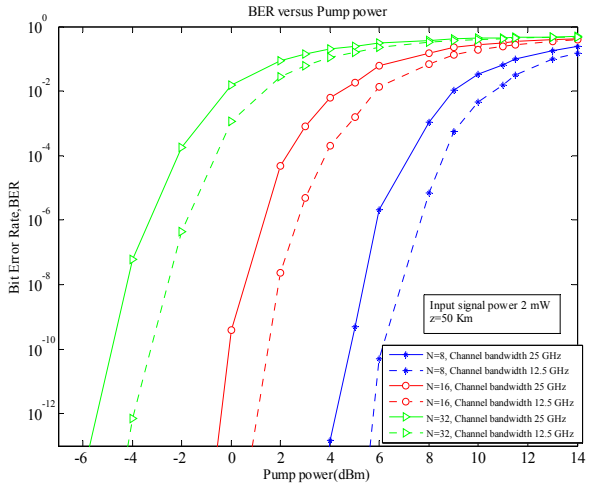


Fig.4. Plots of BER versus pump power at z=50 Km

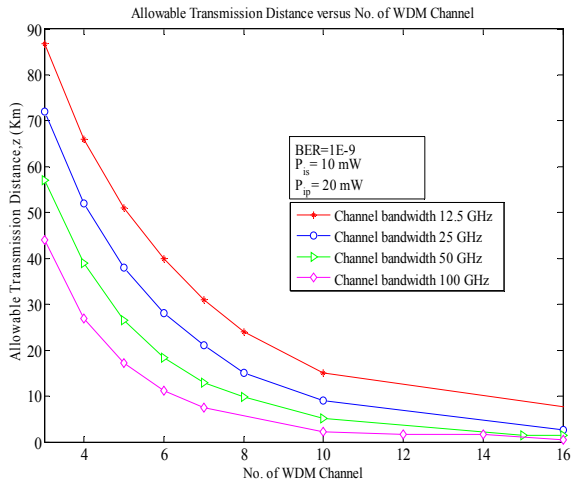


Fig.5. Plots of BER versus total number of WDM channel at a BER 10^{-9}

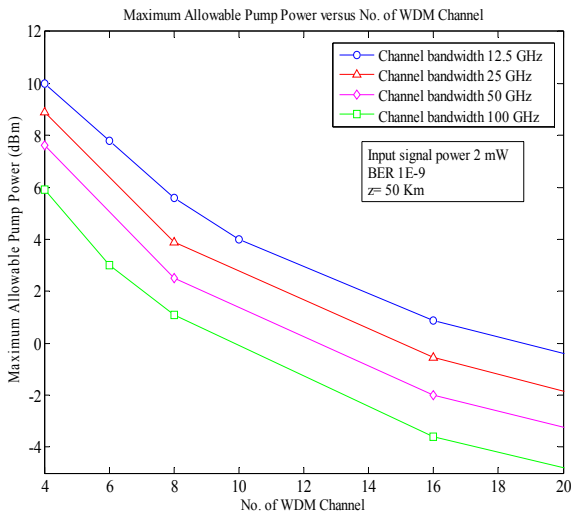


Fig.6. Plots of maximum allowable pump power versus no. of WDM channel for various channel bandwidth at BER 10^{-9}.

The plots of maximum allowable pump power versus total number of WDM channel for various channel bandwidth is shown in Fig.6 for a given input signal power 2 mW and fiber length of 50 Km at BER 10^{-9}. For example it is found that The maximum allowable pump power is 1 mW, 3 mW, 4 mW , 6 mW for channel bandwidth 100 GHz, 50 GHz, 25 GHz and 12.5 GHz respectively for fiber length of 50 Km at BER 10^{-9}.

V. CONCLUSION

Analysis is carried out for a WDM transmission system with Raman amplifier with direct detection receiver considering the effect of ASE noise and crosstalk due to Raman scattering. The bit error rate results are evaluated for difference WDM system parameters, pump power and channel bandwidh. It is found that the effect of ASE noise is significantly deteriorates the BER performance at high pump power and for distance is limited by ASE and crosstalk.

ACKNOWLEDGMENT

This work is carried out in the department of Electrical and Electronic Engineering of Bangladesh University of Engineering Technology (BUET) as a part of Ph.D. dissertation.

REFERENCES

[1] C. Headly and G. Agrawal, " Raman amplification in fiber optical communication systems", Elsevier, 2005.
[2] D.N. Christodoulides and R.B. Jander, " Evolution of stimulated Raman crosstalk in wavelength division multiplexed systems," IEEE photon. technol. Lett., vol.8, pp. 1722-1744,1996.
[3] Anamika and V. Priye, "XPM and SPM induced crosstalk in WDM system employing distributed Raman amplifier for DPSK and OOK modulation format," Optical Fiber Technology, ELSEVIER, vo.19, no.2, pp. 75-82, 2013.
[4] F. Juanni and Z. Ningbo, " The effect of SRS to DWDM optical system", International Conference on Electrical and Control Engineering, pp.2340-2342, 2010.
[5] Anamika and V. Priye, " Power penalty in WDM system due to stimulated Raman crosstalk," Proceedings of International Conference on Fiber Optics and Photonics, IIT Guwahati, SPIE Selected Proceedings, vol. 8173, December 11–15, 2010
[6] D. Dehan and G. Eisenstein., " The properties of amplified spontaneous emission noise in saturated fiber Raman amplifiers operating with CW signals", Journal of Optics Communications, ELSEVIER, vol. 236, no.4-6, 2004.
[7] U.R. Duarte, J.B. Rosolem and M. A. Romero , "A long-reach WDM-PON feeder employing EDFA-Raman amplification and ASE injected R-SOAs for optical-wireless access networks ," Microwave Photonics (MWP), 2012 International Topical Meeting on, pp.248-251, 2012
[8] A. H. Beshr, M. H. Aly and A. K. A. Seoud, "Amplified spontaneous emission noise power in distributed Raman amplifiers", International Journal of Scientific and Engineering Research, vol.3, no.5, pp.1-3, 2012.
[9] J. Shifeng , B. Bristiel, Y. Jaouen, P. Gallion, and E. Pincemin, "Bit-Error-Rate Evaluation of the distributed Raman amplified sransmission Systems in the presence of double rayleigh backscattering noise", IEEE Photonics Technology Letters, vol.19, no.7, pp. 468-470, 2007.
[10] J.Bromage, P.J.Winzer, L.E. Nelson, M.D. Mermelstein, C.Horn, and C.H. Headley, " Amplified spontaneous emission in pulse-pumped Raman amplifiers, " IEEE Photon. Technol. Lett., vol. 15, no.6, pp.667-669, June. 2003.
[11] S. R. Moon, S. H. Yoo and C.H. Lee, " Effect of noise distribution in a system seeded by a spectrum-sliced ASE", Journal of Lightwave Technology, vol. 32, no.12, pp.2272-2276, 2014.
[12] F.H. Tithi and S. P. Majumder, "Bit Error Rate Performance Limitations due to Raman Amplifier Induced Crosstalk in a WDM Transmission system," Journal of optical communication, Germany, 2015.
[13] F.H. Tithi and S. P. Majumder, "Performance limitations of a WDM coherent optical transmission system due to Raman amplifier induced crosstalk", The 18th international conference on computer and Information Teachnology (ICCIT), 2015.

Fast Real and Reactive Power Flow Control of Grid-Tie Photovoltaic Inverter

R. Sadnan,[1*] and Md. Ziaur Rahman Khan [1]

[1]Dept. of EEE, Bangladesh University of Engineering and Technology, Dhaka, Bangladesh
[*] rabayet.sadnan@gmail.com

***Abstract*— A fast power flow control algorithm for a grid tie Photovoltaic inverter is presented here. The proposed method has the merits of design simplicity. The algorithm is fast and has the virtue of quick response to the change of required power output for the PV inverter. Current mode asynchronous sigma delta modulation is implemented to ease the controlled sensitivity to the grid current distortion. This process also opens a window to reduce the harmonics in the current. The simulation is performed in MATLAB/Simulink. A comparison is provided to observe the response behaviour of the algorithm.**

***Index Terms*—Photovoltaic inverter, reactive power control, real power, grid-tied inverter, Current mode asynchronous sigma delta modulation, ac main cycle.**

I. Introduction

Photovoltaic (PV) power generation is an epitome of green renewable energy source. Chief impetus for wandering different aspects of PV generators is the forecasted dearth of fossil fuel. Because of the high cost of the batteries, the most efficient approach is to use grid-tied PV system without the energy storage facilities [1] instead of isolated PV system. Grid-tied PV systems need to supply reactive power as well as real power to the grid, which is very important for the stability of a power system. This feature of the PV system plays important role to instigate more exploration of controlling power injection [2], [3].

Due to the PV system's imperative nature, different control strategies for regulation of reactive and active power has been developed for grid tied inverter [4], [5]. Maknouninejad proposed a new control scheme that enables inverter of the PV system to absorb a little active power from the grid, regulate its DC bus voltage within limits, and inject the desired level of reactive power, even when the PV power is unavailable [4]. This control scheme enables us to operate only in VAR mode. A new theory of instantaneous power in single phase circuits which operate in transient states, is presented in [5]. The theory is based on Hilbert transform. In [6], Bojoi implemented a sinusoidal signal integrator along with the p-q theory to reduce the controlled sensitivity to the grid voltage distortion. Phase locked loop (PLL) for adjusting injected power is proposed by Q. Zhang in [7]. Discrete Fourier transform PLL is used by Liu in [8], to reach more precise reactive power control.

Robust pole placement technique, determined using the linear matrix inequality with D-stability criteria, is used to design the controller by Sampaio [9]. These methods require a powerful processor or software platform to process, as these methods encompasses complex mathematical expression and more intricate steps.

Sun and Liu in [10], [11] proposed a simplified method for controlling the output power of the inverter. The reference current is controlled here and the method is scalar in nature. This scalar current reference generation (CRG) is much easier than the previously used methods [12]. Though it lowers the computational burden compared to the previous practises, many complicated functions are still required, which limits the burden too.

For further reduction of computational burden, Chang [12] proposed an effective output power calculation method using only two sampled current value within one ac mains cycle. The method can commendably measure the injected active and reactive power. It uses iterative method to generate reference current for the inverter pulse generators. Though this method is simpler in nature, but the iteration process makes the convergence time consuming.

Here a new fast and effective algorithm for CRG block is proposed. To ease the controlled sensitivity to the grid current distortion, current mode asynchronous sigma-delta modulation (CMASDM) is adopted for the PV inverter's output current control strategy [13].

In this paper, the methodology is outlined and the proposed algorithm is provided in section II. In section III, the simulated results and the comparisons with the available literatures is provided. The observations from the results are described and concluding remarks are drawn in section IV.

II. Methodology

The proposed controller model consists of the simplified output power calculation, power adjustment block - which uses the new fast algorithm. These first two blocks generate the necessary reference for the output current of inverter. Finally, the last block of the model is CMASDM described as in [13]. This block receives the reference current, generated from the previous block. And this CMASDM block then produces required gate pulses for the switches of the inverter.

A. Output Power calculation

The output power of inverter can be calculated very easily. Two current samples at two selected points within one ac cycle can be very handful to measure both supplied power and var. The grid voltage is assumed to be purely sinusoidal for this method, so that the output power can be directly measured by injected sinusoidal current [12].

The active output power of the inverter can be determined or calculated as

$$P_{out} = \frac{1}{2} V_m I_m \cos\theta = \frac{1}{2} V_m i_{ac}(t_p) \quad (1)$$

Similarly, the reactive output power (var) of the inverter can be determined as:

$$Q_{out} = \frac{1}{2} V_m I_m \sin\theta = \frac{1}{2} V_m i_{ac}(t_q) \quad (2)$$

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

Where,

$v_{ac}(t) = V_m \sin(\omega t)$

$i_{ac}(t) = I_m \sin(\omega t - \theta)$

t_p = Time when the voltage is maximum.

t_q = Time when the voltage is crossing zero.

So, from above equation (1) and (2), the active and reactive power of the inverter can be calculated by two sampled current values. This method reduces tremendous amount of calculative load for the processor. And is very simple way to calculate injected active and reactive power.

B. Power adjustment block

In [12], the author, et al. proposed an iterative method to adjust the required power. The conceptual diagram is given below in fig. 1.

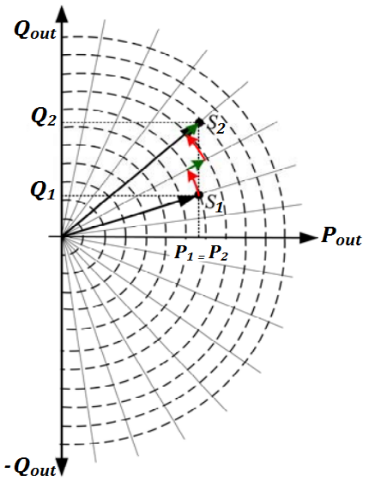


Fig. 1 Conceptual diagram for operating point of PV inverter

To move from point S1 to S2, number of steps are needed in the iterative method. It can be easily seen from the above figure. So the time required for power adjustment is high. A new method is proposed in this paper. The method is to move from point one operating point to another directly, instead of by small steps and adjustments. Here, for S2 point,

$$\theta_{new} = \tan^{-1}\left(\frac{P_{REF}}{Q_{REF}}\right) \quad (3)$$

And, let the change in magnitude of the current due to the change in operating point is ΔI_m

$$\Delta I_m = \frac{2P_{REF}}{V_m}\left[\frac{1}{\cos\theta\cos\Delta\theta - \sin\theta\sin\Delta\theta} - \frac{1}{\cos\theta}\right] \quad (4)$$

Where,

P_{REF} = Reference active power

Q_{REF} = Reference reactive power

θ_{new} = Phase angle of current for P_{REF} and Q_{REF}

θ = Initial phase angle of current

$\Delta\theta = \theta_{new} - \theta$

So, the new magnitude of current becomes,

$$I_{m_new} = I_m + \Delta I_m \quad (5)$$

The reference current generation from v_{ac}, i_{ac}, P_{ref}, Q_{ref}, can be shown in a simple block diagram in fig. 2.

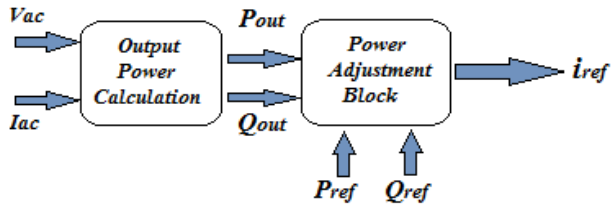


Fig. 2 Basic block diagram of reference current generation

C. CMASDM

Now, by using equation (3) and (5), the reference voltage for ASDM, the pulse generation block, can easily be produced [13]

$$|\bar{v}_{ref}| = \frac{V_{CC}}{V_{DC}}\left(|\bar{v}_{ac}|^2 + \omega^2 L^2 I_{m_new}\right)^{1/2} \quad (6)$$

$$\angle\theta_{ref} = \tan^{-1}\left(\frac{\omega L\, I_{m_new}}{|\bar{v}_{ac}|}\right) \quad (7)$$

Where, $\omega = 2\pi f$ and, L is the output inductance of the inverter. The basic block diagram of CMASDM is shown below.

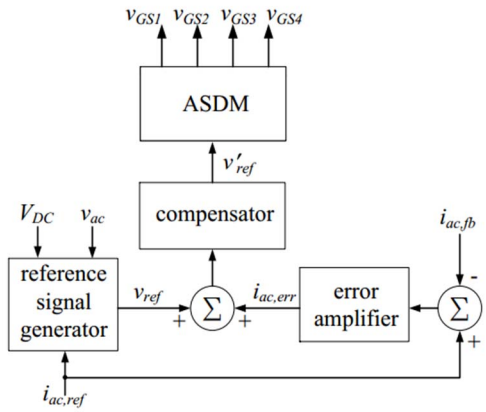


Fig. 3 Basic block diagram of working principle of CMASDM [13]

Here, the four pulse signals are generated by the control of v_{ref}. The pulses generate the required output current from the inverter. The relation between the output current and v_{ref} is [13]

$$i_{ac} = \frac{k\,.\,v_{ref} - v_{ac}}{j\omega L} \quad (8)$$

Where, k is a constant, which depends on the CMASDM system, and L is the output inductance. So, this block generates the desired pulses for reference power and var, and thus the required output current can be produced just within one cycle.

III. Results and Discussions

For simulation, at first, it has been assumed that a 1 kVA inverter is supplying power at 0.8 lagging power factor. So, P = 0.8 pu and Q = 0.6 pu The whole simulation is done by maintaining per-unit method for simplicity. Three cases have been considered for the comparison. Firstly, the real power assumed constant, and only reactive power is changed to 0.4 pu then the reactive power kept constant, while the real power was reduced to 0.65 pu and then finally both the active and reactive powers are changed. For these three cases, it was clear that the proposed algorithm is way faster than the iterative method. The simulation results are shown below. All the cases are started from the lagging 0.8 p.f. point, i.e., P = 0.8 pu and Q = 0.6 pu.

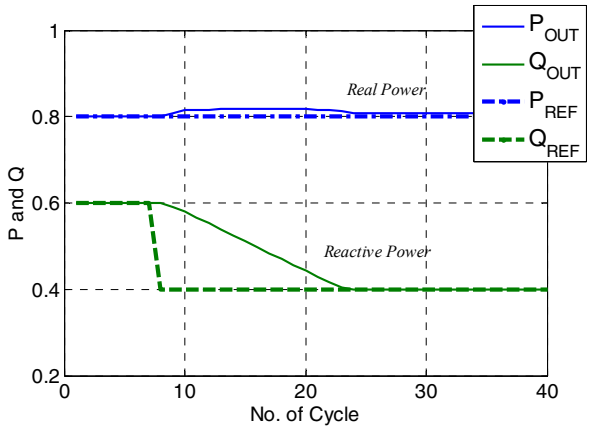


Fig. 4a Real (P) and reactive (Q) power for case I using iterative algorithm

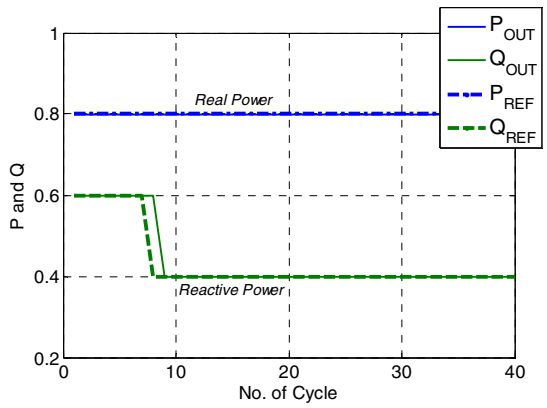


Fig. 4b Real (P) and reactive (Q) power for case I using proposed algorithm

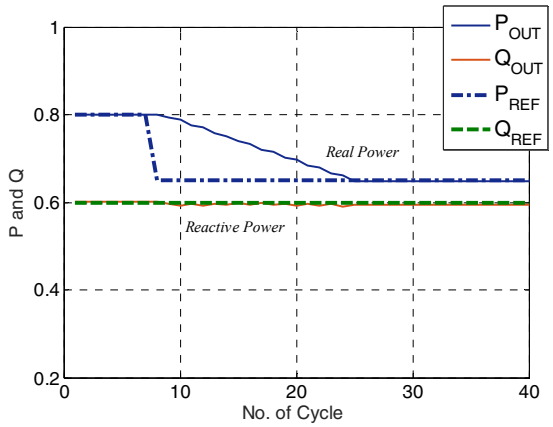


Fig. 5a Real (P) and reactive (Q) power for case II using iterative algorithm

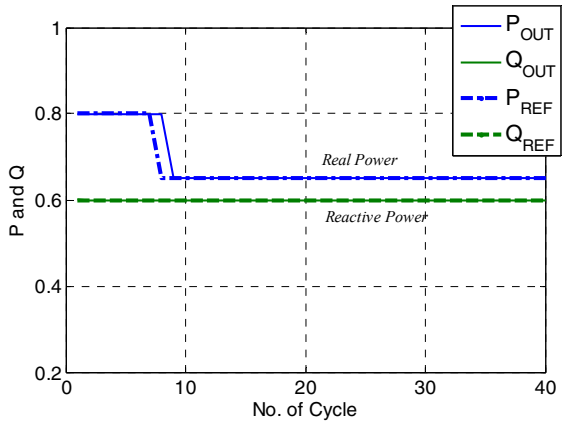


Fig. 5b Real (P) and reactive (Q) power for case II using proposed algorithm

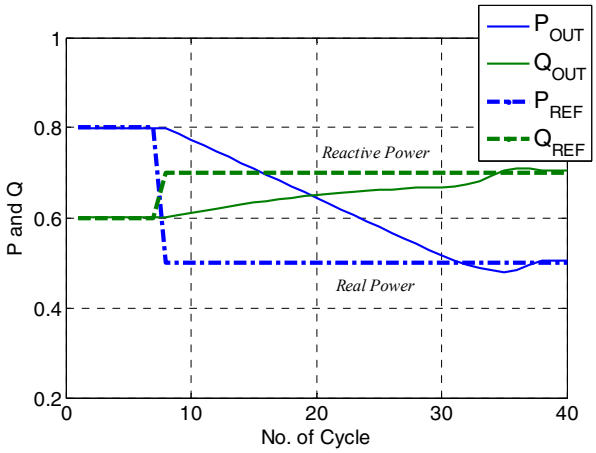


Fig. 6a Real (P) and reactive (Q) power for case III using iterative algorithm

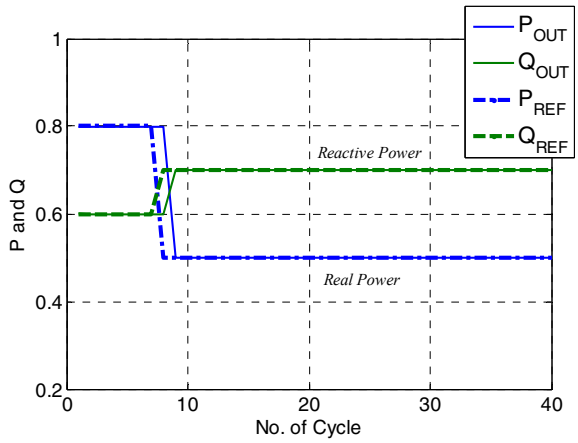


Fig. 6b Real (P) and reactive (Q) power for case III using proposed algorithm

A. Case I:

In this case, the real power P is kept constant, but the reactive power Q is changed from 0.6 pu to 0.4 pu. The proposed power adjustment method changes the power just after one ac cycle (Fig. 4b), But when the iterative method is used, the output power is adjusted to the reference power after almost 17 cycles. also, there occurs a transient in the real power output (Fig 4a).

B. Case II:

For case II, the reactive power output (Q var) is kept constant at 0.6 pu. But the reference of real power P is changed abruptly from 0.8 to 0.65 pu. This rapid change in real power can be adjusted very shortly by proposed fast algorithm (Fig 5b), but the iterative method takes almost 18 ac cycles to produce stable output. It can be seen (Fig 5a) that the proposed algorithm takes only one main cycle here as like as the previous case.

C. Case III:

For this case, we assumed that both the real and reactive power (P and Q) has to be altered. The new P changes to 0.5 pu, and the Q increased to 0.7 pu. For this case, the iterative method needs more than 30 cycles to stabilize the output (Fig 6a). And again, there are ripples in the power waves. But the proposed algorithm needs one cycle again (Fig 6b). And the ripple in the wave is very much undetectable.

For above three cases, a brief comparison between two methods is represented in Table I.

TABLE I

COMPARISON OF TWO ALGORITHMS FOR DIFFERENT CASES

Case	Power Adjusting Time (ms)		Ripple in Power during changing	
	Iterative method	Proposed method	Iterative method	Proposed method
Case I	340	20	High	Very low
Case II	360	20	High	Very low
Case III	600	20	High	Very low

IV. CONCLUSION

This paper proposes a new fast and effective algorithm for generating reference current instead of slow iterative process along with the CMASDM for single phase grid-tied PV inverter. In this paper, real and reactive power control of PV inverter using new algorithm is evaluated and compared with the iterative method of control. The proposed fast algorithm can significantly increase the response time of PV inverter to generate desired active and reactive power, without altering the simplicity of calculation and moreover, reduction in computational burden is still present. Computer simulations for comparing proposed algorithm and iterative method are shown to verify the efficacy of suggested new algorithm.

REFERENCES

[1] Li, Xiaodong, and Ashoka KS Bhat. "A utility-interfaced phase-modulated high-frequency isolated dual LCL DC/AC converter." IEEE Transactions on Industrial Electronics 59, no. 2 (2012): 1008-1019.

[2] Cagnano, Alessia, Enrico De Tuglie, Marco Liserre, and Rosa A. Mastromauro. "Online optimal reactive power control strategy of PV inverters." IEEE Transactions on Industrial Electronics 58, no. 10 (2011): 4549-4558.

[3] Sinsukthavorn, Worpong, Egon Ortjohann, Alaa Mohd, Nedzad Hamsic, and Danny Morton. "Control strategy for three-/four-wire-inverter-based distributed generation." IEEE Transactions on Industrial Electronics 59, no. 10 (2012): 3890-3899.

[4] Maknouninejad, Ali, Nasser Kutkut, Issa Batarseh, and Zhihua Qu. "Analysis and control of PV inverters operating in VAR mode at night." InInnovative Smart Grid Technologies (ISGT), 2011 IEEE PES, pp. 1-5. IEEE, 2011.

[5] Saitou, Makoto, and Tosihisa Shimizu. "Generalized theory of instantaneous active and reactive powers in single-phase circuits based on Hilbert transform." In Power Electronics Specialists Conference, 2002. pesc 02. 2002 IEEE 33rd Annual, vol. 3, pp. 1419-1424. IEEE, 2002.

[6] Bojoi, Radu Iustin, Leonardo Rodrigues Limongi, Daniel Roiu, and Alberto Tenconi. "Enhanced power quality control strategy for single-phase inverters in distributed generation systems." IEEE Transactions on Power Electronics 26, no. 3 (2011): 798-806.

[7] Zhang, Qi, Xiang-Dong Sun, Yan-Ru Zhong, Mikihiko Matsui, and Bi-Ying Ren. "Analysis and design of a digital phase-locked loop for single-phase grid-connected power conversion systems."IEEE Transactions on Industrial Electronics 58, no. 8 (2011): 3581-3592.

[8] Liu, Liming, Hui Li, Zhichao Wu, and Yan Zhou. "A cascaded photovoltaic system integrating segmented energy storages with self-regulating power allocation control and wide range reactive power compensation." IEEE Transactions on Power Electronics 26, no. 12 (2011): 3545-3559.

[9] Sampaio, Leonardo P., Moacyr AG de Brito, Guilherme de A. e Melo, and Carlos A. Canesin. "Grid-tie three-phase inverter with active power injection and reactive power compensation." Renewable Energy 85 (2016): 854-864.

[10] Sun, Pengwei, Chuang Liu, Jih-Sheng Lai, and Chien-Liang Chen. "Grid-tie control of cascade dual-buck inverter with wide-range power flow capability for renewable energy applications."IEEE transactions on power electronics 27, no. 4 (2012): 1839-1849.

[11] Liu, Chuang, Pengwei Sun, Jih-Sheng Lai, Yanchao Ji, Mingyan Wang, Chien-Liang Chen, and Guowei Cai. "Cascade dual-boost/buck active-front-end converter for intelligent universal transformer." IEEE Transactions on Industrial Electronics 59, no. 12 (2012): 4671-4680.

[12] Chang, Chia-Hsi, Yu-Hui Lin, Yaow-Ming Chen, and Yung-Ruei Chang. "Simplified reactive power control for single-phase grid-connected photovoltaic inverters." IEEE Transactions on Industrial Electronics 61, no. 5 (2014): 2286-2296.

[13] Chen, Yaow-Ming, Chia-Hsi Chang, and Kuan-Yu Liu. "Grid-tied inverter with current-mode asynchronous sigma-delta modulation." In 2009 IEEE Energy Conversion Congress and Exposition, pp. 928-933. IEEE, 2009.

On TDMA based Hybrid Channel MAC Protocol for Underwater Sensor Networks

Md. Ibrahim Ibne Alam[1] and Md. Farhad Hossain[2]
Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology
Dhaka-1000, Bangladesh
[1]amitu91191@gmail.com, [2]mfarhadhossain@eee.buet.ac.bd

***Abstract*—Underwater sensor networks play a great role in research areas by gathering knowledge about the inaccessible parts of the underwater. The sensory nodes in the network aid us by performing real time data acquisition which is needed to forecast different natural calamities. However communication in underwater with high success rate is quiet challenging because of the nature of the environment. Due to the high attenuation of EM wave in water, large amount of power is needed to communicate with EM waves. Therefore acoustic waves are mainly used for underwater communication. However, the combination of the EM and acoustic waves can be a better choice for that purpose. Here we have proposed a combination of both the EM and acoustic waves for our MAC protocol. The model uses the EM in the control channel whereas the acoustic wave carries the data. The control channel uses TDMA to avoid any control packet collision and thus making sure that no data packet collision will occur. To our best knowledge this type of proposal with EM wave has not yet been stated anywhere else. The simulation results show that the proposed scheme achieves higher throughput than other MAC schemes examined in this paper.**

***Index Terms*—Electromagnetic wave, Medium Access Control, Time Division Multiple Access, Underwater Sensor Networks.**

I. Introduction

Underwater Sensor Networks are networks composed of underwater nodes with the ability to sense, communicate and process signals. MAC protocols with good performance and efficiency are mandatory for these nodes. Underwater research, exploration and the prediction of natural calamities depends solely on the information send by these nodes. However due to the nature of the environment in underwater we cannot use same protocols for communication in underwater as we use in terrestrial communication. The EM waves suffer very high attenuation in underwater and a wise choice is to use acoustic waves as communication medium. Even if want to use EM waves in communication, nodes with very high power ratings are needed [1], [2].

Thorough researches on MAC protocols for underwater communication started in early 2000s. The ALOHA technique is very popular for underwater communication. In this technique any node with a data packet can send the data immediately without any delay. Though this method increases the likelyhood of collison, the simplicity and low traffic nature of underwater has kept it favorable still now. In [1] the authors mathemtically analyzed ALOHA and slotted ALOHA for underwater accoustic networks. They have shown that slotted ALOHA doesn't work that much better than ALOHA because of the high propagation delay in underwater. Also In [3] the authors showed through simulation that slotted ALOHA can not perform well due to spatio-temporal uncertainty. Also they proposed for guard bands which dramatically increases the performance of the protocol. To use the high propagation delay in advantage Chridchoo et al.[4] proposed two schemes which performed really better than the ALOHA-Carrier Sense (CS). The proposed schemes were ALOHA-Collision Avoidance (CA) and ALOHA-Advanced Notification (AN). Our proposal compares with these two protocols mainly.

In the another research, Peleato and Stojanovic in [5] proposed a handshaking method that decreases collision to almost zero and achieves very high throughput along with saving power consumption. It uses a short packet after the CTS to avoid collision if there becomes a possibility of it. Otherwise only the CTS is sent. In [6] the writers proposed scalable localization of the nodes by Mobility prediction. They used 3 types of nodes for their proposed scheme. These nodes after going through some algorithm can predict their future mobility and can give us a good idea about its localization. To minimize the power consumption by the nodes, Syed et al. [7] proposed T-Lohi MAC scheme which uses a short tone to wake up nodes at the correct time to receive data. The scheme aimed at fighting the space time uncertaitny and deafness of the nodes. For that purpose it uses a tone based reservation mechanism. In [8] the authors suggested concurrent bidirectional data packet exchange to improve the data transmission effieciency. They adopted a packet bursting idea, where a sender–receiver pair can exchange multiple rounds of bidirectional packet transmissions. The whole idea was given the name bidirectional-concurrent MAC with packet bursting (BiC–MAC). In recent works, Liu et al. [9] gave the idea of utilizing the spatial correlation of underwater mobile sensor nodes to estimate the long dynamic propagation delays. Which eventually resulted in better time synchronization of the nodes. Inspite of all the protocols proposed they cannot perform as well as the theoretical result. Therefore Zhu et al. [10] in their works examined the practical feasibility of the MAC protocols. They found out after extensive research that the time sharing MAC protocols are quite promising.

Our work is based on combining the EM and Acoustic channel in a single platform. Within our best knowledge this approach in underwater communication is a novel one. In this MAC protocol two channels, one based on EM wave and another with the acoustic wave are being proposed. As the EM waves attenuate rapidly in underwater and needs a huge amount of power to propagate longer distances, it is used for controlling the nodes only. Very short control packets can perform this task. On the other hand the acoustic channel is doing its traditional task of carrying the data packets. This scheme uses TDMA technique in the control channel where time slots are assigned for every node to reserve channel and a time slot is dedicated for the confirmation message of successfully received packet. The nodes can hear each other

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

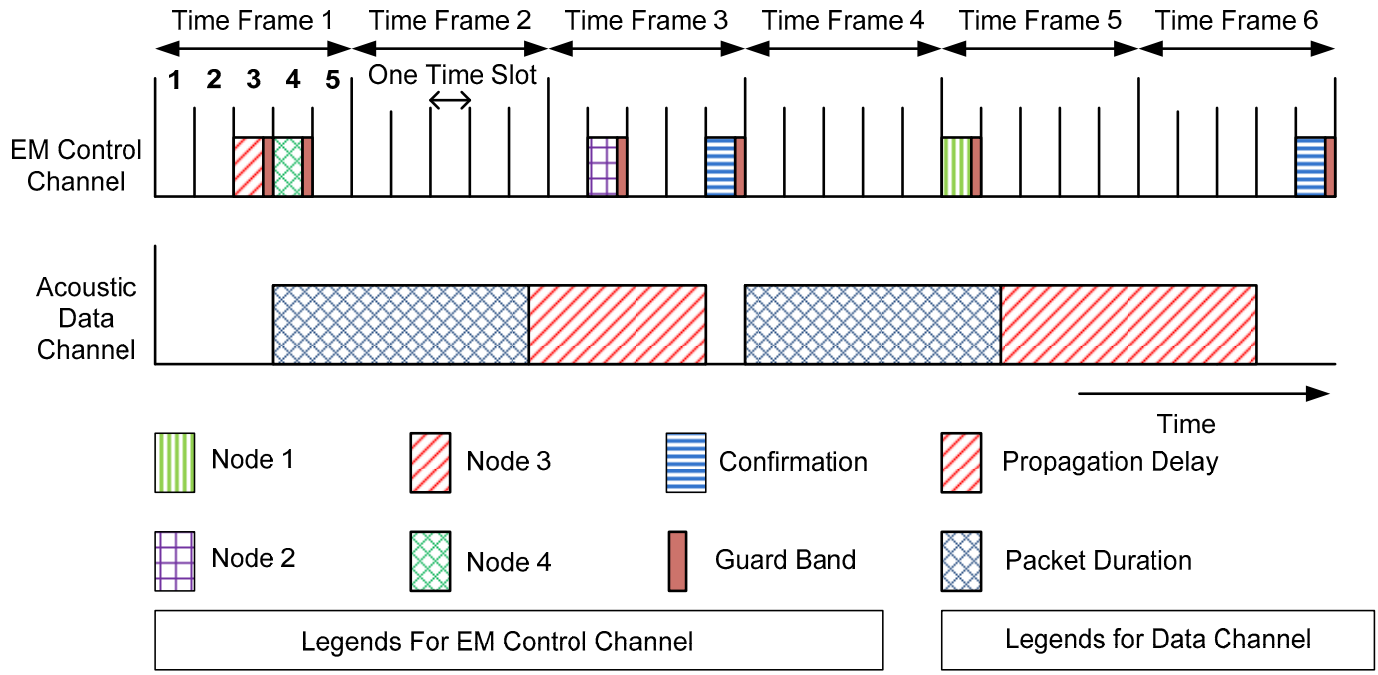


Fig. 1: Time Series Diagram

and maintain a table that keeps the requests sequence of the nodes along with the current state of the data channel if it is idle or carrying packet. When the data channel is free, the transmitter node that is first in the table starts transmitting packet through the channel. This arrangement thus avoids all the collision between data and thus making a reliable MAC protocol.

The remainder of this paper is arranged as follows. Section II gives the overview of the system model of the proposed scheme with the node distribution and time series diagram. In section III we present the simulation results along with their comparison with other MAC protocols. In section IV we summarize all our achievements and future plans in the conclusion part.

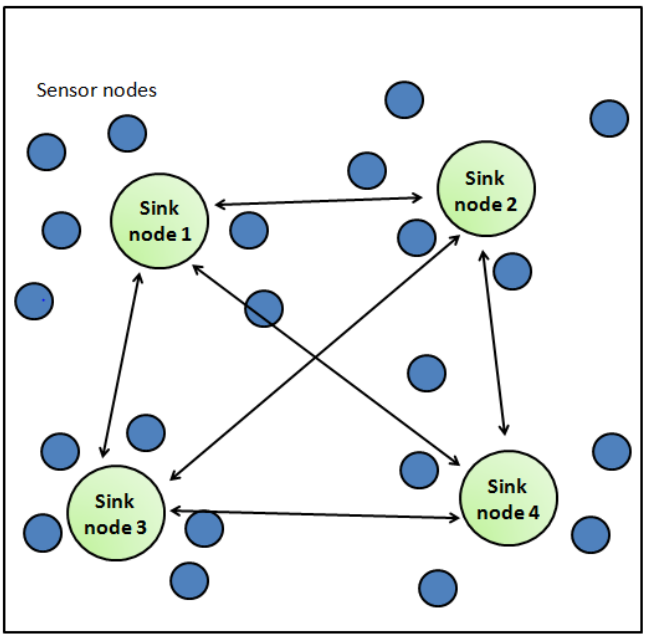


Fig. 2: Network topology

II. System Model

In the proposed network the nodes are static and are scattered in a uniform distribution within an area. The nodes have the ability to simultaneously control both half duplex antennas of EM and acoustic waves. The packet generation rate of any node is Poisson distributed and the intended receiver is chosen randomly with equal probability. In fig. 2 the node distribution underwater is depicted. The sink nodes are our concerned nodes which have the ability to communicate with each other and they are within one single hop distance of each other. The sensor nodes (the smaller ones) communicate only with the sink nodes closer to them. The sink nodes use EM channel as the control channel and acoustic channel for data channel. We are focusing our study on communication between the sink nodes only. The EM channel is TDMA distributed with (n+1) slots per frame. Where n is the total number of nodes in the area. The last slot in the time frame is for data channel status if it is idle or not. When the nodes have data to transmit they ask for reservation through their respective time slot, i.e. number 1 time slot is for node 1. Every node maintains a table where the requests being made by the nodes are kept sequentially. Also the nodes keep the status of the data channel if it free or being used. As this scheme provides a collision free environment, the power consumption for EM wave can be seen as a replacement for the power which would have been wasted in collisions of data packets.

For better illustration on how the system works, a time diagram is given in fig. 1. It can be seen that there are two channels, namely the EM based control channel and acoustic based data channel. To avoid complexity a 4 node system is selected. So the one time frame will consists of 5 time slots, where the first 4 are for the 4 nodes and the last one is for confirmation message of receiving a packet. It is assumed that at first the data channel is free and all the nodes know this. For the scenario depicted in fig. 1, node 3 has the first data packet to transmit. Therefore it asks for reservation through the control channel. It is assumed that all the nodes will hear this reservation call. Consequently all the nodes will update their table of reservation sequence and node 3 will be written at the

top of the table. Since the data channel is free, node 3 will start transmitting data immediately after the reservation call. As the other nodes knew that the channel was free they will assume that node 3 has started transmitting data just after the reservation call. So they will delete node 3 from their tables. Meantime the nodes with intended data to transmit will continue their reservation through their time slots. All the nodes will update their table to maintain the serial of the nodes with transmission requests.

Now continuing through the time series diagram it can be observed that the next contender to transmit data is node 4, so it will make a reservation through its respective time slot (slot 4). All the sink nodes hearing it will update their table. Since the data channel is now occupied node 4 will have to wait for the confirmation message before it can transmit. Whereas in this waiting period node 1 makes a data sending request through the next time frame using time slot 1. Now this request will be queued up in the table of all the nodes. It can be seen that the first data packet with its propagation delay and packet duration reaches its destination before time slot 5 of the time frame 3. Time slot 5 is the confirmation slot for sending signal that the data packet has been received. Thus when the respective receiver receives a packet it transmits a confirmation message though the timeslot 5. This in return ensures that all the nodes get the information that the data channel is now free to use for transmission. Now the first transmitter in the queue, which is node 4 starts transmitting its packet. In parallel the nodes update their table by removing node 4 from the table and shifting the list up one line. Now node 1 will be at the top of the transmission queue table. This process goes on for the whole time series.

III. Simulations and Results

A. Simulation model

Following the theoretical analysis presented in section II, we created the code for finding the total offered load vs. normalized throughput graphs. The network is made up with a 4 node static system with every distributed uniformly through a 3000m by 3000m area. All the nodes have fixed data packet rate of 2400 bps. The control packet size and control packet data rate for the EM channel is made variable. Also the data packet length of the acoustic channel is made variable. Every time slot is in the EM channel is made up with control packet duration plus a guard band time of (0.00015s). With 4 nodes, the total time frame contains (4+1) number of time slots. Thefirst 4 time slots are used by the nodes to reserve channel and the last time slot is to transmit the confirmation message of receiving a data packet. The simulation is run for 100 seconds and for a total traffic load of 0.1 to 1.6. The Simulation time for every traffic load is made inversely proportional to the generation rate to decrease computational time. The total traffic load is normalized for 2400 bits packet. Thus the total traffic load can be calculated as;

$$\text{Total Load} = (\text{Packet length}/2400) \times \text{Generation rate} \quad (1)$$

For the first part of our simulation, we choose three sizes of data packets for the data channel, namely 2400 bits, 4800 bits & 9600 bits. The control packet is fixed with 32 bits and the control channel data rate is 480 bps. The simulation is run to find the throughput for different traffic loads. The output of this simulation is then compared with the throughput of some established MAC protocols. Then in the second part, the simulation is done by varying the control packet rate from 120 bps to 720 bps with increment of 120 bps per step. The data packet size, data rate and control packet size is maintained to

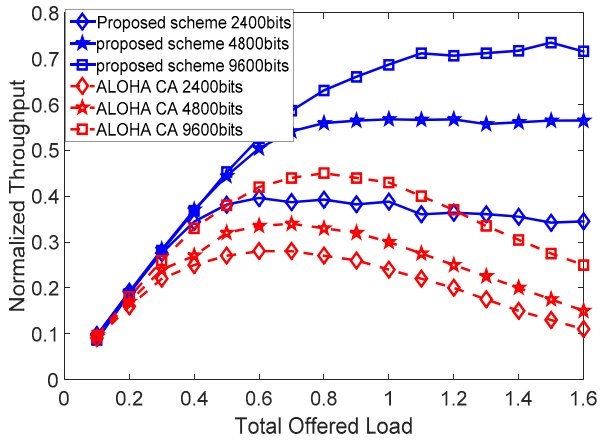


Fig. 3: Throughput comparison of proposed scheme with ALOHA CA

be 9600 bits, 2400 bps and 32 bits respectively. After that the code is modified to vary the control packet size with fixed control packet rate of 480 bps. In the last part of the simulation the code is ran to find the average waiting time of the nodes to transmit data for different traffic loads. The average waiting time for a node is the difference between the packet generation time and the start of the packet transmission time. This simulation was done with 3 different sizes of control packets with 9600 bits data packet.

B. Simulation Results

In figure 2, the comparison between the proposed schemes throughput with the throughput of ALOHA-CA is done. It is noteworthy that our hybrid channel is very much superior to the ALOHA-CA protocol. Also it can be seen that data packet with length 4800 bits of the suggested model outperforms the ALOHA-CA 9600 bits data packets. The highest throughput for ALOHA-CA with 9600bits data packets is 0.45 whereas the proposed method with 9600 bits has around 0.7 throughputs for high total offered load.

Now in fig. 3 the comparison is made with a very efficient MAC protocol namely ALOHA-AN. ALOHA-AN analyses the surrounding data signals and can notify the nodes of probable collision beforehand. However with the little help from the EM waves, the proposed simplistic approach have surpass even ALOHA-AN. It is notable from the figure that ALOHA-AN with data packet of size 2400bits bits cannot perform as well as the hybrid channel scheme. Though for low offered load ALOHA-AN with 9600bits have slightly higher throughput than our proposal, at high load our proposed model outperform ALOHA-AN quite significantly. The ALOHA-AN used here has a lag time of 1 second which gives the best performance for it.

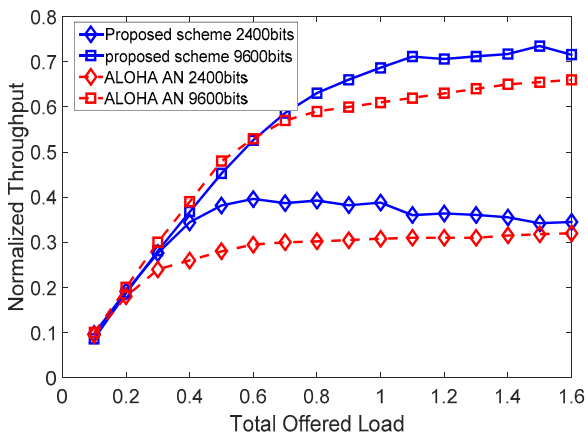


Fig. 4: Throughput comparison of proposed scheme with ALOHA AN

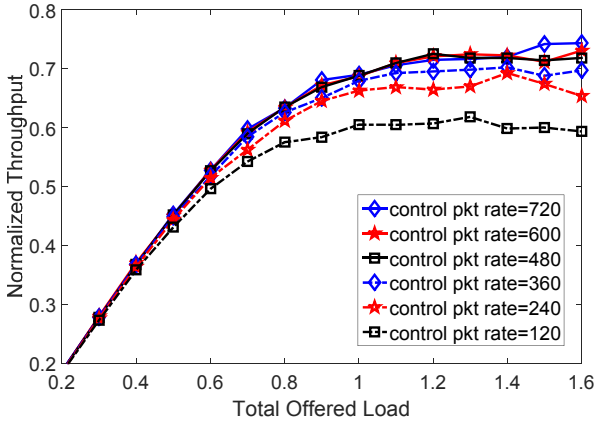


Fig. 5: Throughput of the hybrid channel for different control packet rate

With the next plot (fig. 4), the throughput variation with the change of control packet data rate is depicted. It can be seen that if the data rate gets lower than 240 bps with 32 bits control packet, then the proposed scheme performs worse than the ALOHA-AN. However it outperforms the other ALOHA even with very low bps, like 80-100 bps. Another thing that is noteworthy is that the hybrid channel has consistent throughput for high loads due to its no collison environment. It can be easily observed that with higher data rate higher throughput is achievable, however after 480 bps the throughput increment becomes quite low. The simulation done next is almost the counter part of the previous one that is changing the control packet size (Fig. 5). The variations in throughput for different control packet sizes for a fixed 480 bps control packet rate can be noticable in the figure. The performance degrades with the increment of length in control packet and becomes inferior than ALOHA-AN. It is important to note that with the doubling of packet size the throughput decreases almost constantly for packets larger or equal to 16 bits.

The last figure (fig. 6) in our paper portraits that with the increment of control packet size with fixed control rate of 480 bps, slight increment in the average waiting time of the nodes to transmit their respective data packets is observed. This is due to the fact that bigger control packet means bigger TDMA slot which in fact makes the nodes to wait longer if they want to reserve the channel. So making the control packet smaller or making the control channel data rate faster can minimize the waiting time and make the system optimum.

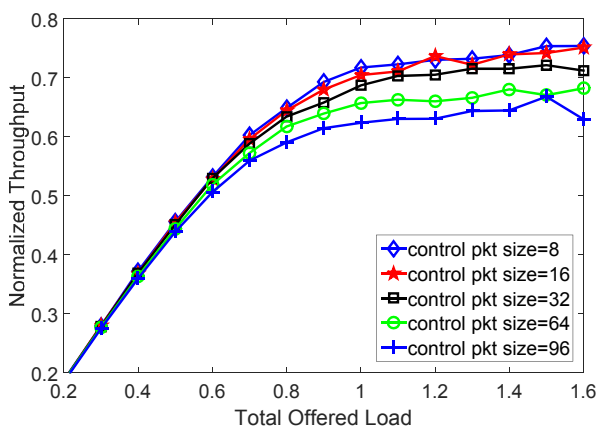


Fig. 6: Throughput of hybrid channel for different control packet size

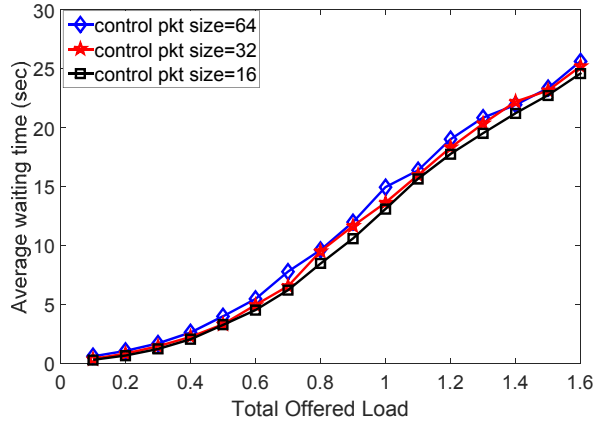


Fig. 7: Average waiting time of a transmitting node for different control packet size

IV. Conclusion

In this paper, a novel MAC protocol based on the combination of EM and Acoustic channel is proposed. The EM wave has been used for control channel and the acoustic wave for data communication. The scheme works on a channel reservation based system, with reservation calls made through the control channel. It has been observed that for different data packet sizes, the throughput remains significantly high and becomes consistent for high offered load. By varying the control packet length along with the control channel data rate, throughput can be increased even further. The system has shown higher throughput compared to some other protocols. Our future plan includes the analytical modelling of the system along with the effect of environment in it. Detailed analysis on power consumption, multi-hop networking and node idle time utilization will be focused in our upcoming works.

References

[1] L.F.M. Vieira, J. Kong, U. Lee, and M. Gerla, "Analysis of ALOHA Protocols for Underwater Acoustic Sensor Networks," *Proc. ACM International Workshop on UnderWater Networks (WUWNet),* 2006.

[2] L. Liu, S. Zhou, and J.-H. Cui, "Prospects and Problems of Wireless Communication for Underwater Sensor Networks," *Wireless Comm. and Mobile Computing,* vol. 8, pp. 977-994, 2008.

[3] A. A. Syed, W. Ye, J. Heidemann, and B. Krishnamachari, "Understanding Spatio-Temporal Uncertainty in Medium Access with ALOHA Protocols," P*roc. ACM WUWNet'07*, pp. 41–48, 2007.

[4] N. Chirdchoo, W. S. Soh and K. C. Chua, "Aloha-Based MAC Protocols with Collision Avoidance for Underwater Acoustic Networks," *Proc. IEEE International Conference on Computer Communications (INFOCOM)*, Anchorage, AK, pp. 2271-2275, 2007.

[5] B. Peleato and M. Stojanovic, "Distance Aware Collision Avoidance Protocol for Ad-hoc Underwater Acoustic Sensor Networks," *IEEE Communications Letters*, vol. 11, no. 12, pp. 1025-1027, Dec 2007.

[6] Z. Zhou, Z. Peng, J. H. Cui, Z. Shi and A. Bagtzoglou, "Scalable Localization with Mobility Prediction for Underwater Sensor Networks," *IEEE Transactions on Mobile Computing*, vol. 10, no. 3, pp. 335-348, Mar 2011.

[7] A. A. Syed, W. Ye and J. Heidemann, "T-Lohi: A New Class of MAC Protocols for Underwater Acoustic Sensor Networks," *Proc. IEEE Conference on Computer Communications (INFOCOM)*, Phoenix, AZ, 2008.

[8] H. H. Ng, W. S. Soh and M. Motani, "A Bidirectional-Concurrent MAC Protocol With Packet Bursting for Underwater Acoustic Networks," *IEEE Journal of Oceanic Engineering*, vol. 38, no. 3, pp. 547-565, July 2013.

[9] J. Liu, Z. Zhou, Z. Peng, J. H. Cui, M. Zuba and L. Fiondella, "Mobi-Sync: Efficient Time Synchronization for Mobile Underwater Sensor Networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 2, pp. 406-416, Feb 2013.

[10] Y. Zhu, Z. Peng, J. H. Cui and H. Chen, "Toward Practical MAC Design for Underwater Acoustic Networks," *IEEE Transactions on Mobile Computing*, vol. 14, no. 4, pp. 872-886, Apr 2015.

An Efficient Cooperative MAC Protocol for Enhancing QoS of IEEE 802.11e EDCA in Saturated Conditions

Jaya Roy,[1] Md. Abubakar Siddik,[1] Shafika Showkat Moni[2] and Mohammad Shah Alam[1]

[1]IICT, Bangladesh University of Engineering and Technology
[2]CSE, Rajshahi University of Engineering and Technology
*joyarbg@gmail.com

***Abstract*— Differentiating channel access probability based on only traffic priority can ensure QoS statistically in enhanced distributed channel access (EDCA) of IEEE 802.11e. However, it may seriously affect the overall performance in presence of degraded channel condition among distant stations. In this paper, we propose a novel channel access scheme which exploits the benefit of cooperative communication in order to mitigate the adverse effect of bad channel condition while maintaining deterministically prioritized channel access for high priority traffic. Unlike EDCA, in our proposed scheme whenever a station with degraded channel condition wins the channel contention, our scheme searches for a prospective helper node to assist the sender in transmission. Eventually by exchanging sequence of busy tones and a special frame, it selects an optimal helper which has data with equal or higher priority as well as supports significantly higher data rate for cooperative transmission. Simulation results demonstrate that our proposed protocol enhances the QoS of the network while improving the overall performance of EDCA.**

Index Terms—**IEEE802.11e; Cooperative Communication; EDCA; Quality of Service.**

I. Introduction

The popularity of wireless technology during the last decades led to a great progress in the field of telecommunications. The expansion of electronic devices such as laptops, smart phones, tablets with social media, different VoIP application software reinforces the need of multimedia services. The most popular standard used for wireless communication is the IEEE 802.11 which designs a medium access control (MAC) protocol for sharing the wireless medium among different wireless node [1]. But it may not satisfy the increasing demand for real time application support. However multimedia applications require some quality of service support such as error rates, bit rate, throughput, transmission delay, availability, jitter, etc. To support multimedia applications with QoS requirements in the IEEE 802.11 MAC protocol, the IEEE 802.11e has been standardized [2]. The distributed access mechanism of IEEE 802.11e is called EDCA (Enhanced Distributed Channel Access), it supports Quality of Service by introducing service differentiation. Different types of traffic are assigned with different priorities based on their QoS requirements, and service differentiation is introduced by using a different set of medium access parameters for each priority.

According to the Binary Exponential Backoff (BEB) algorithm used in WLAN, a low priority or low data rate station may win contention process and thus may occupy the channel for long times. This will degrade the network performance by introducing access delay. The introduction of service differentiation in IEEE 802.11e partially mitigates this problem by allowing high priority stations to access first.

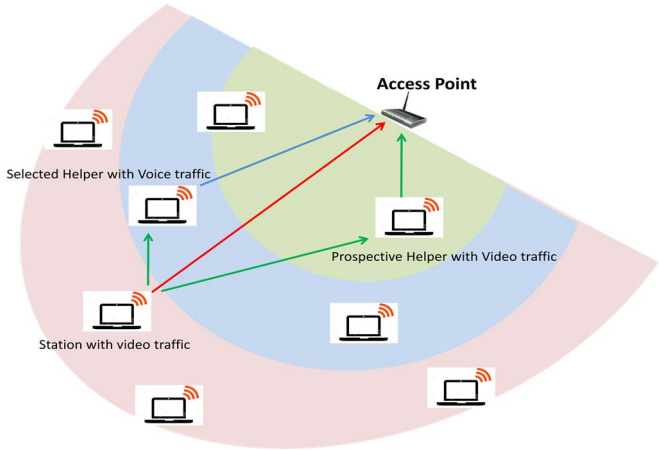


Fig. 1. Proposed Cooperative MAC protocol showing helper selection based on traffic priority.

To resolve this problem we propose an efficient EDCA based cooperative MAC protocol for wireless networks. In our protocol we adopt the concept of cooperation to enhance the network performance when a station with low or moderate data rate wins the contention and have high or low priority traffic to send, the wining station searches a node which have better channel condition than itself. Our proposed protocol is mainly focused on two main features to enhance the QoS of WLAN: Link Quality Enhancement and High Priority Traffic Accommodation. We choose a helper that supports at least two times more data rate than the direct link between source and destination plus a helper that has equal or above priority traffic than source in its queue. After selecting the optimal helper the sender than transmits its data to the helper. The helper than relay the data packets of sender to the destination and if there are enough time remains in the TXOP then the helper piggybacks its own data packets to the destination without reservation. This greatly decreases the reservation overhead and effectively improves cooperation efficiency and enhances the network performance.

This paper is organized as follows: The related work is presented in section II, section III provides the basic protocol design; in section IV simulation or analytical results are shown; finally we draw a conclusion in section V.

II. Related Work

To provide QoS at MAC layer several studies adopt EDCA for wireless communication. EDCA offers traffic differentiation to provide QoS guarantee in the MAC layer. In

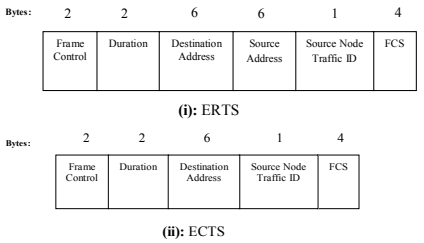


Fig. 2. Frame Formats of (i): ERTS and (ii): ECTS

DC-MAC protocol [3] the authors concentrate on the retransmission process of collided packets by using cooperation. In this protocol they consider two queues at each station and there is a partner queue for each traffic class to support cooperative retransmission. This protocol is only used for retransmission process but the QoS enhancement of the network is not applicable for this protocol. To reduce the overhead, and improve the performance of the network, the QC-MAC protocol [4] extends the CoopMAC protocol [9] by using TXOP and Block ACK parameters of EDCA. It can select a helper like [9] with little modification in RTS CTS frame format and send data packets to the helper one by one instead of a single data packet to the relay for cooperative communication. After receiving the data packets from the helper, the destination then sends a Block ACK to acknowledge the reception of all data packets to reduce the overhead. To ensure QoS guarantee in DCF mode QoS-Guaranteed MAC (QC-MAC) protocol [5] is proposed, where each sender sends busy tone or senses the channel according to the binary digits of its virtual ID. After this highest priority sender wins the contention and then it adopts the cooperation process for data transmission and piggybacks also. In this protocol service differentiation of helper node is not taken into account. To provide guaranteed channel access to multimedia applications a simple scheme called DPCA protocol [6] is proposed. This proposed protocol provides deterministically prioritized channel access for high-priority ACs and avoid collisions caused by low-priority ACs. This protocol does not consider the unfairness issue at the low priority ACs. For contention resolution several studies based on busy tone are proposed in [7], [8]. Several studies on cooperation [9], [10] have been studied to find out the selection of optimal helper. The earlier cooperative MAC protocols for EDCA do not take into account the hidden helper node problems for protocol design. We use here busy tone techniques to contend among helpers and also consider the hidden helper node problems in account. Because in the hidden terminal environment, hidden stations cannot sense the other stations' busy tones so that lower priority stations cannot sense the presence of higher priority stations.

To resolve these problems we propose a novel cooperative MAC protocol for IEEE 802.11e EDCA. Our concentration in this paper is to ensure QoS at the MAC layer with cooperation for EDCA. The novel idea of our proposed protocol is to find out a helper node in the basis of priority class and link rate. Our proposed cooperative protocol for EDCA outperforms the standard EDCA by choosing an optimal helper for cooperation that enhance QoS and network performance.

III. BASIC PROTOCOL DESIGN

In IEEE 802.11e protocol all contending stations gets a chance to access the channel to mitigate the channel fairness issue. But when a station with low or moderate link rate between receiver and itself wins the channel contention then it

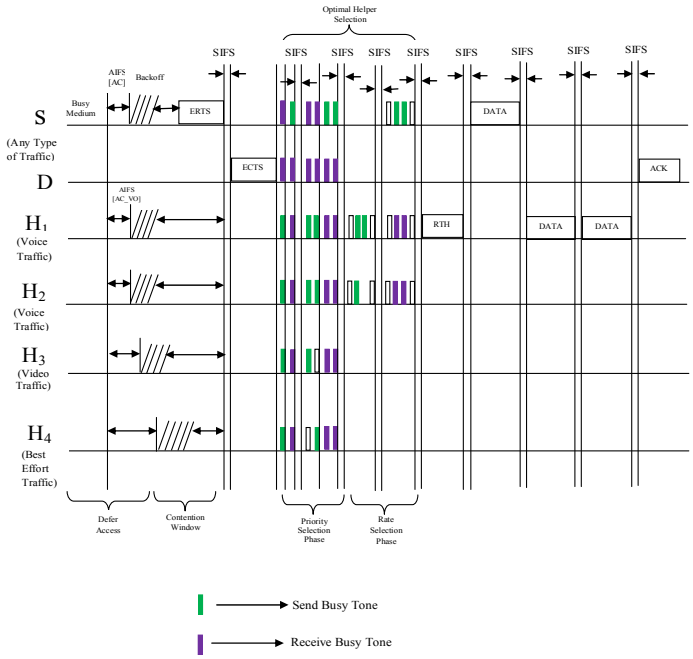


Fig. 3. Cooperative MAC for EDCA (When no collision among RTH frame)

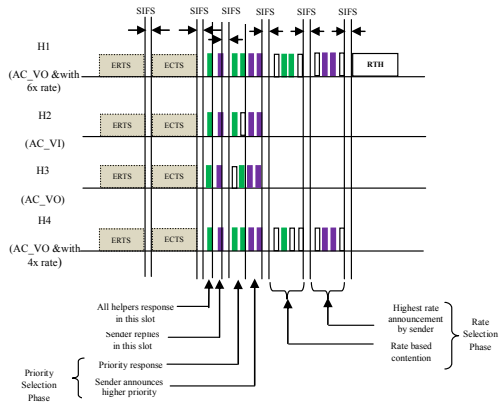


Fig. 4. Helper Selection on the Basis of Traffic Class Priority and Link Rate

will degrade the network performance and affect the QoS of the network. Our proposed protocol based on IEEE 802.11e EDCA improves the network performance with the help of cooperative communication.

A. EDCA CoopMAC

In IEEE 802.11e EDCA, traffic with different priorities is mapped into several Access Categories (ACs) at each station. For each AC, an enhanced variant of the DCF, called an enhanced distributed channel access function (EDCAF), contends for TXOPs using a set of EDCA parameters: the Arbitration Inter Frame Space (AIFS[AC]), minimum Contention Window size (CWmin[AC]), maximum Contention Window size (CWmax[AC]) [3]. Based on these parameters each station contends for access the channel independently for each AC. Our proposed protocol selects a helper to assist the sender with high priority traffic of its own and also has a high average data rate to transfer data through the helper.

1) When the source node has a packet to transmit to the destination, it senses the channel for idleness. If the channel remains idle for the AIFS duration, the source then backs off for a random duration. Once the back off counter reaches zero, the source sends the ERTS packet

TABLE I
SIMULATION PARAMETERS

Parameter	*Voice*	*Video*	*Best Effort*	*Background*
AIFS(µs)	41	57	89	121
CWmin	3	7	15	15
CWmax	7	15	1023	1023
TXOP(ms)	1.504	3.008	0	0

TABLE II
SIMULATION PARAMETERS

Parameter	*Value*
Data Rate	IEEE 802.11g
SIFS(µs)	16
Busy Tone Time(µs)	2

(at 1 Mb/s) to the destination for channel reservation. ERTS is called Enhanced RTS with slightly modification of RTS by adding 1 byte extra field for sender traffic class ID.

2) If the ERTS packet is correctly received and decoded at the destination node, it responds with a ECTS packet after the SIFS duration. ECTS is called Enhanced CTS which is also a modified version of CTS. The frame formats for ERTS and ECTS are shown in Fig1.
3) Each node that receives both ERTS and ECTS packets we can call them as a helper node. But in our proposed protocol we choose two criteria for selecting an optimal helper- First of all the node must have traffic equal or greater priority than sender node and secondly the rate through helper is at least two times or more than the direct rate between sender and receiver. The helper nodes that fulfil those requirements can participate for the next helper contention process.
4) The helper contention process (HCP) starts after SIFS time interval. The HCP starts by sending a busy tone to the source from all contending helper node. This busy tone is an indication to the source that at least one helper presence in the network. Otherwise the sender assumes that no helper presence in the network and it starts sending data through direct transmission. After receiving the busy tone from helper nodes the sender also respond with a busy tone as permission for starting cooperation process. The HCP is divided into two parts one for priority selection and another for rate selection stage-
   To select higher priority helper each node sends busy tones as like their access class (AC) ID. For example, a voice traffic class (VO_AC) has an ID 3, then it sends busy tone according to the binary representation of 3(11), for voice ID is 2 and it sends like (10), for best effort it sends 1(01). The sender then announces/repeats the busy tone of higher priority. After listening the announcement from sender all low priority nodes refrain from the contention. The remaining higher priority nodes contend for next rate selection phase.
   For higher rate helper selection each high priority nodes first calculate the average rate for relaying and then divide by the direct rate between sender and receiver. It finds that how many times the node is better than the direct link for relaying. All contending helper nodes send busy tones according to their link rate times. For example if a node supports six times better than the direct link than it sends busy tone like 0110 (binary representation of six). As a confirmation the sender also announces/repeats the tone of the high rate supporting nodes. This announcement is necessary to avoid hidden node problem among helper nodes.
5) After the above contention, winning node sends RTH frame to initiate the cooperation. If multiple helpers win the contention, then the RTH packets collide and after waiting SIFS+t_h time the helper nodes aware about the collision and resend the RTH packet again. After the collision each node starts a backoff window to resend the RTH packet.
6) After successful reception of the RTH packet the sender sends its data packets to the helper. The successful helper then relays these packets to the destination. If there are enough time remains in the TXOP the helper then piggybacks its own data packets to the receiver. The receiver on receiving all the data packets from helper then acknowledges of the successful reception to both the helper and sender.
7) In case that no relay feedback is received during the initial stage, then the source directly transmits to the destination without any helper contention phase.
8) The frame formats of ERTS and ECTS packets in Fig 2 and Fig 3 shows the proposed protocol description. The helper selection process is depicted on Fig 4.

IV. EVALUAION

In this paper, we evaluate the performance of the proposed protocol by using OMNET++ simulator, and make performance comparison with the legacy EDCA protocol. In the simulation, we take the throughput for single voice and video data and average packet delay to evaluate multiple access performance.

A. Simulation Setup

In this simulation, all the stations are randomly distributed in an area and the Access Point (AP) is located at another part from all stations. In this simulation we consider a topology where stations are randomly distributed around the AP. Packets are transmitted at different rates, depending on the distance between the sender and the receiver. The stations generate different types of traffic: voice, video, best effort and background. Table II shows the EDCA access parameters for different traffics. According to these parameters all stations contend for channel access. The following performance metrics are evaluated through the simulation-

B. Simulation Results

First of all we evaluate the throughput performance of the network for single voice and video traffic. Then we evaluate the delay for single video traffic in the network.

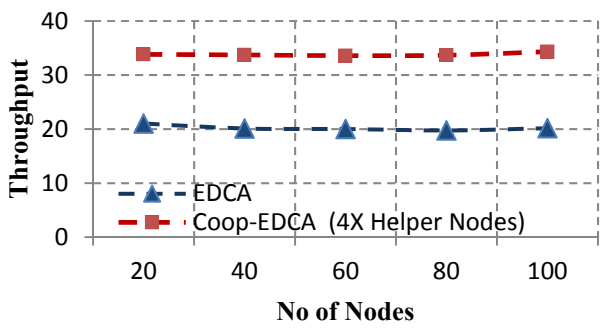


Fig. 5. Throughput vs No of Nodes for Voice Traffic.

Fig. 5 shows the saturation throughput for voice traffic achieved by both the IEEE 802.11e EDCA and our proposed cooperative EDCA with 4times data rate (4X) helper nodes in a wireless network. Our proposed protocol performs better throughput performance because the transmission rate via a 4X helper node is always a better option. Our protocol outperforms the IEEE 802.11e EDCA in case of voice traffic.

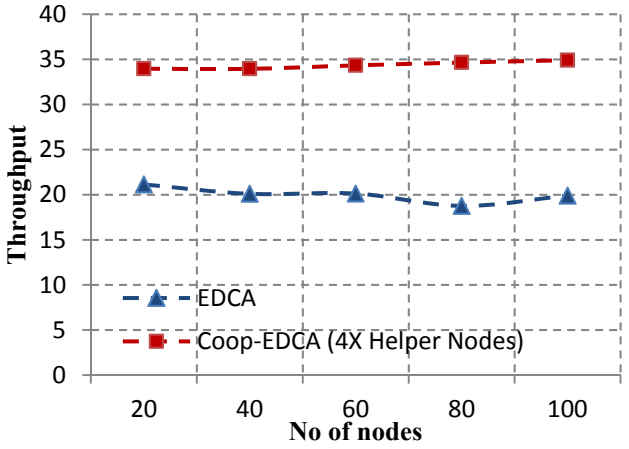


Fig. 6. Throughput vs No of Nodes for Video Traffic.

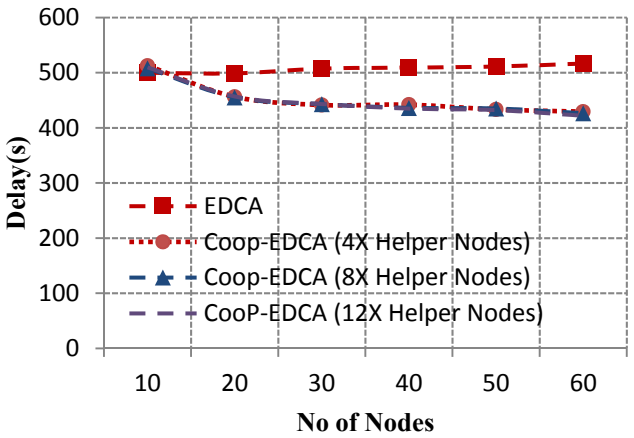


Fig. 7. Delay vs No of Nodes for Video Traffic.

Fig. 6 shows the saturation throughput for video traffic achieved by both the IEEE 802.11e EDCA and our proposed cooperative EDCA with 4X helper nodes in a wireless network. Our proposed protocol performs better throughput performance because the transmission rate via a 4X helper node is always a better option. Our protocol outperforms the IEEE 802.11e EDCA in case of voice traffic.

Fig. 7 shows the delay performance comparison of our proposed protocol for 4X, 8X and 12X helper nodes with the IEEE 802.11e EDCA. Our proposed protocol improves the delay performance, because the link through the optimal helper is much faster than the direct link between sender and the receiver. The helper with best link rate cooperates with the sender to transmit the data packets to reduce the delay. Thus later reduce the delay and improves the overall performance of the network.

V. Conclusions

IEEE 802.11e EDCA is adapted to statistically prioritized channel access among wireless stations. Thus, EDCA partially satisfy the QoS requirements in practice. In this paper, we are trying to enhance the QoS requirements of IEEE 802.11e EDCA with cooperative communication. Our proposed protocol searches a prospective helper node to assist the sender if the transmission rate between sender and receiver is very poor. The proposed protocol finds an optimal helper node on the basis of priority class and link rate to enhance the network performance. Simulation results show that our proposed protocol performs better in both throughput and delay performance thereby enhancing the QoS of the network.

References

[1] Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE Std. 802.11 part-11, Jun. 1999.

[2] IEEE Std. 802.11e, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Medium Access Control (MAC) Quality of Service Enhancements," 2005.

[3] Guo, T., Carrasco, R. A. and Woo, W. L., "Differentiated cooperative multiple access for multimedia communications over fading wireless networks," IET Communications, vol. 3, no. 6, pp. 1005-1015, June 2009.

[4] Kim, S., "Cooperative MAC protocol based on TXOP and block ACK in IEEE 802.11e WLANs", International Journal of Engineering and Technology (IJET), vol. 6, no. 5, Oct-Nov 2014.

[5] Chang, X., Liu, K. and Shu, B., "A QoS Guaranteed Cooperative MAC Protocol for Wireless Networks", Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications (MAPE), 2013 IEEE 5th International Symposium.

[6] Kim, S., Huang, R. and Fang, Y., "Deterministic Priority Channel Access Scheme for QoS Support in IEEE 802.11e Wireless LANs", IEEE Transactions on Vehicular Technology, vol. 58, no. 2, Feb. 2009.

[7] Zhou, B., Marshall, A. and Lee, T. H., "A k-Round Elimination Contention Scheme for WLANs," IEEE Transactions on Mobile Computing, vol. 6, no. 11, Nov. 2007, pp. 1230-1244.

[8] Li, Z. and Liu, K., "A multiple-recipient-based cooperative MAC protocol for wireless ad hoc networks", Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications (MAPE), 2013 IEEE 5th International Symposium.

[9] Liu, P., Tao, Z., Narayanan, S., Korakis, T. and Panwar, S. S., "CoopMAC: a cooperative MAC for wireless LANs," IEEE Journal on Selected Areas in Communications, vol. 25, no. 2, pp. 340-354, Feb. 2007.

[10] Shan, H., Cheng, H. T., and Zhuang, W., "Cross-Layer Cooperative MAC Protocol in Distributed Wireless Networks", IEEE Transactions on Wireless Communications, vol. 10, no. 8, August 2011.

[11] Yao, Y. C., Wen, J. H., Weng, C. E., "The Performance Evaluation of IEEE 802.11e for QoS Support in Wireless LANs", Wireless Pers Commun (2013) 69:413-425.

[12] Lee, Y., "Throughput Analysis Model for IEEE 802.11e EDCA with Multiple Access Categories", Journal of Applied Research and Technology, vol. 11, August 2013.

An Improved Color Image Steganography Technique in Spatial Domain

Saikat Mondal, Rameswar Debnath, Borun Kumar Mondal
Computer Science and Engineering Discipline, Khulna University
Khulna-9208, Bangladesh
saikatcsebd@gmail.com

Abstract—**Steganographic techniques are used to transmit secret information using a carrier file without visibility during communication. In this paper, we proposed an improved RGB image steganographic technique for secured communication between two authorized parties. This technique embeds the information within 2nd to 8th bit position of blue (B) or green (G) component or both of a pixel throughout the blue and green channels; one bit hidden information is placed in one byte. The insertion bit-position is determined by using a hash function, which has a random nature to select the bit position within 2nd to 8th. The method is secure and not vulnerable to intentional attack. The experimental result shows that the high capacity with better peak signal-to-noise (PSNR) ratio is obtained by our proposed method than that of the existing methods.**

***Index Terms*—steganography, spatial domain, hash function, robustness.**

I. Introduction

Data security in the last few decades has gained a wider audience for the growing possibilities of communication of huge data in the computer network. But the problem of unauthorized copying now-a-days is of great concern especially to the music, film, book and software publishing industries. Thus, we need a secured and trust-worthy communication way for transmitting confidential data. Cryptography is one of effective solution to protect data from the unauthorized users. Many cryptographic techniques have been developed in order to secure the information. But cryptography suffers from a number of drawbacks- notably the fact that the mere presence of an encrypted message might be cause for suspicion in itself [11], [12]. Another drawback of cryptography is the limitations that have been enforced by certain governments, which is particularly significant when cryptography is to be used by remote users. Many governments have created laws to either limit the strength of cryptographic systems or to prohibit it altogether [13], [14]. Primarily due to law enforcement's fear of not being able to gain intelligence by information interception [14], the use of cryptography and sometimes even the possession of a cryptographic algorithm are illegal in certain countries [15].

Steganography is the art and science of writing hidden messages in such a way that no one apart from the intended recipient knows the existence of the message. The purpose of steganography is to conceal the presence of message within seemingly harmless carriers or cover during communication [1]. There are various kinds of carrier-file formats have been used for message concealment, but digital images are the most prominent and famous because of their regularity on the internet.

Several methods have been proposed based on LSB [4], [5], [6] substitution methods. These methods cannot provide high security since the insertion of hidden data in LSB is easily predictable and extractable. Several methods have been proposed where embedding is carried out some higher order bits [8], however, the robustness of the methods is still poor.

In [9], Khalaf and Sulaiman established a technique, known as the segmenting and hiding data method. In this method, the message is concealed into two RGB channels depending on the third channel. Every part is hiding individually in two channels of a RGB image. One of the channels is utilized as index path to the subsequent two channels by calculating the number of the digits of 1 within that index path. However, this index path is not applied in a proficient way.

An RGB channel based image steganography technique is proposed in [10]. In this method, a bit of hidden information is inserted within 3rd to 8th bit position of either blue component or green component or red component of a pixel. A hash function and a secret key are used to select the data embedding position inside the color byte. In this method, only a single bit can be inserted into a color byte by changing two LSBs of both the blue and green components. Even if data is not inserted into a pixel, both LSBs of blue and green components are changed. The payload capacity as well as the PSNR is lower because two lower significant bits are changed when secret data is inserted.

In our technique, we proposed a modification of the method in [10] and some other extensions that increase data capacity and computational complexity. In this method, we insert a bit of hidden information within 2nd to 8th bit position of either green or blue component or both of a pixel. We also use a hash function and a secret key to select the insertion bit position. In every insertion, LSB of either blue component or green component is changed whereas the method in [10] changes two LSBs of both blue and green components. Although two LSBs of both blue and green components of a pixel are changed when data is not inserted but these cases are fewer than that of the method in [10].

The rest of this paper is organized as follows: Section II describes our methodology in details. Experimental results are described in Section III. In Section IV, we conclude this paper.

II. METHODOLOGY

A pixel of a color image is a combination of three components (R, G and B) of 8 bits each in a 24-bits color scheme and the message is a bit stream of its ASCII value. Insertion position inside a color component depends on the value of a hash function. The bit stream is inserted in the blue channel and green channel. After inserting the hidden

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

information in the cover image which is called stego-image is passed to the recipient by unsecured communication network. After receiving the stego-image, the recipient will retrieve the secret information from the stego-image. The overview of the proposed image steganography technique is shown in the Figure 1.

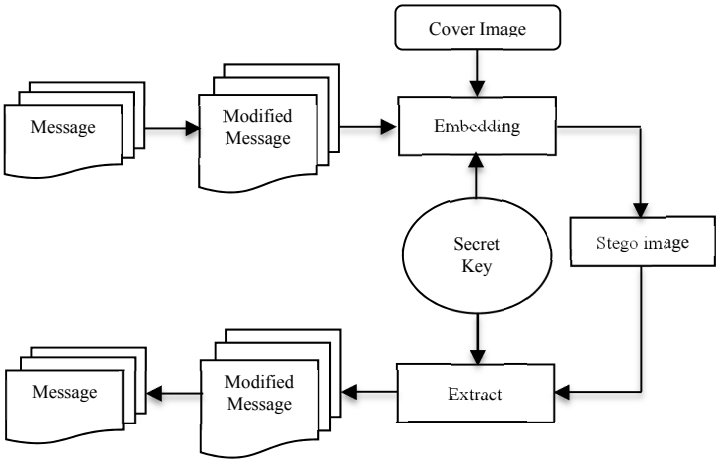


Fig. 1: Overview of the proposed system

A. Data Insertion Technique

For each byte of data from the confidential message we interchange 1st bit with 8th bit, 2nd with 7th, 3rd with 6th and 4th with 5th. Then each bit of the bit stream is inserted one after another into the blue and green channels from the beginning to the end. The inserted location of blue and green components (bytes) is determined randomly within 2nd to 8th position a hash function. If the value of the calculated location (by hash function) in blue component and the bit that has to insert are same then '0' is set to the LSB position of that blue component. Otherwise, '1' is set to the LSB position of that blue component. The process is run until the bit stream is finished or the blue components are finished. If the blue components are finished but message bit stream still remained to embed then same embedding process is run on green channel. Embedding into blue channel gets priority as the change of LSBs of the blue channel is not detectable by human eye.

Data Embedding Algorithm

//**M** = Bit Stream of Message, $\mathbf{I_Z}$ = Image Size, $\mathbf{S_K}$ = Secret //Key, **B** = Binary matrix of the blue channel of the cover //image, **G** = Binary matrix of the green channel of the //cover image, **R** = Binary matrix of the red channel of the //cover image

Step-1: Convert the message into 1D binary matrix, **M**
Step-2: **For** each byte of **M**
Interchange 1st bit with 8th bit, 2nd with 7th, 3rd with 6th and 4th with 5th
After finishing the interchange, a modified bit stream ($\mathbf{M_M}$) is found to embed
Step-3: Calculate, $\mathbf{R_0}$ = No of 0s in Red channel of the current pixel of cover image and $\mathbf{R_1} = \mathbf{S_K} - \mathbf{R_0}$
Step-4: Calculate, $\mathbf{P_B} = ((Iz + \mathbf{R_1}) \%7) + 1$
Step-5: **If** the bit value of $\mathbf{P_B}$ position of the current pixel in blue component and current bit of $\mathbf{M_M}$ to embed are equal **then**
0 is set to the LSB position of that blue component
Else
1 is set to the LSB position of that blue component
Step-6: Repeat the steps 3 to 5 until the bit stream is finished or the blue components are finished
Step-7: **If** the blue components are finished but message bit stream still remained to embed **then** go to the starting pixel of cover image
Step-8: Calculate, $\mathbf{R_2}$ = No of 1s in Red channel of the current pixel of cover image and $\mathbf{R_3} = \mathbf{S_K} - \mathbf{R_2}$
Step-9: Calculate, $\mathbf{P_G} = ((Iz + \mathbf{R_3}) \%7) + 1$
Step-10: **If** the bit value of $\mathbf{P_G}$ position of the current pixel in green component and current bit of $\mathbf{M_M}$ to embed are equal **then**
0 is set to the LSB position of that green component
Else
1 is set to the LSB position of that green component
Step-11: Repeat the steps 8 to 10 until the bit stream is finished

B. Data Extracting Technique

Extracting is the reverse process of embedding technique. In order to recover the secret information from stego-image, LSBs of blue and green components will determine the existence of hidden bit on those components.

Data Extracting Algorithm

//**M** = Bit Stream of Message, $\mathbf{I_Z}$ = Image Size, $\mathbf{S_K}$ = Secret //Key, **B** = Binary matrix of the blue channel of the stego //image, **G** = Binary matrix of the green channel of the // stego image, **R** = Binary matrix of the red channel of the // stego image, $\mathbf{M_M}$ = Modified bit stream

Step-1: **If** the LSB of blue component of the current pixel is 0 then
Calculate, $\mathbf{R_0}$ = No of 0s in Red channel of the current pixel of stego image and $\mathbf{R_1} = \mathbf{S_K} - \mathbf{R_0}$
Calculate, $\mathbf{P_B} = ((Iz + \mathbf{R_1}) \%7) + 1$
Pick the bit from the current blue component of the position $\mathbf{P_B}$ and insert into $\mathbf{M_M}$
Else
Go to the next pixel
Step-2: Repeat the above step until the bit stream is finished to extract or the blue components are finished
Step-3: **If** the blue components are finished but message bit stream still remained to extract **then** go to the starting pixel of stego image
Step-4: **If** the LSB of green component of the current pixel is 0 then
Calculate, $\mathbf{R_2}$ = No of 1s in Red channel of the current pixel of stego image and $\mathbf{R_3} = \mathbf{S_K} - \mathbf{R_2}$
Calculate, $\mathbf{P_G} = ((Iz + \mathbf{R_3}) \%7) + 1$
Pick the bit from the current green component of the position $\mathbf{P_G}$ and insert into $\mathbf{M_M}$
Else
Go to the next pixel
Step-5: Repeat the above step until the bit stream is finished to extract
Step-6: Get the modified bit stream, $\mathbf{M_M}$
Step-7: **For** each byte of $\mathbf{M_M}$
Interchange 1st bit with 8th bit, 2nd with 7th, 3rd with 6th and 4th with 5th

After finishing the interchange, bit stream **M** is found

Step-8: Get the confidential message from bit stream **M**

III. EXPERIMENTAL RESULTS

In this paper, we have shown experimental results of our proposed method and compared with the method in [10]. Here, three benchmark RGB images are used named Lena, Baboon and Pepper for experimental result analysis. The size of each image is 512 X 512. MATLAB R2010a is used to implement the proposed method and the method in [10]. Table 1 shows the PSNR values and percentage of pixels of the cover image required for embedding a message of size 15064 bytes. In general, the value of PSNR below 30dB shows the low quality and more than 40 dB shows high quality stego-image. From Table I, we see that high quality of stego-images is obtained by our proposed method for all images and the average PSNR value is 67.87.

For measuring the Peak-Signal-to-Noise Ratio (PSNR) the following formula is used:

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{C^2{}_{max}}{\mathrm{MSE}}\right)$$

Where, Cmax holds the maximum possible pixel value in the cover image and MSE denotes Mean Square Error which is defined as:

$$\mathrm{MSE} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}(S_{ij} - C_{ij})^2$$

Where, i and j are the image co-ordinates, m and n are the dimensions of the image, S and C are stego image and cover image respectively.

TABLE I
PSNR VALUES OF PROPOSED SYSTEM AND EXISTING SYSTEM [10]

Images of Size 512X512	Data Embedded (bytes)	PSNR of proposed System(dB)	PSNR of existing System[10] (dB)
Lenna	15064	66.36	53.77
Baboon	15064	67.98	53.76
Pepper	15064	69.29	53.80
Average	15064	67.87	53.78

TABLE II
CAPACITY COMPARISON BETWEEN PROPOSED SYSTEM AND EXISTING SYSTEM [10]

Cover Image	Data Embedded (bytes)	Pixel holds data of proposed System (%)	Pixel holds data of existing System [10] (%)
Lena	15064	91.45	85.07
Baboon	15064	97.32	85.66
Pepper	15064	89.94	83.05
Average	15064	92.90	84.59

From the results in Table 2, it is found that greater percentages of pixel are used for information concealing that shows the data capacity. Sometimes steganography suffers from active or some passive attack as well. A steganographic technique is required to secure it from both the attacks. The proposed method is extremely safe as it introduces message in variable position in the color components and data is less susceptible to accidental attack. The selection process of channel, pixel and position by our proposed method adds to the complexity of steganalysis. Moreover, message is modified by bitwise operation before embedding which is very fast but also added computational complexity. From Fig. 2, we can observe the fact that there is no visual artifact in the stego-images; they look same as the corresponding original cover images from they are generated. Fig. 3 and 4 show the comparison of our proposed methods and the method in [10].

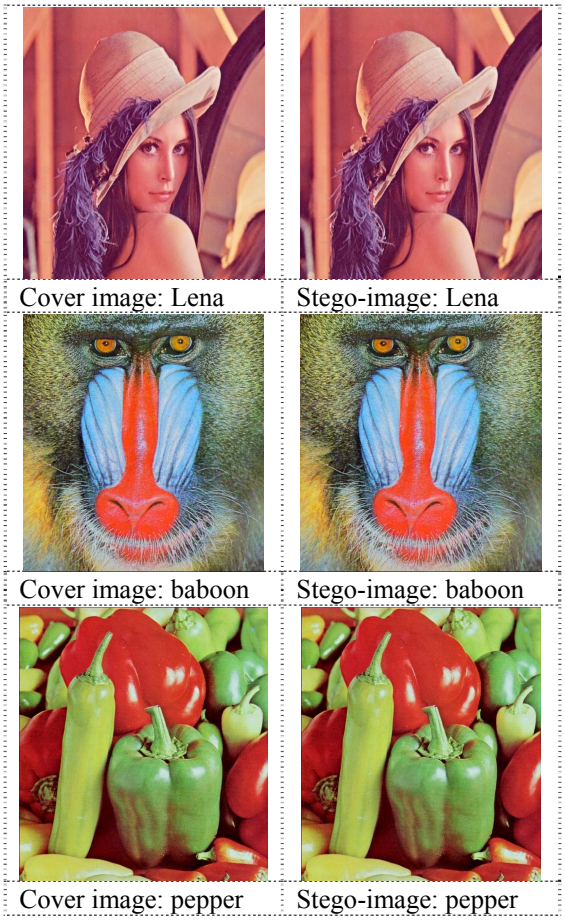


Fig 2: Cover images and their Stego-images

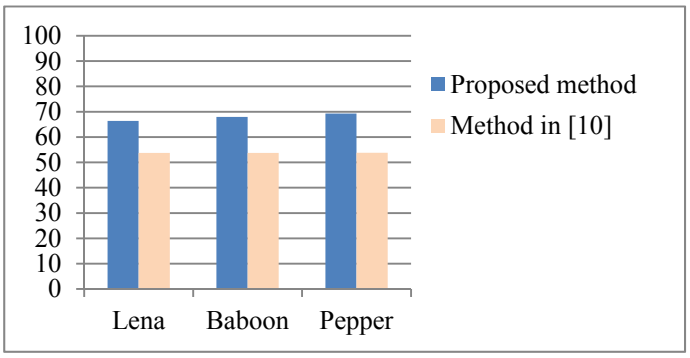


Fig. 3: Comparison of PSNR values of the proposed method and the method in [10]

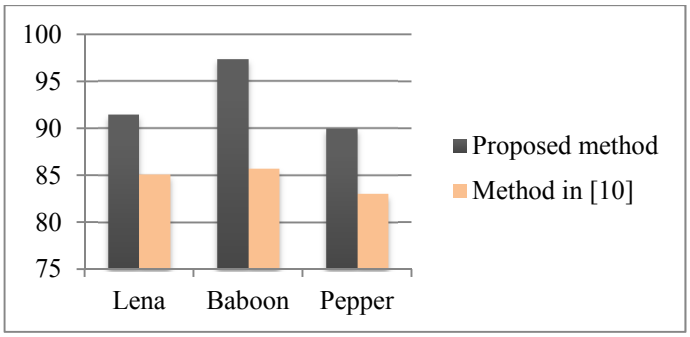


Fig. 4: Comparison of PSNR values of the proposed method and the method in [10]

IV. CONCLUSION

Steganography establishes the secrete communication between two authorized parties hiding the existence of information using a carrier file. In this paper we proposed an improved steganography technique in spatial domain where hidden bits are embedded into variable position within 2nd to 8th bit position of blue then green (if necessary) components of pixels of the cover image. The proposed method is safe against intentional and unintentional attacks as it modifies the raw message to distort the meaning of the secret message and inserts the modified message bits into blue channel mostly and green channel making very low image distortion. High PSNR value ensures that it is awfully difficult for the unlawful users to recognize the changes in stego-image.

REFERENCES

[1] Ramanpreet Kaur, Prof. Baljit Singh "Survey and analysis of various steganographic techniques," International Journal of Engineering Science & Advance Technology. Vol. 2 , issue 3, pp. 561 – 566, May-June 2012.

[2] N.F. Johnson and S. Jajodia, "Exploring steganography: Seeing the unseen ," Computer, vol. 31, no. 2, pp. 26-34 Feb 1998.

[3] F.A.P Peticolas, R.J. Anderson and M.G.Kuhn, "Information hiding –A survey," Proceedings of IEEE, pp. 1062-1078, July 1999.

[4] Na-I Wu, "A study on data hiding/or gray-level and binary images ," Accessed on March 2009.

[5] Junija M. and Sandhu P.S., "Designing of robust stegangraphic technique based on LSB insertion and encryption," Proceedings of International Conference on Advances in Recent Technologies in Communication and Computing, pp. 302-305, 2009.

[6] Juneja M. and Sandhu P. S., "An improved LSB based steganograhy technique for RGB color images," International Journal of Computer and Communication Engineering, vol. 2, no. 4, July 2013.

[7] Adnan Gutub, Mahmoud Ankeer, Muhammad Abu- Ghalioun, Abdulrahman Shaheen, and AleemAlvi, "Pixel indicator high capacity technique for RGB image based steganography," WoSPA 2008 – 5th IEEE International Workshop on Signal Processing and its Applications, University of Sharjah, Sharjah, U.A.E. March 2008.

[8] Kaur M, Gupta S., Sandhu P. S. and Kaur J., "A dynamic RGB intensity based steganography scheme," World Academy of Science, Engineering and Technology, vol. 67, pp. 833-838, 2010.

[9] Emad T. Khalaf, NorrozilaSulaiman,"Segmenting and hiding data randomly based on index channel," International Journal of Computer Science Issues, vol. 8, no. 1, issue 3, May 2011.

[10] Anupam K. Bairagi, Saikat Mondal and Rameswar Debnath, "A robust RGB channel based image steganography technique using a secret key," Proceedings of the 16th International Conference on Computer and Information Technology, pp. 81-87, Khulna, Bangladesh, March, 2014.

[11] M. Shirali-Shahreza, "Stealth Steganography in SMS," Proceedings of IFIP International Conference on Wireless and Optical Communication Network, pp. 316-321, 2006.

[12] E. Kawaguchi and R. O. Eason, "Principles and applications of BPCS steganography," Proceedings of SPIE International Society for Optical Engineering, pp. 464-473, 1999.

[13] J. Krinn,"Introduction to steganography," 2000, http://rr.sans.org/covertchannels/steganography.php.

[14] F. S. Grodzinsky, K. Miller and M. J. Wolf, "The ethical implications of the messenger's haircut: Steganography in the digital image. In: Himma K. E. (ed.) Internet Security: Hacking, Counter Hacking and Society," Jones and Bartlett Publishers, pp. 205, 2007.

[15] B. Dunbar, "A detailed look at steganographic techniques and the use in an open-system environment," SANS Institute, Information Security Reading Room, pp. 3.

Non-intrusive Objective Evaluation of Speech Quality in Noisy Condition

Md. Rafidul Islam, Md. Ashequr Rahman, Md. Numan Hasan, A. N. M. Shahriyar Hossain, Ahmed Nazim Uddin, Mohammad Ariful Haque*

Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka-1205, Bangladesh
*arifulhoque@eee.buet.ac.bd

Abstract— **It is very difficult, if not impossible, to obtain a clean reference signal of a noisy speech recorded in a practical environment. As a result, intrusive methods that evaluate the quality of speech signal with the help of a clean reference signal has little value in real world applications. In this paper, we investigate the effectiveness of data-driven non-intrusive method for assessing quality of speech without using clean reference signal. In the proposed method, a support vector machine based classifier is trained using a labelled dataset and then the classifier provides speech assessment rating on unknown speech signals. The obtained results have been evaluated against the intrusive PESQ score. The results indicate that the proposed technique performs better than the state-of-the-art non-intrusive methods on the same test data set.**

Index Terms—**Speech quality, objective measure, non-intrusive method, support vector machine.**

I. Introduction

Speech is the most important bio signal that gives the most efficient medium for human interaction. Scientists and engineers are providing their utmost efforts for speech processing with a view to developing telecommunication applications, diseases and emotion detections, speaker recognition, native speaker analysis and so on. For all these cases, proper evaluation of speech quality is an important issue.

The quality of speech can be evaluated by human or machine. Human listening and quality grading costs both money and time. So, it is not always feasible and recommended though this process is reliable. Rating by machine is of two types: intrusive and non-intrusive evaluation [1-7]. The first one needs clean speech to compare with the distorted one to give quality index. But, clean data is not available most of the time. In some applications, there may have scopes for creating clean reference speech but it requires special arrangement, costly equipment and laboratory environment. As a result, non-intrusive methods are developed in order to assess the quality of speech without the help of a clean reference signal.

Non-intrusive methods are of two types. In the first approach, a quasi-clean signal is generated using a suitable speech enhancement algorithm. Then using this quasi-clean signal as reference, the quality of the recorded noisy speech is evaluated similar to an intrusive method [8]. The second non-intrusive method is effectively a machine learning process in which a suitable classifier is trained using a labelled dataset and then the classifier is used to provide speech assessment score on unknown speech signal. A number of machine learning based non-intrusive techniques are available in the literature that uses different feature set and classifiers [9]. For example, the Low Complexity Speech Quality Assessment (LCQA) method [10] uses 44 statistical features computed from pitch period and 10th order LPC coefficients in order to train a GMM classifier. The Low-Cost Intelligibility Assessment (LCIA) method [11] reduces the feature set by a correlation based feature selection and principal component analysis (PCA). Non-Intrusive Speech Assessment (NISA) method [12] uses 116 features estimated from short and long term speech segments and a tree-based classifier is trained for speech assessment.

In this paper, we present a machine learning based method for non-intrusive speech quality evaluation. The novelty of our proposed method comes from using a distinguished feature set and support vector machine (SVM) classifier for speech evaluation. We have used only 7 features such as spectral flatness, spectral skewness, spectral slope, kurtosis, skewness, entropy and 1st MFCC coefficient to train the SVM system. We have implemented both regression and classification models for providing continuous and discrete valued scores, respectively, on the test speech. The continuous labelling of speech utterances was done by intrusive PESQ scores and discrete labelling was done by several human evaluators following the MOS rating scale. Publicly open speech dataset NOIZEUS [13] is used in the study. The results show that the correlation coefficient of the proposed method is 0.9334 on NOIZEUS dataset, which is 7.3% better than ITU-T standard recommendation P.563.

II. Method

The objective of the proposed method is to provide an objective speech quality score for each test utterance, based on a set of features extracted from the speech signal. The overall structure of the proposed speech evaluation method is depicted in Fig. 1.

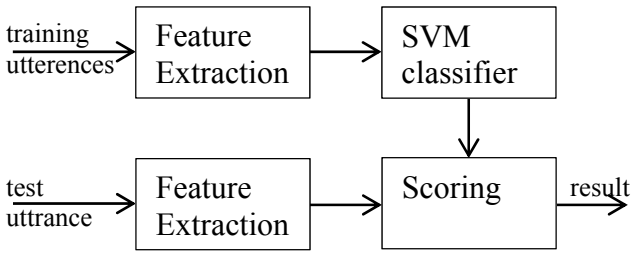


Fig. 1 Framework of the proposed speech evaluation technique

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

In what follows, we present the description of different sub-blocks in more details.

A. Feature extraction

Non-intrusive speech quality evaluation systems are based on the extraction of a suitable feature set from the speech signal. This feature set aims at capturing the low-level aspects of the speech signal structure that are likely relevant to human quality judgment. In this section we discuss features that we have selected in our proposed method.

The first feature that we have used is spectral flatness or tonality coefficient. This feature quantifies how noise-like a sound is, as opposed to being tone-like. The spectral flatness is calculated by dividing the geometric mean of the power spectrum by the arithmetic mean of the power spectrum,

$$Flatness = \frac{\sqrt[N]{\prod_{n=0}^{N-1} x(n)}}{\frac{\sum_{n=0}^{N-1} x(n)}{N}}$$

where, x(n) represents the magnitude of bin number n. The next set of features that we use is spectral and temporal skewness, kurtosis and entropy defined as

$$skewness = E\left[\left(\frac{S-\mu}{\sigma}\right)^3\right]$$

where, S is the random variable denoting speech signal, μ is the mean and σ is the standard deviation.

$$kurtosis = \frac{E[(S-\mu)^4]}{(E[(S-\mu)^2])^2}$$

$$Entropy = -\sum_{i=1}^{N} P(S_i)\log_{10} P(S_i)$$

where, $p(S_i)$ is the probability of occurring S_i. It is reported that both the skewness and kurtosis are high for clean speech and the corresponding entropy is low. The other features that we have used in our evaluation system are spectral slope and 1st MFCC coefficient.

B. Labelling of dataset

In order to train a classifier, we need a labelled data set indicating the ground truth of speech quality. In this work, speech utterances were labelled by intrusive PESQ measures as well as non-intrusive human evaluation. PESQ provides a score in the range of 1 to 5, where 1 is unacceptable and 5 is excellent. The same speech database was re-labelled with the help of human evaluators. The objective of this re-labelling is to evaluate the effectiveness of proposed method to imitate human judgement of speech quality. The human evaluation of speech quality is also marked between 1 and 5 following "ITU-T recommendation P.800: Methods for subjective determination of transmission quality" as shown in Table 1.

TABLE I
NUMERICAL VALUES OF HUMAN EVALUATION SCORE

Rating	Speech quality	Level of distortion
5	Excellent	Imperceptible
4	Good	Just perceptible, but not annoying
3	Fair	Perceptible and slightly annoying
2	Poor	Annoying, but not objectionable
1	Bad	Very annoying and objectionable

C. SVM regression and classification model

Nowadays, SVM is widely used in the purpose of solving machine learning problems for both binary and multiclass classifications. Regression problems can also be taken care of with it. As stated earlier, we performed 5-class multiclass classification for emulating human evaluation and regression analysis for PESQ estimation.

SVM model is an algorithm by which data are mapped in such a way that data of different types are separated by a clear gap which is as large as possible. Let, there is a set of labelled training examples of length N which can be written as Tr = {(x_i, y_i)}, where i=1, 2, ...,N and x_i is a vector of feature length and y_i is either 1 or -1. Now, for an unknown example, binary classification can be done by,

$$f(x) = sign\left(\sum_{i=1}^{N} c_i\, y_i K(\boldsymbol{x_i}, \boldsymbol{x})\right) + b$$

where c_i is the Lagrange Multiplier , K(x_i , x) is known as kernel function, b is the bias term. If hyper plane is not simple, then SVM has the option to use nonlinear kernels. In our task, we used Radial Basis Function (RBF) kernel which is given by,

$$K(\boldsymbol{x}, \boldsymbol{y}) = exp(-\gamma\|\boldsymbol{x}-\boldsymbol{y}\|^2) \qquad \gamma > 0$$

where ||.|| is known as Euclidian L_2-Norm and spread of the kernel is signified by γ. Now, multiclass classification can be done by mapping the single multiclass problem into multiple binary classification problems and the common approach is known as one-versus-all method in which an example is assigned to that class for which the classifier's output is the highest.

In SVM regression (SVMR) model, firstly, input vector **x** is mapped onto an *m*-dimensional feature space by some predefined (nonlinear) mapping. Then, a linear model is constructed with the feature space. The model is given by,

$$f(\boldsymbol{x}, w) = \sum_{j=1}^{m} \omega_j g_j(\boldsymbol{x}) + b$$

where $g_j(x)$, j= 1, 2, ..., m represents a set of nonlinear transformations, and b is the bias term. The quality of estimation is measured by a loss function *L(y, f(x,ω))* defined as

$$L(y, f(\boldsymbol{x}, \omega)) = \begin{cases} 0 & , if\ |y - f(\boldsymbol{x}, \omega)| \le \epsilon \\ |y - f(\boldsymbol{x}, \omega)| & , otherwise \end{cases}$$

III. RESULTS

In this section, we present simulation results to evaluate the effectiveness of our proposed algorithm and provide comparative results with state-of-the-art methods. The performance indices for regression and classification tasks are correlation coefficient and accuracy, respectively.

A. Description of the speech dataset

The noisy speech corpus of 960 speech utterances was obtained from NOIZEUS database. The database contains 30 clean speech utterances each passed through 8 different types of noise namely airport, babble, car, exhibition, restaurant,

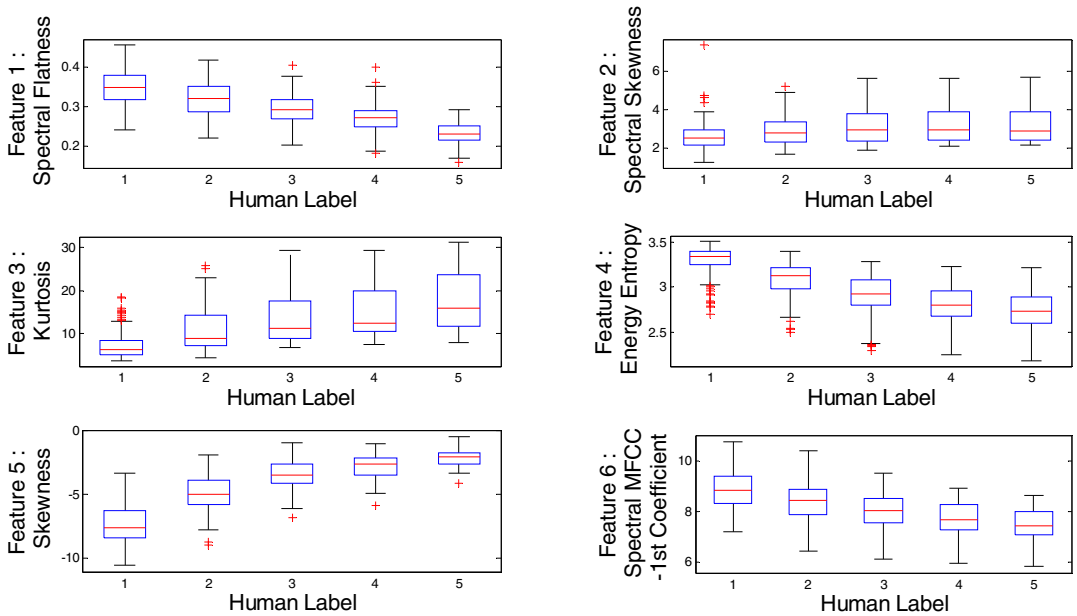


Fig. 2 Box Whisker plot of the proposed features

station, street and suburban train at 4 SNR levels (0 dB, 5 dB, 10 dB and 15 dB) each. Noise signals for degradations of the speech utterances of the database were taken from the AURORA database. Thus, 960 speech utterances are available for 32 different degradation conditions of noise. 70 percent of the data set is used for training the classifier and the rest 30 percent was used for testing.

B. Effectiveness of the proposed feature set

Fig. 2 presents the Box Whisker plot of the proposed features. It is seen that the features have discriminating capabilities for objective evaluation of speech quality. Spectral flatness is high for inferior-quality speech and it gradually decreases for better speech. To the contrary, skewness and kurtosis increases as the speech quality gets better. Entropy and 1st MFCC coefficient is inversely proportional to human evaluation speech quality.

C. Speech evaluation results

Table II presents the correlation coefficients between actual PESQ scores and estimated PESQ scores obtained using our proposed method. We also provide the correlation coefficient reported in different published results on the same NOISEUS database. It is seen that our proposed method surpasses the reported techniques by significant margin.

TABLE III
CORRELATION COEFFICIENT BETWEEN ACTUAL AND ESTIMATED PESQ SCORES

Methods	**Correlation coefficient**
Proposed	0.933
Ref [8]	0.906
Ref [9]	0.869

Now we evaluate the performance of our proposed technique against human assessment. It is not known very accurately that how the human auditory system as well as the brain processes speech utterances to award a quality score to it. As a result, the subjective scores do not correlate well with simple and intuitive objective measures such as mean square error, segmental SNR, spectral distortion, Itakura-Saito distance etc. However, we got 75.52% accurate classification results using our proposed method. So it can be said that our method highly correlates with human judgement.

IV. CONCLUSIONS

The non-intrusive assessment of speech quality is a challenging problem and has diverse applications. A machine learning based technique was proposed for this task, using SVM classifier trained with a small number of feature set. The intrusive PESQ algorithm as well as human evaluation was used to label a large database comprising of speech degraded by additive noise. The proposed method achieved a correlation coefficient of 0.933 with actual PESQ score, outperforming the baseline algorithms and the industry standard P.563 algorithm. The proposed method also gives highly correlated results with human assessment. As a result, the proposed method can be used for evaluating speech quality in a non-intrusive manner.

REFERENCES

[1] ITU-T Rec. P.830, "Subjective performance assessment of telephone band and wideband digital codecs," 1996.
[2] ITU-T Rec. P.800, "Methods for Subjective Determination of Transmission Quality," 1996.
[3] O. Au and K. Lam, "A novel output-based objective speech quality measure for wireless communication," Signal Process. Proc, 4th Int. Conf., vol. 1, pp. 666–669, 1998.
[4] P. Gray, M. Hollier, and R. Massara, "Non-intrusive speech quality assessment using vocal-tract models," in Proc. IEE Vision, Image and Signal Process., vol. 147, no. 6, 2000, pp. 493–501.
[5] J. Liang and R. Kubichek, "Output-based objective speech quality," IEEE 44th Vehicular Technology Conf., vol. 3, no. 8-10, pp. 1719–1723, 1994.
[6] T. Falk, Q. Xu, and W.-Y. Chan, "Non-intrusive GMM based speech quality measurement," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., vol. 1, 2005, pp. 125–128.

[7] G. Chen and V. Parsa, "Bayesian model based non-intrusive speech quality *evaluation*," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., vol. 1, 2005, pp. 385–388.

[8] W. Zhou and Q. He, "Non-intrusive speech quality objective evaluation in high-noise environments," Signal and Information Processing (ChinaSIP), 2015 IEEE China Summit and International Conference on, Chengdu, 2015, pp. 50-54.

[9] R. K. Dubey and A. Kumar, "Non-intrusive speech quality assessment using several combinations of auditory features," Int J Speech Technol, March 2013, Volume 16, Issue 1, pp 89–101.

[10] V. Grancharov, D. Y. Zhao, J. Lindblom, and W. B. Kleijn, "Low complexity, nonintrusive speech quality assessment," IEEE Trans. Audio Speech Lang. Process., 14 (6), 1948–1956. doi:10.1109/TASL.2006.

[11] D. Sharma, G. Hilkhuysen, N. D. Gaubitch, P. A. Naylor, M. Brookes, M. Huckvale, "Data driven method for non-intrusive speech intelligibility estimation," Proceedings of the 2010 European Signal Processing, Conf. (EUSIPCO). Denmark

[12] D. Sharmaa, Y. Wangb, P. A. Naylor, M. Brookes, "A data-driven non-intrusive measure of speech quality and intelligibility," Speech Communication 80 (2016) 84–94

[13] NOISEUS databse. [Online]. Available: http://www.utdallas.edu/~loizou/speech/noizeus/

High Efficient and Stable Ultra-Thin CdTe Solar Cell with a Potential Copper Telluride BSF

Mrinmoy Dey[1, 2*], Maitry Dey[3], M. A. Matin[1, 2] and Nowshad Amin[4, 5]

[1]Department of Electrical and Electronic Engineering, Chittagong University of Engineering and Technology, Chittagong-4349, Bangladesh

[2]Renewable Energy Laboratory (REL), Chittagong University of Engineering and Technology, Chittagong-4349, Bangladesh

[3]Department of Electrical and Electronic Engineering, Premier University, Chittagong-4203, Bangladesh

[4]Solar Energy Research Institute (SERI), The National University of Malaysia, 43600 Bangi, Selangor, Malaysia

[5]Department of Electrical, Electronic and Systems Engineering, The National University of Malaysia, 43600 Bangi, Selangor, Malaysia

*mrinmoycuet@gmail.com

Abstract-**For higher efficiency, low-cost and cell stability polycrystalline cadmium telluride (CdTe) is recognised as the potential solar cell material. In this work, the cell performances (*Jsc*, FF, *Voc*, conversion efficiency and temperature stability) of ultra-thin CdTe solar cell was examined by numerical analysis utilizing AMPS (Analysis of Microelectronic and Photonic Structures) simulator. In the proposed cell, reduction of absorber layer was done and observed that 1µm absorber layer is enough for acceptable range of cell conversion efficiency. The possibility of this ultra-thin CdTe absorber layer was investigated, together with 100 nm Cu_2Te back surface field (BSF) layer to reduce the barrier height in the valence band and to minimize the recombination losses at the back contact of the proposed cell. It was found that the proposed ultra-thin cell have conversion efficiency of 18.68% (*Jsc* = 21.47 mA/cm^2, FF = 0.85, *Voc* = 1.02 V) without BSF and with 100 nm Cu_2Te BSF conversion efficiency increased to 22.51% (*Jsc* = 24.26 mA/cm^2, FF = 0.875, *Voc* = 1.06 V) with only 0.8 µm of CdTe absorber layer. Moreover, the normalized efficiency of the proposed cell linearly decreased with the increasing operating temperature at the gradient of -0.16%/°C, which indicated better stability of the proposed CdTe solar cell.**

Keywords— ***Solar cell, CdTe cell, Ultra-thin, Cu_2Te BSF, AMPS.***

I. INTRODUCTION

Thin film Cadmium Telluride (CdTe) is the potential candidate for efficient, cost-effective and stable solar cells. The CdTe has a direct band gap of 1.45 eV which is very close to the optimum band gap for solar cells and it has a high absorption co-efficient over 5×10^5/cm. Long-term stable performance and high efficiency has been exposed by polycrystalline CdTe thin film solar cells under AM1.5 illumination for terrestrial usage. There are challenges for the improvement of the cell efficiency, material usages and the thermal stability of the solar cell though CdTe based solar cells are currently entering mass production.

CdTe thin film solar cell with thickness of approximately 2µm will absorb nearly 100% of the incident solar radiation and flexibility with regards to the method of manufacture is another advantage of the CdTe technology. Using a variety of low cost techniques, such as close-spaced-sublimation (CSS), physical vapour deposition (PBD), chemical bath deposition (CBD) and magnetron sputtering polycrystalline layers of a CdS/CdTe cell can be deposited.

At standard solar spectrum, the maximum theoretical efficiency for CdTe is about 29%. In 2001, Xuanzhi Wu reported a record efficiency of 16.5% for CdS/CdTe solar cell [1]. Many researchers concluded that without any suitable replacement of polycrystalline n-CdS (2.42 eV) window layer and it is the best suited hetero-junction partner with p-CdTe absorber layer. This work motivates ultra-thin CdS/CdTe solar cells with the insertion of BSF layer for higher conversion efficiency. The formation of a stable, low resistance, non-rectifying back contact to p-CdTe is one of the major and critical challenges of efficient and stable CdTe solar cells. Cu is used as a back contact material as usual, but it leads to the efficiency degradation with time due to Cu diffusion to the front contact which causes shunting effect responsible for instability [2]. It was found that the insertion of suitable BSF was more effective when numerically applied to the champion and baseline case CdTe solar cells with thinner absorber layer. As a result, BSF was inserted to the proposed cell and found very much promising results to implement of the ultra-thin CdTe solar cell.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

The contact problem at the back of CdTe cells is to overcome by typical approach either reduce the barrier height or moderate its width by heavily doping of extra layer such as BSF in between the p-CdTe and final metal back contact. Among these, at back surfaces the use BSF layers has been found quite simple and effective. Zinc telluride (ZnTe) is the most familiar BSF material used in CdS/CdTe solar cell. Many researchers are trying to explore new potential BSF to increase the conversion efficiency of CdS/CdTe solar cells. These are Sb_2Te_3 as discussed by Bätzner et al. [3], As_2Te_3 as discussed by Romeo et al. [4] and PbTe as discussed by Matin and Dey [5]. The specific BSF material preferred in this work is Copper Telluride (Cu_2Te) to investigate its hidden potentiality for CdTe solar cell.

In this work, numerical analysis was done by utilizing AMPS simulator [6] to explore the possibility of ultra-thin CdS/CdTe cell with Cu_2Te BSF to improve the short circuit current density (*Jsc*), the open circuit voltage (*Voc*) and fill factor (FF) of the CdTe solar cells. It was showed from numerical analysis that less than 1µm CdTe absorber layer with ITO/Zn_2SnO_4 front contact and Cu_2Te as BSF are suitable materials for high efficient (22.51%) ultra-thin and stable CdS/CdTe cell.

II. MODELING AND SIMULATION

Numerical modeling and simulation were done in this work to investigate the possibilities of ultra-thin CdTe absorber layer and validity of physical structures of proposed cell. Recently, AMPS-1D program is used by more than 200 groups worldwide to design solar cell [7]. The effect of Cu_2Te BSF insertion on ultra-thin CdTe solar cell has been explored by using the AMPS-1D simulator. The baseline structure of CdTe cell was utilized to estimate the highest efficiency of CdS/CdTe solar cell and it was modified to investigate the possibility of efficient ultra-thin solar cells with suitable BSF inclusion. The baseline structure of CdTe solar cell proposed by Gloeckler [8] was initially modified by adding an extra layer of Zn_2SnO_4 between SnO_2 (TCO) and CdS layer. In this design, the first modification of typical CdS/CdTe cell was to reduce CdS layer to 80nm with Zn_2SnO_4 buffer layer. The second modification was to change the CdTe doping concentration to ~10^{17} cm^{-3} which is now achievable for p-CdTe material. The third modification is to reduce the CdTe absorber layer thickness to the extreme limit for achieving ultra-thin CdS/CdTe solar cell. The fourth modification was the placing of Cu_2Te BSF to reduce the minority carrier recombination loss at the back contact of the ultra-thin CdS/CdTe cell.

Fig. 1 illustrates the conventional structure of CdTe and Fig. 2 depicts the modified proposed ultra thin CdTe solar cell (Glass/SnO_2/Zn_2SnO_4/CdS/CdTe/Cu_2Te/Ni) structure for higher cell performance.

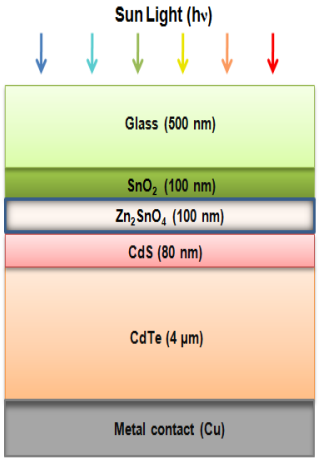


Fig.1 The conventional structure of the CdTe solar cell

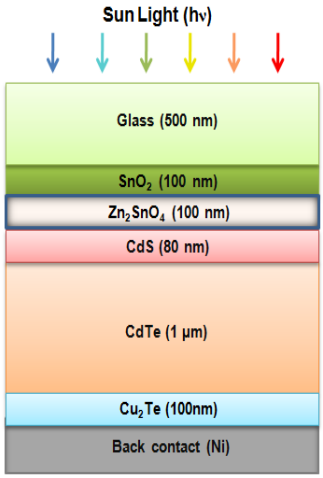


Fig.2 The proposed structure of CdTe solar cell for higher performance

As a result, the four layers of the proposed cell in this analysis are n-Zn_2SnO_4 buffer layer, n-CdS window layer, p-CdTe absorber layer and p-Cu_2Te BSF layer. The CdTe absorber layer thickness was varied from 0.1µm to 10µm and the thickness of other layers were fixed to the usual values. This modeling showed the possibility of getting a highly efficient (around 22%) CdS/CdTe solar cell consists of 1 µm of CdTe, 80 nm of CdS, 100nm of Zn_2SnO_4 and 100 nm of Cu_2Te BSF. In numerical analysis of a particular solar cell model, the number of parameters that can be varied is larger than 50 [9]. Obviously, a problem related to design with 50 variables is too uncertain to solve reliably without the help of computer. It is consequently necessary to minimize the number of variable parameters by fixing many of them at reasonable values. It was a real dispute to select the practically achievable parameters for different layers of the proposed cells. Many of them depend on fabrication techniques and deposition methods and can thus vary even between devices fabricated in the same batch from the same technology. Table I shows the material parameters used in this work, which were selected based on experimental data, literature values, theory, or in some cases reasonable estimations [10].

TABLE I THE MATERIAL PARAMETERS USED FOR THE NUMERICAL ANALYSIS OF THE PROPOSED CELLS

Parameters	n-CdS	p-CdTe	Cu_2Te
Thickness, W (µm)	0.08	0.1-2.9	0.1
Permittivity, ε/ε0	9	9.4	10
Electron mobility, µc (cm^2/Vs)	350	500	500
Hole mobility, µp (cm^2/Vs)	50	60	100
Carrier concentration, n, p (cm^{-3})	1×10^{17}	1×10^{17}	1×10^{21}
Band gap, Eg (eV)	2.42	1.45	1.18
Density of state in conduction band, Nc (cm^{-3})	1.8×10^{19}	7.5×10^{17}	7.8×10^{17}
Density of state in valence band, Nv (cm^{-3})	2.4×10^{18}	1.8×10^{18}	1.6×10^{19}
Electron affinity, χ (eV)	4.50	4.28	4.2

III. RESULTS AND DISCUSSION

A. Optimization of CdTe absorber layer

The starting of this analysis was a three layer device model (SnO_2/CdS/CdTe) of the CdTe baseline case. The improvement of cell conversion efficiency of the proposed cell was achieved due to increased *Jsc,* FF and *Voc*. It is remarkable that in most highly efficient CdS/CdTe solar cells, the CdTe absorber layer is purposely kept around 5 µm or above. Thus, further numerical analysis was done with the proposed modified cell to reduce the thickness of CdTe absorber layer aiming to preserve the usages of absorber CdTe materials and the cost of the fabricated cells in the end.

Encouraged by lower production cost, time and energy, several research groups all over the world have produced CdS/CdTe cells with thinner CdTe absorber layer. In a typical CdTe solar cell, the absorber thickness is usually over 4µm. Thickness reduction from 4µm to less than 1µm could save over 75% of the CdTe material, if the film deposition rate can be kept unchanged, the deposition time and energy could be at least four times lower. Therefore, thinner solar cells without compromising much their performance should lead to lower cost solar cell devices, since they require less materials, less fabrication time and energy. Thus, a new configuration with Cu_2Te BSF could be implemented to restrain the possible recombination losses at the back contact in case of such ultra-thin (~1µm) CdTe absorber layer.

B. Insertion of Cu_2Te BSF

In order to reduce the possible recombination loss and the barrier height at the back contact of ultra-thin (1µm CdTe) cell, a medium band gap (Eg) material Cu_2Te (*Eg* = 1.18 eV) was inserted at the back to reduce back surface recombination rate at the CdTe/Cu_2Te interface. This medium band gap material would act as a BSF to bounce back the carriers (electrons) from the CdTe/Cu_2Te junction and thus would contribute in the enhancement of carriers.

To investigate the hidden potentiality of BSF insertion in between CdTe and back contact, a numerical analysis of the modified cell was done by utilizing AMPS simulator. In this numerical analysis, all the layers of the cell are similar to the previous optimized cell except one additional 100 nm Cu_2Te BSF layer just before the final back contact metal of the proposed cell. Now the modified cell with the entire cell parameters of Table I were simulated along with 1 µm CdTe, 80 nm CdS, 100 nm Zn_2SnO_4 and 100 nm Cu_2Te BSF layers; the AMPS simulated results are shown in Table II to compare the cell performances with and without BSF layer. It is apparent from Table II that the proposed cell without BSF shows conversion efficiency of 18.68%, which is in the acceptable range for CdS/CdTe solar cells. The conversion efficiency increased to 22.41% with the insertion of 100 nm Cu_2Te BSF layer.

TABLE II THE OUTPUT PARAMETERS OF MODIFIED CELLS WITHOUT BSF LAYER AND WITH BSF LAYER

Structure/ parameters	V_{oc} (V)	*Jsc* (mA/cm^2)	FF	Eff. (%)
Base case	1.02	21.470	0.850	18.68
Cu_2Te BSF	1.05	24.335	0.872	22.41

The proposed modified cell structure with Cu_2Te BSF showed higher *Jsc*, FF and *Voc* than the same cell without BSF due to reduced back surface recombination and improved back contact formation with p-CdTe. Simulated J-V characteristics of both the cells are shown in Fig. 3.

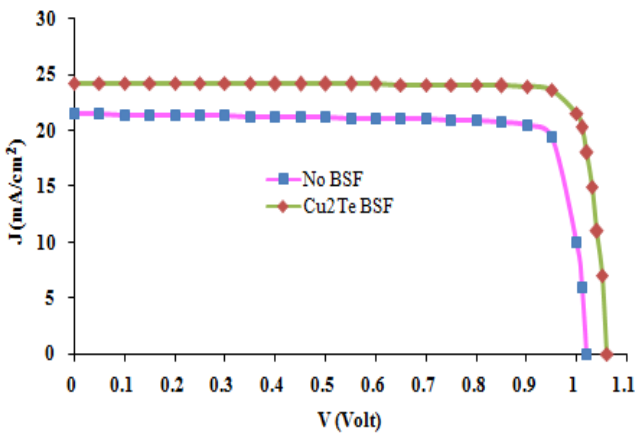


Fig.3 J-V characteristics of the proposed cells

To find the effect of reduction in thickness of CdTe absorber layer with BSF calculations have been carried out for the CdS/CdTe/Cu_2Te configuration. In Fig. 4, the simulation results for variation of the thickness of CdTe absorber layer from 0.1 µm to 2.9 µm with and without 100 nm Cu_2Te BSF layer are shown together for comparison.

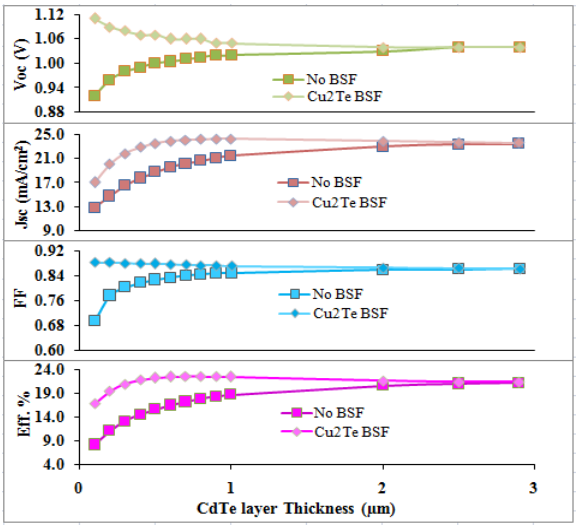


Fig.4 Effect of the CdTe thickness variation for the proposed cells with and without Cu_2Te BSF

The *Voc* and FF show increasing trend as like with decrease of the CdTe layer thickness in the presence of highly doped Cu_2Te BSF, while both FF and *Voc* show decreasing trend for the cell without BSF. The increase of *Voc* and FF both has the largest contribution to improve the efficiency of ultra-thin CdTe solar cell. These results of BSF layer are agreeable to the related published works [10]. Thus, the cell conversion efficiency showed highest value of 22.51% (*Jsc* = 24.26 mA/cm^2, FF = 0.875, *Voc* = 1.06 V) at 0.8 µm of CdTe layer with Cu_2Te BSF.

The stability of the proposed cells at higher operating temperatures can be investigated with Cu_2Te BSF placing to the CdS/CdTe cell. In order to investigate the effects of higher operating temperature on the performances of the cells with and without BSF, simulations were carried out with cell operating temperature ranging from 25°C to 100°C and the simulated results are shown in Fig. 5. It is obvious that without BSF layer and with Cu_2Te BSF layer, the normalized efficiency of the cells linearly decreased with the increase of operating temperature at a temperature coefficient (TC) of -0.18%/°C and -0.16%/°C respectively from Fig. 5. This TC indicates better stability of the cells with BSF at higher operating temperature, which are in good agreement with related published works [10].

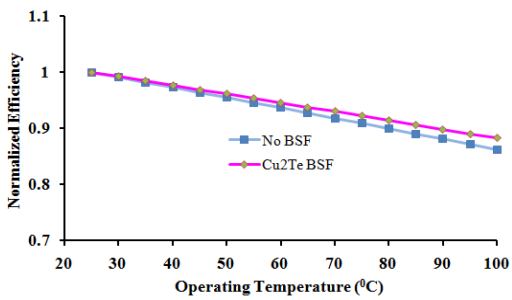


Fig.5 Effects of operating temperature on the cells performance

IV. CONCLUSION

It was found from the CdTe baseline case cell (SnO_2/CdS/CdTe) that 1 µm thick CdTe absorber layer is possible with acceptable range of efficiency of 18.68% (*Jsc* = 21.47 mA/cm^2, FF = 0.85, *Voc* = 1.02 V) without BSF layer. The potential BSF material Cu_2Te was examined for the possible BSF effect with ultra-thin CdTe layer. With a 100 nm BSF layer, the highest efficiency was found where the Cu_2Te BSF with 0.8 µm CdTe absorber layer showed the best conversion efficiency of 22.51% (*Jsc* = 24.27 mA/cm^2, FF = 0.875, *Voc* = 1.06 V). The thermal analysis of the proposed cells showed better stability at higher operating temperatures with a linear TC of -0.16%/°C.

ACKNOWLEDGMENT

This work has been supported by the department of Electrical and Electronic Engineering (EEE) and Renewable Energy Laboratory (REL), Chittagong University of Engineering and Technology (CUET), Bangladesh.

REFERENCES

[1] X. Wu, "High-efficiency polycrystalline CdTe thin-film solar cells," *Solar Energy*, vol. 77, no. 6, pp. 803-814, December 2004.

[2] D. L. Bätzner, A. Romeo, H. Zogg, A. N. Tiwari and R. Wendt, "Development of efficient and stable back contacts on CdTe/CdS solar cells," *Thin Solid Films*, vol. 387, no. 1-2, pp. 151-154, May 2001.

[3] P. D. Demoulin, M. S. Lundstrom, and R. J. Schwartz, "Back-surface field design for n+p GaAs cells," *Solar Cells*, vol. 20, no. 3, pp. 229–236, April 1987.

[4] N. Romeo, A. Bosio, S. Mazzamuto, A. Romeo, and L. Vaillant-Roca, "High efficiency CdTe/CdS thin film solar cells with a novel back contact," in *Proceedings of the 22nd European Photovoltaic Solar Energy Conference*, pp. 1919–1927, Milan, Italy, 2007.

[5] M. A. Matin and Mrinmoy Dey, "High Performance Ultra-Thin CdTe Solar Cell with Lead Telluride BSF," in *Proceedings of the 3rd International Conference on Informatics, Electronics & Vision,* pp. 1-5, Dhaka, Bangladesh, May 2014.

[6] S. J. Fonash, "A Manual for AMPS-ID," *The Centre for Nanotechnology Education and Utilization*, Pennsylvania State University, PA 16802, 2008.

[7] H. Zhu, A. K. Kalkan, J. Hou, and S. J. Fonash, "Applications of AMPS-1D for solar cell simulation," *AIP Conference Proceedings*, vol. 462, pp. 309–314, March 1999.

[8] M. Gloeckler, A. L. Fahrenbruch, and J. R. Sites, "Numerical modeling of CIGS and CdTe solar cells: setting the baseline," in *Proceedings of the 3rd World Conference on Photovoltaic Energy Conversion*, vol. 1, pp. 491–494, May 2003.

[9] M. Burgelman, J. Verschraegen, S. Degrave, and P. Nollet, "Modeling thin film PV devices," *Progress in Photovolaic: Research and Applications,* vol. 12. No. 2-3, pp. 143-153, March 2004.

[10] M. A. Matin, M. M. Aliyu, A. H. Quadery, and N. Amin, "Prospects of novel front and back contacts for high efficiency cadmium telluride thin film solar cells from numerical analysis," *Solar Energy Materials and Solar Cells*, vol. 94, no. 9, pp. 1496–1500, 2010.

Analytical Performance Evaluation of a RF Subcarrier Modulated Optical FSO System under the Influence of Atmospheric Turbulence

Bobby Barua[#1] and S. P. Majumder[#2]

[#1]Ph.D Student, Department of EEE, Bangladesh University of Engineering and Technology, Dhaka, Bangladesh [#2] Professor, Department of EEE, Bangladesh University of Engineering and Technology, Dhaka, Bangladesh
e-mail- bobby@aust.edu

Abstract— **An analytical approach is developed to evaluate the carrier to noise ratio and bit error rate of a free space optical link with RF subcarrier modulation taking into account the effect of weak atmospheric turbulence which is modeled as log-normal distribution. A close form expression for the carrier to noise ratio conditioned on a given turbulence at the output of a direct detection is developed. Average CNR and BER are then evaluated by averaging the results are evaluated in terms of bit error rate and power penalty suffered by the system due to turbulence at a given BER for several system parameters link turbulence variance, system parameters like turbulence variance, link distance, data rate, RF modulation index etc.**

Index Terms- *Amplitude shift keying (ASK) modulation, direct detection, free space optics (FSO), lognormal distribution, RF subcarrier,*

I. Introduction

In recent years, free space optical (FSO) communication has gained significant importance owing to its unique features: large bandwidth, license free spectrum, high data rate, easy and quick deploy ability, less power and low mass requirement. FSO systems (in space and inside the atmosphere) have developed in response to a growing need for high-speed communication systems [1-4]. With the increase in usage of high speed internet, video-conferencing, live streaming etc., the bandwidth and capacity requirements are increasing drastically [1]. This ever growing demand of increase in data and multimedia services has led to congestion in conventionally used radio frequency (RF) spectrum and arises a need to shift from RF carrier to optical carrier [5-8]. Unlike RF carrier where spectrum usage is restricted, optical carrier does not require any spectrum licensing and therefore, is an attractive prospect for high bandwidth and capacity applications [3].

However, as the symbol duration reduces with the increase of data-rate, the systems using single carrier modulation suffer from more severe inter symbol interference *(ISI)* caused by the dispersive fading of FSO channels, and thereby needs more complex equalization [4]. In regard to the FSO system model, however, a significant influence factor is the turbulence-induced channel fading, which impairs the quality of the received signal and link performance improvement. Therefore, it is imperative to apply efficient techniques to mitigate the channel fading caused by the atmospheric turbulence. A number of turbulence mitigation methods for FSO systems have been proposed, such as diversity techniques, adaptive optical technology, robust modulation techniques, etc [5-10]. Again FSO system with optical frequency, amplitude, phase, and polarization modulation are also required to improve the performance that is effected by the atmospheric turbulence. No analytical results are reported with FSO link with RF subcarrier modulation yet.

In this paper, we provide an analytical approach to evaluate the performance of a RF subcarrier modulated free space optical link considering the effect of atmospheric turbulence with ASK modulation scheme. The bit-error rate (BER) performance results are evaluated in the presence of background radiation with weak turbulence assuming that the channel is modeled using lognormal distribution.

II. System Model

Block diagram of the proposed RF subcarrier modulated free space optical link is given in Fig. 1. The mixer of input data and the RF signal is modulated in ASK modulator. The RF modulated signal then transmitted to the optical intensity modulator where electro optic modulation occurs with the laser output. Then the EOIM signal transmitted to through atmospheric turbulent channel. At the receiving end optical signal, received through photodetector. Then the received signal first amplified and passes through the band pass filter to find the RF signal output. Finally the RF signal mixed with the local oscillator signal and demodulated using direct detection demodulator to achieve the data output.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

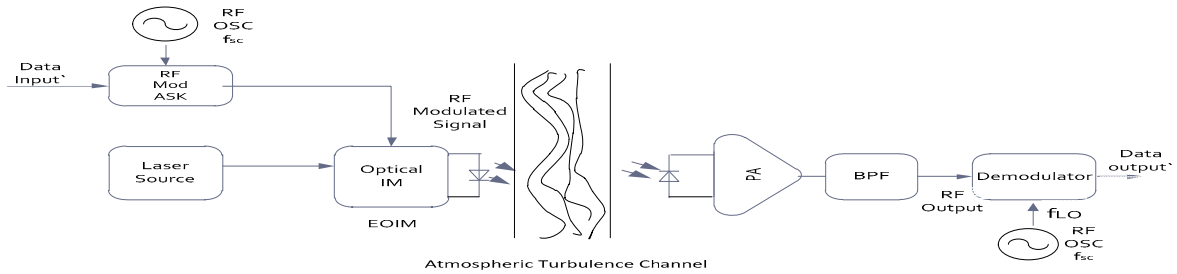


Fig.1: RF Subcarrier Modulated Optical FSO System with Direct Detection Receiver

III. Channel Model

For proper performance analysis of FSO communications it is important to characterize the FSO channel from a communication theory perspective. On that basis it is useful to give a statistical representation of the scintillation, and the reliability of the communication link can be determined if we use a good probabilistic model for the turbulence. There are several probability density functions (PDFs) have already been proposed for the intensity variations at the receiver of an optical link in different publications. Al-Habash et al. [11] proposed a statistical model that factorizes the irradiance as the product of two independent random processes each with a Lognormal PDF. The log-normal distribution is generally used to model the fading associated with the weak atmospheric turbulence regime [5, 9, 11]. This model is mathematically tractable and it is characterised by the Ryotov variance σ_i^2. The probability density function of the received irradiance is given by: [11]

$$p(I) = \frac{1}{\sqrt{2\pi}\,\sigma_i}\frac{1}{I}\exp\left\{-\frac{\left(\ln\left(\frac{I}{I_o}\right)+\frac{\sigma_i^2}{2}\right)^2}{2\sigma_i}\right\} \quad I\geq 0 \qquad (1)$$

Where I is the irradiance at the receiver I_0 is the Signal irradiance without scintillation represents the σ_i^2 Ryotov Variance where is g given by [10]:

$$\sigma_i^2 = 1.23\, C_n^2 k^{7/6} L^{11/6} \text{ and } \sigma_i^2 < 1.2$$

Ryotov Variance C_n^2 is the Refractive index structure Parameter and L represents the Propagation distance k represents the wave number which is $2\pi/\lambda$ and λ represents the operating wave length which is1550 nm.

IV. Theoretical Analysis

The electrical field output of the RF-ASK modulator is given by:

$$e_{sc}(t) = \sum_k a_k\, p(t - kT_b).A_{sc}\, Cos\,(\omega_{sc} t) \qquad (2)$$

where ω_{sc} is the RF angular frequency, A_{sc} is the amplitude of the subcarrier and T_b is the bit period, *p(t)* pulse shape of a bit, a_k is the *k-th* bit.

The RF modulated signal is then used to intensity modulate (IM) then laser source using an electro optic intensity modulator (EOIM).

The electric field output of the intensity modulator is given by

$$e_{opt}(t) = \sqrt{2P_T}\left[1 + k_a.S_{sc}(t)\right]e^{j\omega_c t} \qquad (3)$$

where P_T represents the transmitted laser power and k_a is the intensity modulation index. Now the modulated signal is then transmitted through the FSO channel. The received signal at the receiver is given by:

$$e_r(t) = \sqrt{2P_r}\left[1 + k_a.S_{sc}(t)\right]Ie^{j\omega_c t} + n_b(t) \qquad (4)$$

In the equation P_r is the received power which is equal to $P_T e^{\alpha L}$, α is the attenuation coefficient of the atmospheric channel L represents the link distance, n_b is the background noise and I represents the turbulence induced fading. The photo detector current i_d is given by:

$$\begin{aligned} i_d(t) &= R_d\left|e_r(t)\right|^2 \\ &= 2R_dP_RI\left|1+k_a.S_{sc}(t)\right|^2 \\ &= 2R_dP_RI + 2R_dP_RI\times k_a.S_{sc}(t) + 2R_dP_RI\times\left(k_a.S_{sc}(t)\right)^2 \end{aligned} \qquad (5)$$

where R_d is the responsivity of the photodetector. The RF carrier power at the output of the BPF is given by,

$$C(I) = \frac{(2R_dP_TIk_aA_{sc})^2}{2} \qquad (6)$$

The noise power at the output of the band pass filter is given by,

$$\begin{aligned} N &= \sigma_{sh}^2 + \sigma_{th}^2 \\ &= 2eB[I_s] + \frac{4kT}{R_L}B \\ &= 2eB[R_dP_TI\times k_a.A_{sc}(t)] + \frac{4kT}{R_L}B \end{aligned} \qquad (7)$$

where B is the band width of the BPF, σ_{sh}^2 is the shot noise variance, σ_{th}^2 is the thermal noise variance. Now the Carrier to Noise Power Ratio conditioned on a given turbulence induced fading *I* can be expressed as

$$CNR(I) = \frac{\frac{(2R_dP_RIk_aA_{sc})^2}{2}}{2eB[R_dP_RI\times k_a.A_{sc}(t)] + \frac{4kT}{R_L}B} \qquad (8)$$

So, the conditional BER is then given by,

$$BER(I) = \frac{1}{2}erfc\left[\frac{\sqrt{CNR(I)}}{2\sqrt{2}}\right] \qquad (9)$$

And the average BER,

$$BER = \int BER(I).p(I)dI \qquad (10)$$

Where *p(I)* represents Lognormal PDF.

V. Results and Discussion

In this section the various performance evaluation results which have been obtained by numerically evaluating the mathematical expressions presented in Sections III and IV, we evaluate the bit error rate performance result using Matlab.. The parameters are given in Table I.

Table1: System Parameters used for computation

Parameters	Values
Data rate	1Gbps
Subcarrier no.	1
Laser wavelength λ	1550nm
PIN photodetector Responsivity, R	0.85
RF modulation index, k	0.5 and 1
Link distance, L	3000 m
Received power	-20 to 10 dBm
Subcarrier power	0.5-1.0 watt
Lognormal variance, σ_i^2	$0.1 \le \sigma_i^2 \le 0.7$
Refractive index structure Parameter, C_n^2	$1*10^{-14} m^{-2/3}$

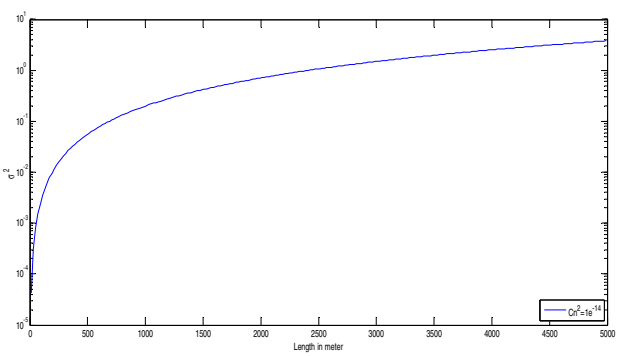


Fig2: Plot of variance vs. link distance

Fig.2 shows the relation between variance and link distance with a particular refractive index structure parameter. From the figure we found that the variance is proportionate to the link distance but after a certain distance (3km) the effectiveness of the variance with link distance became less.

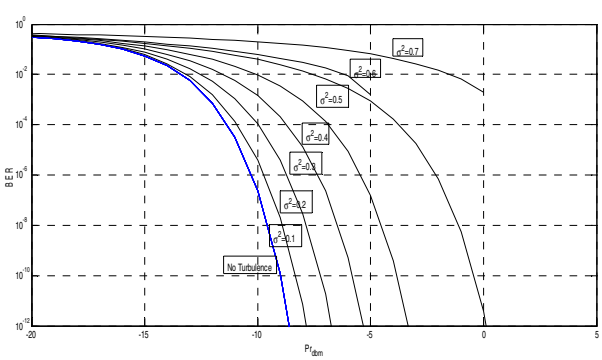


Fig3: Plots of BER vs. Received optical power with turbulence variance.

Fig.3 shows the plots of BER vs. Received power with turbulence variance as parameter. It is observed that the system performance degrades due to the influence of atmospheric turbulence. At higher variance , the required optical power is more at a given data rate.

Now the effect of turbulence can be obtained by plotting the fading variance against the power penalty and the plots are shown in Fig.4. The power penalty represents the increment in received power to achieve a desired BER in a turbulence channel compared to the ideal channel. It is also observed that for same turbulence level the fading penalty is higher for lower BER. For example, when the turbulence variance is 0.5 the fading penalty are ~ 5.5 dB , 6.7 dB and 8.9 dB for BER of 10^{-3}, 10^{-6} and 10^{-9}, respectively.

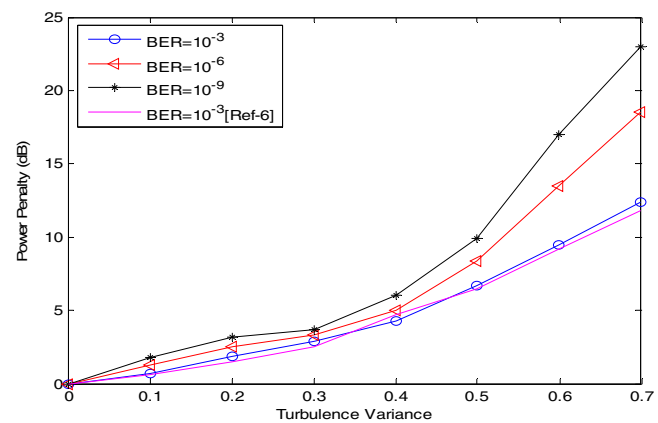


Fig4: Power Penalty vs. Turbulence variance under different BER Condition

Results obtained by simulation [6] are shown in Fig.4. for same parameters for the same turbulence level under a lognormal atmospheric turbulence model. Comparing the results we found that the are very close to that reported by simulation [6]. For example, when the turbulence variance is 0.7 the power penalty is ~ 11.8 dB for BER of 10^{-3} which is close to our analytical result.

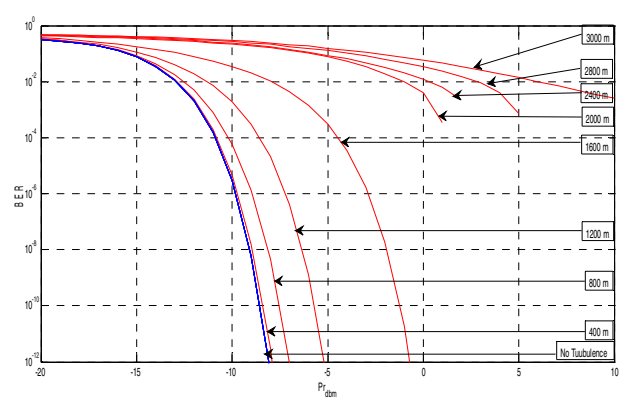


Fig 5: Plots of BER vs. Received power with variation of link distance (meter) for modulation index-1.0

In Fig. 5, the dependence on the link distance L for RF modulation index of 1.0 of the system is illustrated. It is further assumed the operating wavelength is λ = 1550 nm , $C_n^2 = 1\times10^{-14} m^{-2/3}$ and subcarrier power is 1 W. As expected, the

error performance deteriorates as L increases from L = 400m to L = 3000m. Moreover, it is evident that an increase in L from 400m to 2000m results in a more severe performance deterioration than in the case where L increases from 2000m to 3500m. In all cases considered, the analytical results obtained using equation 10.

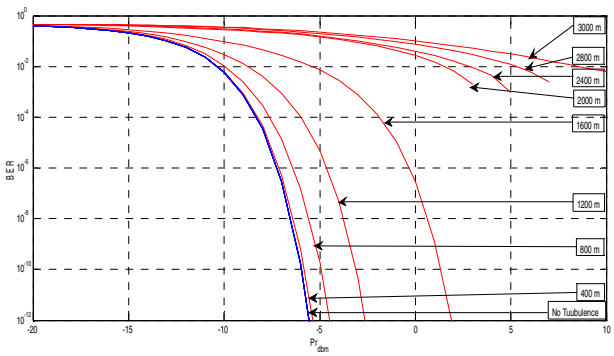


Fig6: Plots of BER vs. Received power with variation of link distance (meter) for modulation index of 0.5

Fig 6 shows the BER characteristics with the variation of link distance for modulation index of 0.5. It is noticed that the system with lower RF modulation index suffers more due to atmospheric turbulence. And the plots of power penalty curve with respect to the link distance at BER 10^{-3}, 10^{-6} and 10^{-9}, are shown in Fig 7. It is noticed from the plots that the requirement of optical power is changed with the variation of modulation index but the corresponding power penalty are remain unchanged.

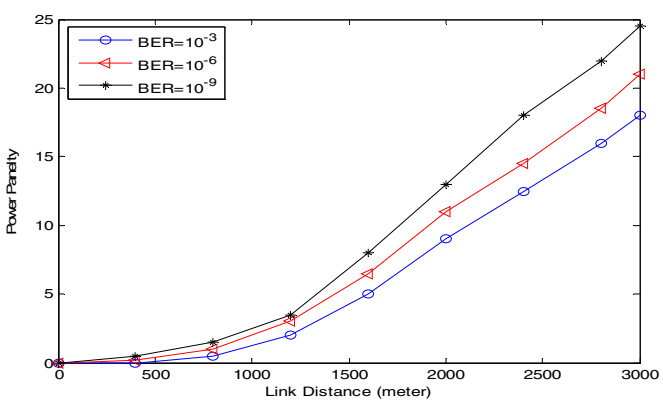


Fig7: Power Penalty vs. Link Distance under different BER Condition

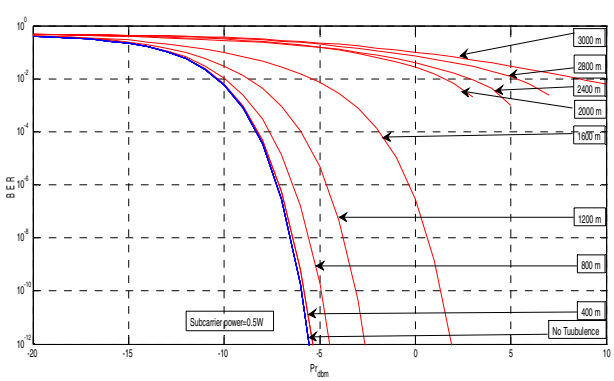


Fig 8: Plots of BER vs. Received power with variation of link distance (m) with subcarrier power 0.5 W.

Next we compare the BER characteristics with the variation of subcarrier power. Fig 8 shows the plots of BER vs. Received power for various link distances with subcarrier power 0.5 W. It is noted that for similar aperture configurations, the BER characteristics is almost similar to the previous one but the only thing is for reducing the value of subcarrier power demands higher power.

VI. Conclusions

An analytical approach is presented to evaluate the bit error rate performance of a RF subcarrier modulated FSO link in presence of atmospheric turbulence which is modeled as lognormal distribution. The results show that system suffers significant power penalty due to atmospheric turbulence which is more prominent at lower RF modulation and lower link distance. The analytical results reported earlier and are found to be in good agreement. The results of this research will find applications in design of RF modulated FSO link for high data rate application.

Acknowledgements

The authors would like to acknowledge with gratitude the support provided by the Dept. of EEE, BUET to carry out this work as a part of Ph.D. research.

References

1. H. Henniger, O. Wilfert," An Introduction to Free-Space Optical Communications'" Journal of Radio Engineering vol. 19, no. 2, June 2010.
2. V. W. S. Chan, "Free-space optical communications," J. Light wave Tech., vol. 24, no. 12, pp. 4750–4762, 2006.
3. S. Das, H. Henniger, ,B. Epple, , C. Moore, , W. Rabinovich, , R. Sova, , D. Young, "Requirements and challenges for tactical free-space" laser comm. IEEE Military Communications Conference . San Diego (USA), 2008, p. 1 – 10.
4. Viswanath, H. Kaushal, V. K. Jain, and S. Kar, "Evaluation of performance of ground to satellite free space optical link under turbulence conditions for different intensity modulation schemes," Proc. SPIE, Free Space Laser Comm. and Atmosph. Prop. (XXVI), vol. 8971, 2014.
5. H. Samimi,; P. Azmi, Performance analysis of adaptive subcarrier intensity-modulated free-space optical systems. IET Optoelectron. 2011, 5, 168–174.
6. X. Tang, S. Rajbhandari; W. O. Popoola, Z. Ghassemlooy, S.S. Muhammad, E. Leitgeb, and G. Kandus, "Performance of BPSK Subcarrier Intensity modulation Free-space optical communications using a Lognormal Atmospheric Turbulence Model" *IEEE Conference* , pp. 17 - 20, July 2010.
7. W.O. Popoola,; Z. Ghassemlooy, BPSK subcarrier intensity modulated free-space optical communications in atmospheric turbulence. J. Light wave Technol. 2009, 27, 967–973
8. W. O. Popoola, Z. Ghassemlooy, J. I. H. Allen, E. Leitgeb, and S. Gao, "Free-space optical communication employing subcarrier modulation and spatial diversity in atmospheric turbulence channel" *Optoelectronics, IET,* vol. 2, pp. 16 - 23, February 2008.
9. Z. Ghassemlooy, W. O. Popoola, and E. Leitgeb, "Free-space optical communication using subcarrier modulation in gamma-gamma atmospheric turbulence," *ICTON '07,* vol. 3, pp. 156 - 160,1 - 5 July 2007.
10. L.C. Andrews,. R.L. Philips, Laser Beam Propagation Through Random Media; SPIE Optical Engineering Press: Bellingham, USA, 2005.
11. M. A. Al-Habash, L. C. Andrews, and R. L. Phillips, "Mathematical model for the irradiance probability density function of a laser beam propagating through turbulent media," Opt. Eng. 40, 1554–1562 (2001).

MTPA and Field Weakening Control of Synchronous Reluctance Motor

S.M.Ferdous,[1] Pablo Garcia,[1] Mohammad Abdul Moin Oninda,[2,*] and Md. Ashraful Hoque[2]

[1]Department of Electrical, Electronics, Computers and Systems Engineering, University of Oviedo, Gijon, Asturias, Spain-33204.

[2]Department of Electrical and Electronic Engineering, Islamic University of Technology, Board Bazar, Gazipur-1704, Bangladesh.
*mamoin@iut-dhaka.edu

Abstract—**An improved algorithm for MTPA (Maximum Torque Per Ampere) and Field Weakening operation which is suitable for implementation with Field Oriented Control, has been proposed for the Synchronous Reluctance motor (SynRel) for high dynamic performance. The proposed method overcomes the problem of excessive current flow into the machine while generating the d-axis and q-axis current commands independently. At the same time to operate the inverter in the linear modulation region (*$0 \le m \le 1$*) to minimize the output torque ripple. The proposed method helps to attain MTPA operation below the base speed to operate with the rated load and a fast speed response in Field Weakening mode. The validity and efficacy of the proposed algorithm is verified by simulation in MATLAB and SIMULINK.**

Index Terms—**Field-weakening; MTPA; FOC; Synchronous Reluctance (SynRel) Motor**

I. Introduction

Due to absence of permanent magnet, Synchronous Reluctance Motor has very poor Torque production capability and relatively low power factor. Hence it is imperative that, the torque production has to be maximized for such kind of motor. MTPA (Maximum Torque per Ampere) is the most common method to maximize the torque production of a motor for a certain amount of current and hence also play a role to improve the overall efficiency [1].

Absence of rotor winding, low torque ripple, low cost and robustness makes a SynRel motor more desirable compared to induction motors [2]-[5]. Moreover absence of permanent magnet gives it the added feature of being cost effective.

Most of the application demands wide speed range of operation which is achieved by Field Weakening. A SynRel Motor is basically a high speed motor with max Speed 2 to 5 times than that of the base speed. Several approaches are available in literature where Field weakening is achieved using the Dynamic Machine equations in d-q reference frame [6]-[8].

This paper represents an Algorithm to incorporate MTPA and FW methods applied together to obtain a high performance SynRel Motor Drives, best suited for Traction application. Basic rotor reference d-q Field Oriented Control equations are used to develop the model.

II. Mathematical Modelling of Synchronous Reluctance Motor in d-q Reference Frame

A. Dynamic Equation

A Synchronous Reluctance Motor does not have any permanent magnet or any winding in the rotor. Hence the motor is modelled in rotor reference frame. The dynamic rotor reference frame equations for d and q axis voltages are-

$$v_d = R_s i_d + L_d \frac{di_d}{dt} - \omega_e L_q i_q \tag{2.1}$$

$$v_q = R_s i_q + L_q \frac{di_q}{dt} + \omega_e L_d i_d \tag{2.2}$$

where, ω_e = Rotor speed in elec. rad/s

The developed electromagnetic torque, in terms of d-q parameters, can be written as –

$$T_e = \frac{3}{2}\frac{P}{2}\left(L_d - L_q\right) i_d i_q \tag{2.3}$$

where, P = No. of poles of the rotor

The dynamic equations for the mechanical subsystem can be expresses as –

$$T_e = J\frac{d\omega_m}{dt} + B\omega_m + T_L \tag{2.4}$$

where, ω_m=Rotor speed in mechanical rad/s, T_L=Load torque

$$\omega_e = \frac{P}{2}\omega_m$$

Two limiting parameters are needed to be addressed for implementing the Field Oriented Control or Vector Control of Synchronous Motor. They are –

$$I_{s,max}^2 = i_d^2 + i_q^2 \tag{2.5}$$

$$V_{s,max}^2 = V_d^2 + V_q^2 \tag{2.6}$$

where,

$I_{s,max}$ = Maximum Allowable Current by the inverter

$V_{s,max}$ = Maximum Inverter voltage (per phase)

$$V_{s,max} = \frac{V_{dc}}{2}$$

where, V_{dc}=Inverter DC link voltage. Here, the inverter is operated under Sinusoidal PWM excitation.

The equivalent circuit for SynRel is shown in Fig.1–

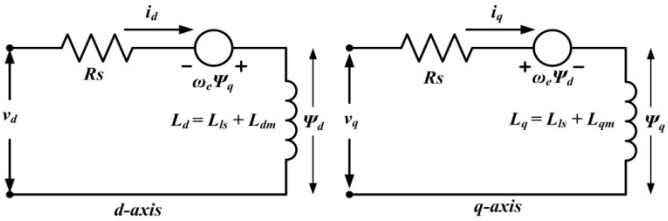


Fig. 1 Equivalent d-q circuit model of SynRel Motor

where, ψ_s=Total Stator Flux

$$\psi_s = \psi_d + \psi_q$$

L_d=d-axis Inductance = $L_{ls} + L_{dm}$

L_q = q-axis Inductance = $L_{ls} + L_{qm}$

L_{ls} = Leakage Inductance

L_{dm} = d-axis Magnetizing Inductance

L_{qm} = q-axis Magnetizing Inductance

B. Determination of Operating Point

The operating point of the motor can be determined by plotting the current and voltage limit equation together in current plane (i_d vs i_q). The current limit equation from (2.5) is an equation of a circle. The radius of the circle is $I_{s,max}$ which represents the maximum allowable current in current plane.

Similarly, the voltage limit equation can also be plotted in the same diagram. To plot it with the same parameters, i.e. as a function of i_d and i_q; the equation can be rewritten neglecting the resistive drops and derivative terms.

This is a valid assumption because at steady state the derivative term are zero and the resistive drops are much smaller than the Back EMF or Speed Induced voltages. Hence the voltage limit equation (2.6) can be written as [3]–

$$V_{s,max}^2 = (-\omega_e L_d i_d)^2 + (\omega_e L_q i_q)^2$$

$$V_{s,max}^2 = \omega_e^2 L_d^2 i_d^2 + \omega_e^2 L_q^2 i_q^2$$

$$1 = \frac{\omega_e^2 L_d^2 i_d^2}{V_{s,max}^2} + \frac{\omega_e^2 L_q^2 i_q^2}{V_{s,max}^2}$$

$$\frac{i_d^2}{\left(\frac{V_{s,max}^2}{\omega_e L_d}\right)^2} + \frac{i_q^2}{\left(\frac{V_{s,max}^2}{\omega_e L_q}\right)^2} = 1 \qquad (2.7)$$

$$\frac{i_d^2}{a^2} + \frac{i_q^2}{b^2} = 1$$

where,

$$a = \frac{V_{s,max}}{\omega_e L_d} \; ; \; b = \frac{V_{s,max}}{\omega_e L_q}$$

This is an equation of an ellipse with centre *(0, 0)*. As $L_d > L_q$ vertex is equal to *'b'* and co-vertex is equal to *'a'*, where $b > a$. Thus this is a vertical ellipse whose major axis lies along i_q and minor axis lies along i_d.

III. MTPA Operation of the Synchronous Reluctance Motor

MTPA or Maximum Torque per Ampere operation is the most preferred mode of operation for any motor operating in Field Oriented Control (FOC). This operation provides the maximum amount of torque for a given operating current.

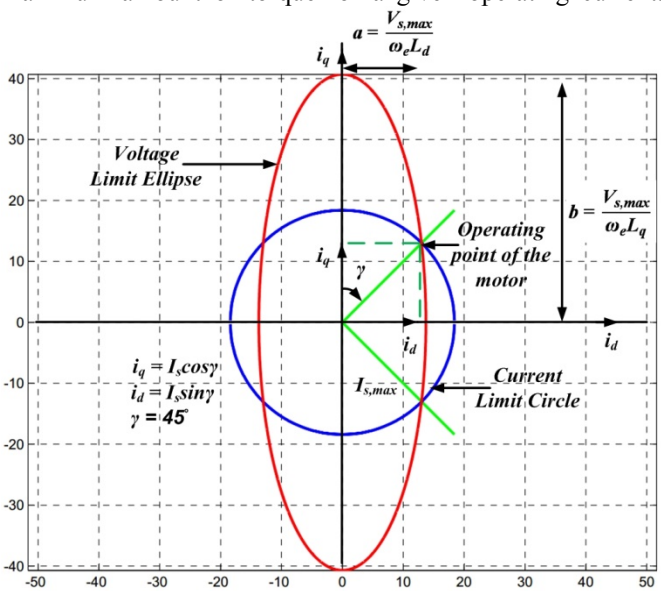


Fig. 2 Operating Characteristics of SynRel Motor

MTPA condition for SynRel Motor can be obtained from the torque equation.

$$T_{em} = \frac{3}{2}\frac{P}{2}(L_d - L_q) i_d i_q$$

$$i_d = I_s sin\gamma$$

$$i_q = I_s cos\gamma$$

$$T_{em} = \frac{3}{2}\frac{P}{2}(L_d - L_q) I_s^2 sin\gamma \, cos\gamma$$

$$T_{em} = K I_s^2 \frac{sin2\gamma}{2}$$

where,

$$K = \frac{3}{2}\frac{P}{2}(L_d - L_q)$$

Condition for MTPA is obtained as-

$$\frac{dT_{em}}{d\gamma} = \frac{K}{2} I_s^2 cos2\gamma \, . \, 2 = 0$$

$$cos2\gamma = 0$$

$$cos2\gamma = cos90°$$

$$\gamma = \frac{90}{2} = 45°$$

$$\gamma_{MTPA} = 45°$$

Thus, the equation for torque under MTPA operation

$$T_{em} = K\frac{I_s^2}{2} sin2\gamma = K\frac{I_s^2}{2}\Big|_{\gamma=45°}$$

$$I_s = \frac{\sqrt{T_{em}}}{\sqrt{K}} = \frac{1}{\sqrt{K}} \times \sqrt{T_{em}} \qquad (3.1)$$

$$i_d = I_s sin45° = \frac{I_s}{\sqrt{2}} = 0.707 I_s \qquad (3.2)$$

$$i_q = I_s cos45° = \frac{I_s}{\sqrt{2}} = 0.707 I_s \qquad (3.3)$$

i.e. $i_d = i_q$

Now, for Negative speed command, reference torque will be negative and hence I_S cannot be determined. Hence a lookup table can be generated in the following approach to generate the current reference I_S^* for all values of T_{em}^*.

$$I_s^* = -\frac{1}{\sqrt{K}} \times \sqrt{|T_{em}|} \qquad \text{for } T_{em} < 0$$

$$I_s^* = \frac{1}{\sqrt{K}} \times \sqrt{T_{em}} \qquad \text{for } T_{em} \geq 0 \qquad (3.4)$$

The current reference I_s^* as a function of T_{em} is shown in the following figure –

The MTPA Control Block diagram is shown in Fig.3. –

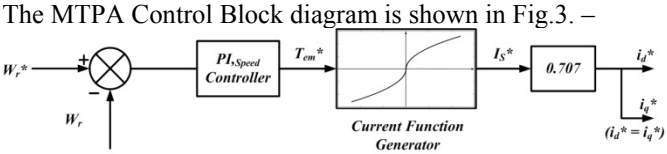


Fig. 3 Close Loop speed control Scheme for MTPA operation

IV. Field Weakening Operation

Field Weakening Equation can be obtained by combining eq. (2.1), (2.2) and (2.6) –

$$V_d = R_s i_d + sL_d i_d - \omega_e L_q i_q$$

$$V_q = R_s i_q + sL_q i_q + \omega_e L_d i_d$$

$$V_{s,\,max}^2 = V_d^2 + V_q^2 = (\omega_e L_q i_q)^2 + (\omega_e L_d i_d)^2$$

$$V_{s,\,max}^2 = \omega_e^2 L_q^2 I_s^2 cos^2\gamma + \omega_e^2 L_d^2 I_s^2 sin^2\gamma$$

$$V_{s,\,max}^2 = \omega_e^2 I_s^2 (L_q^2 cos^2\gamma + L_d^2 sin^2\gamma)$$

$$V_{s,\,max}^2 = \omega_e^2 I_s^2\left[L_q^2 - L_q^2 sin^2\gamma + L_d^2 sin^2\gamma\right]$$

$$sin^2\gamma = \frac{V_{s,\,max}^2 - \omega_e^2 L_q^2 I_s^2}{\omega_e^2 I_s^2\left(L_d^2 - L_q^2\right)}$$

$$\gamma_{FW} = sin^{-1}\sqrt{\frac{V_{s,\,max}^2 - \omega_e^2 L_q^2 I_s^2}{\omega_e^2 I_s^2\left(L_d^2 - L_q^2\right)}} \qquad (4.1)$$

Eq. (4.1) is the governing Field Weakening equation as a function of Motor speed. Using this equation, the angel (γ_{FW}) between d and q axis currents for a particular speed is calculated and hence the reference currents i_d* and i_q* for Field Weakening are generated. It is to be noted that, this equation is to be used only when the speed is greater than base speed. Comparator control logic is needed for this purpose. The control block schematic is shown in Fig.4.

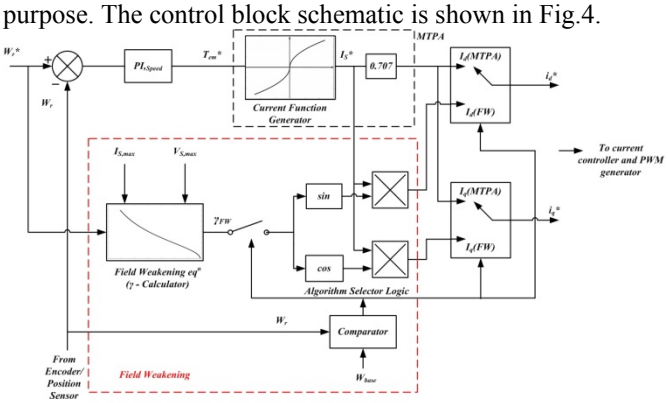


Fig. 4 Field Weakening Control Implementation for SynRel Motor

The operation of MTPA and FW of the motor can be shown together using the i_d vs i_q plot as in Fig.5. OA is the locus for MTPA operation whereas AB is the FW locus. A is the point that corresponds to base speed of the motor and B represents the maximum speed of the motor. The voltage ellipse shrinks as the speed of the motor increases. B is the point where the Current circle and voltage ellipse touches each other and hence this is the maximum speed point. If the voltage further shrinks then there is no point of intersection between them and hence no valid point of operation would exist. So from zero to maximum speed the operation of the motor would traverse through the locus O to A to B, where O to A is MTPA and A to B is FW operation.

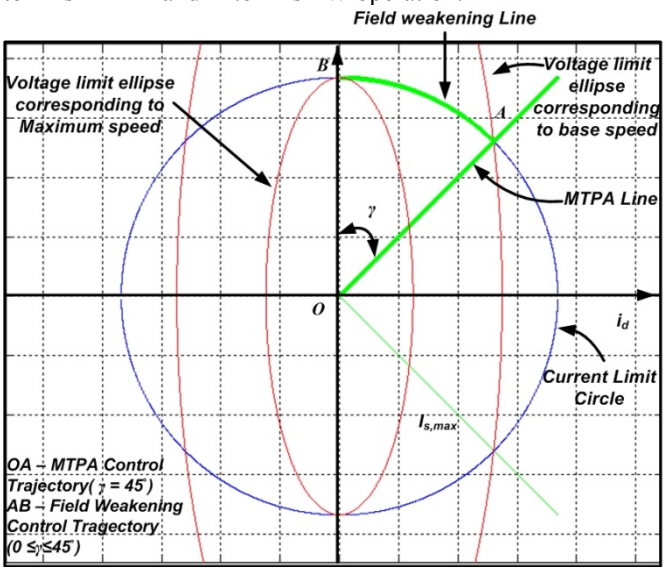


Fig. 5 Region of operation in MTPA and Field Weakening mode

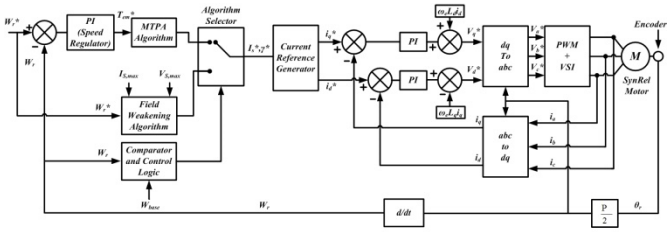


Fig. 6 Overall Schematic of Field Oriented Control of SynRel Motor with MTPA and Field Weakening operation

V. Simulation and Discussion

The developed algorithms are simulated in MATLAB and SIMULINK. The simulation results shown here are divided in to two sections. The first sets of results are for MTPA operation with rated load condition. The second sets of results are for Field weakening operation up to maximum Motor speed with no load condition.

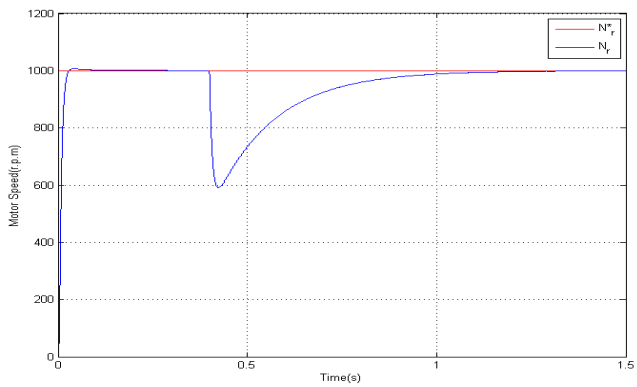
A wide speed range SynRel Motor used for EPS (Electrical Power Steering) application is taken as reference to verify the algorithm as given in [2]. The motor parameters are given in the Appendix.

A. MTPA Operation at Rated Load

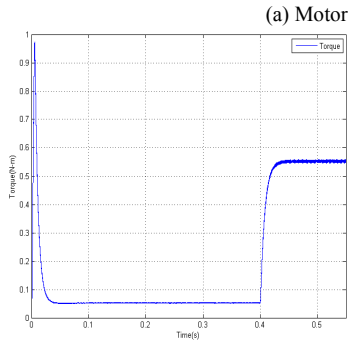
The motor is to operate at its base speed of 1000rpm with rated load of 0.6Nm. The load torque is applied at the instant of 0.4 second. N_r* is the reference speed and N_r is the actual motor speed. The result shows that the speed control loop is capable of tracking the speed precisely as well as compensating the disturbance that occurs due to the application of load torque.

The other parameters like generated Torque, d and q axis currents and three phase Motor currents are also plotted. It is seen that the peak starting torque is around 0.98 Nm which is acceptable for the chosen motor.

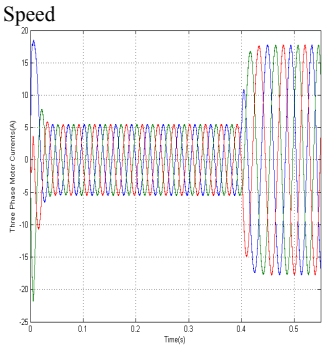
Depending on the peak torque handling capability, the Speed Controller Bandwidth can be chosen to obtain the desired response.



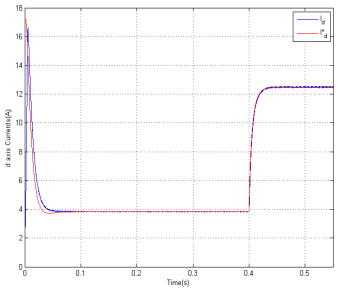
(a) Motor Speed



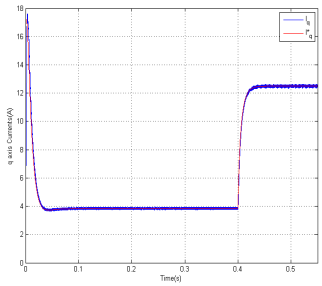
(b) Torque



(c) Three phase Currents



(d) d axis Currents, i_d^* & i_d

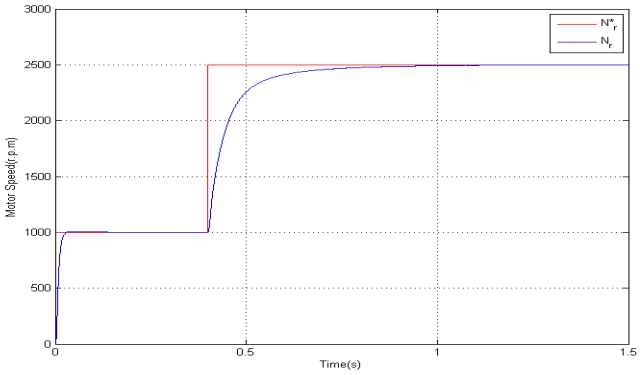


(e) q axis Currents, i_q^* & i_q

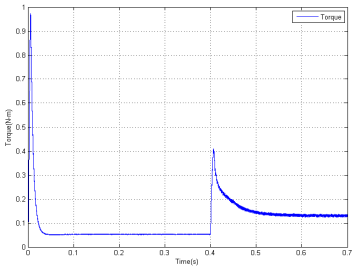
Fig. 7 Motor characteristics and parameters in MTPA operation

B. Field Weakening Operation

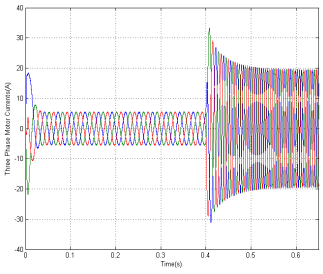
The motor responses under Field weakening operation are as follows-



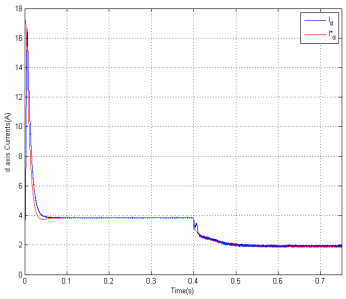
(a) Motor Speed



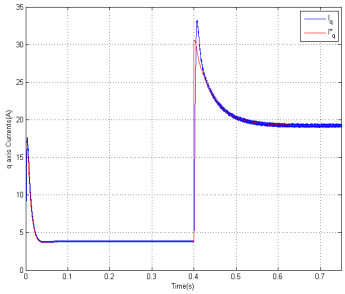
(b) Torque



(c) Three Phase Motor Current



(d) d axis Currents, i_d^* & i_d



(e) q axis Currents, i_q^* & i_q

Fig. 8 Motor characteristics and parameters in Field Weakening operation

In Field Weakening operation there are some overshoot in the transient condition which can be limited by using a saturation block. Here, for this particular motor, the overshoot was allowed because this transient peak power is less than the maximum power of the motor.

VI. Conclusions

The developed algorithms can successfully work on MTPA and Field weakening region. However, two limitations of the algorithm can be pointed out. Firstly, the algorithm is parameter dependent. So any variation in L_d and L_q or R_s will deteriorate the performance. It is quite obvious that the parameters will not remain constant for a SynRel motors as the variation of inductances is quite prominent as a function of motor current. At the same time the speed and current controller performance will also be hampered due to this parametric variation and hence affect the overall operation of the motor. Secondly, the dynamic model is developed neglecting the core loss which may not be neglected in practical implementation. The effect of core needs to be taken in to account and modelled accordingly for high dynamic performance [1]. As a result the developed algorithm of for MTPA will not be valid and a different approach will be needed for MTPA operation of the motor. To overcome the issues associated with these two crucial points can be a very good topic for further future work of the presented paper.

Appendix

TABLE I
Motor Parameters

Parameter	Value
Rated Power	60W
Maximum Power	150W
Rated Torque	0.6Nm
Base Speed	1000rpm
Maximum Speed	2500rpm
Number of Poles	4
Rated Current	18.5A
Inverter DC Link Voltage	12V
Stator Resistance (per phase)	0.11Ω
d-axis Inductance (at rated current)	1.78mH
q-axis Inductance (at rated current)	0.6mH
Current Controller Bandwidth	300Hz
Speed Controller Bandwidth	10Hz

References

[1] E.M.Rashad, T.S.Radwan, M.A.Rahman, "A Maximum Torque per Ampere Vector Control Strategy for Synchronous Reluctance Motors Considering Saturation and Iron Losses", in Proc. IAS, 2004, vol.1, pp.2411-2417.

[2] J. Ahn, S.-B. Lim, K.-C. Kim, J. Lee, J.-H. Choi, S. Kim and J.-P. Hong, "Field weakening control of synchronous reluctance motor for electric power steering," IEEE IET Elec. Power Appl. vol. 1, no. 4, pp.565–570, 2007.

[3] Y. Inoue, S. Morimoto, and M. Sanada, "A Novel Control Scheme for Maximum Power Operation of Synchronous Reluctance Motors Including Maximum Torque Per Flux Control," IEEE Trans. Ind. Appl., vol. 47, no. 1, pp.115–121, 2011.

[4] I.Boldea, L. N. Tutelea, L. Parsa, and D. Dorrell, "Automotive electric propulsion systems with reduced or no permanent magnets: An overview,"IEEE Trans. Ind. Electron., vol. 61, no. 10, pp. 5696–5711, Oct. 2014.

[5] S. Taghavi and P. Pillay, "A sizing methodology of the synchronous reluctance motor for traction applications," IEEE J. Emerg. Sel. Topics Power Electron., vol. 2, no. 2, pp. 329–340, Jun. 2014.

[6] D. A. Staton, T. J. E. Miller, and S. E. Wood, "Maximizing the saliency ratio of the synchronous reluctance motor,"in Proc. Inst. Elect. Eng. Elect. Power Appl., vol. 140, no. 4, pp. 249–259, Jul. 1993.

[7] T.-S. Kwon and S.-K. Sul, "Novel flux-weakening control of an IPMSM for quasi-six-step operation," IEEE Trans. Ind. Appl., vol. 44, no. 6,pp. 1722–1731, Sep./Oct. 2008.

[8] T. M. Jahns, "Flux-weakening regime operation of an interior permanent magnet synchronous motor drive,"IEEE Trans. Ind. Appl., vol. IA-23, no. 4, pp. 681–689, Jul./Aug. 1986..

[9] Seog-Joo Kang, Seung-Ki Sul, "Highly Dynamic Torque Control of Synchronous Reluctance Motor," in Proc. IEEE Power Electronics Specialist Conference, vol. 2, pp. 1793-1797, 1996.

[10] Milutin G. Jovanovit, Robert E. Betz, "Maximum Torque Control of a Sensorless Synchronous Reluctance Motor Drive," inProc. IEEE Ind. Appl., vol. 1, pp. 637-644, 1997.

[11] R. Lagenpist, I. Boldea, T.J.E. Miller, "Sensorless Control of the Synchronous Reluctance Motor," IEEE Tran. on Ind. Appl. vol. 30, no. 3, pp. 673–682, 1994.

Evaluation of Discrimination Power of Facial Parts from 3D Point Cloud Data

Rafiul Amin†, A. Farhan Shams†, S. M. Mahbubur Rahman†, *Member*, *IEEE*, Dimitrios Hatzinakos‡, *Fellow*, *IEEE*
†Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology, Dhaka-1205, Bangladesh
‡Identity, Privacy and Security Institute, Department of Electrical and Computer Engineering
University of Toronto, Toronto, ON, Canada, M5S 2E4
E-mail: {rafiulamin, farhanshams314}@gmail.com, mahbubur@eee.buet.ac.bd, dimitris@comm.utoronto.ca

***Abstract*—Feature selection from facial regions is a well-known approach to increase the performance of 2D image-based face recognition systems. In case of 3D modality, the approach of region-based feature selection for face recognition is relatively new. In this context, this paper presents an approach to evaluate the discrimination power of different regions of a 3D facial surface for its potential use in face recognition systems. We propose the use of weighted average of unit normal vector on the facial surface as the feature for region-based face recognition from 3D point cloud data (PCD). The iterative closest point algorithm is employed for the registration of segmented regions of facial point clouds. A metric based on angular distance between normals is introduced to indicate the similarity between two surfaces of same facial region. Finally, the intra class correlation based discrimination score is formulated to find out the key facial regions such as the eyes, nose, and mouth that are significant while recognizing a person with facial surface PCD.**

I. Introduction

Face recognition has been remained as one of the most challenging research areas in computer vision and pattern recognition. In the literature, there are many approaches that have been adopted to find the best possible automatic face recognition system [1]. Though most of the initial works were based on holistic facial features, the landmark-based features selected from facial regions have also been examined. Examples of the region-based features for whole faces can be found in the methods reported in [2]. Different partitioning schemes to find facial features have also been studied [3]. In some cases, whole face images cannot be obtained or some of the key regions of the faces can be contaminated with heavy noise such as occlusions. Considering those difficulties, the region-based approaches have been considered for face recognition as well as for face detection [4]. For example, the facial parts are used to recognize identity from CCTV footage in computational forensics [5]. Algorithms for recognizing faces after plastic surgery have also incorporated the region-based approach [6]. Li et al. [7] reported an approach for face recognition by matching non-corresponding region of facial images. Although region-based approaches are quite familiar in the intensity-based 2D modality, it is relatively new and yet to be thoroughly studied in case of 3D modalities of facial imaging.

In pursue of accuracy, the 3D modalities such as 3D depth image and 3D voxel data are becoming increasingly popular as compared to 2D pixel data of images. Several works have been reported on face recognition with 3D modality using depth image as well as 3D surface data (see for example, [8] and references therein). The 3D modality has been adopted to construct algorithms that are immune to expression variability. Since the nasal region undergoes less geometric deformation in presence of facial expression, the nasal region-based recognition system has also been reported [9]. It is apparent that a comparative study of recognition capability of different regions of 3D facial surface is necessary to develop an efficient face recognition algorithm.

(a) (b) (c)

Fig. 1: Examples of facial surfaces represented by PCD. (a) Face without occlusion. (b) Face occluded by beard. (c) Partially scanned face.

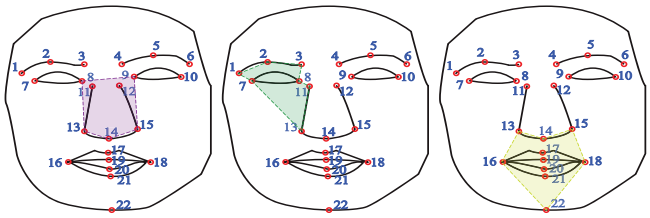
In this paper, a comparison of discriminatory scores of different facial regions is drawn from 3D point cloud data (PCD). A few examples of facial surface represented by PCD is shown in Fig. 1. It can be seen from this figure that the extraction of facial regions becomes challenging due to occlusions or pose variations. Hence, a rigid registration algorithm called the iterative closest point (ICP) is used for the alignment of the surfaces of extracted facial regions represented by PCD. The similarity metric is defined based on alignment of the corresponding unit normal vectors on surface PCD of face regions. Finally, the intraclass correlation coefficient (ICC) of the similarity scores is estimated to obtain the discrimination score of different regions of facial point clouds.

The paper is segmented into several sections. The proposed approach of evaluation of discrimination power is described in Section II. Section III represents the experimental setup, the results obtained, and the validation of the results by comparing with previous works. Finally, Section IV provides conclusion.

II. Proposed Approach of Evaluation

The proposed approach of evaluation follows different steps including the definition of facial regions, segmentation of regions, feature selection, registration, and calculation of discrimination scores. In order to improve the readability of this paper, these steps are described in different subsections.

978-1-5090-2963-1/16/$31.00 ©2016 IEEE



(a) (b) (c)

Fig. 2: Landmarks obtained by algorithm [10], and polygons formed by connecting the landmarks for the three regions, namely, (a) nose, (b) eye, and (c) mouth.

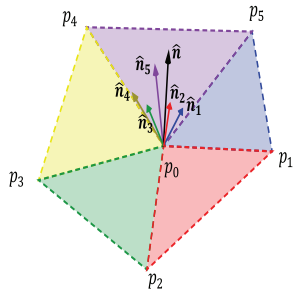


Fig. 3: Normal vectors using local Voronoi mesh for $K = 5$.

A. Defining Different Facial Regions

To compare the discrimination capability of different facial parts, suitable regions on faces are automatically generated from landmarks. In particular, we are inspired by the automatic localization algorithm of landmarks on faces reported in [10]. As we are focused in PCD, the regions are selected by segmenting the facial surface according to certain polygon connecting the landmarks. Specifically, we define three regions, namely, *nose*, *eye*, and *mouth* in terms of 2D polygons formed by connecting certain landmarks, viz., [8, 9, 15, 14, 13], [1, 2, 3, 8, 11, 13, 7] and [12, 14, 15, 18, 21, 16], respectively, as shown in Fig. 2. To select the 2D polygons, the third component of the PCDs representing the vertices are ignored.

B. Segmentation of Facial Surface Represented by PCD

After defining the regions in terms of polygons, it is necessary to segment the surface accordingly. To do so every point in the facial point cloud surface has been checked whether the point belongs to the entire set of points that are within the convex hull of the polygon vertices. Let the total number of points in the PCD of a region be N. First, the polygons are discomposed into all possible triangles. Then, the query point q_i $(i \in \{1, 2, 3, \cdots, N\})$ is checked whether it is within such a triangle. The checking process of a point whether it is within a triangle is performed by using the following steps:

1) The vertices l_1, l_2 and l_3 of a triangle is arranged in anti-clockwise order.
2) It is checked whether the point sets $[l_1 l_2 q_i]$, $[l_2 l_3 q_i]$ and $[l_3 l_1 q_i]$ are arranged in anti-clockwise order. This can be done by determining the area of the triangle formed by each point sets. If the area is positive then their order is anti-clockwise. For example, if (x_1, y_1, z_1), (x_2, y_2, z_2) and (x_3, y_3, z_3) are three corners of a triangle, then the area can be determined as

$$A = \frac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} \tag{1}$$

3) If all these query turned out to be true, then the point q_i is considered to be inside the triangle.

Finally, we take the union of the point sets that are inside the triangles compounding the polygon that forms the face region.

C. Feature of Face Regions

In this paper, the unit normal vector of the 3D point cloud surface has been chosen as facial features considering that the angle between two unit normal vectors can indicate effective discriminatory power. There are several ways through which normal vector on a surface at a given point can be computed. We determine the normal vector at point p_0 by taking the weighted average of the normal vectors on the local mesh triangles of the point. These local Voronoi mesh triangles are found by 2D Delaunay triangulation [11] ignoring the third components of all the points of PCD. Unit normal vector on a surface on specific point can be calculated easily if all the neighboring points of the normal vector can be identified. The local Voronoi mesh can be formed by using these neighboring points. In the mesh, only those triangles that have at least one vertex at point of interest p_0 are counted. All the vertices of the triangles except p_0 are the neighboring points. These neighboring points together compose the local Voronoi mesh. For example, K neighboring points will result in K triangles with common vertices at p_0 as shown in Fig. 3. Let p_i be the ith neighboring point of p_0 and $\overrightarrow{n_i}$ be the normal on ith triangle $(i \in \{1, 2, 3, \cdots, K\})$. Then the weighted average of the unit normal vector $\overrightarrow{n}$ can be computed as

$$\overrightarrow{n} = \sum_{i=1}^{K} w_i \overrightarrow{n_i} \Big/ \sum_{i=1}^{K} w_i \tag{2}$$

where $w_i = \dfrac{|\overrightarrow{p_0 p_i} \times \overrightarrow{p_0 p_{i+1}}|}{2}$ and $w_i \overrightarrow{n_i} = \dfrac{\overrightarrow{p_0 p_i} \times \overrightarrow{p_0 p_{i+1}}}{2}$

D. Registration of Facial Regions

Since the proposed evaluation method incorporates the angle between two normal vectors originated from two different PCD of a face region, the result is susceptible to misalignments of imaging conditions. This problem is tackled by using a suitable registration algorithm. There are many rigid and non-rigid registration algorithms for 3D shapes. In this paper, we implemented the ICP-based [12] rigid registration algorithm to remove the effect of scanning angle. The output of the ICP algorithm provides a 3×3 rotational matrix $\mathbf{R}$ and a 1×3 translation vector $\mathbf{T}$ for the transformation of one point cloud to align to the another after every iteration by minimizing the error metric E given by [12]

$$E = \frac{1}{2} \sum_{i=1}^{\Gamma_r} ||P_i \mathbf{R} + \mathbf{T} - Q_i||^2 \tag{3}$$

where P_i and Q_i are the corresponding vectors formed from ith corresponding points of two different PCDs and Γ_r is the number of points in a PCD. Let $\mathbf{L}$ be the misaligned PCD of size $N \times 3$ representing a facial surface. Then, the registered PCD $\mathbf{M}$ can be found as [12]

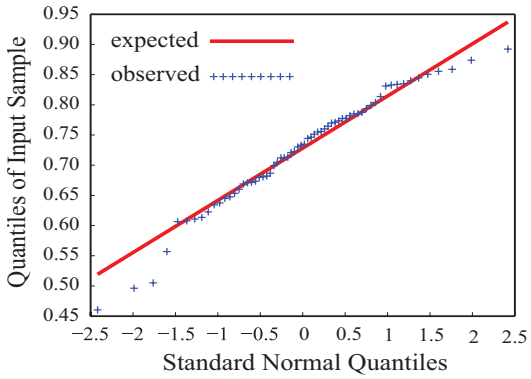


Fig. 4: Sample quantile-quantile plot of similarity metric γ_r.

$$\mathbf{M} = \mathbf{L}\hat{\mathbf{R}} + \mathbf{1}_N\hat{\mathbf{T}} \tag{4}$$

where $\hat{\mathbf{R}}$ and $\hat{\mathbf{T}}$ are obtained by minimizing the error metric of the ICP and $\mathbf{1}_N$ is a row vector of length N with all elements being one.

E. Discrimination Score of Facial Regions

We define similarity metric in terms of the cosine of the angles between the corresponding unit normal vectors of two point clouds. Let there be two point clouds $\mathbf{J}$ and $\mathbf{M}$ representing same region of two faces. If the cloud $\mathbf{J}$ has Γ_r number of points, then there will be equal number of normal vectors for each of the points of a region r ($r \in \{$ *eyes*, *nose*, *mouth* $\}$). Since the clouds are registered, a set of nearest neighbor of Γ_r points can be established between the clouds $\mathbf{J}$ and $\mathbf{M}$. The nearest neighbor is determined by the popular kD-tree search algorithm [13]. Then the similarity metric of a region r in terms of dot product of unit normal vectors $\overrightarrow{n}_J(m)$ and $\overrightarrow{n}_M(m)$ on mth point of the clouds $\mathbf{J}$ and $\mathbf{M}$ can be estimated as

$$\gamma_r = \frac{1}{\Gamma_r}\sum_{m=1}^{\Gamma_r} \overrightarrow{n}_{\mathbf{J}}(m) \cdot \overrightarrow{n}_{\mathbf{M}}(m) \tag{5}$$

Fig. 4 shows that typical quantile-quantile plot of γ_r is almost a straight line, which means that γ_r follows closely a Gaussian distribution. Thus, the discrimination score of a region can be obtained in terms of the ICC of similarity metric γ_r given by [14]

$$\rho_r = \frac{(\phi_{br} - \phi_{wr})}{\phi_{br} + (1-\lambda)\phi_{wr}} \tag{6}$$

where ϕ_{br} and ϕ_{wr} represent the 'between-class mean-square' and 'within-class mean-square', respectively, and are given by

$$\phi_{br} = \frac{\lambda \sum_{\ell=1}^{K} (\mu_r^\ell - \mu_r)^2}{K-1} \tag{7}$$

$$\phi_{wr} = \frac{\sum_{\ell=1}^{K}\sum_{k=1}^{\lambda} (\gamma_r^\ell(k) - \mu_r^\ell)^2}{K(\lambda-1)} \tag{8}$$

where μ_r is the overall mean of the similarity scores considering all the subjects, μ_r^ℓ being the mean of the scores within the class label ℓ, λ is the number of samples for each subject, and K is the total number of subjects.

III. Experiments and Results

To determine the rank of discriminatory facial parts experiments are conducted on database having faces represented by the PCD. The description of the data set of faces, experimental setup, and the results obtained are given in separate subsections.

A. Database

The PCDs of 3D faces from Bosphorus database [15] are used in the experiments. The database contains faces of 104 subjects. There are faces for each subject with variations in terms of expressions, orientation, and occlusion. In order to evaluate the discrimination power of facial regions, we selected only the front view faces that are free of occlusion. The generic dataset of the experiment considers 8 subjects with 8 point clouds per subject chosen randomly from the database. The 24 landmarks provided by the database [10] are used to segment the regions of 3D faces.

B. Setup

The background points were eliminated by taking only the points that have coordinate values greater than a threshold of -2×10^8. In the implementation of ICP algorithm, the maximum number of iteration was chosen as 100, the parameter "extrapolation" was set true to evaluate the direction of iteration, and other parameters were kept same as prescribed in [16]. The kD-tree algorithm was implemented by carrying out search for all the nearest points. The 2D Delaunay triangulation was applied on the XY plane ignoring the third dimension of the points, while estimating the unit normal vector.

C. Results

Table I shows the estimated between- and within-class mean-squares, and discrimination scores ρ_r for different regions of facial surface PCD. It can be seen from this table that the region *eye* has the greatest score among the facial parts. The region *nose* has the second most discrimination score, and the lowest score is shown by the *mouth*. Table I also shows that the region *nose* has the lowest between-and within-class mean-squares, thus indicating the less variability in different faces in nasal region. In other words, the nasal region of faces is immune to expression, the primary cause of distortion in the faces of dataset. In case of region *mouth*, though the value of between-class mean-square is high, the within-class mean-square is also high. This fact indicates that the region *mouth* has the lowest discrimination score in order to distinguish faces of the subjects. Such a result is expected due to the fact that in practice the region *mouth* suffers the most for facial deformations in the presence of expressions. The moderate levels of between- and within-class mean-squares of the region *eye* signify that this region is the best for distinguishing faces as compared to other regions even in the presence of facial distortions due to expressions.

We compare the results obtained in the experiments for the regions of 3D facial surface with that reported for the 2D intensity images of faces [5]. Fig. 5 shows the identification performance of the regions *eye*, *nose*, and *mouth* when the well-known features such as the Eigenface and Adaboost are used. It is seen from this figure that the identification rate from region *eye* is the highest as compared to other regions. These

TABLE I: Discrimination score of regions of 3D facial PCD.

Regions	Between-class mean-square ϕ_{br}	Within-class mean-square ϕ_{wr}	Discrimination score ρ_r
Eye	0.248368	0.008972	0.4879
Nose	0.181765	0.007131	0.4666
Mouth	0.326715	0.016726	0.3983

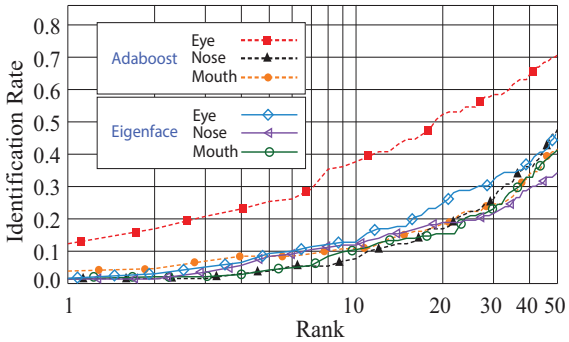


Fig. 5: Identification performance of well-known features of facial regions obtained from 2D intensity images [5].

results ensure that the discrimination power evaluated in this paper for the point cloud of face regions are consistent with the previously reported observations. One question may arise here as to why the discrimination scores of facial point clouds are not so high. A low-level of discrimination score can be explained by the fact that it is really difficult to recognize a person only with one region of face. In practice, the rank - 1 identification rates of facial region-based approaches are less than 25% [5]. Higher identification rates can be attained by fusing or accommodating the interrelation of the features of different facial regions. In few instances of experiments, it is observed that the ICP-based registration process converges towards one of the local minima instead of the global minima. Such registration errors can lead to a low-level discrimination power of facial regions. Fig. 6 shows the examples of incorrect and correct registrations along with their estimated similarity metrics for the facial region *eye* of same and different persons. It is recommended that high cares should have been taken on the registration process for facial region-based recognition, especially in the presence of expression variability.

IV. Conclusion

In this paper, the discrimination powers of regions of a face have been evaluated using the facial surface represented by 3D PCD. Three regions of point clouds, namely, *eye*, *nose*, and *mouth* have been segmented by using the automatically generated landmarks. To compare two point clouds of a facial region, the ICP-based rigid registration algorithm has been employed. The similarity between two PCD is estimated in terms of cosine angles of weighted average of unit normal vectors on 3D facial surface as feature. According to the statistics of similarity metric, the ICC-based scoring has been used to evaluate the significance of facial regions for discriminating the subjects. Experimental results on regions of facial point clouds reveal that the ranking of the facial regions according to their discrimination capability are *eye*, *nose* and *mouth*. This result is consistent with the observations of region-based face recognition conducted on the 2D intensity images of faces.

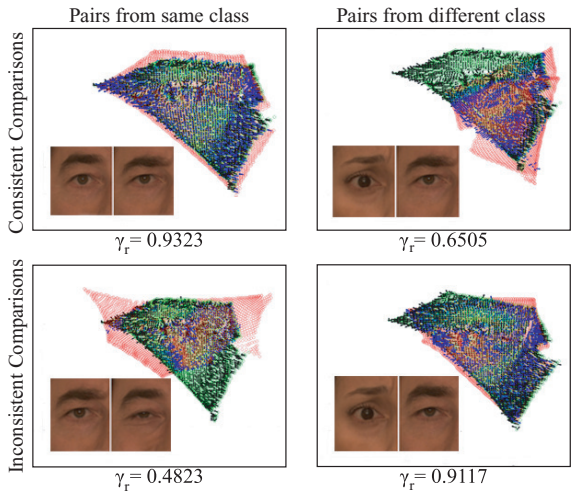


Fig. 6: Examples of correct and incorrect registration for the segmented region *eye* of same and different persons.

References

[1] R. Patel and S. B. Yagnik, "A literature survey on face recognition techniques," *Int. J. Computer Trends and Technology*, vol. 5, no. 4, pp. 189–194, 2013.

[2] S. S. Ngambeki, X. Ding, and M. D. Nachipyangu, "Real time face recognition using region-based segmentation algorithm," in *Int. J. Engineering Research and Technology*, vol. 4, no. 4, 2015, pp. 875–878.

[3] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037–2041, 2006.

[4] S. Yang, P. Luo, C. C. Loy, and X. Tang, "From facial parts responses to face detection: A deep learning approach," in *Proc. IEEE. Int. Conf. Computer Vision*, Santiago, Chile, 2015, pp. 3676–3684.

[5] T. Ali, P. Tome, J. Fierrez, R. Vera-Rodriguez, L. J. Spreeuwers, and R. N. J. Veldhuis, "A study of identification performance of facial regions from CCTV images," in *Lecture Notes in Computer Science : Computational Forensics*, 2015, vol. 8915, pp. 83–91.

[6] M. De Marsico, M. Nappi, D. Riccio, and H. Wechsler, "Robust face recognition after plastic surgery using region-based approaches," *Pattern Recognition*, vol. 48, no. 4, pp. 1261–1276, 2015.

[7] A. Li, S. Shan, X. Chen, and W. Gao, "Face recognition based on non-corresponding region matching," in *Proc. IEEE Int. Conf. Computer Vision*, Barcelona, Spain, 2011, pp. 1060–1067.

[8] Y. Lei, M. Bennamoun, M. Hayat, and Y. Guo, "An efficient 3D face recognition approach using local geometrical signatures," *Pattern Recognition*, vol. 47, no. 2, pp. 509–524, 2014.

[9] K. I. Chang, W. Bowyer, and P. J. Flynn, "Multiple nose region matching for 3D face recognition under varying facial expression," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, no. 10, pp. 1695–1700, 2006.

[10] H. Dibeklioğlu, A. A. Salah, and L. Akarun, "3D facial landmarking under expression, pose, and occlusion variations," in *Proc. IEEE Int. Conf. Biometrics: Theory, Applications and Systems*, Arlington, VA, 2008, pp. 1–6.

[11] B. Delaunay, "Sur la sphère vide. A la mémoire de Georges Voronoï," *Bulletin de l'Académie des Sciences de l'URSS. Classe des sciences mathématiques et na*, no. 6, pp. 793–800, 1934.

[12] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–356, 1992.

[13] J. L. Bentley, "Multidimensional binary search trees used for associative searching," *Communications of the ACM*, vol. 18, no. 9, pp. 509–517, 1975.

[14] P. E. Shrout and J. L. Fleiss, "Intraclass correlations: Uses in assessing rater reliability," *Psychological Bulletin*, vol. 86, no. 2, pp. 420–428, 1979.

[15] A. Savran, N. Alyüz, H. Dibeklioğlu, O. Çeliktutan, B. Gökberk, B. Sankur, and L. Akarun, "Bosphorus database for 3D face analysis," in *Lectrure Notes in Computer Scinece : Biometrics and Identity Management*, 2008, vol. 5372, pp. 47–56.

[16] H. M. Kjer and J. Wilm, "Evaluation of surface registration algorithms for PET motion correction," BSc Thesis, Technical University of Denmark, Denmark, 2010.

Third Harmonic Differential Over Current Relay Based Protection System for Stator Ground Fault in Synchronous Generator

Md. Asif Chowdhury[1], Kazi Taimur Reza[2], Md. Didarul Alam and Enamul Basher[3]
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology, Dhaka, Bangladesh
E-mail: [1]asif.chowdhury9663@gmail.com, [2]tanvir10111@gmail.com, [3]enamul_basher@eee.buet.ac.bd

Abstract— **In this paper, a new method is proposed for 100% protection against the stator earth fault in synchronous generator. Conventional ground fault protection schemes can be applied to any generator but it covers at most 95% of generator stator length. For ensuring 100% protection of synchronous generator stator winding, conventional ground fault schemes are combined with our proposed third harmonic differential over current protection scheme. Moreover, an investigation into the differences between first harmonic fault current and third harmonic fault current is performed. By using proper simulation model, earth faults are simulated at various distant points from the generator neutral and using this simulations, first harmonic fault current and third harmonic fault current characteristics are properly distinguished. Finally, the threshold current of our proposed differential third harmonic over current relay is selected based on the results of the simulations.**

Index Terms— **Stator ground fault, third harmonic fault current, third harmonic voltage, protection.**

I. Introduction

Synchronous generator, major source of commercial electrical energy, needs to be operated safe and sound in order to ensure uninterrupted power supply to end line consumers and also to make certain the stability of existing power system. Reliability of the whole power system and quality of the power supply depend on the successful operation of each individual unit connected to the grid. To protect the power system from major black out each unit has to be checked against for a variety of faults. All faults to synchronous generators can be divided as either insulation failures or abnormal running conditions [1], [2]. Mainly mechanical stresses on the stator winding result in stator winding faults in turbo generators. An insulation failure in the stator winding can result in 3 types of fault: an inter-turn fault, a phase fault or an earth fault, but most common type of fault is the earth fault because most insulation failures eventually connect the winding with the core [1]. The insulation degradation in the windings and atmospheric influences such us moisture, oil or dirt which settles on the coil surfaces outside the stator slots cause stator ground fault. This fault also leads to electrical tracking discharges in the end winding. Eventually it punctures the ground wall [3].

Usually biased differential protection using electro mechanical, solid state and numerical relays are the most common methods that are used for synchronous generator protection [4]. The conventional schemes consisting of differential protection scheme and stator earth fault detection schemes are found to be slow in clearing the faults. As the voltage developed for faults close to neutral is not sufficient to drive the fault current, these faults can't be detected by earth fault relays. Moreover, the differential relay may be tripped in the absence of any earth fault [10]-[11]. To overcome these limitations several techniques have been developed and proposed.

There are some variations in the process of detection of single line to ground faults depending on the magnitude of the grounding impedance. It can be classified into two groups: - low and high impedance grounding. In case of low impedance grounding, a differential relay can detect and provide protection of only about, 95% of the windings. However, for high impedance grounding, ground faults are not normally detectable by the differential relay because the fault current is, usually, less than the sensitivity of the relay. In such case an overvoltage relay (59GN) connected across the grounding resistor has been used to sense the zero sequence voltage. This over voltage relay ensures 95% stator winding protection.

In order to protect 100% of the stator windings against ground fault, supplemental protection schemes of two different techniques have been adopted. These techniques are based on either injection of a low frequency voltage to the neutral or measuring the third harmonic voltage content [5]. Also, some other methods are proposed such as Kalman filtering method, Fuzzy logic controlled neural network based method, Wavelet transform based method, Independent Component Analysis (ICA) based method and etc [6, 12-14].

Generally, the output voltage of the generator is not a pure sinusoidal, rather it is distorted by harmonic components. The harmonic components are found in triple harmonics as the third harmonic, 9th harmonic, 15th harmonic and so on. The triple harmonics are found in all phases with the same phase angle. That's why the common point is not zero as the triple harmonic components do not add up to zero. The magnitude of the third harmonic voltage is comparatively greater than most others triple harmonics [7]. Normally, it is around 1% to 6% of the rated voltage of the generator. Due to the physical design of the generator, this 3rd harmonic voltage is found in the generator. Using digital filter techniques, modern relay technology is capable of measuring and evaluating these small harmonic signals with distinction and reliability. That's why our proposed method is based on third harmonic component of earth fault current.

In this paper, the first harmonic equivalent circuit for stator ground fault is simulated at different positions of the stator winding of the synchronous generator. Then the third harmonic equivalent circuit is also simulated at different positions of the stator winding of the synchronous generator. Later, the difference between the first harmonic fault current and the third harmonic fault current is analysed and a new method for protection against 100% stator ground fault is proposed. Our proposed method is third harmonic differential

978-1-5090-2963-1/16/$31.00 ©2016 IEEE

over current relay and it takes into account the significantly different shape (or characteristics) of third harmonic fault current and first harmonic fault current.

II. NON-FAULT AND FAULTED MODEL OF SYNCHRONOUS GENERATOR FOR THIRD HARMONIC VOLTAGE

Here two models are shown which were developed to represent the stator winding equivalent circuit of a synchronous generator for third harmonic voltage [8]. The first model is the equivalent scheme of the generator under normal condition while the second one is the equivalent scheme of the generator working under faulted conditions.

A. Non-fault model for third harmonic voltage

The equivalent circuit of the synchronous generator during non-fault condition is shown in figure 1.

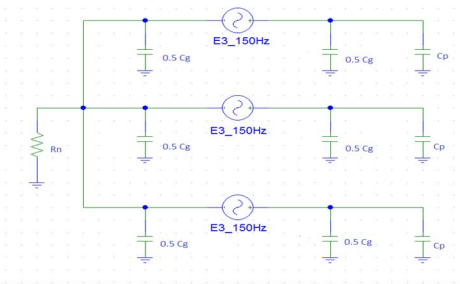


Fig. 1. Equivalent synchronous generator model under non-fault condition

The third harmonic voltage is a zero-sequence quantity. In this model, the third harmonic voltage produced by the synchronous generator is distributed between the terminal and the neutral shunt impedances governing the zero-sequence. In Fig.1, Rn is the resistor between neutral and ground, Cg is the phase capacitance to ground of the generator stator winding, Cp is the total external phase capacitance of the system as seen from the generator, and E3 is the generated third harmonic voltage.

B. Faulted model for third harmonic voltage

The equivalent circuit of synchronous generator for faulted conditions is presented in Fig 2. Here, k is the per unit stator length measured from the neutral to the fault location. The third harmonic voltage produced by the stator winding between the generator neutral and the ground-fault location k is termed as E3n which equals to kE3. The third harmonic voltage between the generator terminal and the ground fault location k is termed as E3t which equals to (1-k)E3.

Moreover, Cn is the phase capacitances to ground of the generator stator winding between the ground-fault location k and the generator neutral and it equals to Cg. The phase capacitances to ground of the stator winding between the generator terminal and the ground-fault location k is termed as Ct which is equivalent to (1-k)Cg.

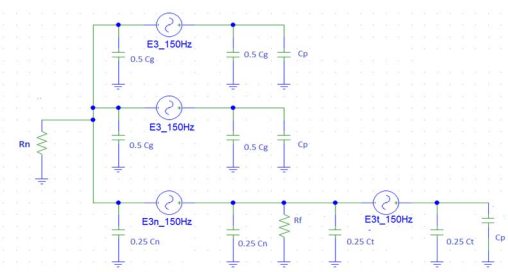


Fig. 2. Equivalent synchronous generator model working under faulted condition

The circuit in Fig. 2 can be solved for finding third harmonic fault current If and third harmonic fault voltage Vf. By applying superposition theorem to all the voltage sources, we can get the equation of Vf.

$$Vf=\frac{2ERRf\omega(0.5Cg+Cp)}{\omega RRf(2.5Cg+0.5kCg+2Cp)-J(R+Rf)}+\frac{kER(1+J\omega 0.5kCgRf)}{R+Rf+J\omega RRf(2Cg+kCg+2Cp)}+\frac{(1-k)ERRf\omega(0.5Cg-0.5kCg+Cp)}{RRf\omega(3Cp+3Cg)-J(R+Rf)} \quad \ldots\ldots\ldots\ldots (1)$$

By dividing the third harmonic fault voltage (Vf) with fault resistance (Rf) , we can find the third harmonic fault current (If).

$$If=Vf/Rf \quad \ldots\ldots\ldots\ldots (2)$$

Here, Vf= Third harmonic Fault voltage

If= Third harmonic Fault current

Cg= Ground side capacitance.

Cp=Load side capacitance.

R= neutral resistance.

Rf= fault resistance.

III. SIMULATION OF INTERNAL FAULT IN SYNCHRONOUS GENERATOR

A power system network consisting of a synchronous generator and a three phase load is chosen for fault simulation and is shown in Fig. 3. Article [9] provides a generator model for internal fault simulation. Using this approach, the generator internal fault is simulated by the addition of two series voltage sources of reverse polarities to each other, in series with each of the faulty phases at the external terminals. The EMF of each of these voltage sources is equal to the EMF E of the sound portion of the windings. Each of the generator sub-transient, transient and synchronous reactances are reduced by a value of x which is equal to the sub-transient reactance of the sound portion of the windings. A reactance x is added externally with each of the three phases for the compensation of the previous reduction. This is illustrated for an internal earth fault as shown in Fig. 4.

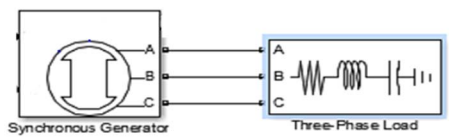


Fig.3. Simulated system model

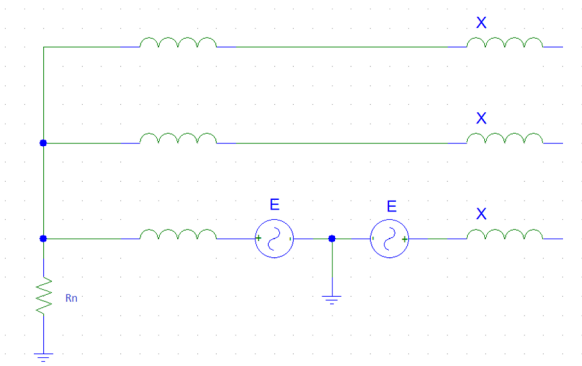


Fig.4. Internal fault equivalent circuit

IV. RESULT

The generator ratings that have been used for simulation purpose in listed in Table I.

TABLE I. GENERATOR RATINGS USED IN SIMULATION

Rated power	850 MVA
Rated rotational speed	3000 rpm
Rated frequency	50 Hz
Rotor type	Round rotor
Nominal voltage (Un)	14.85 kV
Synchronous reactance (Xd)	2.44 p.u.
Transient reactance (Xd')	0.43 p.u.
Subtransient reactance (Xd")	0.25
Zero sequence inductance (X0)	0.13
Negative sequence inductance (X2)	0.24
Zero sequence resistance (R0)	0.0025
Positive sequence resistance (R1)	0.0034
Negative sequence resistance (R2)	0.04
Capacitance to ground of the	0.385 pF
Power factor	0.85
Grounding resistor (Rn)	Rn=1212 ohms
Bus capacitance (per phase)	0.1 pF/phase
Surge capacitor between step-up transformer and circuit breakers capacitance (per phase)	0.25 pF/phase
Step-up transformer capacitance (per phase)	0.2 pF/phase

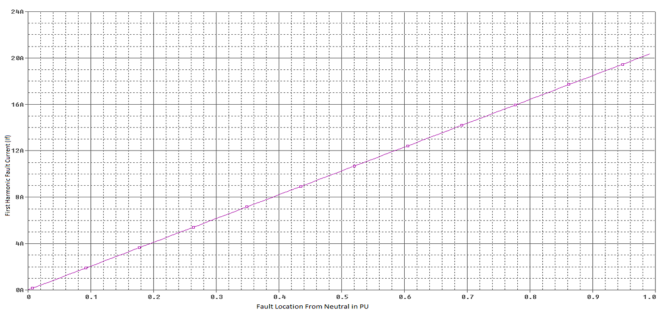


Fig. 5. First harmonic fault current vs. fault location from neutral to terminal.

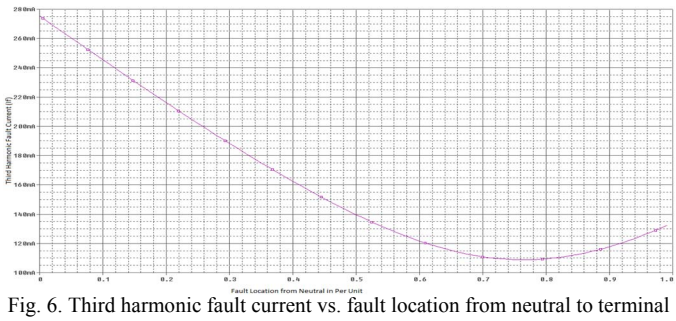


Fig. 6. Third harmonic fault current vs. fault location from neutral to terminal under full load condition.

By simulating the internal fault equivalent circuit model shown in Fig.4 in Orcad PSpice software, the curve in Fig. 5 can be attained. This curve in Fig.5 shows the variation of first harmonic fault current for different fault locations. This curve clearly demonstrates that first harmonic fault current gradually increases from fault locations at neutral to fault locations at terminal.

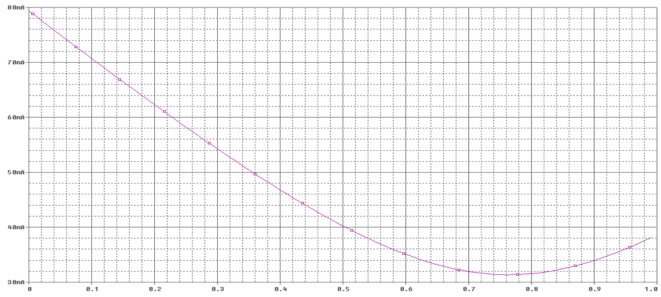


Fig. 7. Third harmonic fault current vs. fault location from neutral to terminal under light load condition.

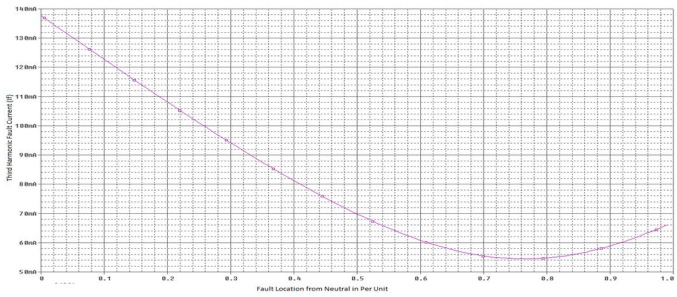


Fig.8. Third harmonic fault current vs. fault location from neutral to terminal under no load condition.

In the case of third harmonic voltage, three load conditions need to be considered [15]. These are full load condition, light load condition and no load condition. For simulating third harmonic equivalent synchronous generator model under faulted conditions, each of these three types of loading conditions is considered. Under full load condition, light load condition and no load condition, we will get the third harmonic current vs. fault location curve as fig. 6, fig. 7, and fig. 8 respectively. From these three figures, it is clear that third harmonic fault current is high near neutral (0-10%) and low in other fault location.

By comparing Fig. 5 and Fig. 6, it is clear that first harmonic fault current is very low near neutral whereas third harmonic fault current is high in that region. For example, for fault location at 10% of the winding near neutral with fault resistance, Rf =1kΩ, we can find that first harmonic fault current is 2.05A. On the other hand, under full load condition, third harmonic fault current at 10% of the stator winding is 245.4 mA. Table II demonstrates the comparison between first harmonic and third harmonic fault current when Rf= 1kΩ.

TABLE II. COMPARISON BETWEEN THIRD HARMONIC AND FIRST HARMONIC FAULT CURRENT

Fault Location from Neutral in Percentage	Third Harmonic Fault Current when Rf=1kΩ			First Harmonic Fault Current when Rf=1kΩ
	Full Load	Light Load	No load	
10%	245.4mA	70.7mA	122mA	2.05A

It is clear from TABLE II that the third harmonic fault current is lower for light load condition than other two conditions.

V. PROPOSED METHOD FOR 100% STATOR GROUND FAULT PROTECTION

Usually, differential relay is used for stator ground fault protection and it gives around 85% protection against this fault. Also, neutral over-voltage protection provides around 95% protection. Hence, only 5-10% of the winding is left unprotected.

As we have found from comparing Fig. 5 and fig. 6 that third harmonic fault current is higher near neutral region than other region whereas first harmonic fault current is lower near neutral region than other region. Moreover, from Fig. 6, Fig. 7 and Fig 8, it is clear that fault current is comparatively low for light load situation. So, if we want to set the tripping condition on the basis of third harmonic current, worst case will be light load condition. Now, a conventional differential protection system is used for detecting fault. Hence our proposed method for detecting fault within 10% of the stator winding is to filter out the third harmonic component of fault current (differential current) and use a sufficient value of this current as the threshold current. The threshold current is fixed based on the allowed maximum earth fault resistance (Rf). For instance, using TABLE II , it can be said that if third harmonic over current relay pickup value is set to 70.7mA then it can protect 10% of the stator winding from neutral for maximum Rf= 1kΩ.

The proposed method is implemented in Matlab-simulink. A stator ground fault is started at 0.1s in phase 'a'. Our proposed third harmonic differential over current relay clears this fault. In fig. 9, phase c current (Ic) is shown and it is evident that the fault is cleared at around 0.15s.

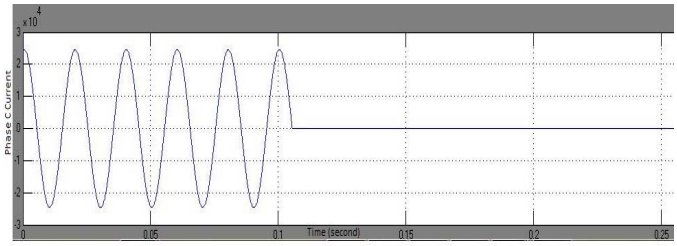


Fig. 9. Phase c current vs. time(s).

VI. CONCLUSION

It is clear that third harmonic fault current gradually decreases as the fault point moves further from neutral and third harmonic fault current is higher for fault locations near neutral (0-10%) than other fault positions. On the other hand, first harmonic current rises as the fault point moves further from neutral. This characteristic is used to develop a new protection algorithm using third harmonic current in combination with first harmonic current. This algorithm uses third harmonic current to detect faults near neutral because the change in third harmonic fault current is significant near neutral. It may use first harmonic fault current to detect faults very distant from neutral.

REFERENCES

[1] Protective Relays Applications Guide. The English Electric Company Limited, Relay Division, Stafford, 1975.

[2] C. J. Mozina. "IEEE Tutorial on the Protection of Synchronous Generators." IEEE Tutorial Course, IEEE Power Engineering Society Special Publ., No. 95 TP102, 1995.

[3] Finney, D., Adamiak, M., Kasztenny, B. (2002), Dynamic testing of generator protection using a model generator platform, 56th Georgia Tech Protective Relaying Conference, pp 1-18.

[4] IEEE Power Engineering Education Committee: IEEE Tutorial on the Protection of Synchronous generators (1995)

[5] H. Ungrad, W. Winker and A. Wiszniewski, "Protection techniques in Electrical Energy Systems", Marcel Dekker, 1995.

[6] N. Yadaiah , N. Ravi, "Fault Detection in Synchronous Generators Based on Independent Component Analysis",Proceedings of the International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) Volume 199 of the series Advances in Intelligent Systems and Computing pp 285-292

[7] SULTAN, A.R. and bin MUSTAFA, M.W., 2013. Ground Fault Protection Methods of a Generator Stator. Przegląd Elektrotechniczny, 89(10), pp.225-229.

[8] Alcántara, R., and F. García. "100% Stator ground fault protection a comparison of two protection methods." Department of Industrial Electrical Engineering and Automation Lund Institute of Technology (2005). pp:20-46

[9] Taalab, A.I., Dawarish, H.A., Kawady, T.A.: ANN - Based Novel Fault Detector for Generator Windings Protection. IEEE Transaction on Power Delivery 14(3), 824-830 (1999)

[10] D. Jones, M.Sc., C. Eng., M.I.E.E., Analysis and Protection of Electrical Power Systems, Pitman Publishing, 1971.

[11] Denis Robertson, "Power system protection", reference manual, reyolle protection, first published in 1982 by Oriel Press Ltd., pp. 104.

[12] Safari-Shad N., and Franklin R.: "Adaptive 100% Stator Ground Fault Protection Based on Third-Harmonic Differential Voltage Scheme", IEEE Transactions on Power Delivery (2015)

[13] Bo, Z.Q., Wang, G.S., Wang, P.Y., Weller, G.: Non-Differential Protection of Geneartor using Fuzzy Neural Networks. In: International Conference on Power System Technology, pp. 1072–1076 (1998)

[14] Gaffor, S.A., Ramana Rao, P.V.: A New Wavelet Based Ground Fault Protection Scheme for Turbo Generators. In: Second International Conference on Industrial and Information Systems, Sri Lanka, pp. 353–355 (2007)

[15] Ilar, M. et al,"Total Generator Ground Fault Protection", Protective Relaying Conference, Georgia Institute of Technology, May 3-4, 1979.