Lecture Notes in Electrical Engineering 911

Amit Dhawan
R. A. Mishra
Karm Veer Arya
Carlos Ruiz Zamarreño *Editors*

Advances in VLSI, Communication, and Signal Processing

Select Proceedings of VCAS 2021

Springer

Lecture Notes in Electrical Engineering

Volume 911

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering—quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Editorial Director (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

USA, Canada

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** This series is indexed by EI Compendex and Scopus databases. ****

Amit Dhawan · R. A. Mishra · Karm Veer Arya ·
Carlos Ruiz Zamarreño
Editors

Advances in VLSI, Communication, and Signal Processing

Select Proceedings of VCAS 2021

Springer

Editors
Amit Dhawan
Electronics and Communication
Motilal Nehru National Institute
of Technology
Allahabad, India

R. A. Mishra
Electronics and Communication
Motilal Nehru National Institute
of Technology
Allahabad, India

Karm Veer Arya
Information and Communication
Technology
ABV-Indian Institute of Information
Technology and Management
Gwalior, India

Carlos Ruiz Zamarreño
Department of EEC Engineering
Universidad Publica de Navarra
Pamplona, Spain

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-19-2630-3 ISBN 978-981-19-2631-0 (eBook)
https://doi.org/10.1007/978-981-19-2631-0

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Contents

About the Editors

Amit Dhawan is Professor and Head of the Department of Electronics and Communication Engineering, MNNIT, India. He received his Bachelors in Electronics and Communication Engineering from Birla Institute of Technology, Mesra, Ranchi, India, and Masters (Control & Instrumentation) and Ph. D. (Electronics and Communication Engineering) from Motilal Nehru National Institute of Technology (MNNIT), Allahabad, Prayagraj, India. His research interest areas are digital signal processing, robust stability, guaranteed cost control, delayed systems, and multidimensional systems. He has published several papers in indexed journals and conference proceedings. He has also served as Reviewer for many international journals and conferences. His biography was listed in Marquis's Who's Who in the World, USA.

R. A. Mishra passed B. Tech. in Electronics Engineering from KNIT, Sultanpur, Avadh University, Faizabad (Uttar Pradesh), in 1989, M.S. degree in Electronic and Control from BITS Pilani, Rajasthan, in 1996, a Ph.D. degree in Electronics and Telecommunication Engineering from Jadavpur University, Kolkata, in 2006. Currently, he is Professor in the Department of Electronics and Communication Engineering at Motilal Nehru National Institute of Technology, Allahabad, India. His area of research is on the topic of device modeling, residue number system-based signal processing, and low-power VLSI design. He has published many research papers in reputed journals and conferences. He has guided eight Ph.D. students, many M.Tech., and B.Tech. students in their thesis and projects.

Karm Veer Arya has received his B.Sc. from Rohilkhand University, Bareilly, India, in 1986. He did his M.E. and Ph.D. from the Indian Institute of Science, Bangalore, India, and the Indian Institute of Technology Kanpur, Kanpur, India, in 1991 and 2007, respectively. He is currently working as Professor at ABV-Indian Institute of Information Technology and Management, Gwalior, India. He has published over 150 research papers in international journals and conferences. He has supervised 11 Ph.D. scholars and 92 Master's students. He has completed several funded R&D projects, and his research interest areas are image processing, biometrics, and information security.

Carlos Ruiz Zamarreño is Professor in the Electrical, Electronic, and Communications Engineering Department at the Public University of Navarra (UPNA), Pamplona-Iruña, Spain. He received an MS degree in Electrical and Electronic Engineering and his Ph.D. in Communications from UPNA, in 2005 and 2009, respectively. He obtained in 2012 a permanent position as Associate Professor at the Electrical, Electronic, and Communications Engineering Department at UPNA. In 2008, he was Visiting Scientist at the Massachusetts Institute of Technology (Boston, MA, USA) under the guidance of Prof. Henry I. Smith. He also performed a postdoctoral research stay at Siemens AG (Munich, Germany) in the group of Prof. Maximilian Fleischer. He is Author of more than 200 scientific journal and conference publications and has participated in 20 different research projects with public and private entities.

Numerical Analysis and Comparison of Mxene Based Plasmonic Sensor

J. B. Maurya and Alka Verma

Abstract In this paper, two dimensional layered Mxene ($Ti_3C_2T_x$) based plasmonic sensor is numerically analyzed and compared with conventional and other plasmonic sensors based on Graphene, MoS_2, and black phosphorene. The angular interrogation method is used to obtain the reflectance curves for all types of plasmonic sensors. Thereafter, the performance of all the plasmonic sensors is obtained in terms of shift in resonance angle, minimum reflection intensity, and beam width at different refractive index of sensing medium. The percent increment in the sensitivity of MoS_2, Mxene, BlackP, and graphene with respect to conventional are; 19.34%, 12.24%, 10.20%, and 4.86%, respectively. Although addition of these nanomaterials increases sensitivity, they degrade the resolution of the sensor. The percent decrement in the resolution of MoS_2, Mxene, BlackP, and graphene with respect to conventional are; 43.45%, 41.23%, 8.15%, and 17.89%, respectively. Although numerical performance of MoS_2 is slightly higher than the Mxene, easier synthesis and higher affinity of Mxene towards biomolecules (single stranded Deoxyribonucleic acid) can replace MoS_2 and other nanomaterials in the biosensing application.

Keywords Sensor · Surface plasmon resonance · Grpahene · MoS_2 · Mxene · Phosphorene · Sensitivity

1 Introduction

The 2D nanomaterials, viz. graphene, MoS_2, black phosphorene (BlackP), Mxenes, etc. have found potential application in different fields due to their extraordinary properties in comparison to bulk materials. These 2D nanomaterials are playing

J. B. Maurya (✉)
Departement of Electronics and Communication Engineering, National Institute of Technology Patna, Patna, Bihar 800005, India
e-mail: jbm.ec@nitp.ac.in

A. Verma
Departement of Electronics Engineering, Institute of Engineering and Rural Technology, Prayagraj, Uttar Pradesh 211002, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_1

vital role in the field of biosensing since last decade. The well-developed conventional non-optical biosensors are polymerase chain reaction (PCR), reverse transcription polymerase chain reaction (RTPCR), enzyme immunoassay (EIA), lateral flow immunoassay (LFIA), serum virus neutralization assay (SVNA), isothermal nucleic acid amplification (INAA), and clustered regularly interspaced short palindromic repeats (CRISPR) [1]. The emerging non-optical biosensors are next generation sequencing (NGS), reverse transcriptase loop-mediated isothermal amplification (RT-LAMP), and recombinase polymerase amplification (RPA) [1]. Among these non-optical biosensors, the PCR and RTPCR are the gold standard and give most accurate results within 2–3 days [1]. Although, PCR and RTPCR are commercially being used but they are limited in terms of sensitivity, selectivity, fast response, etc. as it is seen in the current pandemic due to COVID-19 [1]. Therefore a biosensor is required which can detect the biomolecules with higher accuracy by improving the selectivity, high sensitivity, and fast detection response. These requirements can be fulfilled by optical biosensors viz. colorimetric and refractive index based sensors. The colorimetric sensors are limited in terms of quantification and labelling [2]. In contrast, refractive index based sensors are label free and capable of exact quantification after calibration [3, 4]. The surface plasmon resonance (SPR) sensor is the most explored refractive index based optical sensor. The SPR biosensor has several advantages over non-optical biosensors, viz. higher accuracy, higher sensitivity, real time monitoring, label free, faster, reliable, repeatable, reusable, and requirement of small amount of sample, etc. [5]. Because of these extraordinary properties, SPR sensors can be used in gas sensing, chemical sensing, biosensing, study of biomolecular interactions, etc. [5].

The conventional SPR sensor was first invented in 1968 by two contemporary physicist; E. Kretschmann and A. Otto. The SPR sensor structures developed by them are known as Kretschmann configuration [6] and Otto configuration [7], respectively. The thin film (~50 nm) of plasmonic metals viz. silver, gold, copper, aluminium, etc. are directly deposited on the glass substrate in the Kretschmann configuration, whereas thin air gap has to maintain between substrate and metal in the Otto's configuration. The Otto's configuration suffers from maintaining a thin air gap and loading of sample. Although gold is highest stable among the plasmonic metals, it has poor affinity towards biomolecules and nanomaterials on its surface [8]. In contrast, silver has highest resolution and comparable sensitivity [9].

The conventional SPR sensors have poor affinity towards biomolecules and also suffer from oxidation and corrosion. Therefore, they should be coated by a material which has higher affinity towards the metal as well as towards biomolecules. The 2D nanomaterials have ability to fulfil both of these requirements. Graphene attaches the ssDNA and other biomolecules with the help of π-stacking bonding [10, 11]. MoS_2 binds the ssDNA probe with the help van der Waals force between nucleobases and the basal plane of MoS_2 [12, 13]. The BlackP can be functionalized with cationic polymer poly-L-lysine (PLL) to bind ssDNA on its surface [14–16]. It can also detect NO_2 gas due to its specific absorption in the presence of CO, H_2S, and H_2 [17–19]. The Mxene ($Ti_3C_2T_x$) binds ssDNA after getting functionalized with thereinto, tris(2,2_-bipyridyl)ruthenium(II) (Ru(bpy)$^{32+)}$ [20–22]. Although these nanomaterials increase the sensitivity by increasing the binding of particular biomolecules on their surface, they may degrade other performance parameters of the sensor. As per the best of our knowledge the comparative numerical analysis of these nanomaterials has not accomplished yet. Therefore, we have analysed the role of graphene, MoS_2, BlackP, and Mxene in the SPR sensor based on Kretschmann configuration using silver as a plasmonic metal.

2 Sensor Structure, Principle of Operation, and Mathematical Formulation

The schematic of Kretschmann configured SPR sensor structure is shown in Fig. 1. In this structure, silver thin film is deposited on the base of a triangular prism of BK7 glass. In conventional structure, sensing medium (sample) is in direct contact with this silver thin film. The silver thin film is coated with either of the 2D nanomaterials graphene, MoS_2, BlackP, and Mxene. The refractive indices of prism, silver, graphene, MoS_2, BlackP, Mxene, and sample are 1.5151, $0.0803 + 4.2347i$ [23], $3 + 1.1491i$ [24], $5.08 + 1.1723i$ [25], $3.5 + 0.01i$ [26], $2.38 + 1.33i$ [27], 1.33–1.335, respectively, and thicknesses are semi-infinite, 50 nm, 0.34 nm, 0.65 nm, 0.56 nm, 0.993 nm, and semi-infinite, respectively.

The prism loaded sensor structure is energised by impinging a p-polarized light wave of 633 nm wavelength. This light falls on the prism-metal interface with an incident angle of θ_i. This light generates an evanescent field in the metal which generates surface plasmons wave (SPW) at the metal-sample interface after the incident angle being greater than the critical angle between prism and sample, i.e., θ_i

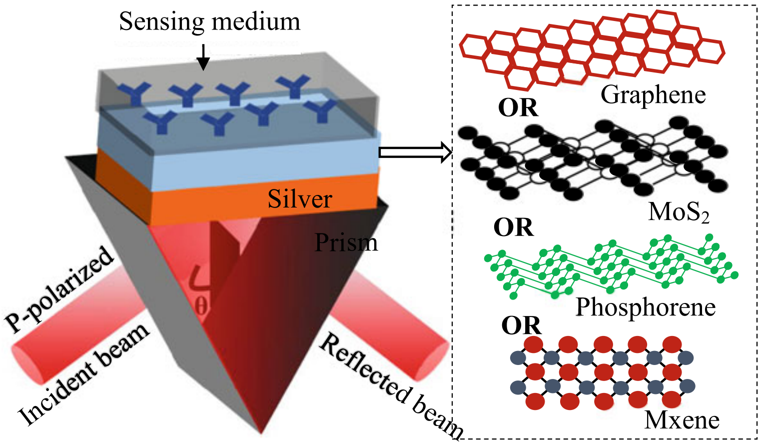


Fig. 1 Schematic of 2D layered nanomaterial based surface plasmon resonance sensor

< $\theta_c = sin^{-1}(n_{sample}/n_{prism})$ [4]. This SPW is generated due to the absorption of incident light in the metal. If $\theta_i < \theta_c$, light will not be absorbed in the metal and hence cannot generate SPW at metal-sample interface. The generated SPW propagates along the metal-sample interface and has evanescent field in the sample and metal in the normal direction to the interface. This evanescent field decays rapidly in the metal layer because of its absorptive nature and penetrates longer in the sensing medium. This evanescent field is highly sensitive to any change in the refractive index of sensing medium if all the other parameters kept constant.

The attenuated total reflected (ATR) light is measured for different incident angle at the opposite face with respect to incident light of the prism. The intensity of this ATR light is plotted with respect to incident angle which is termed as SPR curve or reflectance curve. This reflected curve has a dip at certain value of incident angle. This angle is termed as resonance angle (θ_r). The reflection intensity is minimum at this resonance angle which resembles the maximum absorption of incident light in the metal which ensures the maximum generation of SPW at metal-sample interface. Now, if the refractive index of sample is changed by δn then corresponding change in resonance angle $\delta\theta_r$ is observed. After proper calibration, the quantification of the desired biomolecules is possible by measuring the $\delta\theta_r$ corresponding to δn provided that only desired biomolecules is responsible for the δn.

The reflection intensity of the ATR light can by formulated by Fresnel's coefficients using transfer matrix method for N-layer model [28]. This method is efficient and does not consider any approximation. The thicknesses of the k^{th} layer, d_k, is considered along the normal to the interface (z-axis). The dielectric constant and refractive index of the k^{th} layer is considered as ε_k and n_k, respectively. By applying the boundary condition, the tangential fields at $Z = Z_1 = 0$ are presented in terms of the tangential field at $Z = Z_{N-1}$ as follows;

$$\begin{bmatrix} U_1 \\ V_1 \end{bmatrix} = M \begin{bmatrix} U_{N-1} \\ V_{N-1} \end{bmatrix} \tag{1}$$

where U_1 and V_1 represents the tangential components of electric and magnetic fields, respectively, at the boundary of the first layer and U_{N-1} and V_{N-1} are the corresponding fields for the boundary at N^{th} layer. The M_{ij} presents the characteristics matrix of the combined structure of the sensor, and for p- polarized light it can be given as;

$$M_{ij} = \left(\prod_{k=2}^{N-1} M_k\right)_{ij} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \tag{2}$$

with,

$$M_k = \begin{bmatrix} cos\beta_k & (-isin\beta_k)/q_k \\ -iq_k sin\beta_k & cos\beta_k \end{bmatrix} \tag{3}$$

where,

$$q_k = \left(\frac{\mu_k}{\varepsilon_k}\right)^{1/2} cos\theta_k = \frac{\left(\varepsilon_k - n_1^2 sin^2\theta_1\right)^{1/2}}{\varepsilon_k} \tag{4}$$

and

$$\beta_k = \frac{2\pi}{\lambda} n_k cos\theta_k (z_k - z_{k-1}) = \frac{2\pi d_k}{\lambda}\left(\varepsilon_k - n_1^2 sin^2\theta_1\right)^{1/2} \tag{5}$$

After simplification, one can obtain the reflection coefficient for *p*-polarized light which is given below:

$$r_p = \frac{(M_{11} + M_{12}q_N)q_1 - (M_{21} + M_{22}q_N)}{(M_{11} + M_{12}q_N)q_1 + (M_{21} + M_{22}q_N)} \tag{6}$$

The reflectivity R_p of the defined multilayer structure is given as;

$$R_p = \left|r_p\right|^2 \tag{7}$$

3 Results and Discussion

The comparative analysis of different 2D nanomaterials; graphene, MoS_2, BlackP, and Mxene are done on the basis of reflectance curve at 1.33 and 1.335 refractive index of sensing medium, and variation of performance defining viz. shift in resonance angle ($\delta\theta_r$), minimum reflection intensity (Rmin), and beam width (BW) [29] at different values of refractive index of sensing medium. These performance defining parameters can be used to define main performance parameters viz. sensitivity ($S = \delta\theta_r / \delta n$), resolution ($1/BW$), and quality factor ($Q = S/BW$). For a good SPR sensor, the sensitivity, resolution, and quality factor should be as high as possible. Therefore, shift in resonance angle should be as high as possible for a constant change in refractive index of sensing medium. Further, the beam width should be as low as possible. In addition, the Rmin should be as low as possible which ensures the enhancement in the generation of SPW. Generally, normalized Rmin should be less than 0.2. Other performance parameters which are also important viz. selectivity or specificity but they cannot be determined through numerical analysis.

In Fig. 2, the reflectance curve is plotted for conventional sensor and the sensor with different nanomaterials at 1.33 and 1.335 refractive index of sensing medium. It is observed that the obtained reflectance curves have proper dip type shape. It can be observed that the reflectance dip is obtained at incident angle greater than critical angle between prism and sample, i.e., $\theta_c = sin^{-1}(1.33/1.5151) = 61.3817°$. It is also observed that the Rmin is less than 0.2 for all the reflectance curves which confirms the good generation of SPW. Further, the Rmin for all the reflectance curves are very close to zero except for MoS_2 and Mxene at 1.33 and 1.335 refractive index of sensing

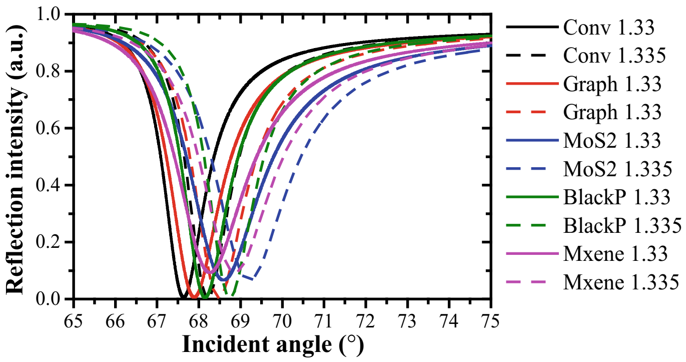
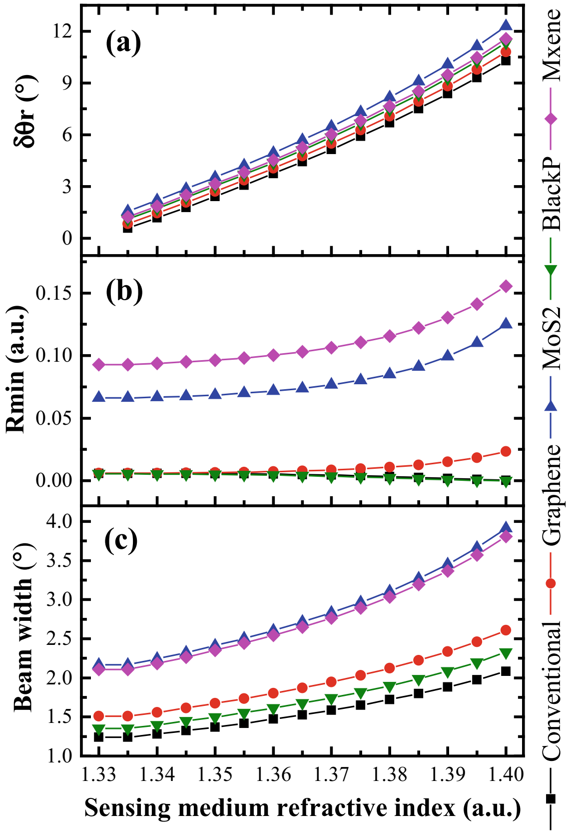


Fig. 2 Variation of normalized reflection intensity with respect to incident angle for different nanomaterials at sensing medium refractive index 1.33 and 1.335

medium. This can be just because of the large thickness and high extinction coefficient of MoS_2 and Mxene [30]. It should be noted that the extinction coefficient of the layer above plasmonic metal is responsible for the damping of surface plasmon [31]. This damping of surface plasmon shallows and broadens the reflectance curve [31]. The damping of surface plasmons increases by increasing the extinction coefficient and thickness of the layer above plasmonic metal. The shallowing and broadening of reflectance curve are responsible for increasing the Rmin and BW, respectively. Therefore, Rmin and BW of reflectance curve of MoS_2 and Mxene are comparatively higher.

Further, in this article, the shift in resonance angle, Rmin, and BW are analysed by varying the refractive index of sensing medium in Fig. 3(a–c), respectively. It can be observed from Fig. 3(a) that the shift in resonance angle increases monotonically with the sensing medium refractive index. Further, the shift in resonance angle is lowest for conventional sensor and highest for MoS_2. The order of shift in resonance angle is; {MoS_2 (12.28) > Mxene (11.55) > BlackP (11.34) > Graphene (10.79) > conventional (10.29)}. Therefore, the percent increment in the sensitivity of MoS_2, Mxene, BlackP, and graphene with respect to conventional are; 19.34%, 12.24%, 10.20%, and 4.86%, respectively. It is very tough to say that shift in resonacne angle follow the order of thickness or refractive index rather it depends on the effective combination of thickness and refractive index of nanomaterials. It can be observed from Fig. 3(b) that Rmin increases monotonicaly with the sesning medium refractive index for all the sesnors except conventional and BlackP. The Rmin for conventional and BlackP is almost constant upto 1.35 beyond which it decreases slightly just because of absence of any layer above metal in the case of conventional and very low extinction coefficient in the case of BlackP. Further, it can be observed that Rmin is lowest for BlackP which is almost equal to conventional and highest for Mxene. The order of Rmin is; {Mxene (0.0928) > MoS_2 (0.0662) > > Graphene (0.0060) > BlackP (0.0056) ≈ conventional (0.0059)}. Further, Rmin for Mxene and MoS_2 are very high with respect to others because of the high thickness and high extinction

Fig. 3 Sensor performance with respect to sensing medium refractive index: **a** Shift in resonance angle, **b** Minimum reflection intensity, **c** Beam width



coefficient. It should be noted that the Rmin is less than 0.2 for considered entire wide range of sensing medium refractive index. It can be observed from Fig. 3(c) that the beam width increases monotonicaly with the sesning medium refractive index for all the sesnors. Further, the beam width is lowest for conventional sensor and highest for MoS_2. The order of beam width is; {MoS_2 (2.17) > Mxene (2.11) > > Graphene (1.51) > BlackP (1.35) > conventional (1.24)}. Although addition of these nanomaterials increases sensitivity, they degrades the resolution of the sensor. The percent decrement in the resolution of MoS_2, Mxene, BlackP, and graphene with respect to conventional are; 43.45%, 41.23%, 8.15%, and 17.89%, respectively.

4 Conclusion

SPR sensor using 2D nanomaterils is studied, anlaysed and compared with conventional SPR sensor. It is found that coating of plasmonic metal with 2D nanomaterials improves the shift in resonace angle which enhances the sensitivity of the sensor but increases the beam width which degrades the resolution and quality factor. The percent increment in the sensitivity of MoS_2, Mxene, BlackP, and graphene with respect to conventional are; 19.34%, 12.24%, 10.20%, and 4.86%, respectively. The percent decrement in the resolution of MoS_2, Mxene, BlackP, and graphene with respect to conventional are; 43.45%, 41.23%, 8.15%, and 17.89%, respectively. Also, nanomaterials increases the Rmin which signifies the loss of surface plasmons inside it due to damping. Further, it is found that Mxene has moderate sensitivty almost equal to BlackP and resolution almost equal to MoS_2. Mxene has highest Rmin which signifies the highest surface plasmon damping inside it due electron energy loss.

References

1. D'Cruz, R.J., Currier, A.W., Sampson, V.B.: Laboratory testing methods for novel severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2). Front. Cell Dev. Biol. **8**, 468 (2020)
2. Yang, T., Luo, Z., Tian, Y., et al.: Design strategies of AuNPs-based nucleic acid colorimetric biosensors. TrAC—Trends Anal. Chem. **124**, 115795 (2020)
3. Chen, Y., Liu, J., Yang, Z., et al.: Optical biosensors based on refractometric sensing schemes: a review. Biosens. Bioelectron. **144**, 111693 (2019)
4. Maurya, J.B., François, A., Prajapati, Y.K.: Two-dimensional layered nanomaterial-based one-dimensional photonic crystal refractive index sensor. Sensors (Switzerland) **18**, 1–7 (2018)
5. Homola, J.: Surface plasmon resonance sensors for detection of chemical and biological species. Chem. Rev. **108**, 462–493 (2008)
6. Kretschmann, E., Raether, H.: Radiative decay of non radiative surface plasmons excited by light. Zeitschrift für Naturforschung A **23**, 2135–2136 (1968)
7. Otto, A.: Excitation of nonradiative surface plasma waves in silver by the method of frustrated total reflection. Zeitschrift für Phys. **216**, 398–410 (1968)
8. Wu, L., Chu, H.S., Koh, W.S., Li, E.P.: Highly sensitive graphene biosensors based on surface plasmon resonance. Opt. Express **18**, 14395–14400 (2010)
9. Maurya, J.B., Prajapati, Y.K.: A comparative study of different metal and prism in the surface plasmon resonance biosensor having MoS_2–graphene. Opt. Quantum Electron. **48**, 280 (2016)
10. Maurya, J.B., Prajapati, Y.K., Singh, V., Saini, J.P.: Sensitivity enhancement of surface plasmon resonance sensor based on graphene–MoS_2 hybrid structure with TiO_2–SiO_2 composite layer. Appl. Phys. A Mater. Sci. Process. **121**, 525–533 (2015)
11. McGaughey, G.B., Gagné, M., Rappé, A.K.: π-stacking interactions. alive and well in proteins. J. Biol. Chem. **273**, 15458–15463 (1998)
12. Maurya, J.B., Prajapati, Y.K., Singh, V., et al.: Improved performance of the surface plasmon resonance biosensor based on graphene or MoS_2 using silicon. Opt. Commun. **359**, 426–434 (2016)
13. Zhu, C., Zeng, Z., Li, H., et al.: Single-layer MoS_2-based nanoprobes for homogeneous detection of biomolecules. J. Am. Chem. Soc. **135**, 5998–6001 (2013)
14. Kumar, V., Brent, J.R., Shorie, M., et al.: Nanostructured Aptamer-functionalized black phosphorus sensing platform for label-free detection of myoglobin, a cardiovascular disease biomarker. ACS Appl. Mat. Inter. **8**, 22860–22868 (2016)
15. Pal, S., Verma, A., Raikwar, S., et al.: Detection of DNA hybridization using graphene-coated black phosphorus surface plasmon resonance sensor. Appl. Phys. **124**, 394 (2018)
16. Pal, S., Verma, A., Prajapati, Y.K., Saini, J.P.: Influence of black phosphorous on performance of surface plasmon resonance biosensor. Opt. Quantum Electron. **49**, 403 (2017)
17. Cui, S., Pu, H., Wells, S.A., et al.: Ultrahigh sensitivity and layer-dependent sensing performance of phosphorene-based gas sensors. Nat. Commun. **6**, 1–9 (2015)

18. Maurya, J.B., Prajapati, Y.K., Raikwar, S., Saini, J.P.: A silicon-black phosphorous based surface plasmon resonance sensor for the detection of NO2 gas. Optik (Stuttg) **160**, 428–433 (2018)
19. Maurya, J.B., Prajapati, Y.K.: Comparative analysis of silicon and black phosphorous as an add-layer in nanomaterial based plasmonic sensor. Opt. Quantum Electron. **51**, 1–11 (2019)
20. Fang, Y., Yang, X., Chen, T., et al.: Two-dimensional titanium carbide (MXene)-based solid-state electrochemiluminescent sensor for label-free single-nucleotide mismatch discrimination in human urine. Sensors Actuators, B Chem. **263**, 400–407 (2018)
21. Kumar, R., Pal, S., Pal, N., et al.: Figure of merit enhancement of Ti3C2Tx-graphene based long-range surface plasmon sensor at telecommunication wavelength. Opt. Quantum Electron. **53**, 218 (2021)
22. Kumar, R., Pal, S., Pal, N., et al.: High-performance bimetallic surface plasmon resonance biochemical sensor using a black phosphorus–MXene hybrid structure. Appl. Phys. A Mater. Sci. Process. 127 (2021)
23. Gupta, B.D., Sharma, A.K.: Sensitivity evaluation of a multi-layered surface plasmon resonance-based fiber optic sensor: a theoretical study. Sensors Actuators B Chem. **107**, 40–46 (2005)
24. Bruna, M., Borini, S.: Optical constants of graphene layers in the visible range. Appl. Phys. Lett. **94**, 031901 (2009)
25. Castellanos-Gomez, A., Agrait, N., Rubio-Bollinger, G.: Optical identification of atomically thin dichalcogenide crystals. Appl. Phys. Lett. **96**, 213116 (2010)
26. Wu, L., Guo, J., Wang, Q., et al.: Sensitivity enhancement by using few-layer black phosphorus-graphene/TMDCs heterostructure in surface plasmon resonance biochemical sensor. Sensors Actuators, B Chem. **249**, 542–548 (2017)
27. Wu, L., You, Q., Shan, Y., et al.: Few-layer Ti3C2Tx MXene: a promising surface plasmon resonance biosensing material to enhance the sensitivity. Sensors Actuators, B Chem. **277**, 210–215 (2018)
28. Yamamoto, M.: Surface plasmon resonance (SPR) theory: tutorial. Rev. Polarogr. **48**, 209–237 (2002)
29. Maurya, J.B., Prajapati, Y.K.: A novel method to calculate beam width of SPR reflectance curve: a comparative analysis. IEEE Sensors Lett. **1**, 1–4 (2017)
30. Maurya, J.B., Prajapati, Y.K.: Influence of dielectric coating on performance of surface plasmon resonance sensor. Plasmonics **12**, 1121–1130 (2017)
31. Maharana, P.K., Jha, R., Palei, S.: Sensitivity enhancement by air mediated graphene multilayer based surface plasmon resonance biosensor for near infrared. Sensors Actuators B Chem. **190**, 494–501 (2014)

A Comparative Performance Analysis of SPR Biosensor Using Metamaterial and Different Metal Oxides

Narendra Pal, Sarika Pal, Yogendra Kumar Prajapati, and J. P. Saini

Abstract We have investigated a comparative performance analysis of surface plasmon resonance (SPR) biosensor on using metamaterial and different metal oxides. Use of 10 nm SiO_2 in the sensor configuration provided the maximum sensitivity of 478.9°/RIU, which is highest among all other metal oxides like ZnO, SiO_2, TiO_2, Al_2O_3, MoO_3 and MgO. Sensor performance is evaluated by calculating SPR sensor performance factors for smaller refractive index variation of sensing medium (1.33–1.34). Sensor performance factors are compared for conventional SPR and proposed sensor with and without using metal oxides. We have used transfer matrix method for this 5-layer Kretschmann configuration based sensor configuration to perform analytical simulations using MATLAB software. Citing exemplary references from literature on metamaterial based SPR sensor for comparison, this work presents very high values of performance factors (Sensitivity = 478.9 °/RIU, detection accuracy = 2.12/°, and quality factor = 141.2 RIU^{-1}) for proposed SPR biosensor on using 10 nm SiO_2 metal oxide.

Keywords Aluminum oxide (Al_2O_3) · Molybdnem trioxide (MoO_3) · Magnesium oxide (MgO) · Metamaterial · Silicon dioxide (SiO_2) · Surface plasmon resonance (SPR) · Sensitivity · Titanium dioxide (TiO_2)

N. Pal (✉)
Department of Electronics and Communication Engineering, Rajkiya Engineering College, Kannauj, Uttar Pradesh 209732, India
e-mail: narendrapal2002@gmail.com

S. Pal
Department of Electronics Engineering, National Institute of Technology, Srinagar, Uttarakhand 246174, India
e-mail: sarikapal@nituk.ac.in

Y. K. Prajapati
Department of Electronics and Communication Engineering, Motilal Nehru National Institute of Technology, Allahabad 211004, India

J. P. Saini
Netaji Subhash University of Technology, New Delhi 110078, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_2

1 Introduction

Surface plasmon resonance (SPR) sensors have employed stupendous attention and popularity among other optical sensors due to its competitive advantage of accurate, low cost, label free, and real time sensing procedures [1, 2]. SPR sensors are tremendously favorable and well suited for sensing of analytes, biomolecules, various chemicals and gases [1]. It rely on phenomenon of refractive index, RI sensing of sensing medium, which get modulated on exposure of biorecognition (BRE) layer to target analyte/biomolecules. The interaction of BRE layer with target analytes alters the sensing medium RI, which can be measured as optical SPR signal in a particular angular range. It is impossible to generate surface plasmons (SPs) on metal surface through direct illumination. So, frequently employed light coupling methods like prism, grating and optical fiber are used for efficient generation of SPs that may help in achieving resonance condition. To achieve SPR condition, wave vector of incident light should be equal to of SPs supported by metal dielectric interface[1]. Photodetector is used to measure reflected intensity after achieving SPR condition. The reflected intensity of p-polarized incident light may be measured at different angle keeping fixed wavelength (angular interrogation) or at different wavelength keeping fixed angle (wavelength interrogation) mode. Angular interrogation for prism coupled SPR sensors employing Kretschmann configuration has already proven its applicability for its simple and realizable geometry [3]. SPR sensors have shown their exceptional potential in context of biosensing e.g. in drug diagnostic, medical, food and environment safety than other optical sensor [4–7]. For biosensing application use of BRE layer in close vicinity to metal binds target analyte kept in aqueous solution. The binding of analyte on BRE layer creates local rise in sensing layer RI, which further modifies propagation constant of SPs and resulting in shifting of dip position of SPR curve [8]. Most popular metals used for SPs generation are mainly gold and silver. Gold shows better stability, biocompatibility for adsorption of analytes with high sensitivity [1]. But, their use in SPR sensor results in wide angular width of SPR curve leading to poor accuracy for detection of analytes. Accurate detection of analyte is very important parameter to evaluate performance of SPR biosensor. Use of silver affords narrower angular width but shows poor chemical stability and also smaller sensitivity [9]. So, using another high RI metal oxide layer over silver can prevent oxidation and improves sensitivity too.

Burgeoning developments in large scale synthesis of 2D nanomaterials and metamaterials have further intensified their implementations in SPR sensor for sensitivity enhancement. Since the discovery of graphene, various 2D layered nanomaterials (transition metal dichalcogenides, antimonene, MXene, and black phosphorus etc.) have been investigated for improvement in performance factors of SPR sensors [8–12]. Metamaterials are also investigated for providing enhanced electromagnetic field at sensing layer interface to get high sensitivity. Moreover, the use of metamaterial in SPR sensor provides better confinement of electromagnetic field by providing stronger localization of field at sensing layer interface. This leads to higher adsorption of biomolecules on BRE layer [9]. Researchers are able to fabricate metamaterials and sensors based on metamaterial too with available advanced fabricating techniques [8, 9, 13]. Firstly, Ishimaru et al. proposed the idea of using metamaterial based SPR sensor for improved performance [14]. Not only 2D nanomaterials, metamaterials high RI metal oxides were also investigated for enhancing SPR sensor performance [15–17]. Metal oxides are known for their wide bandgap, high electron transfer rate, high ionic conductivity and stability in chemical and biosensing applications. Semiconducting metal oxides are engineered nanomaterial that raises the sensitivity of SPR biosensor [15–17]. Recently, our group proposed a SPR biosensor which uses metamaterial and zinc oxide for sensitivity enhancement [18]. Highest sensitivity of 401°/RIU has been achieved on using 5 nm ZnO for the proposed work. But, we have not investigated the comparison of the proposed SPR sensor on using other metal oxides in place of ZnO. So, in this paper we presented a comparative performance of SPR biosensor based on metamaterial on using other metal oxides like Al_2O_3, SiO_2, TiO_2, MoO_3 and MgO. The motivation behind this work is to improve the sensitivity and quality factor of SPR sensor by utilizing the stronger light matter interaction provided with the use of metamaterial for biosensing and to compare the performance of proposed sensor for different metal oxides.

2 Numerical Modeling

2.1 Performance Parameters and Transfer Matrix Method

Following three factors namely S-sensitivity; DA-detection accuracy; and QF-quality factor indicates performance factor of SPR biosensors [9].

$$\mathrm{S} = \Delta\theta_{\mathrm{SPR}}/\Delta \mathrm{n_s} \tag{1}$$

$$\mathrm{DA} = \Delta\theta_{\mathrm{SPR}}/\mathrm{FWHM} \tag{2}$$

where $\Delta\theta_{SPR}$, Δn_s, and FWHM represents SPR angle shift, sensing medium RI shift and angular width of the SPR curve respectively. The FWHM is difference of incident angle measured at 50% reflected intensity, which is evaluated from SPR curve [9]

$$\mathrm{QF} = \mathrm{S/FWHM}\left(\mathrm{RIU}^{-1}\right) \tag{3}$$

SPR curve is plotted between reflection intensity vs. angle, where reflection intensity is calculated using transfer matrix method for 5-layer model. Transfer matrix method, an accurate method without any approximation is extensively used here for proposed 5 layer model [9]. Each k^{th} layer is considered to be a homogenous, nonmagnetic and isotropic media. Thus, reflectivity can be represented as:

$$R_p = |r_p|^2 \tag{4}$$

and Fresnel's reflection coefficient r_p, may be represented as:

$$r_p = \frac{(M_{11} + M_{12}q_5)q_1 - (M_{21} + M_{22}q_5)}{(M_{11} + M_{12}q_5)q_1 + (M_{21} + M_{22}q_5)} \tag{5}$$

where M_{11}, M_{12}, M_{21}, and M_{22} are elements of character matrix M, which is represented as

$$M = \prod_{k=2}^{N-1} M_k = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \tag{6}$$

$$\text{where, } M_k = \begin{bmatrix} \mathrm{Cos}\beta_\mathrm{k} & \frac{-\mathrm{iSin}\beta_\mathrm{k}}{\mathrm{q_k}} \\ -\mathrm{iq_k Sin}\beta_\mathrm{k} & \mathrm{Cos}\beta_\mathrm{k} \end{bmatrix} \tag{7}$$

k varies from 1-N and β_k tells the phase of electromagnetic radiation at k^{th} layer.

$$\beta_\mathrm{k} = \frac{2\pi}{\lambda} \mathrm{n_k Cos}\theta_\mathrm{k}(\mathrm{z_k} - \mathrm{z_{k-1}}) = \frac{2\pi \mathrm{d_k}}{\lambda}\left(\varepsilon_\mathrm{k} - \mathrm{n_1^2 Sin^2}\theta_1\right)^{1/2} \tag{8}$$

$$\mathrm{q_k} = \left(\frac{\mu_\mathrm{k}}{\varepsilon_\mathrm{k}}\right)^{1/2} \mathrm{Cos}\theta_\mathrm{k} = \frac{\left(\varepsilon_\mathrm{k} - \mathrm{n_1^2 Sin^2}\theta_1\right)^{1/2}}{\varepsilon_\mathrm{k}} \tag{9}$$

2.2 Schematic Structure of the Proposed SPR Biosensor

Figure 1 illustrates the schematic structure of proposed SPR biosensor which works at 1000 nm characteristic wavelength. Different layers of the schematic structures are arranged as: BK-7 prism-Silver-Metal oxide-Metamaterial- and Sensing region.

The RI of BK-7 prism considered to be 1.5075 at 1000 nm [9, 18]. The wavelength dependent RI of silver is calculated using Drude model given as

$$n_{Ag}(\lambda) = \left[1 - \frac{\lambda^2 \lambda_c}{\lambda_p^2(\lambda_c + i\lambda)}\right]^{1/2} \tag{10}$$

The λ_c=1.7614 × 10^{-5} m and λ_p=1.4541 × 10^{-7} m are considered for silver, where λ_c, λ_p are measured in m and λ in nm [18]. Then, high RI metal oxide layer having 10 nm thickness is laid over silver to protect it from oxidation and also for sensitivity improvement. The RIs of metal oxide layers Al_2O_3, SiO_2, TiO_2, MoO_3 and MgO are 1.7557, 1.4504, 2.4856, 2.0838 + i0.0028793 and 1.7228 respectively

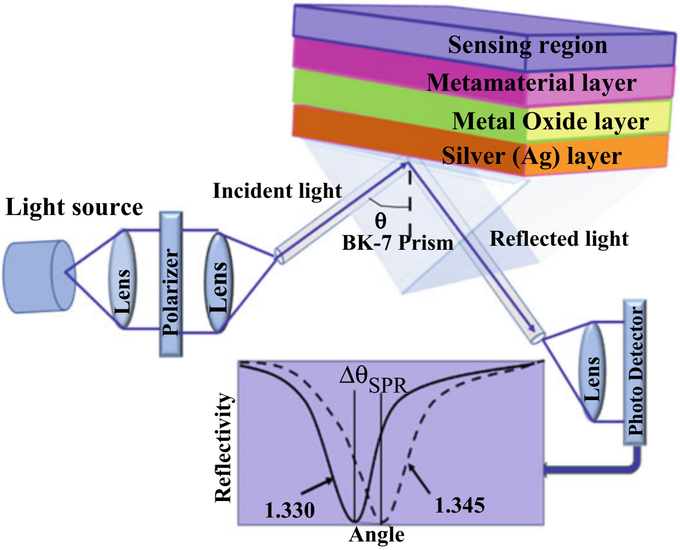


Fig. 1 Schematic structure

[19]. The electric permittivity, ε_{meta} and magnetic permeability, μ_{meta} considered for metamaterial are [–4 + i0.001] and [–2.4 + i0.001] respectively [9, 18]. RI of metamaterial (n_{meta}) can be calculated from following relation, $n_{Meta} = \sqrt{\varepsilon_{Meta} \times \mu_{Meta}}$. Thereafter, 50 nm binding layer and RI = 1.462 is laid over metamaterial for binding of analyte present in aqueous solution [18]. Sensing region containing aqueous solution having 1.33 RI is considered 5th layer of the schematic diagram. Local increase in sensing medium RI, 1.33 + Δn_s is obtained after adsorption of target analyte, where Δn_s tells change in sensing medium RI.

3 Results and Analysis

This section is categorized into three sub-sections further. Section 1 illustrates about thickness optimization of silver layer and metamaterial layer for the proposed SPR sensor configuration. Section 2 discusses the SPR curve analysis plotted for conventional SPR, proposed SPR with and without using metal oxide layer. Finally, Sect. 3 illustrates the sensor performance factor variation with sensing medium RI (1.33–1.34).

3.1 Thickness Optimization of Metal and Metamaterial Layer

Minimum reflectance, $R_{min.}$ is very important parameter for optimization of metal layer thickness, as its minimum value reflects the narrower angular width of the SPR curve. We adopted two thicknesses of metal oxide layers 5 nm and 10 nm and 350 nm metamaterial thickness to optimize silver layer thickness in terms of minimum value of minimum reflectance, R_{min}. for proposed SPR biosensor configuration. Figure 2(a) and (b) illustrates the optimization of silver layer thickness, when different metal oxides are used in proposed design. A clear observation from Fig. 2a about the minimum value of obtained $R_{min.}$ are at 42 nm, 42 nm, 42 nm, 42 nm, 43 nm, and 42 nm respectively of silver for ZnO, Al_2O_3, SiO_2, TiO_2, MoO_3 and MgO at 5 nm thickness.

Similarly, Fig. 2b shows minimum value of R_{min}. for 10 nm metal oxide layers Al_2O_3, SiO_2, TiO_2, MoO_3 and MgO at 42 nm, 42 nm, 41 nm 43 nm and 42 nm of optimized silver thickness respectively. Metamaterial thickness is optimized in terms of maximum sensitivity attained for proposed SPR using 5 nm and 10 nm different metal oxides at optimized silver layer thickness. Figure 3(a) and (b) shows optimization of metamaterial thickness at 5 nm and 10 nm metal oxide layer thicknesses respectively. Figure 3a gives maximum sensitivity of proposed SPR at metamaterial thicknesses 351 nm, 347 nm, 348 nm, 346 nm, 349 nm 347 nm respectively for 5 nm ZnO, Al_2O_3, SiO_2, TiO_2, MoO_3 and MgO. Similarly, Fig. 3b tells optimized metamaterial thicknesses of 346 nm, 348 nm, 345 nm, 350 nm, 346 nm respectively for 10 nm Al_2O_3, SiO_2, TiO_2, MoO_3 and MgO at which maximum sensitivities are obtained. The optimized metamaterial and silver layer thickness will lead to obtain maximum sensitivity and better detection accuracy at 5 nm and 10 nm metal oxide layer thicknesses.

3.2 SPR Curve Analysis

Performance factors of proposed SPR biosensor with and without (w/o) metal oxides and conventional SPR biosensor are discussed in this sub-section. SPR curve sketched in Fig. 4 tells higher resonance angle shifts ($\Delta\theta_{SPR}$) achieved for proposed SPR w/o metal oxides than conventional SPR sensor. Higher $\Delta\theta_{SPR}$ achieved for proposed SPR biosensor w/o metal oxides is due to use of metamaterial which enhances the electromagnetic field at sensing layer interface. All performance factors mentioned in modeling section are calculated from Fig. 4 for proposed SPR w/o metal oxides and conventional SPR. Angular width of the proposed SPR w/o using metal oxides increases with the use of metamaterial due to its higher imaginary value of RI [20]. Figures 5 and 6 indicates SPR curve for proposed work using different metal oxides with 5 nm, and 10 nm thicknesses respectively at $n_s = 1.33$ and1.345. Figures 4, 5 and 6 demonstrates the shifting of SPR angle to higher value due to local change in sensing layer RI (1.345) with binding of analyte on BRE layer. ZnO with 10 nm

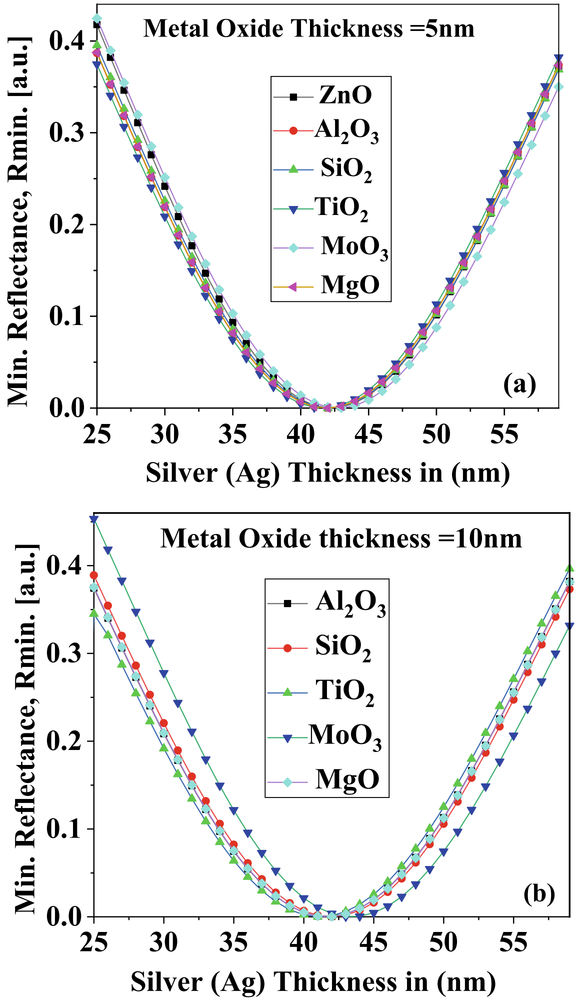


Fig. 2 Thickness optimization of Silver layer at **a** 5 nm metal oxide thickness **b** 10 nm metal oxide thickness

thickness for proposed SPR is not plotted in Fig. 6 because $R_{min.}$ attains very high value at this ZnO thickness.

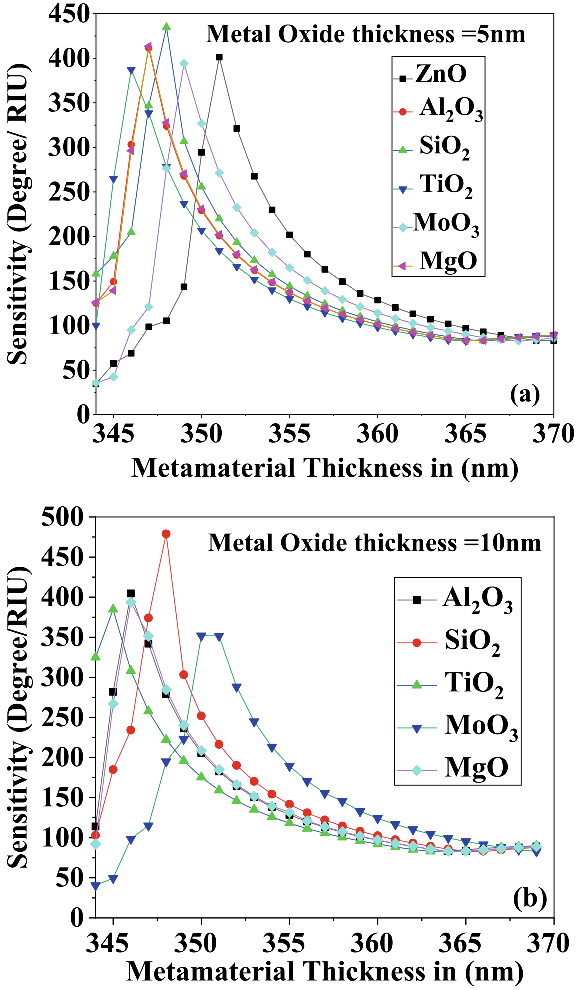


Fig. 3 Metamaterial thickness optimization at **a** 5 nm metal oxide layer thickness **b** 10 nm metal oxide layer thickness

3.3 Sensor Performance Factor Analysis

SPR angle shift ($\Delta\theta_{SPR}$), $R_{min.}$, FWHM, DA and QF are calculated for different metal oxides having thickness 5 nm, 10 nm from Figs. 5 and 6, respectively and mentioned in Table 1.

It can be clearly concluded from Table 1 that best performance (S = 478.9°/RIU, DA = 2.12 /°, QF = 141.2 RIU^{-1}) is obtained for 10 nm SiO_2 metal oxide layer. This sub-section further evaluates the variation of S, DA and QF vs. sensing layer RI lying in range (1.33–1.34) for proposed SPR biosensor using different metal oxide layers

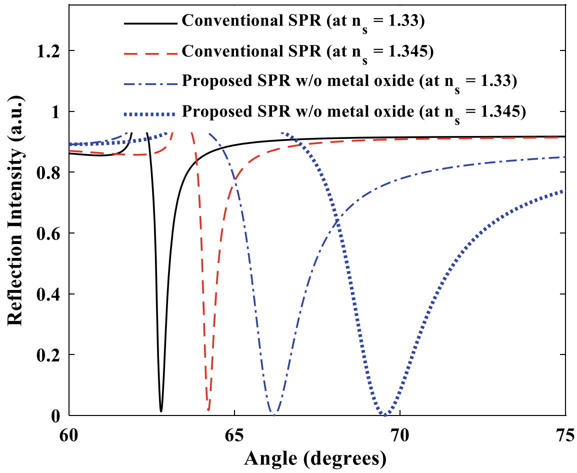


Fig. 4 SPR curve: proposed SPR without (w/o) metal oxides and Conventional SPR sensor

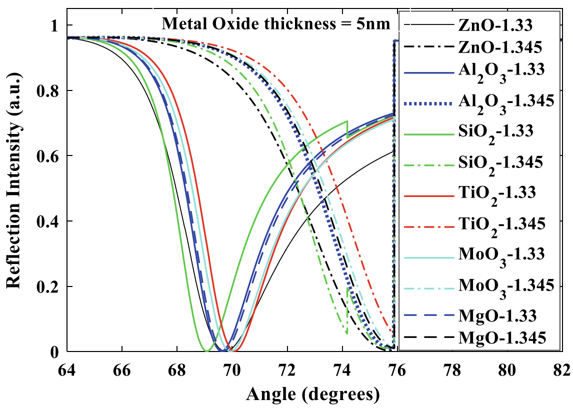


Fig. 5 SPR curve: proposed SPR biosensor for different metal oxide with thickness 5 nm

at 10 nm thickness. Figure 7a suggests sensitivity increases from [311–369.6 °/RIU, 343–418 °/RIU, 346.9–418.3 °/RIU, 401.3–49 °/RIU, 404.3–507.1°/RIU] for Al_2O_3, SiO_2, TiO_2, MoO_3 and MgO respectively. It can be interpreted from Fig. 7a that for smaller RI variation of sensing layer (1.33–1.34), maximum sensitivity variations are obtained for MgO and MoO_3 layer. It is indicated from Fig. 7b that DA decreases from [0.82–0.02 /°, 1.05–0.03 /°, 1.07–0.03 /°, 0.88–0.03 /°, 1.03–0.03 /°] for Al_2O_3, SiO_2, TiO_2, MoO_3 and MgO respectively. Maximum DA is obtained for TiO_2 and then for SiO_2 layer. Similarly, Fig. 7c suggests quality factor decreases from [79.1–57.3 RIU^{-1}, 101.2–76.7 RIU^{-1}, 102.6–78.2 RIU^{-1}, 85–68.8 RIU^{-1}, 99.6–77.1RIU^{-1}] for Al_2O_3, SiO_2, TiO_2, MoO_3 and MgO respectively. Best range of QF is obtained for SiO_2 layer.

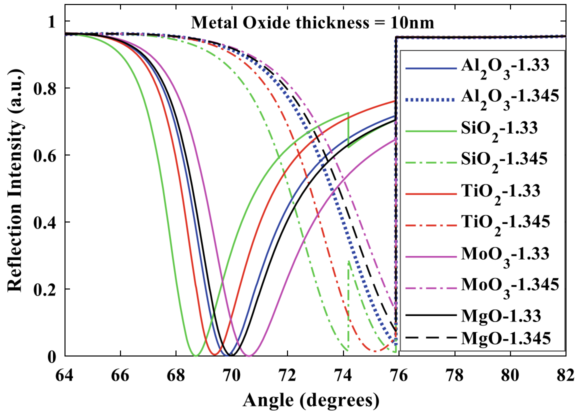


Fig. 6 SPR curve: proposed SPR biosensor for different metal oxide layer with thickness 10 nm

Table 1 Performance factors: SPR angle shift ($\Delta\theta_{SPR}$), FWHM, S, DA and QF for Conventional SPR, Proposed SPR biosensor with and w/o using metal oxides

Sensor configuration			$\Delta\theta_{SPR}$ (°)	FWHM (1/ °)	S (°/RIU)	DA	QF (RIU^{-1})
Conventional SPR			1.429	0.42	95.27	3.40	226.8
Proposed SPR w/o metal oxide			3.347	1.96	223.1	1.708	113.8
Proposed SPR biosensor	ZnO	5 nm	6.021	5.8	401.4	1.038	69.21
	Al_2O_3	5 nm	6.17	3.74	411.3	1.65	110.0
		10 nm	6.07	3.93	404.7	1.54	102.8
	SiO_2	5 nm	6.52	3.56	434.8	1.83	122.1
		10 nm	7.18	3.39	478.9	2.12	141.2
	TiO_2	5 nm	5.81	3.84	387.2	1.51	100.8
		10 nm	5.77	3.38	384.8	1.70	113.8
	MoO_3	5 nm	5.91	4.03	394.3	1.47	97.84
		10 nm	5.28	4.72	351.9	1.12	74.6
	MgO	5 nm	6.20	3.74	413.3	1.66	110.5
		10 nm	5.91	4.06	393.8	1.46	97.0

All above simulation leads to conclusion that best performance of proposed SPR biosensor is obtained on using 10 nm SiO_2 layer than on using other metal oxide layers like Al_2O_3, TiO_2, MoO_3 and MgO. Furthermore, we have compared the sensitivity of some more SPR biosensor using metamaterial with our proposed work in Table 2. It is clearly stated from Table 2 that best values of S, DA and QF are obtained for this proposed work.

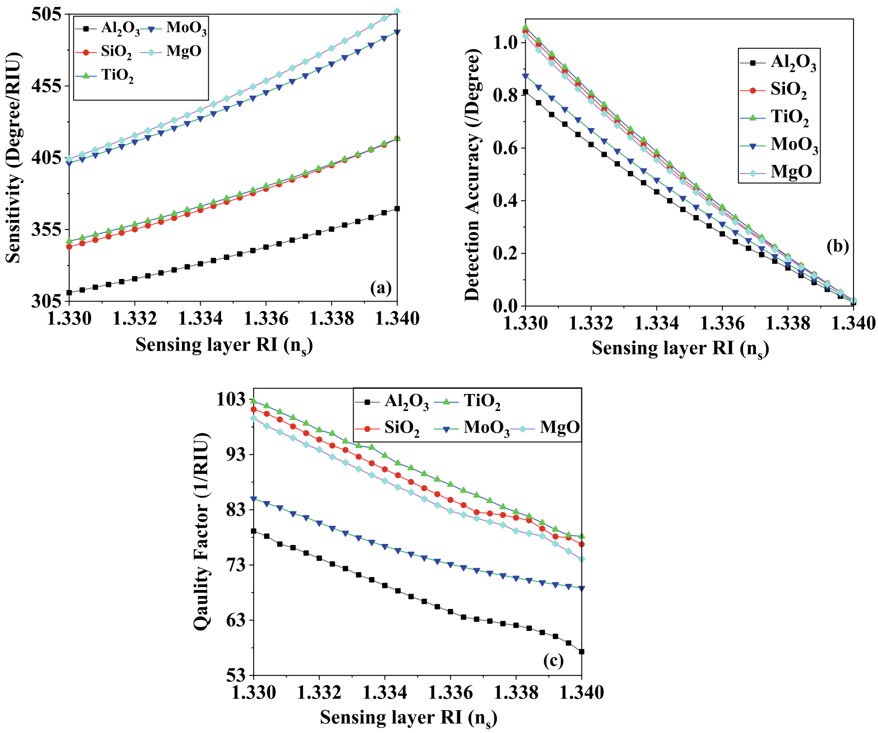


Fig. 7 Shows the graph of **a** sensitivity **b** detection accuracy **c** quality factor vs. sensing layer RI (1.33–1.34)

Table 2 Comparative table indicating performance factors of metamaterial based SPR biosensor

Ref. No	Year	λ(nm)	S	DA	QF
[21]	2016	1000	333.7	–	–
[22]	2016	1000	49.10	17.67	129.9
[23]	2017	738	–	0.373	42
[9]	2018	1000	337.4	0.44	43.03
[24]	2018	632.8	85.5	–	–
[18]	2020	1000	401.4	1.038	69.21
This work	–	1000	478.9	2.12	141.2

4 Conclusion

This work highlights and compares the sensitivity achieved for proposed SPR biosensor on using metamaterial and different metal oxides at 1000 nm characteristic wavelength. A comprehensive performance analysis of proposed work is performed and performance factors are compared to conventional SPR biosensor and proposed work without using metal oxides. Conclusively, considering the numerous advantages of metamaterial and metal oxides utilized in this work for detection of target analytes, the proposed work show full research potential in the field of biosensing. Further in near future, metasurfaces may be used in place of metamaterials to design and develop cost effective and high performance SPR chip for medical diagnostic, drug diagnostics and biosensing purpose.

References

1. Homola, J., Sinclair, S.Y., Gauglitz, G.: Surface plasmon resonance sensors: review. Sens. Actu. B: Chem. **54**, 3–15 (1999)
2. Gupta, B.D., Pal, B.P.: Fiber optical sensors and instruments for bio-science. IETE Tech. Rev. **8**, 19–332 (2015)
3. Singh, A.K., Sharma, S.C.: A fixed detector Kretschmann configuration optical system to study surface plasmon excitations. Opt. Laser Technol. **56**, 56–262 (2014)
4. Firdous, S., Anwar, S., Rafya, R.: Development of surface plasmon resonance (SPR) biosensors for use in the diagnostics of malignant and infectious diseases. Laser Phys. Lett. **15** (2018)
5. Ladd, J., Taylor, A., Jiang, S.: SPR biosensors for food safety. In: Homola, J. (eds) Surface plasmon resonance based sensors, Springer Series on Chemical Sensors and Biosensors, vol. 4, Springer, Berlin (2006)
6. Weiss, M.N., Srivastava, R., Groger, H., Lo, P., Luo, S.F.: A theoretical investigation of environmental monitoring using surface plasmon resonance waveguide sensors. Sens. Actuators, A **51**, 211–217 (1995)
7. Anupam, R.S.: Potentialities and applications of biosensors for on-line environmental monitoring. IETE Tech. Rev. **9**, 146–150 (2015)
8. Pal, S., Pal, N., Prajapati, Y.K., Saini, J.P.: Performance evaluation of SPR biosensor using metamaterial over conventional SPR and graphene based SPR biosensor. In: IEEE Proceedings, SPIN-2018, Amity University, Noida, 22–23 February
9. Prajapati, Y.K., Pal, S., Saini, J.P.: Effect of metamaterial and silicon layers on performance of surface plasmon resonance biosensor in infrared range. SILICON **10**, 1451–1460 (2017)
10. Pal, S., Verma, A., Saini, J.P., Prajapati, Y.K.: Sensitivity enhancement using silicon-black phosphorus-TDMC coated surface plasmon resonance biosensor. IET Opto. **13**, 196–201 (2019)
11. Kumar, M., Pal, S., Verma, A., Prajapati, Y.K., Saini, J.P.: Sensitivity ımprovement of surface plasmon resonance sensor on using BlueP/MoS_2 heterostructure and antimonene. IEEE Sensor Lett. **4**(7), 1–4 (2020). https://doi.org/10.1109/LSENS.2020.3005942.
12. Kumar, R., Pal, S., Verma, A., Prajapati, Y.K., Saini, J.P.: Effect of silicon on sensitivity of SPR biosensor using hybrid nanostructure of black phosphorus and MXene. Superlattices Microstruct. **145**, 106591 (2020)
13. Chen, T., Li, S., Sun, H.: Metamaterials applications in sensing sensors. Sensors **12**, 2742–2765 (2012)
14. Ishimaru, A., Jaruwatanadilok, S., Kuga, Y.: Generalized surface plasmon resonance sensors using metamaterials and negative index materials. Prog. Electromagn. Res. **51**, 139–152 (2005)
15. Ahmadi, L., Hiltunen, M., Hiltunen, J., Aikio, S., Saarinen, J., Honkanen, S., Roussey, M.: Influence of an Al_2O_3 surface coating on the response of polymeric waveguide sensors. Opt. Express **25**, 25102–25112 (2017)
16. Li, W., Zhang, A., Cheng, Q., Sun, C., Li, Y.: Theoretical analysis on SPR based optical fiber refractive index sensor with resonance wavelength covering communication C+L band. Optik **123**, 164696 (2020)
17. Dorozinsky, H., Lyapin, O., Dorozinsky, G., Maslov, V.: Current trends in technology and materials of sensors based on surface plasmon resonance. J. Appl. Mat. Sci. Eng. Res. **2** (2018)

18. Pal, S., Pal, N., Teotia, A., Kumar, D.: Improved sensitivity of metamaterial based SPR biosensor using Zinc Oxide. In: 7th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, pp. 170–175 (2020)
19. https://refractiveindex.info.com
20. Pockrand, I.: Surface plasma oscillations at silver surfaces with thin transparent and absorbing coatings. Surf. Sci. **72**(3), 577–588 (1978)
21. Pal, S., Prajapati, Y.K., Saini, J.P., Singh, V.: Sensitivity enhancement of metamaterial based surface plasmon resonance biosensor for near infrared. Opt. Appl. **46**, 131–134 (2016)
22. Verma, A., Prakash, A., Tripathi, R.: Comparative study of a surface plasmon resonance biosensor based on metamaterial and graphene. SILICON **9**, 309–320 (2017)
23. Cherifi, A., Bouhafs, B.: Sensitivity enhancement of a surface plasmon resonance sensor using porous metamaterial layers. Mater. Res. Express **4**, 125009 (2017)
24. Tewatia, A., Pal, S., Pal, N.: Performance evaluation of surface plasmon resonance biosensor using metamaterial. Mater. Today Proc. **5**, 28384–28391 (2018)

Metal-Insulator-Metal Metamaterial Helical Absorber

Sajal Agarwal and Yogendra Kumar Prajapati

Abstract Metamaterial is a very popular advanced material used for the designing of absorbers. Absorbers are very important component of various communication devices as well as energy harvesting devices also. This study is concentrated to design a metal-insulator-metal type metamaterial. Here the properties of bismuth are utilized to get the high absorbance as well as wide absorbance region. To complete the study, structural parameters of helix are also varied. Through the simulations it is observed that geometrical parameters affect the absorber performance severely and optimization is very important. Proposed absorber has two absorption peaks; one at 490 and other at 1800 nm. Peak absorption for both absorption peaks is approximately 0.8 a.u.

Keywords Absorber · Bismuth · Helix · Metamaterial · Tapering

1 Introduction

Nanotechnology is the most fascinating field of research which is not even limited to only one research area. It extends from simple material physics to computer engineering [1], optical engineering [2], agricultural field [3] etc. These fields are not only bounded to single area however, interdisciplinary applications like microscopy [4], sensing [5], energy harvesting [6] etc. Energy harvesting and communication research areas are very much related to the nano-optics for some components such as antenna [7], signal processing [8], absorber [9] etc. The primary concern of researchers working in this areas is to satisfy current demand while keeping device/component size small, light weight, efficient.

S. Agarwal (✉)
Rajiv Gandhi Institute of Petroleum Technology, Jais, Amethi, India
e-mail: sagarwal@rgipt.ac.in

Y. K. Prajapati
Motilal Nehru National Institute of Technology Allahabad, Allahabad, India
e-mail: yogendrapra@mnnit.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_3

From all the above discussed devices/components, absorber draws wide attention, because of their multidisciplinary application in communication devices [7], solar cells [10], RADARs [11] etc. Absorbers are the optical devices which can absorb all/partial radiation incident upon it. These devices are widely useful in thermo-photovoltaic cells (TPV) to capture solar spectrum and convert it into heat to improve the efficiency [12]. As much as wide spectrum will be absorbed by absorber better the performance for TPV cells. Thus, it is highly vital to design electromagnetic (EM) absorber, absorbs wide range of frequencies [12]. There are a lot of ways through which the desired performance of the absorber would get, however, use of metamaterial is the most used and suitable method. Metamaterial is the artificially engineered material and has tunable properties by changing its geometrical and structural parameters.

First, perfect metamaterial absorber was proposed by N. I. Landy et al. in 2008 [13], which is a split ring structure working in giga hertz range. There are a lot of aspects of metamaterial structure which can be analyzed for the better performance of the overall absorber. Three dimensional structure of metamaterial is one of the most used method for improved absorption because of the better interaction of radiation with the absorber. In 2015 S. Agarwal et al. [12] proposed a helix based absorber for the absorbing application having aluminum (Al) as the constituent metal and displayed that helix metamaterial has wide operating range for very complex geometry. However, Al can get oxidized easily thus, stability issues are there. Recently, A. Ghobadi et al. [14] proposed a metal-insulator-metal metamaterial absorber for two different applications, absorber and filter, using bismuth (Bi). It is observed that metal-insulator-metal structure provide much better absorbance and control over parameters with simple geometry. However, the proposed work has limited range which can further be improved by optimization of the unit cell structure. A lot of studies are published till now using Bi for absorber modeling [15, 16]. Thus, in this study helix based metamaterial is modeled utilizing the properties of Bi in metal-insulator-metal design.

2 Design and Simulation

This study explores the properties of helices for metamaterial designing. Al is used as the substrate having the thickness of 30 nm. Over the metal substrate dielectric material is used as the insulator, i.e. titanium dioxide (TiO_2). Over the dielectric material metal helices are place made of Bi. In each unit cell four helices are used having opposite handed polarity. Finite difference time domain (FDTD) method is used for the modeling of the proposed structure, Lumerical FDTD software is used for the task. Figure 1 has the modeled metamaterial structure.

Modeling of the metals is done using Lorentz-Drude model, which is the most accurate metal modeling method. Opposite handed helix placement ensures the polarization insensitive nature of the proposed structure. For the proposed study, tapered helices are used having large lower diameter and small upper diameter. Tapering of

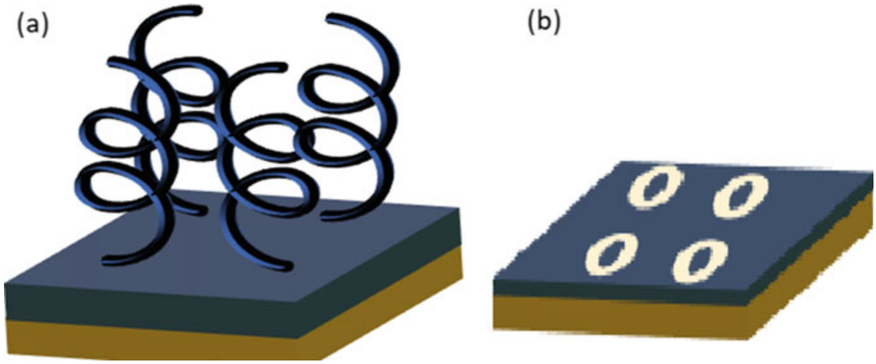


Fig. 1 Modeled helical metamaterial absorber **a** three dimensional view **b** top view

Table 1 Peak absorbance and wavelength for individual metals

Metal	Peak absorbance (a.u.)	Wavelength (nm)
Gold (Au)	0.92	420
Tungsten (W)	0.98	380
Bi	0.98	450
Al	0.90	378
Tin (Sn)	0.91	514

the helices ensures that better absorption of the radiation. For the systematic study geometrical parameters are varied to find the optimized structure of absorber. For the analysis, wide wavelength region has been decided. Since, the intended application of the proposed absorber is thermo-photovoltaic cells thus covering whole solar spectrum varies from visible to far-infrared region is important.

3 Results and Discussion

For the detailed study, three most important parameters are picked, fiber diameter (though the wire helix made), upper and lower radius of the helices, and pitch of the helices. To decide the metal used for helices, individual absorption of various metals are analyzed. Peak absorbance for these individual metals and the corresponding wavelength is given in Table 1. From the Table, it is observed that absorbance of the metals varies and W and Bi provides highest peak absorbance near unity. Thus, Bi is choose to be used as the constituent metal for the helices in the proposed metamaterial. Thickness of the substrate metal and dielectric is fixed, i.e. 30 and 100 nm. For the starter, pitch of the helix is varied from 150 to 200 nm, and the absorption spectrum are plotted in Fig. 2.

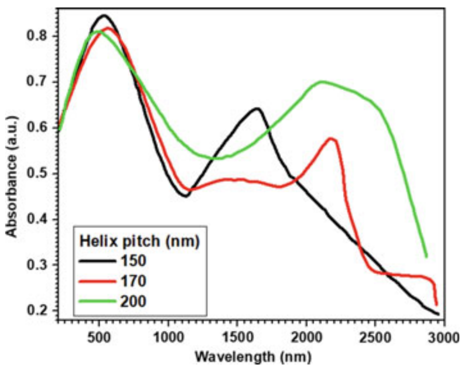


Fig. 2 Absorption spectrum for different pitch of the helix

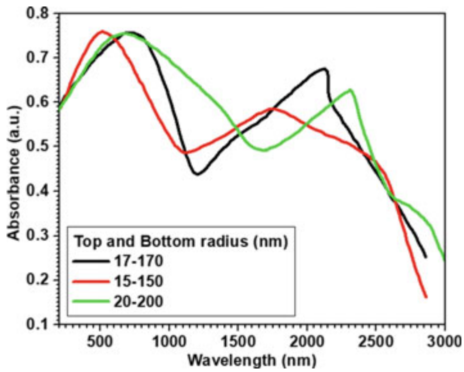


Fig. 3 Absorption spectrum for different top/bottom radius combinations of the helix

It is observed that pitch of 150 nm gives high absorbance, but pitch of 200 nm provides better absorbance for all over the wavelength region. Thus for the further simulations, 200 nm long pitch is considered. Next, top and bottom radius of the helices are varied in combinations with fixed fiber radius of 15 nm and the absorption spectrum are given in Fig. 3.

It is observed that from the radius analysis that higher radius's on both ends provides better absorption however peak absorption is not changed much but it broaden the first peak width. In the end fiber radius is varied to see the effect on absorbance. Absorption spectrum is given in Fig. 4.

Since, it is observed from the above analysis that the absorption depends on the geometrical parameter as well as the constituent material, it is highly appreciated to chose the parameters wisely after thorough analysis. It is observed that two very neat peaks are achieved around 490 and 1800 nm having 0.8 a.u. absorption. However, proposed absorber is not wide band but it can be used for various communication applications according to the interested wavelength band. This is a very initial study but different combinations of materials can be analyzed to further improve the absorbance. Fabrication of the proposed tapered helix metamaterial surface is possible by using direct laser writing technique precisely.

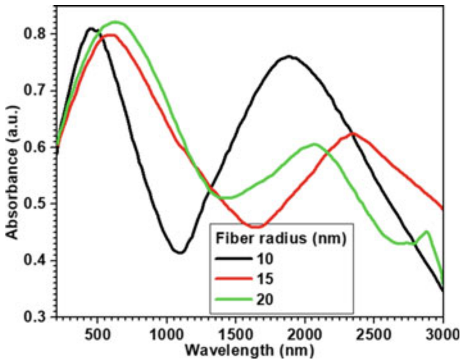


Fig. 4 Absorption spectrum for varied fiber radius

4 Conclusion

This study proposed a very promising absorber metamaterial structure for communication applications. To improve the proposed absorber for wide band wavelength region intertwining of the helices can be done. However, on choosing the helix parameters carefully, absorption bands are widened. The peak absorption found for both peaks is 0.81 and 0.78 a.u. at 490 and 1800 nm wavelength respectively. It is also observed that the constitutional metal of helix would affect the absorption severely.

References

1. Qiu, M., Khisamutdinov, E., Zhao, Z., Pan, C., Choi, J., Leontis, N.B., Guo, P.: RNA nanotechnology for computer design and in vivo computation. Philos. Trans. Royal Soc. A Math. Phys. Eng. Sci. **371**(2000), 20120310 (2013)
2. Chen, H.T., Padilla, W.J., Zide, J.M., Gossard, A.C., Taylor, A.J., Averitt, R.D.: Active terahertz metamaterial devices. Nature **444**(7119), 597–600 (2006)
3. Smith, G.B.: Green Nanotechnology: Solutions for Sustainability and Energy in the Built Environment. CRC Press, Boca Raton (2019)
4. Horcas, I., Fernández, R., Gomez-Rodriguez, J.M., Colchero, J.W.S.X., Gómez-Herrero, J.W.S.X.M., Baro, A.M.: WSXM: a software for scanning probe microscopy and a tool for nanotechnology. Rev. Sci. Instrum. **78**(1), 013705 (2007)
5. Farzin, L., Shamsipur, M., Samandari, L., Sheibani, S.: HIV biosensors for early diagnosis of infection: the intertwine of nanotechnology with sensing strategies. Talanta **206**, 120201 (2020)
6. Abdin, Z., Alim, M.A., Saidur, R., Islam, M.R., Rashmi, W., Mekhilef, S., Wadi, A.: Solar energy harvesting with the application of nanotechnology. Renew. Sustain. Energy Rev. **26**, 837–852 (2013)
7. Shchegolkov, D.Y., Azad, A.K., O'hara, J.F., Simakov, E.I.: Perfect subwavelength fishnetlike metamaterial-based film terahertz absorbers. Phys. Rev. B **82**(20), 205117 (2010)
8. Chen, J., Wang, S.T.C.: Nanotechnology for genomic signal processing in cancer research-A focus on the genomic signal processing hardware design of the nanotools for cancer research. IEEE Signal Process. Mag. **24**(1), 111–121 (2007)

9. Yang, J., Gong, C., Sun, L., Chen, P., Lin, L., Liu, W.: Tunable reflecting terahertz filter based on chirped metamaterial structure. Sci. Rep. **6** (2016)
10. Yin, W.J., Shi, T., Yan, Y.: Unusual defect physics in CH3NH3PbI3 perovskite solar cell absorber. Appl. Phys. Lett. **104**(6), 063903 (2014)
11. Li, M., Xiao, S., Bai, Y.Y., Wang, B.Z.: An ultrathin and broadband radar absorber using resistive FSS. IEEE Antennas Wirel. Propag. Lett. **11**, 748–751 (2012)
12. Agarwal, S., Prajapati, Y.K.: Analysis of metamaterial-based absorber for thermo-photovoltaic cell applications. IET Optoelectron. **11**(5), 208–212 (2017)
13. Landy, N.I., Sajuyigbe, S., Mock, J.J., Smith, D.R., Padilla, W.J.: Perfect metamaterial absorber. Phys. Rev. Lett. **100**(20), 207402 (2008)
14. Ghobadi, A., Hajian, H., Gokbayrak, M., Butun, B., Ozbay, E.: Bismuth-based metamaterials: from narrowband reflective color filter to extremely broadband near perfect absorber. Nanophotonics **8**(5), 823–832 (2019)
15. Sadecka, K., Toudert, J., Surma, H.B., Pawlak, D.A.: Temperature and atmosphere tunability of the nanoplasmonic resonance of a volumetric eutectic-based Bi_2O_3-Ag metamaterial. Opt. Express **23**(15), 19098–19111 (2015)
16. Cheng, X.S., Ahmad, H., Harun, S.W.: Broadband ASE source using bismuth-based erbium-doped fibers in double-pass set-up. Microw. Opt. Technol. Lett. **52**(7), 1636–1638 (2010)

Compact Circularly Polarized Microstrip Fed Wide Slot Antenna for C-Band Applications

Sunidhi Dubey, Sudhanshu Verma, and Saurabh Singh

Abstract The paper presents a compact circularly-polarized micro-strip fed wide slot antenna (WSA) of size 25×25 mm^2 for C-band applications. The anticipated design consists of a horizontal slotted patch, a wide C-shape slot and a stub fed by 50 Ω micro-strip feed-line. The corners of the wide slot are chamfered symmetrically and a small stub is introduced in the ground. The micro-strip feed is placed under the extended horizontal stub protruding from ground to generate the two orthogonal modes. The slotted patch and a chamfering in the ground are responsible for enhancing the impedance bandwidth (IBW) and attaining wide axial ratio bandwidth (ARBW). The proposed WSA achieved 10-dB return loss of 96.64% (3.23–9.27 GHz) and ARBW of 63.45% (3.69–7.12 GHz).

Keywords Circularly polarized · Slot antenna · Axial ratio · C band

1 Introduction

In recent time, printed slot micro-strip antennas are in great demands for wireless communication systems as it exhibits the properties of less bulky, cost effective, low profile, wide bandwidth and easy fabrication process. The circularly polarized antenna is popular in wireless communication systems because CP signals are good at penetration and bending around obstacles and are more repellent to signal distortion due to cold weather conditions for reflection, absorption, multi path, LOS and also it is much effective to establish and maintain communication links. The principle of operation for CP antenna is to stimulate two orthogonal modes of equivalent amplitude and phase quadrature. The fraction of the electromagnetic spectrum assigned for satellite transmissions in the 4–8 GHz frequency range is referred to as C-band. Many satellite communications broadcasts, Wi-Fi networks, cordless telephones, and some surveillance and weather radar systems all operate in the C band [1].

S. Dubey (✉) · S. Verma · S. Singh
Department of Electronics and Communication Engineering, Madan Mohan Malaviya, University of Technology, Gorakhpur, Uttar Pradesh, India
e-mail: dubeysunidhi1211@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_4

The CP can be achieved by various techniques that are mentioned in literature [2–8]. In [2], for CP, the antenna's excitation is done by L-shaped strip with a narrow end, in series connection to a microstrip-feedline situated diagonally of the square slot. The achieved $S_{11} < -10$-dB is 48.41% (1686–2763 MHz) and ARBW is 42.22% (1700–2610 MHz). In [3], a micro-strip fed two coinciding ellipse shaped slots and inserting two orthogonal stubs in the feed are new methodologies to stimulate two orthogonal vectors. The attained impedance BW is 60.16% (4.3–8 GHz) and axial-ratio < 3-dB bandwidth of 40.74% (4.3–6.5 GHz). In [4], a micro-strip fed two overlapped annular slots are aimed to produce a wide circularly polarized bandwidth. The achieved 10-dB is 61.53% ranging from (2.25–4.25 GHz) and 3-dB ARBW of 56.60% (1.9–3.4 GHz). In [5], implanting dual L-ground strips on reverse corners of a square slot antenna for CP radiation. The attained impedance bandwidth is 62.51% ranges (1600–3055 MHz) and axial ratio bandwidth of 27.39% (2300–3030 MHz). In [6], a circularly polarized slot antenna is intended with a slot comprising of multiple circular sectors (MCS). The achieved impedance bandwidth is of 64.7% (2.06–4.03 GHz) and 3-dB axial ratio bandwidth of 61.5% (2.14–4.04 GHz). In [7], the circular polarization can be accomplished by introduction of asymmetric agitation on the lower left of an open slot which is fed by a tuning stub. The 10-dB return loss achieved is 111.15% (2.13–7.46 GHz) and the 3-dB AR bandwidth can attain 27.02% (3.2–4.2 GHz). In [8], for broadband CP generation, the micro-strip feed is placed under the extended horizontal stub protruding from ground. The obtained impedance matching bandwidth is of 90.2% ranging (3.5–9.25 GHz), and AR bandwidth of 40% ranging (4.6–6.9 GHz).

This paper presented a microstrip-fed compact wide-band circular polarised wide slot antenna, which consist a horizontal multiple slot radiating patch with wide C-shape slot having symmetric chamfered at the opposite corners and a stub excited by a feed of 50 Ω microstrip line. The corners of the wide slot are chamfered symmetrically and a stub is introduced in the ground plane. The micro-strip feed is placed under the extended horizontal stub protruding from ground to generate the two orthogonal modes. The multiple slots patch and a chamfering in the ground are responsible for enhancing the impedance bandwidth and attaining wide axial-ratio bandwidth. The designing and simulation of anticipated wide slot antenna is performed on Ansys electronic desktop tool (version 16.2).

2 Antenna Design

The designed geometry of wide slot antenna is showed in Fig. 1. FR4 is used as substrate of dielectric constant 4.4, loss tangent of 0.02 and 1.6 mm thickness to print the antenna. This antenna is composed of horizontal slotted radiation patch fed by 50 Ω micro-strip feed-line, a wide square slot antenna having symmetric chamfering to realize the circularly polarized radiation. The right edge of the ground has a stub that projects into the slot's center. To minimize linkage between the ground and the feed-line, a gap on the ground plane is provided above the feedline. The compact size of antenna is $25 \times 25 \times 1.6$ mm^3. The microstrip feedline is etched on the top of FR4 substrate having width of 3 mm. The x-axis of slotted patch is of 5 mm and y-axis is 12.75 mm. The width of the slot in the patch is of 0.5 mm and strip size is 1 mm. To realize broad band CP operation, the micro-strip line-fed is placed on the right edge of the antenna under the protruded stub.

The development of the proposed WSA is represented in Fig. 2. To elucidate the CP performance of the WSA, three antennas evolution is discussed. In first step, Ant.1 consists of a simple slot and a short stub on the top of microstrip feedline towards –*y* axis direction as discussed in [8]. In second step of evolution, Ant.2 comprises of a horizontal slotted patch. In third step, Ant.3 consists of chamfered corners of the square slot.

The resultant graph of impedance bandwidths and AR bandwidths of the three antennas are shown in Fig. 3. In first step, the ant.1 has wide IBW and ARBW which radiate in frequency band of 87.5% (3.6–9.2 GHz) and 37.93% (4.7–6.9 GHz) as showed in Fig. 3. In second step, by slotting of radiation patch IBW and ARBW is shifted towards left side in frequency range of 92.52% (3.20–8.71 GHz) and

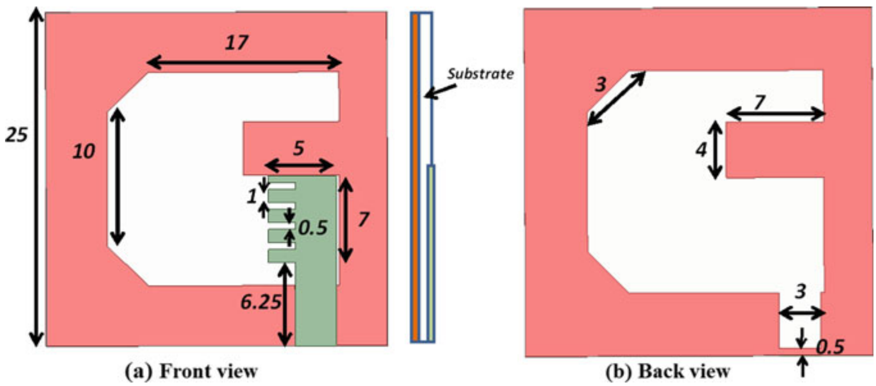


Fig. 1 Antenna geometry of proposed WSA

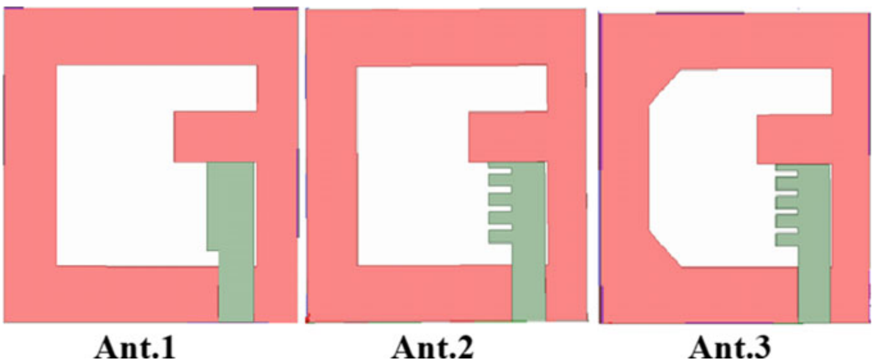


Fig. 2 Steps to realize the proposed WSA (1–3)

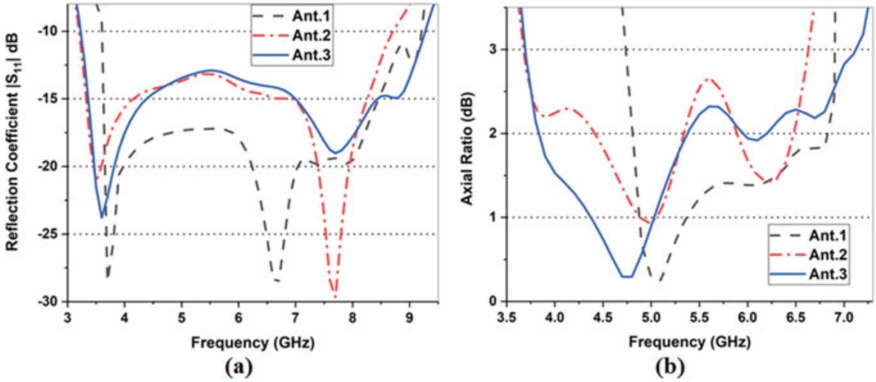


Fig. 3 Simulated results of evolution steps: **a** Reflection Coefficient $|S_{11}|$ dB. **b** Axial Ratio of three antennas

57.22% (3.68–6.63 GHz) as shown in Fig. 3. Further, ant.3 enhances the impedance bandwidth up-to 96.64% (3.23–9.27 GHz) and achieves a 63.45% wide axial ratio bandwidth ranging (3.69–7.12 GHz) as shown in Fig. 3.

3 Results and Discussion

The simulated results of WSA are analyzed on Ansys electronic desktop tool (version 16.2) with a solution frequency of 3.5 GHz. The simulation result of reflection coefficient of proposed WSA is depicted in Fig. 4 which shows that the proposed WSA is radiating in (3.23–9.27 GHz) with fractional bandwidth of 96.64%. The obtained resonance peak is at 3.59 GHz. The simulated axial ratio of proposed WSA is shown in Fig. 5 which depicts that the proposed WSA is radiating in frequency band (3.69–7.12 GHz) having fractional bandwidth of 63.45%. Thus, the proposed WSA is a wide band antenna for C-band applications.

Figure 6 shows the flow of current distribution at 4.7 GHz of the proposed WSA which depicts that the concentration of current is maximum in the patch. The distribution of current shows a centered frequency which is also used for radiation pattern. The current flows at different phase angle such as 0°, 90°, 180° and 270°. It depicts that at 0° the flow of current is towards downward, at 90° flow of current is in $-x$-direction, at 180° the current streams is in upward direction and at 270° the direction of current is $+x$ direction which shows the phenomena of CP radiation which is left hand circular polarized (LHCP) at positive z direction.

Figure 7 illustrates the simulated gain of the proposed WSA which shows that the peak gain of proposed WSA is 4.78 dBi at 4 GHz. The simulated result of the radiation pattern of the anticipated C-shaped slot antenna at resonance peak is exemplified in Fig. 8. The radiation pattern is obtained in both *xz* and *yz* planes and LHCP is emitted in a positive *z*-direction, while RHCP is emitted in a negative *z*-direction.

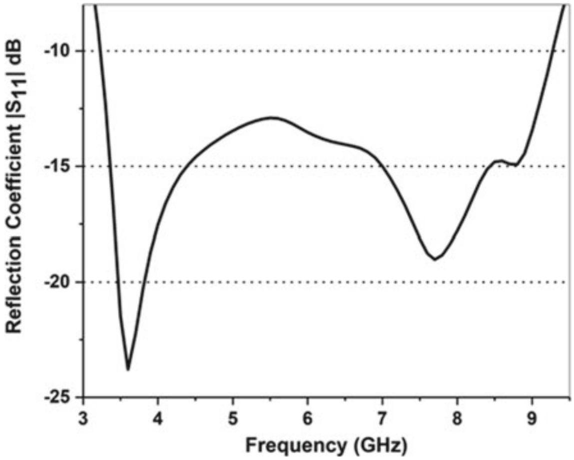


Fig. 4 Simulated reflection coefficient $|S_{11}|$ dB of proposed WSA

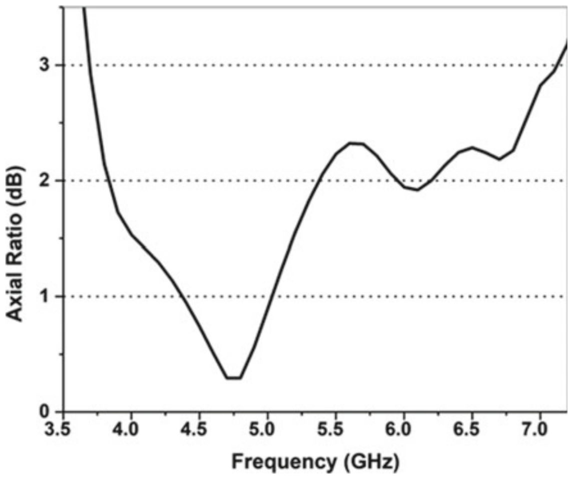


Fig. 5 Simulated axial ratio of proposed WSA

From the radiation pattern it is stated that the proposed C-shape slot antenna radiates bi-directional. There is a minor tilt in the path of the concentrated radiation. This is mainly due to the irregular construction of the proposed C-shaped wide slot antenna which causes the path of concentrated radiation to slightly shift towards the $+x$ and $+y$ directions in the xz plane and yz plane, respectively. It can be stated that the antenna is radiating bi-directionally with opposed circular polarization. The RHCP is attained for $z < 0$ while LHCP is attained for $z > 0$.

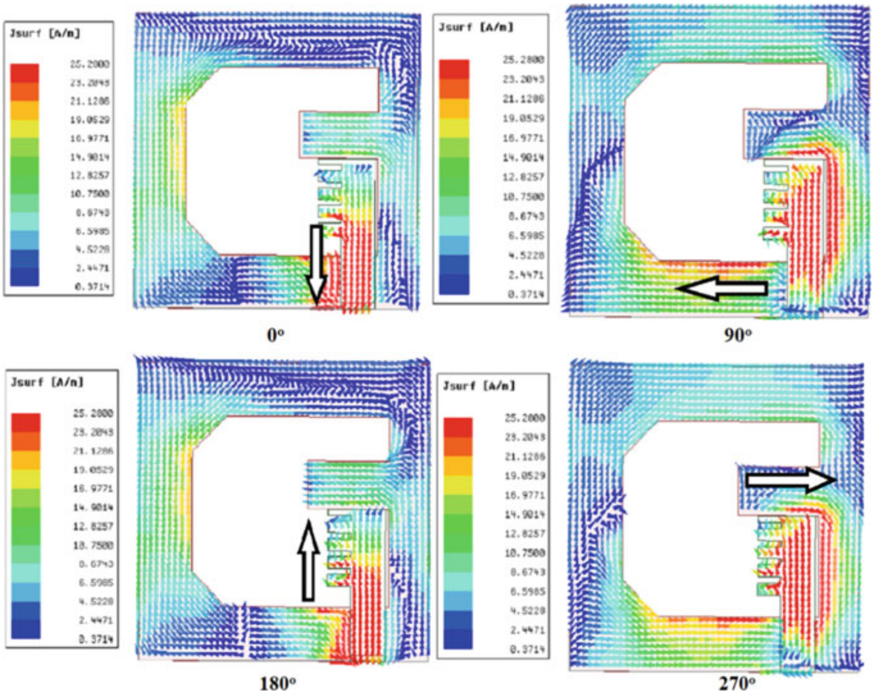


Fig. 6 Distribution of current in proposed WSA at 4.7 GHz

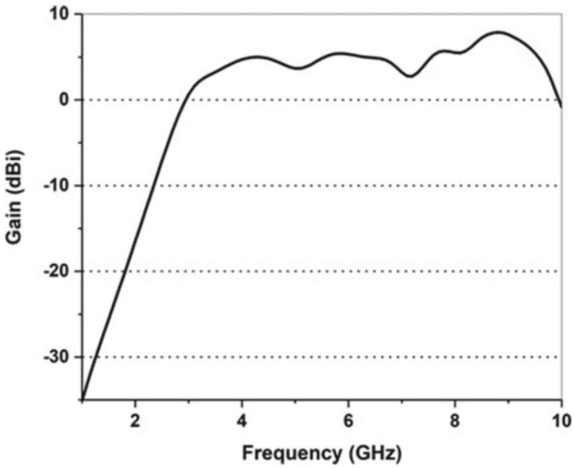


Fig. 7 Simulated gain of proposed WSA

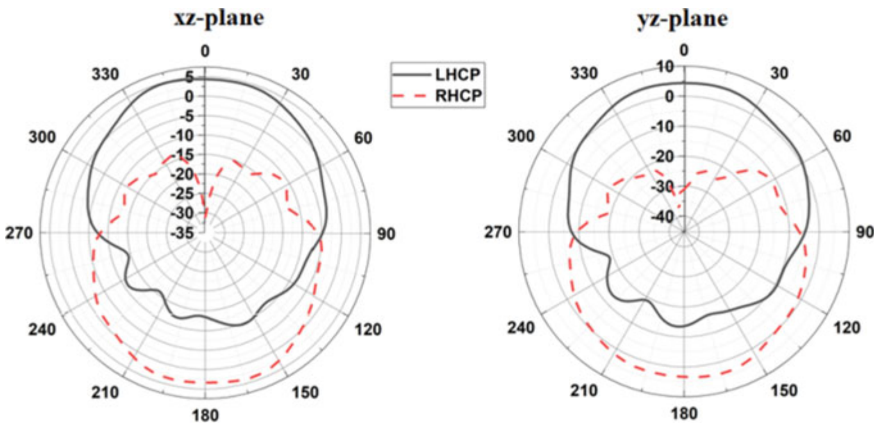


Fig. 8 Simulated radiation pattern of proposed WSA at 4.7 GHz

Table 1 Comparative study of proposed WSA and the alike ones

Ref.	IBW (GHz)	ARBW (GHz)
[9]	17.4% (3.05–3.63)	5.4% (3.25–3.43)
[10]	40% (4.8–7.2)	31.84% (5.15–7.1)
[11]	62% (3.6–6.85)	49% (3.6–5.93)
[12]	84% (3.25–8)	41.3% (4.41–6.67)
Proposed work	96.64% (3.23–9.27)	63.45% (3.69–7.12)

Table 1 represents the BWs of the designed WSA antenna and preceding designs. In contrast, the overall results depicts that the proposed wide slot antenna is having wide impedance BW and ARBW.

4 Conclusion

A compact antenna with simple structure of size 25 × 25 × 1.6 mm^3 is anticipated in this paper. The proposed WSA design comprises of a slotted patch with chamfered wide square slot and a 50 Ω excitation by micro-strip line feed. The micro-strip feed is placed under the extended horizontal stub protruding from ground to generate the two orthogonal modes. The multi-slot patch and a chamfering in the ground are responsible for the improvement of impedance bandwidth and attaining wide axial ratio bandwidth. The proposed WSA achieved 10-dB return loss of 96.64% (3.23–9.27 GHz) and ARBW of 63.45% (3.69–7.12 GHz) and obtained a gain of 4.78 dBi at 4 GHz with bi-directional radiation pattern. The proposed WSA radiate at resonance peak and is well-suited for C-band applications.

References

1. Gao, S., Luo, Q., Zhu, F.: Antenna Theory-Analysis and Design. Wiley (2014)
2. Han, T.-Y.: Broadband circularly polarized square-slot antenna. J. Electromagn. Waves Appl. **22**, 549–554 (2008)
3. Mohammadi, S., Nourinia, J., Ghobadi, C., Pourahmadazar, J., Shokri, M.: Compact broadband circularly polarized slot antenna using two linked elliptical slots for C-band applications. IEEE Antennas Wireless Propag. Lett. **12**, 1094–1097 (2013)
4. Chang, T.N.: Wideband circularly polarised antenna using two linked annular slots. Electron. Lett. **47**(13), 737–739 (2011)
5. Sze, J.-Y., Chang, C.-C.: Circularly polarized square slot antenna with a pair of inverted-L grounded strips. IEEE Antennas Wireless Propag. Lett. **7**, 149–151 (2008)
6. Yeung, S.H., Man, K.F., Chan, W.S.: A bandwidth improved circular polarized slot antenna using a slot composed of multiple circular sectors. IEEE Trans. Antennas Propag. **59**(8), 3065–3070 (2011)
7. Jan, J.-Y., Pan, C.-Y., Chiu, K.-Y., Chen, H.-M.: Broadband CPW-fed circularly polarized slot antenna with an open slot. IEEE Trans. Antennas Propag. **61**(3), 1418–1422 (2013)
8. Ellis, M.S., Zhao, Z., Wu, J., Ding, X., Nie, Z., Liu, Q.H.: A novel simple and compact microstrip-fed circularly polarized wide slot antenna with wide axial ratio bandwidth for C-band applications. IEEE Trans. Antennas Propag. **64**(4), 1552–1555 (2016)
9. Hao, S.S., Chen, Q.Q., Li, J.Y., Xie, J.: A high-gain circularly polarized slotted patch antenna. IEEE Antennas Wireless Propag. Lett. **19**(6), 1022–1026 (2020)
10. Ahdi Rezaeieh, S., Abbosh, A., Antoniades, M.A.: Compact CPW-fed planar monopole antenna with wide circular polarization bandwidth. IEEE Antennas Wireless Propag. Lett. **12** (2013)
11. Ullah, U., Koziel, S.: A broadband circularly polarized wide-slot antenna with a miniaturized footprint. IEEE Antennas Wireless Propag. Lett. **17**(12) (2018)
12. Nosrati, M., Tavassolian, N.: Miniaturized circularly polarized square slot antenna with enhanced axial-ratio bandwidth using an antipodal Y-strip. IEEE Antennas Wireless Propag. Lett. **16**, 817–820 (2016)

Cross Age Face Generator: A Generative Adversarial Networks (GANs) Based Approach

Prathamesh V. Barve and Amit D. Joshi

Abstract The interpretation of various features in the world can be carried out with the help of computer vision. Cross age face generation is one of the emerging field in computer vision with great significance to solve different identity related problems. Different deep learning approaches aids computers to yield effective results in the area of computer vision. This work focuses on generation of person's face at different age in individual's lifespan. Cross age face generation helps in solving various problems related to identity of an individual, such as the recognition of a missing person, identification of criminals, etc. Preservation of the identity of a person is very important and challenging task while creating images at different ages. This work proposes a system using Conditional Generative Adversarial Networks and Convolutional AutoEncoder model. The Convolutional AutoEncoder model is used for identity preservation. Conditional Generative Adversarial Networks are used for generating images of a person in different ages. The system is evaluated on the IMDB-Wikicleaned dataset with 120 K images and Cross-Age Celebrity Dataset with 163 K images. Various tests are performed to quantify the results. The proposed solution gives better results in terms of identity preservation, age estimation error and frechet inception distance parameters. Identity preservation with the proposed system. is 96.82% with 2.75 as age estimation error.

Keywords Conditional generative adversarial networks · Deep convolutional generative adversarial networks · Convolutional auto encoder · Deep learning

1 Introduction

Cross age face generation is the way towards delivering a picture with the same characteristics applying age alterations or transformations [1]. It is also termed as

P. V. Barve (✉) · A. D. Joshi
Department of Computer Engineering and IT, College of Engineering Pune, Pune, India
e-mail: barvepv19.comp@coep.ac.in

A. D. Joshi
e-mail: adj.comp@coep.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_5

age progression, face age synthesis, face aging [2]. Cross age face generation deals with the problem of change in the appearance of an individual person by keeping the identity intact. Cross age face generation techniques have pulled in tons of researchers to incorporate a face having significant age changes and identity preservation in different domains. For an instance, it is useful to identify a long time missing person who can have age effects on face or a victim by generating the person's face in the target age [3].

Many methods have been proposed to add age progression to face. Some of them includes conventional methods, rule based methods and methods with deep learning models. Conventional face ageing models can be generally separated into physical model methodologies and prototype methodologies. Physical model processes shape and texture of every age group [2]. The prototype approach targets building a face prototype for different age groups and reduces the texture differences in the prototypes and the given image. However, these types of models are completely dependent upon manual rules that ignores the personalized data [4]. However, once the difference between input and desired ages is large like 35–40 years, the changes in the shape of the face or texture of the skin becomes very pronounced and face generation becomes very challenging [5]. Also, factors like living styles, generic plastic surgery and lack of sizable labelled information, etc. play an important role in making the cross age face generation a challenging task. Such kind of problems can be solved with the help of Generative Adversarial Networks (GANs).

1.1 GANs

GANs based model is an unsupervised or semi-supervised deep learning model that has two neural networks. GANs generates synthetic data in the same domain as that of training data. These two different parts of GANs are known as generator and discriminator [6]. The latent vector of random noise values (z) is given as input to the generator that generates new fake data G(z). This generated data is indistinguishable from the real data. The discriminator works as a classifier that classifies the generated data as machine generated or real in terms of D(G(z)) = x. Both the networks are trained in adversarial way. The loss function for the model is given as follows [6].

$$[\min_G \max_D V(D, G) = E_x[log(D(x))] + E_z[log(1 - D(G(z)))]] \tag{1}$$

There is a competition between the generator and the discriminator. The discriminator is trained better than the generator, followed by classification of the data generated by generator correctly as machine-generated data. It penalizes generator with increasing loss due to which generator train itself to generate better fake samples and vice versa. The ultimate goal of training of networks is to minimize the loss for the generator, so that it can generate more realistic images.

As time is a continuous variable, face transformation should be demonstrated as a continuous process. It can be very difficult to do so without having ground truth character-specific data like different images of a person over his lifespan. It is very challenging to collect such type of training data [7]. An unsupervised learning nature of GANs model solves the problem of data. The image of a person in any age is sufficient to train the GANs based model. GANs process the input image and target age. It generates the image of same person in the target age.

As discussed earlier, generator takes a latent vector of random noise (z) as an input and generates the image in the domain of training samples. The characteristics of the image generated by the generator can't be controlled in simple vanilla GANs. In the given problem, the generator must generate the image of the given target age. There are different types of GANs. The normal classification include Deep colvolutional GANs (DCGANs), non-saturating GANs (NSGANs), Least squares GANs (LSGANs), Wasserstein GANs (WGANs), Wasserstein GANs-gradient penalty (WGANs-GP), deep regret analytic GANs (DRAGANs), boundary equilibrium GANs (BGANs) [8]. This work attempts to solve the discussed problem using CGANs.

1.2 Conditional Generative Adversarial Networks (CGANs)

In CGANs, the features of synthesized image can be controlled by giving target mode or category as an input to the generator and the discriminator. The latent vector of random noise (z) and target label vector (y) is given as an input to the generator. The generator processes the input image and generates the image in the given target mode. The generated image, real data, and target label vector are given to the discriminator. Discriminator classifies the image as a machine generated or real image as well as it classifies whether the generated image belongs to the target category or not. The CGANs model is trained using the following loss function [9].

$$[\min_G \max_D V(D, G) = E_x[log(D(x|y))] + E_z[log(1 - D(G(z|y)))]] \tag{2}$$

Along with the latent vector, a one-hot encoded vector of target age is given to the CGANs model. Hence, the CGANs model generates the image in the target age. In the proposed system, the input image is given by the user along with the target age. The target age can be accommodated in the system using CGANs. As discussed earlier, CGANs generates an image from a latent vector of random noise (z). The input image can't be given as an input to the CGANs. This problem is solved using Convolutional AutoEncoder [10].

1.3 Convolutional AutoEncoder

Convolutional AutoEncoder is used to create an optimized latent vector of the given input image [10]. Encoder and decoder are two neural networks used in Convolutional AutoEncoder. The input image is given to the encoder that is transformed into an n-dimensional latent vector. This latent vector is passed to the decoder as an input which converts it back into the original image. Different convolutional layers are used in the encoder network to down-sample the image into an n-dimensional latent vector. Contra-convolutional layers are used in the decoder network to up-sample the latent vector to create the original image.

Various techniques and models are elaborated in the Sect. 2. Research gaps between these techniques are discussed. To overcome these gaps the system is proposed in the Sect. 3. Implementation details for the system are given in the Sect. 4. The results of the system are evaluated on different performance metrics. The analysis of these results is given in the Sect. 5.

2 Literature Survey

The conventional models can be categorized into two categories. Physical model methodologies and prototype methodologies. Physical model processes shape and texture of every age group. For instance Suo et al. published a compositional model for facial aging. This model incorporates ageing features such as a change in skin and texture, change of face shape and the changes in hairstyles, etc. [2]. However, these methods are expensive computationally. Other model-based face progression methods incorporate, active appearance model, support vector regression and implicit function [5, 11]. Previous models focused on creating separate models for detecting each ageing feature like wrinkles, face sub region and craniofacial change. However, the AAM model uses a small dataset FGNET that contains images of a subject in different poses, lightning and environmental conditions that makes the training more challenging [2].

Recently a coupled dictionary learning model is published by Shu et al. It uses the dictionary bases to encode the face ageing patterns. Every two adjacent dictionaries are trained jointly. However this method still does not perform well in reconstructing the aged face [7].

Wang et al. proposed a Recurrent Face Ageing (RFA) framework using Recurrent Neural Network (RNN). The model has two layers having gated recurrent unit. The bottom layer transforms an image to a latent vector and it is decoded to a respective aged face through top layer. These procedures require adequate age successions as the training data, which is a drawback of this strategy [4].

Recently the GANs are proven better in many image related applications including super-resolution [12], pix2pix [13], CycleGANs [14], inpainting [15], visual manipulations on the images [16]. Various variants of GANs have been proposed like DCGANs [17], WGANs [18] etc. An empirical analysis of various types of GANs is given by Kokate et.al. [8].

Contextual Generative Adversarial Nets (C-GANs) is proposed by Liu et al. It takes ages as a condition to gain proficiency with ageing effects [19]. Nevertheless, existing methods need the paired samples as an input, i.e. face images of an individual at different age groups and some methods also require paired images over a long range of age span. Collection of such kind of data is challenging task. Antipo et al. presented the method for face ageing using CGANs. The major emphasize of this work is to preserve the individual's face identity by presenting an "identity-preserving" optimization of GANs latent vectors [20]. Song et al. have proposed a methodology for face ageing using dual conditional GANs. These methods have gained improved results than previous methods but the encoder is trained using a synthetic dataset [22]. This can be improved by using the original dataset while training the encoder.

3 Proposed Methodology

The proposed system incorporates two deep learning models - Convolutional AutoEncoder and Conditional Generative Adversarial Networks. Each of these models have two networks. Convolutional AutoEncoder consists of encoder network and decoder network. In CGANs generator and discriminator are two different networks. The decoder and discriminator are used to train the encoder and generator respectively. The input to the system is an image and the target age in which the image should be transformed. To train the CGANs, we must first convert the image into latent vector using encoder. For this purpose, the Convulutional AutoEncoder is trained before the training of CGANs model.

3.1 Training of AutoEncoder

In Convolutional AutoEncoder Model, convolutional layers are used in the encoder network to down-sample the image into an n-dimensional latent vector. Contra-convolutional layers are used in the decoder to up-sample the latent vector to create the original image. In the proposed methodology, 100-dimensional latent vector is considered. This size of latent vector can held all the important features of the input image [10]. These features has a significant part in the transformation of image into the target age.

There are four convolutional blocks and two fully connected layers in the encoder network. Every convolutional block has a convolutional layer, batch normalization layer and a LeakyReLU activation function. LeakyReLU function improves the training process in the initial phase as it activates even for less suitable data. All convolutional layers have the kernel of size of 5*5 along with stride two. After training

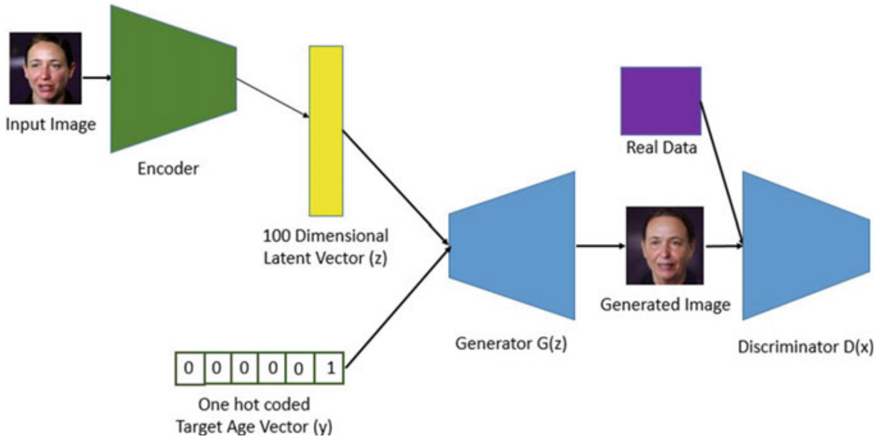


Fig. 1 System Architecture

of the Convolutional AutoEncoder, the encoder network is used while training the CGANs model.

3.2 Training of CGANs

As shown in Fig. 1, the System consists of three neural networks: Encoder, Generator, and Discriminator. The input and the output of the system are facial images with 64*64*3 resolution.

The trained encoder converts the input image into 100-dimensional latent vector. This vector along with the one-hot coded target age vector is passed to the generator. After processing the input, it generates 64*64*3 resolution image of a person in the target age category. The discriminator takes this 64*64*3 image generated by the generator as an input along with a one-hot coded vector of the target age category. After processing the input, the sigmoid layer gives output either 0 representing the generated image does not belong to the target age category or 1 representing the generated image belongs to the target age category. The system is expressed as following object function.

$$[\min_G \max_D V(D, G) = E_x[log(D(x|y))] + E_z[log(1 - D(G(Enc(z)|y)))]] \quad (3)$$

The generator consists of two dense layer blocks and three convolutional blocks. Each dense layer block has a dense layer, a a dropout layer and a batch normalization layer. Every convolutional block consists of up-sampling layer, convolutional layer with kernel size of 5*5. The discriminator consists of four convolutional blocks and one dense layer. Every convolutional block consists of a convolutional layer, batch normalization, and a LeakyReLU activation function. Each convolutional layer has a kernel size of 3*3 and two strides. The activation function for the last dense layer is sigmoid.

4 Experimental Setup

This section focuses to describe the details of experimentation environment and performance metrics considerations.

4.1 Data Collection

For training the system, two different datasets have been used. IMDB-Wikicleaned dataset with having 120,000 images [23] and Cross Age Celebrity Dataset (CACD) with having 163,446 images [24]. In both the datasets each data item consists of four different attributes. These attributes include date of birth, year in which the photo was taken, location of the face in the image and path to the image. The following equation is used to obtain age of a person.

$$[Age = PhotoTakenYear - BirthYear] \tag{4}$$

All the images are cropped using the face location given in the dataset. Reshaping is performed to convert the images into 64*64*3 size. All the images are then classified into age categories as specified in the Table 1. The dataset is split into 90:10 ratio for training set and testing set respectively. IMDB-Wikicleaned dataset is splitted with 108,000 images and 12,000 images in training set and testing set respectively. Similarly, CACD is splitted with 147,000 images in training set and 16,000 images in testing set.

Table 1 Age categories classification

Categories	Age Span
Category 1	0 to 18
Category 2	19 to 29
Category 3	30 to 39
Category 4	40 to 49
Category 5	50 to 59
Category 6	60+

4.2 Implementation Details

For training both the models, Nvidia DGX workstation has been used which has four NVIDIA TESLA V100 GPUs with 32 GB RAM each. Along with GPUs it has high performing Intel Xeon E5-2968 20-core CPU. It took 26 h to train both the models on the DGX workstation. First, the Convolutional AutoEncoder model is trained on the training dataset. The trained encoder is used in the training of the proposed system. The model is trained on the training dataset for 20,000 epochs. The adam optimizer is used to optimize the networks. For training the system on the GPU, Tensorflow latest GPU jupyter docker is used. GPU jupyter docker removes all hardware and software dependencies while building the system.

4.3 Performance Metrics

Following performance metrics are used to discuss the obtained results.

Age Estimation Error: To calculate the age estimation error the Face++ tool is used. It measures the difference of age distributions of both real and generated images in each age category. For generated images in each category, the age is estimated. The mean of the predicted ages is calculated for every category. The mean is then compared with the actual age. The difference between both of them is referred as age estimation error. A lower value of age estimation error indicates a more accurate prediction [25].

Frechet Inception Distance (FID): FID calculates the difference between two feature vectors calculated for real and generated images. FID score indicates the quality of image generated by the generator. The incepctioon v3 model with 1000 output neurons is used to calculate the FID. Inception score does not differentiate between generated and real images. FID removes this drawback. The score incorporates the quality as well as diversity of generated images. Lower FID score indicates greater similarity between real and generated images [27].

Identity Preservation: Face comparison experiments have been performed using the Face++ tool. Two images of a person in different age categories are provided as an input to the Face++. The output indicates the probability of detecting same person in the given input images. The average is taken for results for every pair of images for a single person. The percentage is averaged for each category separately [28].

5 Results and Discussions

Figure 2 indicates the genrated faces of different ages for an input image taken from IMDB-Wikicleaned dataset. Similarly, Fig. 3 indicates generated faces of different

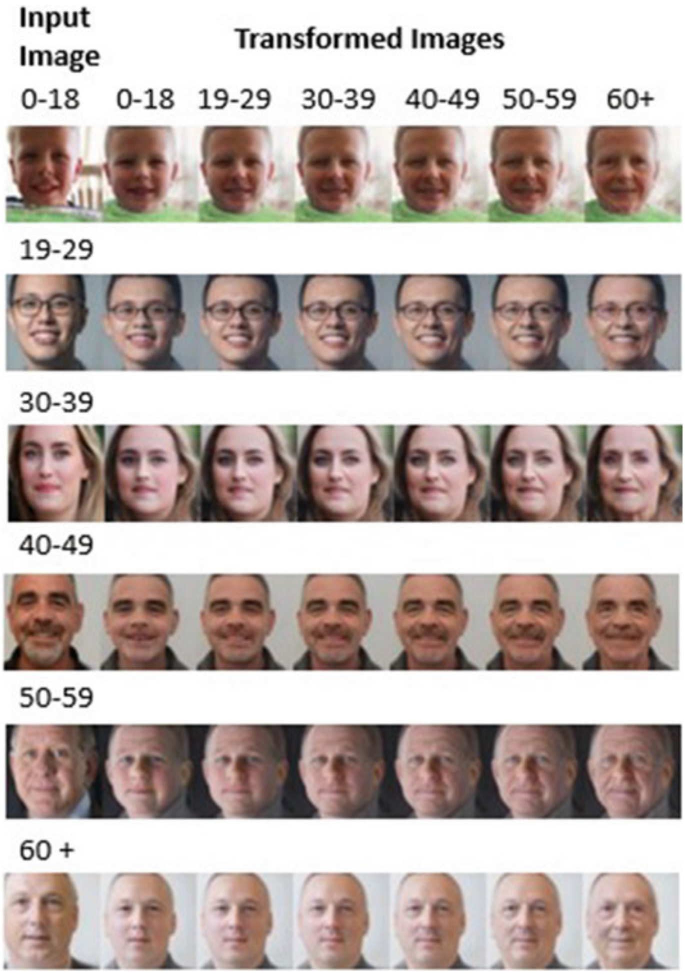


Fig. 2 Results generated by CGANs on IMDB-Wikicleaned dataset. The input image (1st column of each row) conditioned over different age categories (columns)

ages for an input image taken from CACD. For both Fig. 2 and Fig. 3, each row corresponds to the input image in a different age category and each column corresponds to six age categories as mention in Table 1.

Figure 2 and Fig. 3 shows that as the age increases, the shape of the face of the input image is transformed. The CGANs add some wrinkles on the forehead and other parts of the face. The significant difference can be seen around the jaw as per the age changes. The system takes care of the beard as well as it decreases the beard for 0–18 and 60+ category.

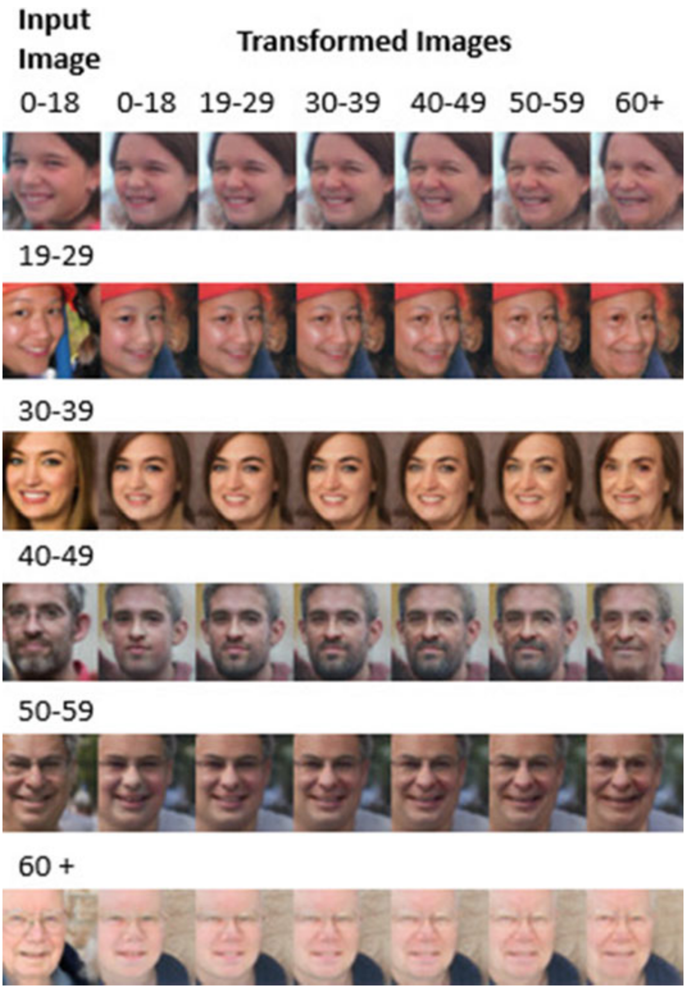


Fig. 3 Results generated by CGANs on CACD. The input image (1st column of each row) conditioned over different age categories (columns)

5.1 Age Estimation Error

The Face++ APIs is used to calculate the age estimation error. For IMDB-Wikicleaned dataset, age estimation error is lowest for category 0–18 category. The average error recorded is 2.35. For CACD, the lowest error recorded is 0.78 for category 60+ while average age estimation error is 3.10. Age estimation error has an opposite trend than FID because the age for extreme cases can easily be identified through the dominant features possessed by the images from these categories. Wrinkles, wide jaw, double

Table 2 Quantitative analysis

Dataset name	Age category	FID	Age estimation error	Identity preservation (%)
IMDB-Wikicleaned dataset	0–18	10.88	**0.82**	94.76
	19–29	8.5	2.67	97.86
	30–39	**3.5**	2.99	**99.76**
	40–49	5.9	3.45	99.29
	50–59	9.1	2.34	97.20
	60 +	12.5	1.87	96.27
	Average	**8.40**	**2.35**	**97.52**
Cross-Age Celebrity dataset	0–18	15.65	1.34	94.43
	19–29	13.97	3.54	95.23
	30–39	**8.21**	4.98	97.78
	40–49	12.87	4.09	**99.87**
	50–59	11.43	3.87	96.23
	60 +	18.51	**0.78**	95.11
	Average	**13.44**	**3.1**	**96.44**

chin are the dominant features in the 60+ category while fuller cheeks, no facial hairs represent the 0–18 age category.

5.2 FID Score

For IMDB-Wikicleaned dataset, FID score is the lowest for category 30–39 with value 3.5. The average FID score recorded is 8.4. For CACD, the lowest FID recorded is 8.21 for same category as that of IMDB-Wikicleaned dataset i.e. 30–39 while average FID score is 13.44. In an individual's lifespan, an person's face changes drastically in the first 20–25 years and after 50 years. But in between, no drastic changes happens in the person's face. That's why the FID score for middle four categories is lower than extreme two categories.

5.3 Identity Preservation

For IMDB-Wikicleaned dataset, the highest percentage is 99.76% for category 30–39. The avrage percentage is 97.52%. For CACD, the highest percentage recorded is 99.87% for category 40–49 while average percentage is 96.44%. As face changes drastically in the first 20–25 years, the changes slow down after it. Hence the verification for those categories is accurate than the first category.

Table 3 Comparative analysis

Methods	Identity preservation (%)	fAge estimation error
Age-cGANs	85.83	8.54
IPCGANs	96.90	3.54
CycleGANs	96.69	3.15
Proposed method	**96.82**	**2.75**

Table 2 shows the quantitative results for the performance parameters age estimation error, FID score and identity preservation.

5.4 Comparative Analysis

The results are compared with the latest methods like age-cGANs [20], IPCGANs [29] and face age progression using cycleGANs [30]. Identity preservation is the most important and challenging task while performing the cross age face generation. The exsisting works uses various methods for this task. The age-cGANs implemented Pixelwise Optimization and Identity-Preserving Optimization for identity perservation. IPCGANs uses age classifier along with identity preservation module. In both the works, specific formulae are used for the task which reduces the results. In both age-cGANs and IPCGANs, CGANs is used for transforming the image. In face age progrssion using cycleGANs, cycleGANs are used to transform the image from one domain to another domain. This model cannot generate the output image for the specified target age.

In this proposed work, identity preservation is implemented by deep learning base method using Convolutional AutoEncoder. This technique improves the results compare to existing approaches in terms of identity preservation and age estimation error. Table 3 describes the comparison between proposed work and the existing approaches. The encoder is trained using real data rather than synthetically generated data. This improves the latent vector optimization that aids to the better generation of the images. CGANs performs better that cycleGANs for transforming the image.

IPCGANs originally was trained on five age groups, similarly cycleGANs was trained on four age groups. For fair comparisons, the age groups are increased to six categories as mentioned in Table 1. The Face++ tool is used to calculate the identity preservation percentage for the results. Around 10,000 input images are randomly chosen from IMDB-Wikicleaned Dataset and CACD for quantifying the results.

The use of AutoEncoder for creating latent vectors of input image enhances the identity preservation for the individual. This improves the similarity between output images and the input image that results into better quantitative results than existing approaches.

6 Conclusion and Future Work

This work proposes a system to solve the problem of cross age face generation. The system consists of Convolutional AutoEncoder and Conditional Generative Adversarial Networks. The encoder is used for converting the input image into 100-dimensional latent vector. The generator uses this latent vector to generate images for an individual into all age categories. As Convolutional AutoEncoder is trained before the training of the CGANs model, real data can be used while training the system rather than synthetically generated data. This helps to improve the system performance in a real world. The system is evaluated on two datasets and quantitative results are measured. The model performs better on the age categories 30–39 and 40–49 than other age categories. It has achieved the highest 96.82% identity preservation with age estimation error as 2.75. The comparative analysis shows that the proposed model gives better outcomes than recent existing approaches. In this work, the focus is to generate cross age faces but it can also be used for other applications like changing the gender, changing the skin color or race of an individual in the given image. However, the performance for the 0–18 category is lowest compared with all other categories. This can be improved in order to achieve the state-of-the art results. The synthetically generated data can be used to increase the images in the dataset for this category. The super resolution GANs can be used to improve the quality of the output image. The pretrained models like styleGAN2 can be incorporate into the networks to achieve the better quality.

Acknowledgements We are thankful to the Department of Computer Engineering and Information Technology, College of Engineering Pune for providing the high computing GPU server facility procured under TEQIP-III (a world bank project) for the dissertation work.

References

1. Fu, Y., Guo, G., Huang, T.S.: Age synthesis and estimation via faces: a survey. IEEE Trans. Pattern Anal. Mach. Intell. **32**(11), 1955–1976 (2010)
2. Suo, J., Zhu, S.C., Shan, S., Chen, X.: A compositional and dynamic model for face aging. IEEE Trans. Pattern Anal. Mach. Intell. **32**(3), 385–401 (2009)
3. Hill, C.M., Solomon, C.J., Gibson, S.J.: Aging the human face-a statistically rigorous approach (2005)
4. Wang, W., et al.: Recurrent face aging. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2378–2386 (2016)
5. Lanitis, A., Taylor, C.J., Cootes, T.F.: Toward automatic simulation of aging effects on face images. IEEE Trans. Pattern Anal. Mach. Intell. **24**(4), 442–455 (2002)
6. Goodfellow, I., et al.: Generative adversarial nets. In: Advances in Neural Information Processing Systems, pp. 2672–2680 (2014)
7. Shu, X., Tang, J., Lai, H., Liu, L., Yan, S.: Personalized age progression with aging dictionary. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 3970–3978 (2015)
8. Kokate, P., Joshi, A.D., Tamizharasan, P.S.: An empirical comparison of generative adversarial network (GAN) measures. In: Hura, G.S., Singh, A.K., Siong Hoe, L. (eds.) Advances in

Communication and Computational Technology. LNEE, vol. 668, pp. 1383–1396. Springer, Singapore (2021). https://doi.org/10.1007/978-981-15-5341-7_105
9. Mirza, M., Osindero, S.: Conditional generative adversarial nets. arXivpreprint arXiv:1411.1784 (2014)
10. Cheng, Z., et al.: Deep convolutional autoencoder-based lossy image compression. In: 2018 Picture Coding Symposium (PCS). IEEE (2018)
11. Patterson, E., Sethuram, A., Albert, M., Ricanek, K., King, M.: Aspects of age variation in facial morphology affecting biometrics. In: 2007 First IEEE International Conference on Biometrics: Theory, Applications, and Systems, pp. 1–6. IEEE (2007)
12. Ledig, C., et al.: Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4681–4690 (2017)
13. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1125–1134 (2017)
14. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2223–2232 (2017)
15. Pathak, D., Krahenbuhl, P., Donahue, J., Darrell, T., Efros, A.A.: Context encoders: feature learning by inpainting. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2536–2544 (2016)
16. Zhu, J.-Y., Krähenbühl, P., Shechtman, E., Efros, A.A.: Generative visual manipulation on the natural image manifold. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9909, pp. 597–613. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46454-1_36
17. Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434 (2015)
18. Arjovsky, M., Chintala, S., Bottou, L.: Wasserstein gan. arXivpreprint arXiv:1701.07875 (2017)
19. Liu, S., et al.: Face aging with contextual generative adversarial nets. In: Proceedings of the 25th ACM International Conference on Multimedia, pp. 82–90 (2017)
20. Antipov, G., Baccouche, M., Dugelay, J.L.: Face aging with conditional generative adversarial networks. In: 2017 IEEE International Conference on Image Processing (ICIP), pp. 2089–2093. IEEE (2017)
21. Shu, Z., Yumer, E., Hadap, S., Sunkavalli, K., Shechtman, E., Samaras, D.: Neural face editing with intrinsic image disentangling. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5541–5550 (2017)
22. Song, J., Zhang, J., Gao, L., Liu, X., Shen, H.T.: Dual conditional GANs for face aging and rejuvenation. In: IJCAI, pp. 899–905 (2018)
23. Antipov, G., Baccouche, M., Berrani, S.A., Dugelay, J.L.: Apparent age estimation from face images combining general and children-specialized deep learning models. In: Proceedings of Computer Vision and Pattern Recognition Workshops, Las Vegas, USA (2016)
24. Chen, B.-C., Chen, C.-S., Hsu, W.H.: Cross-age reference coding for age-invariant face recognition and retrieval. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8694, pp. 768–783. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10599-4_49
25. Yang, H., Huang, D., Wang, Y., Jain, A.K.: Learning face age progression: a pyramid architecture of GANs. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 31–39 (2018)
26. Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., Chen, X.: Improved techniques for training GANs. In: Advances in Neural Information Processing Systems, pp. 2234–2242 (2016)
27. Heusel, M., et al.: GANs trained by a two time-scale update rule converge to a local Nash equilibrium. arXiv preprint arXiv:1706.08500 (2017)

28. Liu, Y., Li, Q., Sun, Z.: Attribute-aware face aging with wavelet-based generative adversarial networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11869–11878 (2019)
29. Wang, Z., et al.: Face aging with identity-preserved conditional generative adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2018)
30. Sharma, N., Sharma, R., Jindal, N.: An improved technique for face age progression and enhanced super-resolution with generative adversarial networks. Wirel. Pers. Commun. **114**, 2215–2233 (2020)

Synthesis of Stochastic Functions in Presence of Varied Degrees of Correlation

Shyamali Mitra, Debojyoti Banerjee, Sayantan Banerjee, and Mrinal Kanti Naskar

Abstract Correlation is an intriguing topic in stochastic computing and often leads to efficient solutions when applied to certain class of problems. There are various works to formalize the description of *Stochastic Function* (SF) implemented using correlation in numbers. In the present work, a heuristic mathematical study has been conducted using the properties of correlation interaction in *Stochastic Numbers* (SNs). Relevant theorems are developed based on the findings with a motto to illustrate the necessary mathematical background on correlation injection in SNs. We further conceptualized the idea to synthesize complex Boolean functions consisting of n variables using simplistic approach.

1 Introduction

Correlation in stochastic numbers is an important property that must be taken into account while designing probabilistic circuits [1–4]. Varied degrees of correlation between SNs arise from the way numbers are generated from LFSRs (Linear Feedback Shift Registers) [5]. Therefore, to generate two uncorrelated sequence usually two independent LFSRs are used. Whereas, to generate positively correlated sequence, numbers are generated from same LFSR. It is interesting to note that, correlation in SNs is an obvious phenomenon and can altogether change the functionality implemented by the circuit [6]. There can be varied degrees of correlation that can exist between two SNs and is primarily estimated with Stochastic Correlation Coefficient (SCC). As for example, at $SCC = 0$ two SNs are uncorrelated, for $SCC = +1$ two SNs are positively correlated and for $SCC = -1$, two SNs are negatively correlated. With these three crisp values of SCCs, it is interesting to note that this always leads to some different functionalities when each of these SCCs are given individual treatment. Correlation between two bitstreams has been identified

S. Mitra (✉) · D. Banerjee · S. Banerjee
Department of IEE, Jadavpur University, Kolkata, India
e-mail: shyamalimitra.iee@jadavpuruniversity.in

M. K. Naskar
Department of ETCE, Jadavpur University, Kolkata, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_6

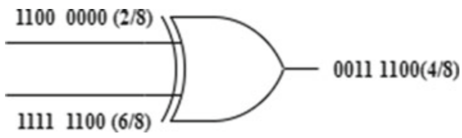


Fig. 1 XOR gate as absolute subtractor when inputs are positively correlated

as a major source of inaccuracy in certain stochastic circuits [7]. But elaborate discussion on the effect of these correlations on SNs and analysis of *Stochastic Function* (SF) is still pending that grabs much of researcher's attention.

In Stochastic computing correlation indicates that the bitstreams generated by LSFR [5] or SNG [8] inherit some sort of dependence between them (cross correlation) or between the bits of the same bitstream (auto correlation). Correlation in stochastic circuits was identified as an inaccuracy caused by a pair of bitstream, e.g., multiplication operation using an AND gate. In recent researches correlation has been identified as a means to generate useful stochastic functions requiring less hardware [6]. To quantify the correlation between input bitstreams X and Y, SCC which is analogous to the similarity coefficient [9] is defined as,

$$SCC(X,Y) = \begin{cases} \frac{p_{X\cap Y} - p_X \cdot p_Y}{min(p_X, p_Y) - p_X p_Y}, \; p_{X\cap Y} > p_X p_Y \\ \frac{p_{X\cap Y} - p_X \cdot p_Y}{p_X p_Y - max(p_X + p_Y - 1, 0)}, \; \text{otherwise} \end{cases} \tag{1}$$

where, $p_{x\cap y}$ is obtained by performing bitwise AND operation between X and Y. Other generalized way of representing the SCC is:

$$SCC(X,Y) = \begin{cases} \frac{N_{11}.N_{00} - N_{01}.N_{10}}{N.min(N_{11}+N_{10}, N_{11}+N_{01}) - (N_{11}+N_{10})(N_{11}+N_{01})}, \; N_{11}.N_{00} > N_{01}.N_{10} \\ \frac{N_{11}.N_{00} - N_{01}.N_{10}}{(N_{11}+N_{10})(N_{11}+N_{01}) - N.max(N_{11} - N_{00}, 0)}, \; N_{11}.N_{00} < N_{01}.N_{10} \end{cases}$$

N_{11}, N_{10}, N_{01} and N_{00} are the respective overlaps of X and Y. Let, $X = 11001111$, and $Y = 01001111$, then, $SCC = +1$. But, if $X = 11110001$, $Y = 11111100$, the corresponding $SCC = 0.2$. In this case, not every 1 in Y is influenced by the presence of 1 in that particular position in X. There is overlapping of $0's$ in X and $1's$ in Y as well as $1's$ in X and $0's$ in Y. Thus, the pair of bitstream is positively correlated to a certain degree. But correlation has also been found to influence the circuits' behaviour in a positive way as well [10].

For inputs with two extreme values, the measure of SCC becomes indeterminate. In both these cases, it is impossible to change the undefined SCC value with the help of any external circuit such as correlator. Consider two SNs, $X = 00000000$ and $Y = 11111111$. Logic operations on these numbers will produce output that will stick to the boundary values itself, either 0 or 1. Attempts to change the correlation between these numbers will result in a change in probability value which is undesired. Using a correlator circuit such as [6] will not be able to alter the degree of correlation between X and Y because only grouping of one kind of bit-pair (here 01) is possible and we lose the leverage of pairing other three bit pairs. For any degree of correlation, we can write the output p_z as a linear equation [11], given as:

$$p_z = (1 + SCC)F_0 - SCC.F_{-1}, \quad \forall\, SCC < 0 \tag{2}$$

$$p_z = (1 - SCC)F_0 + SCC.F_{+1}, \quad \forall\, SCC > 0 \tag{3}$$

where, F_0, F_{-1}, $F{+}1$ are the functions realized with SCC values of 0, –1, +1.

In the present work, a detailed study on correlation has been carried out that can unfold its potential to classify and identify different functions based on given inputs. Also, a strong mathematical foundation is presented to comprehend the nature of manipulation that can be done on SNs in presence of different correlations. We also explore the implications of the definitions and properties of correlation to generate Stochastic Functions (SFs) consisting of 2 variables. We further extended this idea to compute SFs for higher order variables using the proposed algorithms of correlation injection. The contribution in the present work can be precisely stated as:

– We demonstrated various theorems to lay a strong mathematical foundation on the theory of correlation induction in SNs.
– A correlation based framework has been developed to synthesize functions of 2 variables. The idea has been extended to synthesize complex function of n variables.

2 Correlation-Sensitive Logic Elements

There are some correlation-sensitive logic blocks e.g., AND, XOR and OR gate whose functionalities change drastically when correlation is altered.

i) AND gate: Arithmetic multiplication can be done using an AND gate when SCC between bitstreams is assumed to be 0 as shown in Fig. 2. When $SCC = 1$, AND gate computes minimum of two numbers. When $SCC = -1$, theSFimplemented by AND gate is $p_z = max(p_x + p_y - 1, 0)$. Figure 3 shows different SFs implemented at different correlation status.

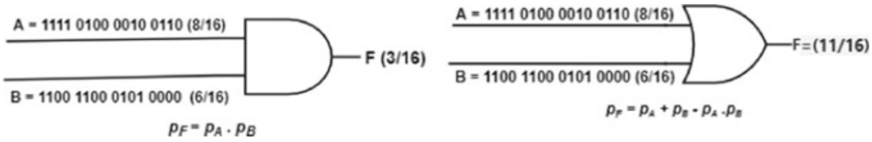


Fig. 2 SFs implemented by AND and OR gate at $SCC = 0$

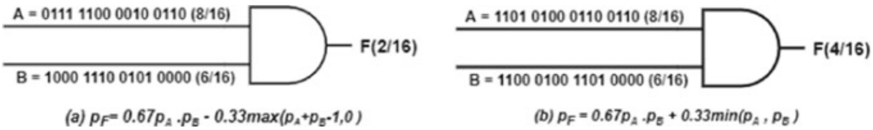


Fig. 3 SFs implemented by an AND gate at **a**) $SCC = -0.33$, **b**) $SCC = +0.33$

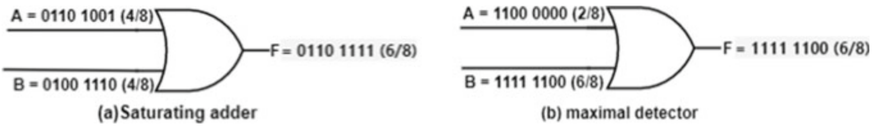


Fig. 4 SFs implemented by an OR gate at **a**) $SCC = -1$, **b**) $SCC = +1$

ii) XOR gate: One interesting property of XOR gate is that it acts as absolute subtractor when inputs are positively correlated (see Fig. 1). Implementing the same function using binary inputs increases hardware complexity [12].
iii) OR gate: OR gate acts as a maximum detector at maximally positive correlation while it acts as a saturating adder at maximally negative correlation as shown in Fig. 4.

3 Theorems on Correlation Intervention in Stochastic Numbers

In Stochastic Computing the value of SN can be represented in terms of probabilities which is given by the number of $1's$ in the bitstream. Consider a Boolean network consisting of k inputs $x_1, x_2, ...x_k$. This is realized by a sequence of gates where inputs to each gate in the network is obtained from the outputs of the previous gates. For a single output function, the output is usually obtained from the output node of the last gate of the Boolean network. Suppose $n-$ bit SNs are inputs to Boolean Logic network where all the stochastic logic manipulations are valid. The output from the logic network is $n-$ bit SN and the logic operations within the network are given in terms of intermediate SFs.

3.1 Complement of an SN

The complement of an SN, B having probability p_B can be obtained by replacing p_B with $1 - p_B$.

3.2 Correlation Status Between Two SN's: A Case Study

Let us suppose, SN A and SN B be two real-valued random variables having Bernoulli distribution, where the probability of having a $'1'$ in their bitstreams are represented by p_A and p_B respectively. It has been studied [11] that correlation between two SNs can be successfully represented by corresponding SCC. The SCC between two SNs, A and B may be viewed as the difference between $A \wedge B$ and product of the individual probabilities p_A and p_B (when numbers are uncorrelated in nature). The values of SCC between two stochastic numbers are scaled from $+1$ to -1. For two positively correlated SNs, $SCC = +1$ confirms maximum overlap of $11's$ and $00's$. For two negatively correlated SNs, $SCC = -1$ results in minimum overlap of $11's$ and $00's$ or maximum overlap of $10's$ and $01's$. When $SCC = 0$, an equilibrium between all four combinations of overlaps (00, 01, 10, 11) can be observed between the SNs. For uncorrelated numbers $SCC = 0$ signifies that, $A \cap B = p_A p_B$. The values of SCC can be found out using definition of Eq. 1.

From the definitions of SCC given in Eq. 1, six possible situations may arise when two SNs are either positively or negatively correlated.

Property I: When two SNs A and B are positively correlated there are three possible correlation dependencies between A and B.

- $\bar{A}$ and $\bar{B}$ are positively correlated.
- $\bar{A}$ and B are negatively correlated.
- A and $\bar{B}$ are negatively correlated.

Example 1: Let A and B are two positively correlated SNs with $p_A = \frac{3}{8}$ and $p_B = \frac{4}{8}$, represented by the bitstreams 11100000 and 11110000 respectively, then from the definition of SCC [11], following three cases may arise:

1. $\bar{A}$ is given by 00011111 and $\bar{B}$ is given by 00001111, then from the definition of SCC, then $SCC_{AB} = +1$ and hence they are positively correlated.
2. $\bar{A}$ is given by 00011111 and B is given by 11110000, then $SCC_{AB} = -1$ and hence they are negatively correlated.
3. Similarly A is given by 11100000 and $\bar{B}$ is given by 00001111, then $SCC_{AB} = -1$ and hence they are negatively correlated.

Property II: Similarly, when two SNs A and B are negatively correlated then correlation can be defined in three ways:

- $\bar{A}$ and $\bar{B}$ are negatively correlated.
- $\bar{A}$ and B are positively correlated.
- A and $\bar{B}$ are positively correlated.

Example 2: Let SN A and SN B are negatively correlated with $p_A = \frac{3}{8}$ and $p_B = \frac{2}{8}$ and are represented by bitstreams 11100000 and 00000011 respectively. From the definition of SCC [11], following three cases are considered:

1. $\bar{A}$ is given by 00011111 and $\bar{B}$ is given by 11111100. Thus, corresponding $SCC = -1$ and hence they are negatively correlated.
2. Similarly, $\bar{A}$ is given by 00011111 and B is given by 00000011, the corresponding $SCC = +1$ and hence are positively correlated.
3. Similarly A is given by 11100000 and $\bar{B}$ is given by 11111100, the corresponding $SCC = +1$ and hence are positively correlated.

The following two theorems would comply the preceeding statements.

Theorem I: *If SCC between A and B is SCC_{AB}, then SCC between A and $\bar{B}$ is given by $-SCC_{AB}$, provided $SCC_{AB} > 0$.*

Proof: We assume that $p_A > p_B$ and $SCC_{AB} > 0$ without the loss of generality. Now, $p_{A \cap B} = \frac{N_{11}}{n}$, $p_A = \frac{N_{11}+N_{10}}{n}$, $p_B = \frac{N_{11}+N_{01}}{n}$. For $p_A > p_B$ and $SCC_{AB} > 0$, $min(p_A, p_B) = p_B$. Thus, from the definition of SCC,

$$SCC_{AB} = \frac{\frac{N_{11}}{n} - \frac{N_{11}+N_{10}}{n} \cdot \frac{N_{11}+N_{01}}{n}}{\frac{N_{11}+N_{01}}{n} - \frac{N_{11}+N_{10}}{n} \cdot \frac{N_{11}+N_{01}}{n}}$$

$$SCC_{AB} = \frac{nN_{11} - (N_{11} + N_{10})(N_{11} + N_{01})}{(N_{11} + N_{01})(N_{00} + N_{01})} \tag{4}$$

Also, $n = (N_{11} + N_{10} + N_{01} + N_{00})$. Substituting the value of n we get,

$$SCC_{AB} = \frac{N_{00} \cdot N_{11} - N_{01} \cdot N_{10}}{(N_{11} + N_{01})(N_{00} + N_{01})} \tag{5}$$

If $SCC_{AB} > 0$, then, $N_{11}N_{00} > N_{10}N_{01}$. In order to find SCC between A and $\bar{B}$, we need to find whether $p_{A \cap \bar{B}}$ is greater than $p_A p_{\bar{B}}$ or not.

$$p_{A \cap \bar{B}} = \frac{N_{10}}{n}, \quad p_{\bar{B}} = \frac{N_{00} + N_{10}}{n}, \; p_{A \cap \bar{B}} - p_A \cdot p_{\bar{B}} = N_{10} \cdot N_{01} - N_{11} \cdot N_{00}$$

From the expression, $p_{A \cap \bar{B}} - p_A p_{\bar{B}}$ is < 0. Hence,

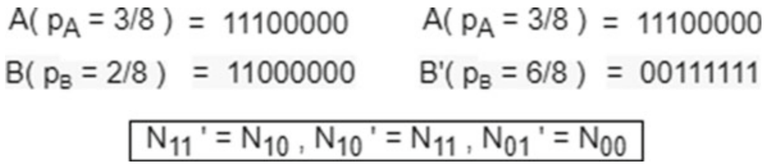


Fig. 5 A schematic diagram showing different bit overlaps in A and B

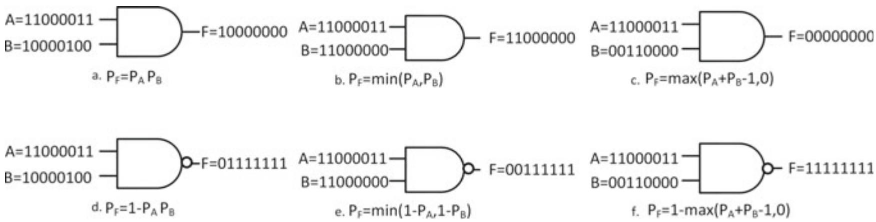


Fig. 6 SFs implemented by AND and NAND logic at different correlation status; **a** and **d** at $SCC = 0$; **b** and **e** at $SCC = +1$; **c** and **f** at $SCC = -1$

$$SCC_{A\bar{B}} = \frac{p_{A\cap\bar{B}} - p_A \cdot p_{\bar{B}}}{p_A \cdot p_{\bar{B}} - max(p_A + p_{\bar{B}} - 1, 0)} = \frac{N_{10} \cdot N_{01} - N_{11} \cdot N_{00}}{(N_{11} + N_{10}) \cdot (N_{00} + N_{10}) - (N_{10} - N_{01}) \cdot n}$$
$$= \frac{N_{10} \cdot N_{01} - N_{11} \cdot N_{00}}{N_{01} \cdot (n - N_{10}) + N_{11} \cdot N_{00}} = \frac{N_{10} \cdot N_{01} - N_{11} \cdot N_{00}}{(N_{11} + N_{01})(N_{00} + N_{01})}$$

which is equivalent to $-SCC_{AB}$.

Theorem II: *If SCC between A and B is SCC_{AB}, then SCC between SN $\bar{A}$ and SN $\bar{B}$ is given by SCC_{AB}, if $SCC_{AB} > 0$ is satisfied.*

Proof: With reference to Theorem I, it is understood that when SCC between A and B is greater than 0, SCC_{AB} can be written as:

$$SCC_{AB} = \frac{N_{00} \cdot N_{11} - N_{01} \cdot N_{10}}{(N_{11} + N_{01})(N_{00} + N_{01})} \tag{6}$$

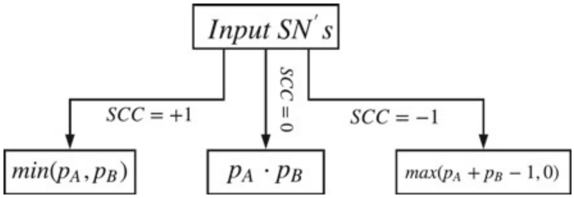
In order to find out the SCC between $\bar{A}$ and $\bar{B}$, the probabilities of the bitstreams required to represent the expression is given as:

$$\begin{gathered} p_{\bar{A}\cap\bar{B}} = \frac{N_{00}}{n}, \quad p_{\bar{A}} = \frac{N_{00} + N_{01}}{n}, \quad p_{\bar{B}} = \frac{N_{00} + N_{10}}{n}, \\ p_{\bar{A}\cap\bar{B}} - p_{\bar{A}} \cdot p_{\bar{B}} = \frac{N_{11}}{n} - \frac{N_{01} + N_{00}}{n} \cdot \frac{N_{10} + N_{00}}{n} = \frac{N_{11} \cdot N_{00} - N_{10} \cdot N_{01}}{n^2} \end{gathered} \tag{7}$$

Since, $SCC_{AB} > 0$ and $N_{11}N_{00} > N_{10}N_{01}$,we can write,

$$SCC_{\bar{A}\bar{B}} = \frac{p_{\bar{A}\cap\bar{B}} - p_{\bar{A}} \cdot p_{\bar{B}}}{min(p_{\bar{A}}, p_{\bar{B}}) - p_{\bar{A}} \cdot p_{\bar{B}}} \tag{8}$$

Fig. 7 A schematic diagram of functions implemented by two SN's with different correlation status



As $p_A > p_B$, $(1 - p_A) < (1 - p_B)$ which further says $p_{\bar{A}} < p_{\bar{B}}$. Thus,

$$SCC_{\bar{A}\bar{B}} = \frac{p_{\bar{A}\cap\bar{B}} - p_{\bar{A}} \cdot p_{\bar{B}}}{p_{\bar{A}} - p_{\bar{A}} \cdot p_{\bar{B}}} \tag{9}$$

$$SCC_{\bar{A}\bar{B}} = \frac{N_{11} \cdot N_{00} - N_{10} \cdot N_{01}}{(N_{11} + N_{01}) \cdot (N_{00} + N_{01})} \tag{10}$$

which is equivalent to SCC_{AB} in Eq. 6.

It can also be proved that if SCC between A and B is SCC_{AB}, then SCC between $\bar{A}$ and B is given by $-SCC_{AB}$, provided $SCC_{AB} > 0$. The three cases of Property II can also be proved using a similar foregoing approach. An example showing different bit overlaps between A and B given in Fig. 5.

3.3 Correlation Interaction Between Two SN's

Different correlation injections between two SNs always lead to different functions being implemented as shown in Fig. 6. Three cases of correlation interactions between SNs can be investigated and is shown in Fig. 7.

i) Case I: Two SNs are uncorrelated: Whenever two SNs are generated from two independent LFSRs, they are uncorrelated, i.e., $SCC = 0$ and the SF is obtained by replacing variables in the Boolean expression with their corresponding probabilities i.e., A with p_A and B with p_B. Thus, for the Boolean expression $F = A \cdot B$, the SF, denoted by p_F is given as $p_F = p_A p_B$.

Example 3: If SN A with $p_A = \frac{4}{8}$ represented by 11000011 and SN B with $p_B = \frac{2}{8}$ represented by 10000100 are given inputs to an AND gate, the output F can be given by 10000000 with $p_F = \frac{1}{8}$, which can be also obtained by multiplying p_A with p_B i.e., $p_F = p_A \cdot p_B = \frac{1}{8}$.

ii) Case II: Two SNs are positively correlated: Whenever two SNs are generated from same LFSR, they are positively correlated, the SF corresponding to the Boolean function is obtained as, $p_F = min(p_A, p_B)$.

Theorem III: *If two SNs, A and B having probabilities p_A and p_B are generated from same LFSR, such that they exhibit positive correlation between them, the SF corresponding to Boolean AND operation is given as $p_F = min(p_A, p_B)$.*

Proof: Let A' and B' be two 8 bit binary numbers and $A' \geq B'$ without the loss of generality. The 256-bit SNs for A and B can be generated corresponding to A' and B' by comparing two binary numbers with the LFSR for 256 clock pulses. The output produced in this comparison is a stream of binary 1s and 0s depending upon the relative value of the LFSR and input numbers. However, we have assumed that $A' \geq B'$, so if a 1 is produced in B while comparing the content of LFSR with B' it implies that A bitstream will definitely contain a 1. However, reverse is not true. So, the nature of LFSR dictates that, wherever there are $1's$ in B, there must be $1's$ in A, to confirm the maximum overlap of B with A. When we perform multiplication of two such SNs we will get B. The converse can be proved by taking $B' \geq A'$.

Example 4: If A with $p_A = \frac{4}{8}$ is represented by 11000011 and B with $p_B = \frac{2}{8}$ is represented by 11000000 are given inputs to AND gate, the output SN F is given by 11000000 with $p_F = \frac{2}{8}$. Thus, $p_F = min(p_A, p_B) = min(\frac{4}{8}, \frac{2}{8}) = \frac{2}{8}$.

iii) Case III: Two SNs are negatively correlated: Let two SNs, A and B having probabilities p_A and p_B are negatively correlated, then SF is obtained from the Boolean AND function of the two variables as, $p_F = max(p_A + p_B - 1, 0)$.

Corollary 1: *If two SNs A and B having probabilities p_A and p_B are negatively correlated, then the SF corresponding to the Boolean AND function is given as, $p_F = max(p_A + p_B - 1, 0)$.*

Proof: Let A and B be two SNs with $SCC = -1$. Thus, A and $\bar{B}$ are positively correlated with $SCC = +1$. This can be established by considering the bitwise flipping of SN B. So, minimal overlap between A and B is converted to maximal overlap between A and $\bar{B}$. Now, $A \cdot B = A(1 - \bar{B})$. Thus, $p_F = p_A - min(p_A, 1 - p_B)$.

We know that,

$$min(p_A, p_B) = \frac{(p_A + p_B) - |p_A - p_B|}{2} \tag{11}$$

$$max(p_A, p_B) = \frac{(p_A + p_B) + |p_A - p_B|}{2} \tag{12}$$

Hence, from Eq. 11

$$p_F = p_A - \frac{(p_A + (1 - p_B)) + |p_A - (1 - p_B)|}{2}$$

$$p_F = \frac{(p_A + p_B - 1) + |p_A + p_B - 1|}{2}$$

$$p_F = \frac{((p_A + p_B - 1) + 0) + ||p_A + p_B - 1)| - 0|}{2}$$

Table 1 SFs implemented by different logic elements for different values of SCC

Gates	$SCC = +1$	$SCC = 0$	$SCC = -1$
AND	$min(p_A, p_B)$	$p_A \cdot p_B$	$max(p_A + p_B - 1, 0)$
OR	$max(p_A, p_B)$	$p_A + p_B - p_A \cdot p_B$	$min(p_A + p_B, 1)$
NAND	$max(1 - p_A, 1 - p_B)$	$1 - p_A \cdot p_B$	$1 - max(p_A + p_B - 1, 0)$
NOR	$min(1 - p_A, 1 - p_B)$	$(1 - p_A) \cdot (1 - p_B)$	$max(1 - p_A - p_B, 0)$
EXOR	$\|p_A - p_B\|$	$p_A + p_B - 2 \cdot p_A \cdot p_B$	$1 - \|p_A + p_B - 1\|$
EXNOR	$1 - \|p_A - p_B\|$	$1 - (p_A + p_B - 2 \cdot p_A \cdot p_B)$	$\|p_A + p_B - 1\|$

$$p_F = max(p_A + p_B - 1, 0) \tag{13}$$

Example 5: Consider A with $p_A = \frac{4}{8}$ represented by 11000011 and B with $p_B = \frac{2}{8}$ by 00110000, then output is 00000000 with $p_F = \frac{0}{8}$. Validating the result with *Corollary* 1, we get $p_F = max(p_A + p_B - 1, 0) = max(\frac{4}{8} + \frac{2}{8} - 1, 0) = 0$.

4 A Correlation Based Framework for Minterm Calculation

4.1 Synthesizing Functions of Two Variables

In case of two variables, the method that is followed for determination of the function executed by the boolean expression when stochastic number is given as an input to the circuit is described as follows:

1. Calculate the SCC between the two numbers.
2. Three extreme cases are followed up; $SCC = +1, 0, -1$.
   - If $SCC = +1$ then the function can be described using Theorem III.
   - If $SCC = -1$ then the function implemented is described in Corollary 1.
   - If $SCC = 0$ then the function implemented by the given expression is described in Sect. 3.3.
3. For any intermediate correlation between numbers the SF implemented for a given Boolean expression is a linear combination of the two functions implemented in the extreme cases and is given in Eq. 2,3.

The method may be used to validate the SFs mentioned in Table 1 for a two variable boolean expression when the correlations between them are 0, +1 or –1.

Example 6: When two SNs, A and B are positively correlated $SCC = +1$, the SF implemented by the minterm $A \cdot \bar{B}$ is given by $max(p_A - p_B, 0)$. This can be illustrated using the following logic.

If A and B are positively correlated, then A and $\bar{B}$ are negatively correlated as described in Sect. 3.2. If B is represented by p_B, then $\bar{B}$ is represented by $1 - p_B$, and the function can be deduced from Sect. 3.3 as $p_Z = max(p_A + (1 - p_B) - 1, 0) = max(p_A - p_B, 0)$.

Example 7: From Table 1, when SN A and SN B are negatively correlated, the SF implemented by the minterm $\bar{A} \cdot B$ is given by $min(1 - p_A, p_B)$. This can be deduced by the above mentioned method.

If A and B are negatively correlated then $\bar{A}$ and B are positively correlated as described in Sect. 3.2. If A is represented by the probability p_A then $\bar{A}$ is represented by the probability $1 - p_A$ the function implemented using Sect. 3.3 as $p_Z = min(1 - p_A, p_B)$.

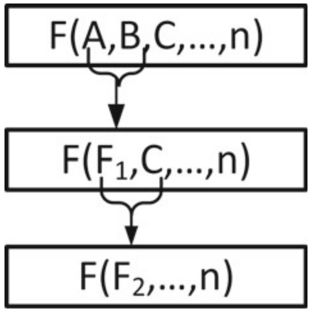
The other cases mentioned in Table 1 may also be verified in a similar manner.

4.2 Synthesizing Functions of More Than 2 Variables

In this section, we develop stochastic functions for complex Boolean functions using a correlation-based framework for correlation-sensitive stochastic logic circuits. Two methods are proposed that compute SFs corresponding to complex Boolean functions consisting of n variables. Successive variable approach deals with logic variables in a rather straight forward manner. Whereas the second approach is selective towards choosing variables based on their correlation status.

I) Successive Variable Approach: In case of three or more variables the calculation of the minterms becomes a rigorous process and can be solved programmatically as the variables may be related each other with various degrees of SCC among them. To solve minterms consisting of more than two variables, we adopt a bottom-up approach. For the purpose of analysis and implementation of SFs, we first consider any two variables at a time and compute the SCC between them. Next, find the function implemented by them to compute the result. Further, calculate the next output using the previous output and the SCC with the next variable and continue till the last variable is encountered to arrive to the resultant SF. A schematic to construct SFs using this approach is shown in Fig. 8.

Fig. 8 A schematic diagram for computing resultant SF for a function with more than two variables, where F_1 and F_2 are calculated according to Method I



Example 8: To construct SF for the Boolean function, $F = A \cdot \bar{B} \cdot C$, where, A with probability $p_A = \frac{5}{8}$ is given as 11111000, B with $p_B = \frac{3}{8}$ is given as 11100000, and C with $p_C = \frac{4}{8}$ given as 11110000.

Consider first two variables of the Boolean expression and let us denote the intermediate Boolean expression as F_1 and the SF implemented by them as SF_1. Since, A and B are positively correlated, A and $\bar{B}$ are negatively correlated [Theorem I], the function implemented by $A \cdot \bar{B}$ is $SF_1 = max(p_A - p_B, 0) = max(\frac{5-3}{8}, 0)$ which gives $p_{F_1} = \frac{2}{8}$. This can be validated by bitwise operation of $A \cdot \bar{B}$ which is 00011000($\frac{2}{8}$). In the next step, we compute the correlation between F_1 and C, which dictates that they are uncorrelated i.e., $SCC = 0$. Thus, SF implemented by this combination is $p_{F_1} \cdot p_C$ which is $\frac{2}{8} \times \frac{4}{8} = \frac{1}{8}$. The result can be validated by computing $F_1 \cdot C$ which is 00010000($\frac{1}{8}$).

II) A Selective-Based Approach to Positively or Negatively Correlated Variables: When dealing with complex functions consisting of n variables the successive variable approach may not be that efficient to construct SFs. The selective approach deals with analyzing SFs by first decomposing input variables that exhibit either positive or negative correlations between them. In order to compute the function implemented by the given minterms the following steps may be followed to implement the given boolean function:

1. If the variables (say A,B,C) are positively (or negatively) correlated, compute $A \cdot B$, $\bar{A} \cdot \bar{B}$, $B \cdot C$, $\bar{B} \cdot \bar{C}$, $A \cdot C$ and $\bar{A} \cdot \bar{C}$. This operations does not change the nature of correlation between the numbers.
2. If none of the above operations remain then compute the other operations.

Algorithm 1 is shown to depict the flow of computation of SFs when the consecutive variables are either positively or negatively correlated. Figure 9 shows a schematic diagram for computing the required function implemented for more than two SNs. The final output is calculated using the output of the last gate via intermediate logic within the Boolean network according to the rules 1,2 of selective-based approach of correlation interaction between SNs.

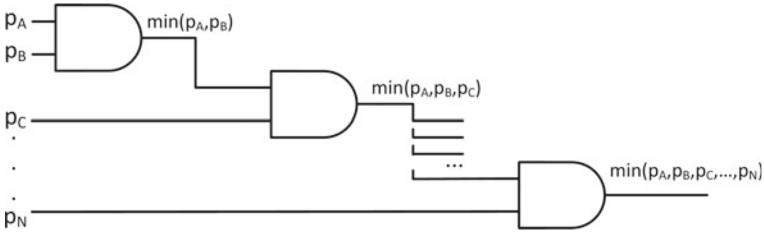


Fig. 9 A sample circuit to find resultant SF for a function with more than two variables, where F_1 and F_2 are calculated according to Method II

Example 9: We can compute Example 8 using selective approach. SN's A, B and C are assumed to have similar values as in the previous example.

Instead of computing $A \cdot \bar{B}$, we first compute $A \cdot C$. From the nature of the bit-stream it is evident that, SN A and SN C are positively correlated with $SCC = +1$. Thus $A \cdot C$ gives us $min(p_A, p_C)$ i.e., $min(\frac{5}{8}, \frac{4}{8}) = \frac{4}{8}$. Let us denote the intermediate result as F_2. Clearly, F_2 has same nature as SN A or SN C. F_2 is positively correlated with SN B and hence negatively correlated with SN $\bar{B}$. Thus, function implemented by $F_2 \cdot \bar{B}$ is $F = max(p_{F_2} - p_B, 0)$ i.e., $max(\frac{4-3}{8}, 0) = \frac{1}{8}$, which is identical with the result as in the previous case.

The selective approach towards variables let us infer that whether the variables are having uncomplemented or complemented, whose SCCs are already known, can be transformed into functions that calculates either minimum or maximum of the input values. It also gives us flexibility to simplify functions using De Morgans' Theorem.

Theorem IV: *If $p_{A_1 \cap A_2 \cdots \cap A_n} > p_{A_1 \cap A_2 \cdots \cap A_{n-1}} \cdot p_{A_n}$, then SCC between $F_1 = A_1 \cap A_2 \cdots \cap A_{n-1}$ and A_n is $+1$.*

Proof: We know that,

$$SCC_{F_1 \cap C} = \frac{p_{F_1 \cap C} - p_{F_1} \cdot p_C}{min(p_{F_1 \cap C}, p_C) - p_{F_1} \cdot p_C}$$

and $p_{F_1 \cap C} \leq p_C$. Hence, $min(p_{F_1 \cap C}, p_C) = p_{F_1 \cap C}$

$$SCC_{F_1 \cap C} = \frac{p_{F_1 \cap C} - p_{F1} \cdot p_C}{p_{F_1 \cap C} - p_{F_1} \cdot p_C} = 1$$

It can be observed that the above condition holds when $n(A) \subset n(B) \subset n(C)$, which further dictates $p(A) > p(B) > p(C)$ and $SCC = 1$ (for $n = 3$).

Let A with $p(A) = \frac{4}{8}$ and B with $p(B) = \frac{5}{8}$ with SCC = –1 and C with $p(C) = \frac{4}{8}$ is negatively correlated with B, such that C is positively correlated with A as shown in Fig. 10.

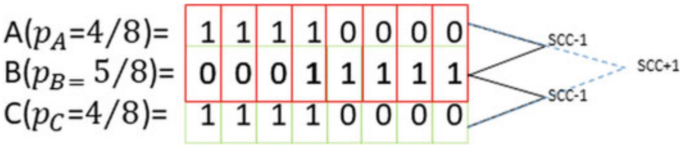


Fig. 10 A schematic diagram for computing SF of two remote variables

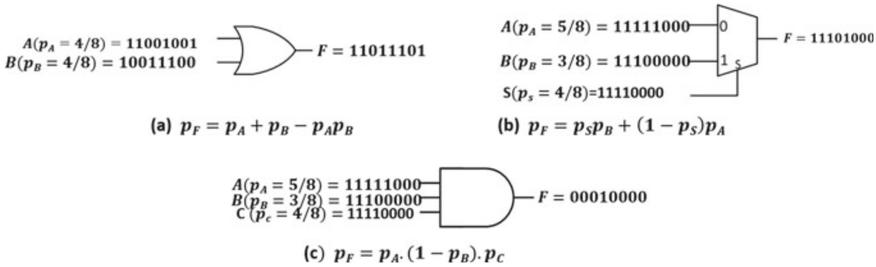


Fig. 11 Some standard Boolean functions and their corresponding SFs

Conversely, it can be told that if we have SCCs specified between any two consecutive variables, then stochastic numbers corresponding to the input variables with desired probability can be generated.

4.3 Computing SFs Corresponding to Complex Boolean Functions

Algorithm 1 reveals that SFs can be directly computed from the Boolean expressions having n number of variables. In Fig. 11 some basic stochastic logic circuits are shown to implement standard SFs. Boolean expressions are first simplified using De Morgan's Theorem and then directly applied to compute SFs using the theorems of correlation intervention as given in Sect. 3.

Example 10: Consider $F(A, B, C) = ABC$ with $p_A = \frac{5}{8}$ is represented as 11111000, B with $p_B = \frac{3}{8}$ is represented as 00000111 and $p_C = \frac{4}{8}$ is represented as 00001111. Thus $SCC_{AB} = -1$ and $SCC_{AC} = -1$. Thus, $F = ABC = A(1 - \bar{B})(1 - \bar{C}) = A(1 - \bar{B} - \bar{C} + \bar{B}\bar{C}) = A - A\bar{B} - A\bar{C} + A\bar{B}\bar{C}$. The corresponding $SF = p_A - min(p_A, \bar{p_B}) - min(p_A, \bar{p_C}) + min(p_A, \bar{p_B}, \bar{p_C}) = \frac{5}{8} - \frac{5}{8} - \frac{4}{8} + \frac{4}{8} = 0$.

Algorithm 1. Generating SF using selective approach for ($n > 2$).

1: **Input**:Any Boolean function consisting of n variables $F(A, B, C, ..., n)$.
2: **Output**:Stochastic Function (SF) corresponding to F.
3: Start with any variable say, A as reference variable.
4: Computation of SCC between A and B
5: **if** $SCC_{AB} = +1$ and next variable is uncomplemented (say B), **then**
6: $SF = min(p_A, p_B)$
7: **else**
8: $SF = p_A - min(p_A, p_B)$
9: **end if**
10: **if** $SCC_{AB} = -1$ and variable is uncomplemented (say B), **then** SF $= max(p_A + p_B - 1, 0)$
11: **else** $SF = p_A - max(p_A + p_B - 1, 0)$
12: **end if**
13: Steps 5 to 16 are used to evaluate the overall SF pairwise till the last variable n of $F(A, B, C, ..., n)$ is encountered .
14: **return** SF

Example 11: This method of simplification can be verified for a 2 input OR gate where the inputs are uncorrelated to produce $Z = A + B$. Using De-Morgan's theorem $Z = \overline{\bar{A} \cdot \bar{B}}$. Thus, $p_Z = 1 - (1 - p_A)(1 - p_B) = p_A + p_B - p_A p_B$.

Consider two uncorrelated bitstreams $A = \frac{4}{8} = 11001001$ and $B = \frac{4}{8} = 10011100$. Thus, $Z = 11011101(p_Z = \frac{6}{8})$. p_Z calculated using the function is $p_Z = \frac{4}{8} + \frac{4}{8} - \frac{4}{8} . \frac{4}{8} = \frac{6}{8}$ matches with the bitwise OR operation of SNs.

Example 12: Consider a Boolean function described by $Z = \bar{S} \cdot A + S \cdot B$. The corresponding SF. to the Boolean function can be obtained by first simplifying the function with the help of De-Morgan's Laws. Thus, $Z = \bar{S} \cdot A + S \cdot B = \overline{\overline{\bar{S} \cdot A + S \cdot B}} = \overline{\overline{\bar{S} \cdot A} \cdot \overline{S \cdot B}}$.

Let S, A and B are represented by the SNs $11110000(\frac{4}{8})$, $11111000(\frac{5}{8})$ and $11100000(\frac{3}{8})$ respectively. From the nature of bitstreams it is evident that A, B and S are positively correlated. Thus, the SF corresponding to the Boolean expression can be constructed as shown in Fig. 12. The output is $11101000(= \frac{4}{8})$ which is exactly similar with the calculated result.

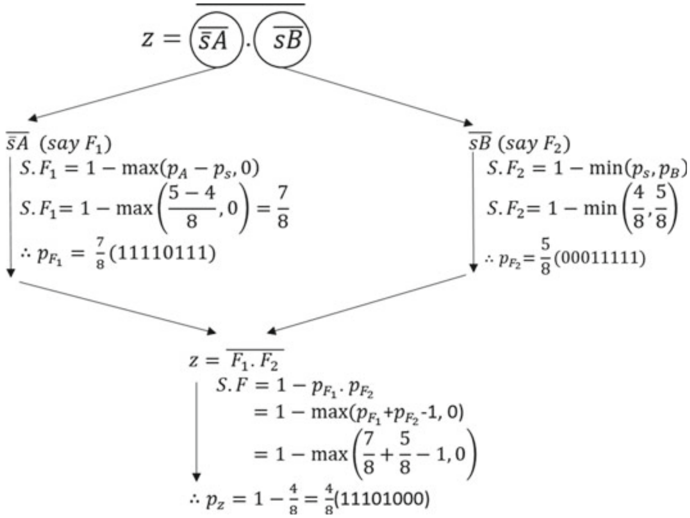


Fig. 12 The flowchart of an example illustrating Method II

5 Conclusion

Correlation is an important aspect in designing accurate stochastic circuits. This paper portrays deductive based reasoning on varied degrees of correlation between two stochastic numbers. It investigates the propagation of correlation to construct SFs of complex Boolean Functions. It is observed that different correlation assumptions leads to realization of different SFs. This property is harnessed to develop two algorithms that help us to generate SFs in a simplistic approach. The pros and cons of the algorithms are discussed that allow users to evaluate SFs in an efficient way. Overall the paper lays a strong mathematical background to understand the operations between two SNs in the light of SCC. In future, we will try to extend the concept to compute SFs corresponding to a Boolean function where intermediate values of SCC coexist.

References

1. Alaghi, A., Hayes, J.P.: Survey of stochastic computing. ACM Trans. Embed. Comput. Syst. (TECS) **12**(2s), 1–19 (2013)
2. Gaines, B.R.: Stochastic computing systems. In: Tou, J.T. (eds.) Advances in Information Systems Science. Advances in Information Systems Science, pp. 37–172. Springer, Boston, MA (1969). https://doi.org/10.1007/978-1-4899-5841-9_2

3. Mitra, S., Banerjee, D., Naskar, M.K.: A low latency stochastic square root circuit. In: 34th International Conference on VLSI Design and 20th International Conference on Embedded Systems, VLSID 2021, Guwahati, India, 20–24 February 2021, pp. 7–12. IEEE (2021). https://doi.org/10.1109/VLSID51830.2021.00006
4. Mitra, S., Santosh, K.C., Naskar, M.K.: Niblack binarization on document images: area efficient, low cost, and noise tolerant stochastic architecture. Int. J. Pattern Recognit. Artif. Intell. **35**(4), 2 154 013:1–2 154 013:25 (2021). https://doi.org/10.1142/S0218001421540136
5. Anderson, J.H., Hara-Azumi, Y., Yamashita, S.: Effect of LFSR seeding, scrambling and feedback polynomial on stochastic computing accuracy. In: 2016 Design, Automation & Test in Europe Conference & Exhibition (DATE), pp. 1550–1555. IEEE (2016)
6. Lee, V.T., Alaghi, A., Ceze, L.: Correlation manipulating circuits for stochastic computing. In: 2018 Design, Automation & Test in Europe Conference & Exhibition (DATE), pp. 1417–1422. IEEE (2018)
7. Chen, T.-H., Hayes, J.P.: Analyzing and controlling accuracy in stochastic circuits. In: 2014 IEEE 32nd International Conference on Computer Design (ICCD), pp. 367–373. IEEE (2014)
8. Jeavons, P., Cohen, D.A., Shawe-Taylor, J.: Generating binary sequences for stochastic computing. IEEE Trans. Inf. Theory **40**(3), 716–720 (1994)
9. Ahlgren, P., Jarneving, B., Rousseau, R.: Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. J. Am. Soc. Inform. Sci. Technol. **54**(6), 550–560 (2003)
10. Budhwani, R.K., Ragavan, R., Sentieys, O.: Taking advantage of correlation in stochastic computing. In: 2017 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1–4. IEEE (2017)
11. Alaghi, A., Hayes, J.P.: Exploiting correlation in stochastic circuit design. In: 2013 IEEE 31st International Conference on Computer Design (ICCD), pp. 39–46. IEEE (2013)
12. Kanoh, T.: Absolute value calculating circuit having a single adder, uS Patent 4,953,115, 28 August 1990

The Impact of Fin Shape Variation on Accumulation-Mode Bulk FinFETs' Analog and RF Performance for Different Dielectrics

Mihir Vaghela and R. K. Nagaria

Abstract Multi-gate FinFETs can be scaled down to 20 nm and beyond because to their non-planar 3D structure, which significantly improves performance. However, every change in the cross-sectional geometry of the fin has an effect on device performance. The influence of different fin cross-sectional shapes on accumulation mode bulk FinFETs with thin fins and short channel length has been investigated in this work. Different important device performance parameters are evaluated and for different fin shapes, including ON-current (ION), OFF current (IOFF), ratio of ON/OFF current, Threshold voltage (Vth), swing (SS), drain-induced barrier lowering (DIBL), Total gate capacitance (Cgg). Two dielectrics, Hafnium Oxide (HfO2) and Lanthanum Oxide (La2O3) are taken into consideration for simulation of device. The geometry of the fin cross-section has a significant influence on the device's performance, according to the research. The device with decreased fin top width showed a ~20% drop in DIBL and a ~9% reduction in SS, indicating an improvement in. At the cost of a drop in ON current, a nearly tenfold reduction in leakage current was found, along with a two and a half times rise in the Ion/ratio. In order to have superior and acceptable Analog/RF applications, an optimum fin structure for the bulk FinFET is also produced.

Keywords FinFET · Fins · DIBL · Fin shape

1 Introduction

The MOSFET has been aggressively downscaled in order to meet the demands of more sophisticated and very dense circuits. Downscaling, on the other hand,

M. Vaghela (✉) · R. K. Nagaria (✉)
MNNIT Allahabad, Prayagraj, UP, India
e-mail: vaghelamihir28@gmail.com

R. K. Nagaria
e-mail: rkn@mnnit.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_7

causes short channel effects (SCEs), which have a negative impact on device performance [1, 2]. The OFF-state current of any ordinary silicon MOSFET with channel length in nano meters increases as well. Reduced gate oxide thickness to counteract these effects increases gate leakage current, which is undesirable. Lower-resistance, Higher-K dielectrics and metal-based gate electrodes are also required for nano devices to continue to grow. FinFET devices which are having 3D fin shaped channels and multiple gates have been studied a lot by academics in recent years because they have the potential to alleviate these major scaling challenges [3–6] (Sikarwar et al. 2013). FinFETs are already being considered as the most useful option at the 20 nm technology node and below because of their less short channel effects (SCE) and very good gate control in as compared to usual MOSFETs, according to this research. They also have a higher ratio and current per unit area than planar devices. Because of all these benefits of FinFETs are used in a huge number of applications such as high-speed digital ICs, analogue ICs, SRAMs, DRAMs, and so on. While if we look at another side, the junctionless transistors have the same doping type and concentration in the source, drain, and channel regions resulting in an easy and simple production process.

This thin layer is required for complete carrier depletion during the device's OFF state. However, if the semiconductor's doping level is less, it can easily empty the carriers when device is off, while source and drain resistance increases. While looking at another aspect, if the device's doping is high enough to allow a substantial quantity of current to flow during the ON states of the device, the ability to make it off is compromised. Because current flows through the majority of the channel rather than the surface in junctionless MOSFETs, it is less responsive to the oxide–semiconductor contact. There are lots of advantages of junctionless MOSFETs over ordinary devices which include safe keeping device from short channel effects, higher scalability, much improved subthreshold swing (SS), and better drain-induced barrier lowering (DIBL), among others. JL FinFET devices have been reported to show better results than Inversion mode FinFETs in CMOS circuits [7]. Combining the benefits of both technologies is necessary to reap the benefits of both. FinFET architecture and Junctionless Transistor technologies are becoming viable options.

Fin shape has been demonstrated to have a significant impact on FinFET performance [3, 6]. Many literatures on fin shape, however, have primarily concentrated on calculating the effect of fin shape variation on short channel effects [3, 8–10] and have detailed some leakage current variation findings. FinFETs characteristics, both being ON and OFF, are studied extensively and described in relation to fin-edge roughness and work function driven variation [11]. As The FinFET's performance is mostly determined by the fin, it is believed that the shape of the fin cross section will be determined by geometry. Its other parameters such as RF and analog parameters should be affected as well. As a result, there is room for inquiry into device performance in cross-section form modification for application in circuits that combine RF and analogue signals.

This study illustrates the effects of fin shape modification on the device's RF/Analog performance. The effect of variation of top fin width geometry on its performance for analog/RF applications is evaluated using a 3D simulation, that too for multiple dielectrics with different K values. The following is a breakdown of the paper's structure. Starting Sect. 1 with introduction, Sect. 2 follows with the device geometry and simulation model. Section 3 examines the findings of detailed simulations. The overall conclusions are drawn in Sect. 4.

2 Device Geometry and Simulation

Figure 1 shows a schematic representation of a Junctionless Accumulation Mode (JAM) bulk FinFET device for n-channel mode of operation. The shown simulation of 3D FinFET device was done for 20 nm channel length and 15 nm Fin height for better accuracy. Bottom Fin width for the device was set at 15 nm, while the top fin width ranges from 3 to 15 nm. Figure 2 shows same device with triangular (3 nm) top fin width. At this scale, a high-k gate dielectric (HfO2) was utilized as an insulator to reduce leakage (Off) current of gate induced due to quantum tunneling near the gate oxide [12, 13]. Table 1 lists other significant parameters of the device used in this study.

The effective oxide thickness (EOT) of standard SiO2 dielectrics must be less than 1 nm for devices using 22 nm technology, resulting in very high leakage current due to the quantum tunneling phenomenon. As a result, dielectric material with a higher value of dielectric constant (k-value) can inhibit the flow of this leakage (Off) current while maintaining the necessary effective oxide thickness despite having a thicker

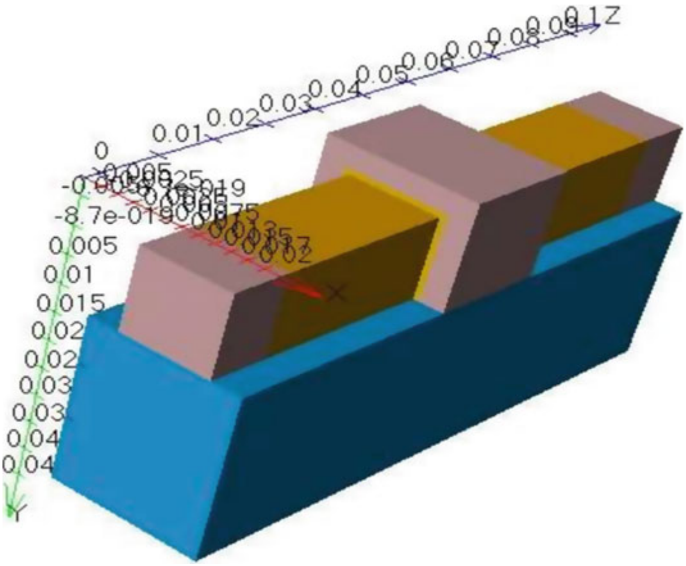


Fig. 1 3-D JAM Bulk FinFET schematic with rectangular (15 nm) top fin width

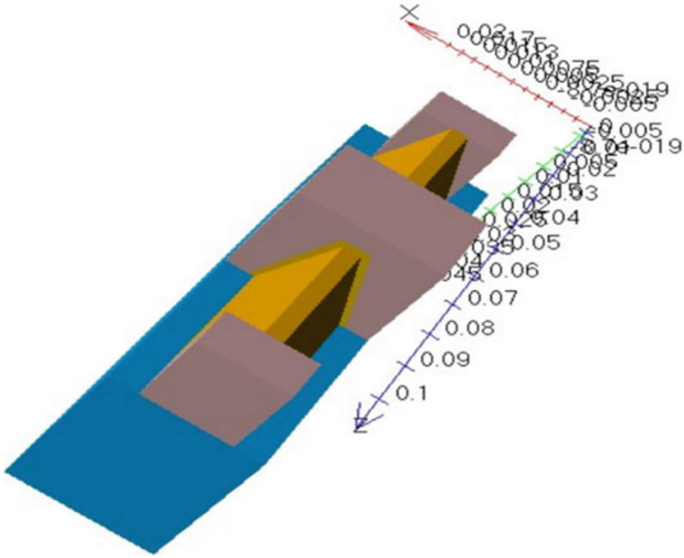


Fig. 2 3-D JAM Bulk FinFET schematic with triangular (3 nm) top fin width

Table 1 Basic parameters of the device

Parameter	Value
Fin height	15 nm
Bottom fin width (Wbottom)	15 nm
Gate length	20 nm
Top fin width (Wtop)	3–15 nm
Equivalent oxide thickness	1 nm
Gate oxide material	HfO_2 & La_2O_3
Gate work function (W.F)	4.65 eV
Channel doping (cm^{-3})	1×1028
Drain supply voltage (Vds)	0.9 V
Source/drain doping (cm^{-3})	1×1021

physical dimension. A big k value, on the other hand, causes very huge fringing effect in the drain as well as source regions, which is undesirable.

Hafnium oxide (HfO2) with a K value i.e. dielectric constant of 22 have excellent insulating and capacitance properties. When biased in the saturation area, the device with HfO2 spacer performs substantially better in terms of drain current and trans conductance than the device with SiO2 spacer [13]. Another dielectric with even higher value of K is chosen for our study, which is La2O3 (Lanthanum oxide). It has value of K = 30. Overall work is done by simulating device for both HfO2 and La2O3. Molybdenum, with a work function of 4.65 eV, is being used as a material for gate electrode in our research. Using nitrogen implantation followed by thermal annealing, the work function of Molybdenum can be varied through a large range (4.5–4.8 eV).

Silvaco ATLAS [14], a common TCAD device simulation programme, was used to simulate the device under investigation. Two files should be made available to ATLAS for majority of simulations to be carried out. A file containing ATLAS' directives is the first text file. The other file which is used as an input to ATLAS is a structure file, which specifies the overall structure of the device we want to create. Three different type of files are generated as output of ATLAS. First output file is the file generated during run-time. This file shows the simulation's progress as well as any errors or warnings. Second file type which we get an as output file is the log file, and it accumulates all voltages and currents of all the terminals of the device through the device analysis. The third file which we get in form of the output file is knows as solution file, and it contains multi-dimensional (2D and 3D) data about the values variables in the device under simulation at a specific bias point. The drift–diffusion equations for electrons and holes were solved in the simulation. The velocity saturation effect was studied using the FLDMOD (field dependent mobility) model, which provides an electric field dependent model for electrons. The Fermi–Dirac distribution model without impact ionization was used in the simulation due to the high channel doping concentration. The electric field dependent carrier model was employed in the simulation setup. A band gap narrowing effect may occur due to a highly doped channel, which was addressed using the Bandgap Narrowing Model (BGN). To appropriately simulate the bipolar current gain, this model (BGN) was required. To account for leakage currents caused by heat generation, the simulation model employs Shockley–Read–Hall (SRH) and Auger recombination/generation. The Newton and Gummel method is utilized for numerical computations, and a temperature of 300 K has been set in the simulation. For all of the simulation work here, the same model parameters are used with varied fin structures.

3 Results and Discussion

Width of the Fin at the bottom is kept constant at 15 nm in this study, whereas top fin width varies from 15 nm (rectangular) to 3 nm (Triangular). Recent fin shape research is primarily based towards determining the impact of fin shape on short channel effects and leakage current. Various critical electrical characteristics such as Threshold voltage (Vth), SS, DIBL, ON current (Ion), Off current (Ioff), Ion/Ioff ratio and Gate Capacitance (Cgg) were used to investigate the influence of the nonrectangular channel geometry. The input transfer characteristics (Ids-Vgs) and output transfer characteristics of the various structures have been plotted in Figs. 3 and 4 respectively. The predicted series resistance is larger when the fin top width is less. As a result, as the fin top width is reduced, Ids lowers. The constant current approach was used to calculate the device's threshold voltage for various fin top widths from the Ids-Vgs curves.

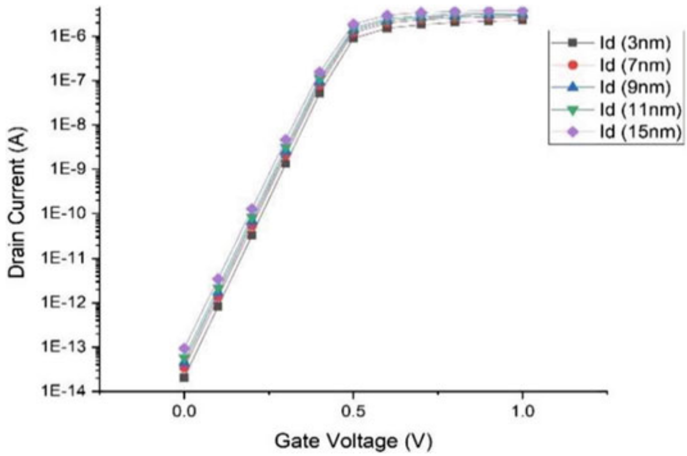


Fig. 3 Input transfer characteristics (Id-Vgs) at $V_D = 0.9$ V

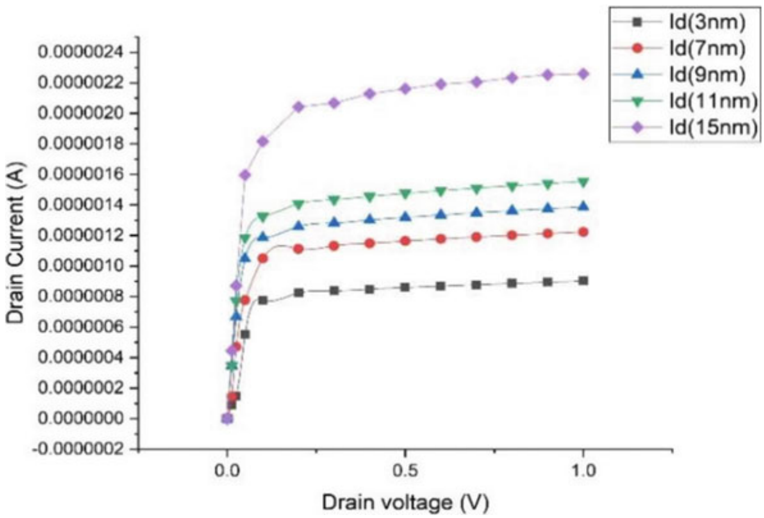


Fig. 4 Output transfer characteristics (Id-Vds) at $V_G = 0.5$ V

Variation of threshold voltage with respect to change in top fin width is plotted in Fig. 5. The graph shows that when the top fin width lowers, the threshold voltage rises. The fluctuation in resistance across the device and its leakage current appears to effect Vth shift. The narrow fin width impact causes Vth to increase for a FinFET as the top fin width reduces. Here similar nature is seen for both the dielectrics with La2O3 having threshold a little higher. It seems manageable to build FinFETs with multiple value of threshold by altering fin shape inside a single device. On current for the device is shown in Fig. 6. ON Current is the active current when device is in the ON state. When the device is on, it is the current flowing between drain and source. The drain current Ids obtained when Vgs = 1 V and Vds = 0.9 V is referred to as ON current. As evident from Id(log)-Vgs graph, as top fin width decreases, ON current decrease.

We want ON current to be as high as feasible for any device, but we are not getting the results we want by decreasing the top fin width. This drop in current is comprehensible when we consider the additional advantages of reducing top fin width, which we will discuss further down. When the top fin width was changed from 15 to 3 nm, the ON current was reduced from 4 to 2.25 A, a 44% reduction.

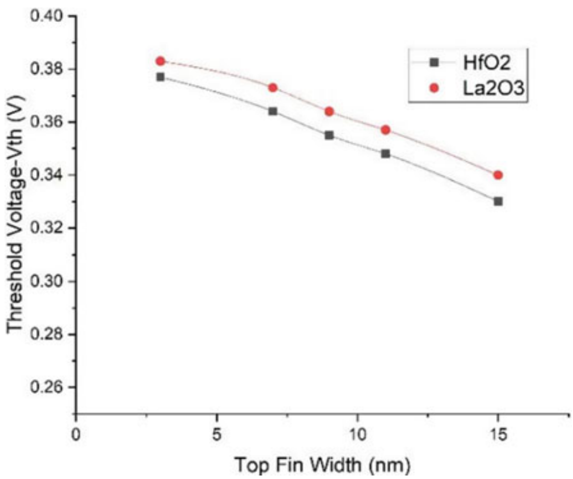


Fig. 5 Change in threshold voltage—Vth, with respect to variation in fin top width of the device

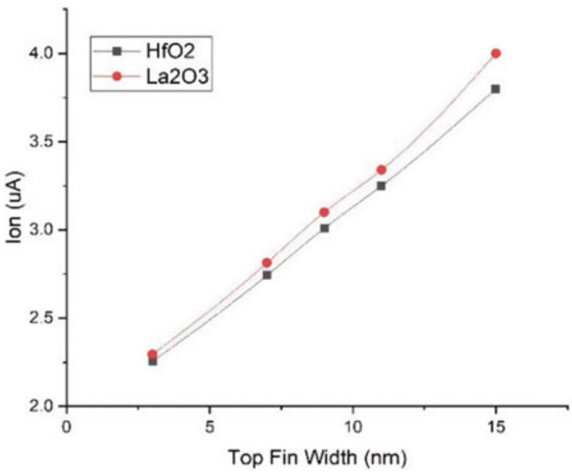


Fig. 6 Change in ON current—Ion, with respect to variation in fin top width of the device

OFF current is leakage current. We want the leakage current of any device to be as low as possible. With increase in top fin width, Threshold voltage increases. Now as the threshold voltage increases, subthreshold (leakage) current decreases which is in turn OFF current. Figure 5 shows change in value of OFF current with respect to variation in top fin width. We can observe from Fig. 7 that as top fin width decreases, OFF current decreases which is analogous to our discussion in the last paragraph. We want our device to have zero leakage current or as little leakage current as feasible and this expected behavior is noticed as the top fin width of our FinFET is reduced. As the top fin width changes from 15 to 3 nm, the current value drops approximately tenfold (10^{-13} to 10^{-14}). By keeping Vgs = 0.0 V and Vds = 0.9 V, the OFF current Ids is produced.

Figure 8 demonstrates the change in the device's ION/IOFF current ratio with respect to variation in fin top width of the device. It describes the amount of difference between the ON and OFF currents of the device under study. We want the ON current to be maximum and the OFF current to be minimum, hence the ratio of Ion/Ioff should be as high as possible. It should, in theory, be infinite.

We can see that the ratio of Ion/Ioff grows as the top fin width decreases. As the top fin width changes from 15 to 3 nm, the ratio increases by around two and a half times. From Figs. 4 and 5, decrease in IOFF is more compared to decrease in ION so overall ratio of ION/IOFF increases here. As a result, a decrease in the value of ON current came at the expense of a far greater decrease in the value of OFF current. Overall, the Ion/Ioff ratio was found to be in the 108 range.

The subthreshold swing (SS) and the drain induced barrier lowering (DIBL) are extensively employed to evaluate devices' vulnerability to SCEs because they quantify the resistance of channel control to the rising of the drain voltage. Figure 9 shows plot of DIBL with respect to variation in top fin width. As the top fin width ranges from 15 to 3 nm, the value of DIBL drops by about 20%. If we compare it to

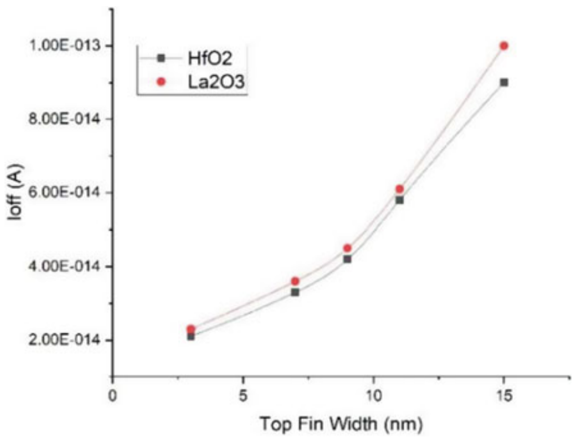


Fig. 7 Change in OFF current—Ioff, with respect to variation in fin top width of the device

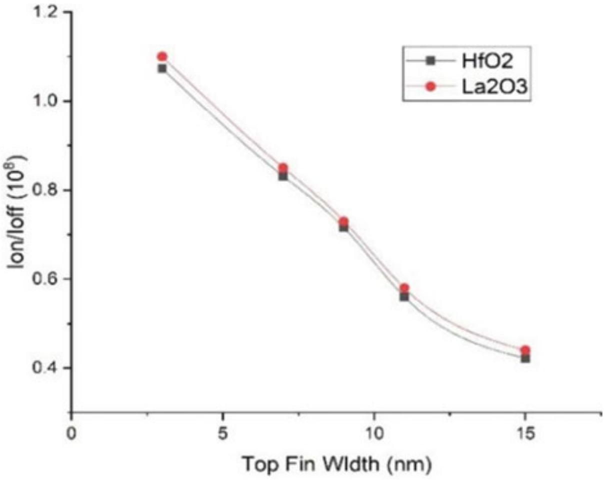


Fig. 8 Change in Ion/Ioff ratio with respect to variation in fin top width of the device

the threshold voltage curve, we can see that when DIBL increases, the control over the gate voltage of the channel diminishes, and the threshold voltage of the device decreases as well. So as DIBL increases, threshold voltage decreases and vice versa. This nature can be clearly seen from the curve of DIBL and threshold.

Figure 10 demonstrates value of subthreshold swing with respect to variation in top fin width. We can observe from the graph that SS reduces as top fin width decreases. It is essentially a leakage current measurement. Subthreshold swing should be kept to a bare minimum. Because it measures leakage current, a lower subthreshold swing value indicates better channel control, whereas a lower leakage current indicates improved Ion/Ioff current. As the top fin width is changed from 15 to 3 nm, the value

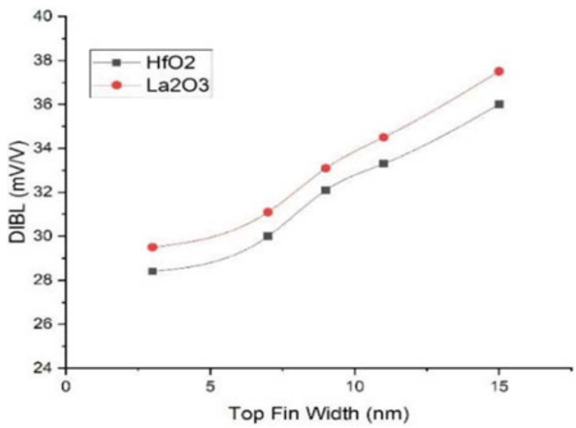


Fig. 9 Change in DIBL with respect to variation in fin top width of the device

of SS drops by about 9%. As SS decreases, the leakage current reduces, the value of Ioff decreases, and the ratio of Ion/Ioff increases. All of the previous results are analogous to this in terms of the variance in top fin width.

For varied fin top widths, Fig. 11 depicts the total gate capacitance (Cgg) as a function of gate voltage (Vgs). The total gate capacitance (Cgg) is lower for FinFETs with smaller Fin top width, as seen in Fig. 10. When the fin top width is lowered from 15 to 3 nm, the gate capacitance of the triangular fin is reduced by about 6.5%, as shown in Fig. 11.

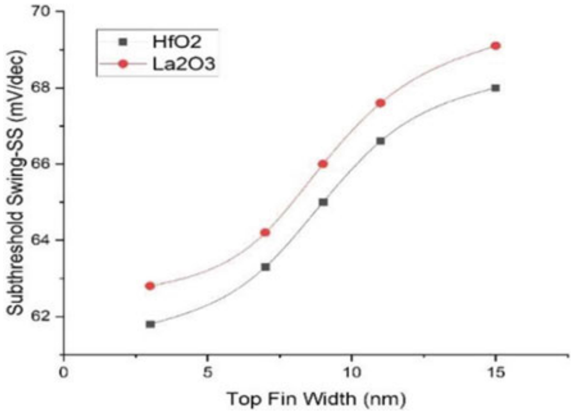


Fig. 10 Change in SS with respect to variation in fin top width of the device

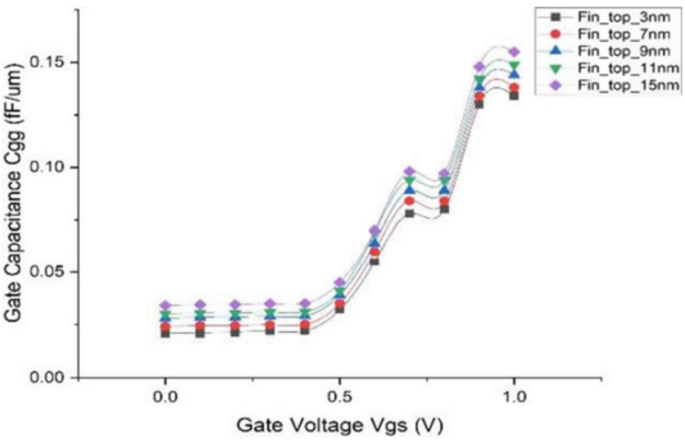


Fig. 11 Change in gate capacitance—Cgg with respect to variation in fin top width of the device

4 Conclusion

The device's RF/analog performance was assessed in this study to investigate the influence of fin shape variation. Calculating several parameters of the device with varied fin forms was used to make a comparison. Device was made using two dielectrics, HfO2 (Hafnium Oxide) and La2O3 (Lanthanum Oxide) Based on the findings, we can conclude that, rather than scaling down a FinFET's non-planar 3D structure, adjusting its top fin width has a substantial impact on its device properties, allowing us to achieve better outcomes with additional benefits. When the device's fin shape was changed from a rectangular fin with a width of 15 nm to a triangular fin with a fin top width of 3 nm, SCEs improved by roughly 20% in terms of DIBL and 9% in terms of SS. Other criteria showed a significant impact as well. At the expense of reduced ON current, a nearly tenfold reduction in leakage current was reported, along with a two and a half times rise in the Ion/Ioff ratio. The ideal worth of top fin width for this model ought to be kept at 7 nm, while the width of bottom fin width ought to be kept at 15 nm according to the observations. Fin top width variation can be used to in the applications like multiple thresholds and low-leakage FinFET designs. SOI FinFET can be employed and results can be compared.

References

1. Crupi, F., Alioto, M., Franco, J., Magnone, P., Togo, M., Horiguchi, N., Groeseneken, G.: Understanding the basic advantages of bulk FinFETs for sub- and near-threshold logic circuits from device measurements. IEEE Trans. Circuits Syst. II **59**(7), 439–442 (2012)
2. Xu, W., Yin, H., Ma, X., Hong, P., Xu, M., Meng, L.: Novel 14-nm scallop-shaped FinFETs (S-FinFETs) on bulk-Si substrate. Nanoscale Res Lett **10**(249), 1–7 (2015)
3. Dubey, S., Kondekar, P.N.: Fin shape dependent variability for strained SOI FinFETs. Microelectron Eng **162**(16), 63–68 (2016)
4. Gaynor, B.D., Hassoun, S.: Fin shape impact on FinFET leakage with application to multithreshold and ultralow-leakage FinFET design. IEEE Trans. Electron Devices **61**(8), 2738–2744 (2014)
5. Hsu, T.-H., Lue, H.-T., Lai, E.-K., Hsieh, J.-Y., Wang, Z.-Y., Yang, L.-W., King, Y.-C., Yang, T., Chen, K.-C., Hsieh, K.-Y., Liu, R., Lu, C.-Y.: A high-speed BE-SONOS NAND flash utilizing the field enhancement effect of FinFET. In: IEDM Technical Digest, pp. 913–916 (2007)
6. Sahay, S., Kumar, M.J.: Diameter dependence of leakage current in nanowire junctionless field effect transistors. IEEE Trans. Electron Devices **64**(3), 1330–1335 (2017)
7. Guin, S., Sil, M., Mallik, A.: Comparison of logic performance of CMOS circuits implemented with junctionless and inversionmode FinFETs. IEEE Trans. Electron Devices **64**(3), 1366–1374 (2017)
8. Duarte, J.P., Paydavosi, N., Venugopalan, S., Sachid, A., Hu, C.: Unified FinFET compact model: modelling trapezoidal triplegate FinFETs. In: International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Glasgow, pp. 135–138 (2013)
9. Yu, Z., Chang, S., Wang, H., He, J., Huang, Q.: Effects of Fin shape on sub-10 nm FinFETs. J. Comput. Electron. **14**(2), 515–523 (2015)
10. Zhang, J., Si, M., Lou, X.B., Wu, W., Gordon, R.G., Ye, P.D.: InGaAs 3D MOSFETs with drastically different shapes formed by anisotropic wet etching. In: 2015 IEEE International Electron Devices Meeting (IEDM), Washington, DC, pp. 15.2.1–15.2.4 (2015)

11. Seoane, N., et al.: Comparison of fin-edge roughness and metal grain work function variability in InGaAs and Si FinFETs. IEEE Trans. Electron Devices **63**(3), 1209–1216 (2016)
12. Manoj, C.R., Rao, V.R.: Impact of high-k gate dielectrics on the device and circuit performance of nanoscale FinFETs. IEEE Electron Device Lett. **28**(4), 295–297 (2007)
13. Sachid, A.B., Chen, M.-C., Hu, C.: FinFET with high-j spacers for improved drive current. IEEE Electron Device Lett. **37**(7), 835–838 (2016)
14. ATLAS: ATLAS user manual. Silvaco Int., Santa Clara (online). http://www.silvaco.com (2011)
15. Raskin, J.-P., Chung, T.M., Kilchytska, V., Lederer, D., Flandre, D.: Analog/RF performance of multiple gate SOI devices: wideband simulations and characterization. IEEE Trans. Electron Devices **53**(5), 1088–1095 (2006)

Active Metal Coated Dual-Side Polished Plasmonic PCF Biosensor

Shivam Singh, Sanat Kumar Pandey, and Y. K. Prajapati

Abstract A novel two side polished PCF biosensor is proposed to detect the analyte whose refractive index (RI) varies from 1.30 to 1.34. The plasmonically active metal gold is coated on both the polished surfaces to produce plasmons. The strong adhesion of gold nanoparticles with fiber is supported by titanium dioxide (TiO_2) deposited on flat surfaces. Finite element method (FEM) is used to carry out the numerical analysis. After structure simulation, an intensive coupling has been observed at the interface (meta-dielectric) that produced very high resolution, amplitude and wavelength sensitivities of 5×10^{-6} RIU, 335 RIU^{-1} and 34,000 nm/RIU. In addition, the proposed senor has been also tested for the little fluctuations in various material and geometrical parameters.

Keywords Gold · Refractive index · Sensitivity TiO_2

1 Introduction

Nowadays, surface plasmon resonance (SPR) based sensors are effectively employed in various fields such as biosensing, environmental monitoring, and diagnosis in medical fields [1]. SPR takes place when a TM-polarized wave strikes at metal–dielectric interface. Due to this, the generated evanescent field near the interface excites the metal charge particles. The oscillation of these electrons responsible for the generation of surface plasmon polariton (SPP) waves propagated along the interface. These waves are sensitive to change in refractive index (RI) of the medium.

S. Singh (✉)
Department of Electronics and Communication Engineering, Aditya Engineering College, Surampalem, Andhra Pradesh 533437, India
e-mail: shivams20@gmail.com

S. K. Pandey
Department of Electronics Engineering, Government Polytechnic, Lucknow, UP 226016, India

Y. K. Prajapati
Department of Electronics and Communication Engineering, Motilal Nehru National Institute of Technology Allahabad, Prayagraj, UP 211004, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_8

So, by changing RI of the medium, one can verify the fluctuations in resonance wavelength [2]. In recent years, optical fiber or prism coupling is used for SPR sensing mechanism [3]. In prism coupling method, metal is applied on the prism's base. SPR occurs only after the wave vector matching of SPP and incident waves. But, to attain this condition, different optical and mechanical components are needed which makes the implementation complex and costly. To overcome these limitations, a combination of PCF with SPR has been widely used. This combination gives various advantages like single mode operation, low propagation loss, and high non linearity. To achieve the low losses and high sensitivity, we can alter the geometrical parameters such as pitch size, number of holes, hole's diameter, length of fiber, and lattice arrangements.

Apart from this, plasmonic metal selection also plays a vital role in PCF sensor designing. Gold, silver, aluminium and copper, are generally used metals for SPR based sensors [4]. But, gold is mostly preferred because it offers a large resonant shift, responsible for high sensitivity. And, it is chemically inert metal. However, poor adhesion with silica is a major demerit of gold [5]. To overcome this problem, a thin layer of titanium dioxide (TiO_2) coated on the fiber before metal deposition would be the best solution. TiO_2 acts as an adhesive layer that forces the metal nanoparticles to attach well with the smooth fiber surface. Using aforesaid approach, in 2017, A. A. Rifat et al. [6] proposed a plasmonic biosensor whose RI resolution and sensitivity was 9.21×10^{-6} RIU and 46,000 nm/RIU, respectively.

There are two sensing mechanisms used in SPR based sensors. One is based on internal sensing that requires query sample to be applied inside the tiny air hole. This sensing mechanism is not much friendly for real time sensing because emptying or filling the sample inside tiny air hole is quite arduous. Another one is external sensing in which plasmonic metals or liquid applied on the fiber's exterior portion [7, 8]. Recently, D-shaped structures become more prominent due to its geometrical advantages. Moreover, the metal deposition is easier on their flat surfaces [9, 10]. In 2019, Gangwar et al. [11] proposed a D-shaped PCF sensor and achieved high RI resolution and sensitivity of 3.6×10^{-6} RIU and 30,000 nm/RIU, respectively.

F. Wang et al. [12] demonstrated dual D-shaped RI sensor. Gold metal layer was deposited on plain surfaces to boost up the coupling strength. Hence, the resolution and sensitivity were optimized to 7.41×10^{-6} RIU and 13,500 nm/RIU. Likewise, S. Wang et al. [13] reported a twin-core based sensor with metal coated on both side-polished surfaces that strengthen the coupling phenomenon. Hence, the obtained resolution and sensitivity of this sensor were 1.3×10^{-5} RIU and 8000 nm/RIU.

In this article, SPR based biosensor is proposed employing gold as an active metal. The structure possess dual core that offers more than one coupling channels. Rather than adopting a conventional single sensing channel, we have created two microopenings at bottom and top portion of the fiber. Due to this, four sensing channel channels have been formed that enhances the sensing feasibilities. These channels are quite sensitive to the change in its surrounding medium RI. Moreover, one of the major advantages of introducing micro opening is that metal is coated only on the selected portion of the fiber rather than coating on the entire flat surface. A thin TiO_2 layer has been applied to all the four channels to compensate the poor adhesion of gold nanoparticles with fiber. By doing so, the overall sensing performance got improved. Moreover, the geometrical and material parameters have been varied to observe the sensing fluctuations of the proposed sensor.

2 Structure Modelling

The cross section of the propounded PCF model is depicted in Fig. 1. The PCF geometry is arranged in a hexagonal lattice composed of four small air holes of diameter 0.75 μm and five large air holes of diameter 1.6 μm. In order to strengthen the coupling phenomenon, dual core channel has been formed by inserting a single large air hole at the middle of the fiber. The remaining four large air holes are placed in horizontal and vertical direction forming symmetry with the middle large air hole. Moreover, distance between the adjacent air holes (Λ) has been taken as 3.1 μm. Thin layers of TiO_2 and gold having thicknesses 5 and 40 nm has been deposited to all the four flat and narrow channels. In addition, the exterior boundary of PCF has been protected with a perfectly matched layer in order to avoid scattering emission.

The fabrication of the proposed model can be done following few major steps. The initial step is bundling of thin, thick capillaries and solid rods with each other utilizing stack and draw technique [14]. After structure formation, the bottom and top circular fiber portion can be polished to certain extent so that they become flat [15]. Then, TiO_2 and gold nanolayers can be easily deposited using the well-known method called chemical vapor deposition [16].

The numerical analysis can be initiated choosing silica as a host material. And, its refractive index can be estimated as follows [17].

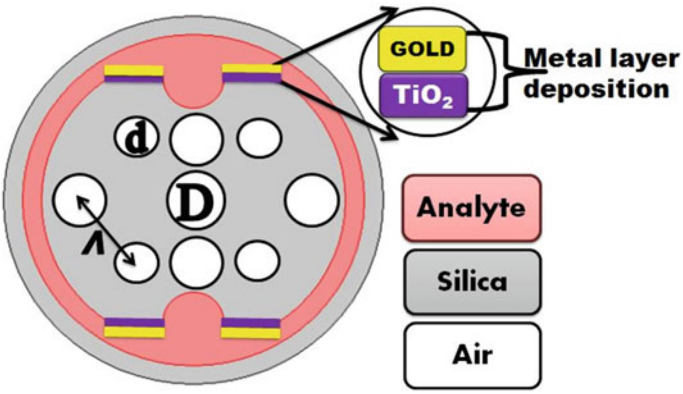


Fig. 1 Front view of the two side polished PCF

$$n_{silica}^2 = 1 + \sum_{k=1}^{3} \frac{G_k \lambda^2}{\lambda^2 - H_k^2} \tag{1}$$

where, λ denotes operating wavelength. G_k and H_k are the coefficients and can be find out using [17].

The refractive index of TiO_2 can be expressed as follows [18].

$$n_{T_i}^2 = 5.91 + \frac{0.2441}{\lambda^2 - 0.0843} \tag{2}$$

Drude model is mainly involved for the estimation of gold's dielectric constant and can be expressed as follows [19].

$$\varepsilon_{Au} = 1 - \frac{\lambda^2 \lambda_c}{\lambda_p^2(\lambda_c + i * \lambda)} \tag{3}$$

where, λ_c and λ_p are collision and plasma wavelength.

The confinement loss (CL) while the wave propagating in the fiber core can be find out as follows [20].

$$CL(dB/cm) = \frac{40\pi}{\lambda * \ln 10} * Im(n_{eff}) * 10^4 \tag{4}$$

3 Results and Discussion

This segment deals with simulation outcome that shows modal field distributions, loss spectrum and dispersion relation. Also, the loss peak fluctuation with the variation in geometrical parameters has been briefly discussed. The even *x*-polarized electric field is shown in Fig. 2 (a) in which the whole energy (yellow color) distributes only in twin core channel, called core mode field distribution. Figure 2 (b-c) shows that the light energy (red color) totally transfers from dual core to interfaces of the top and bottom flat surfaces, called SPP mode. Figure 2 (d) depicts the equal distribution of light energy for both the interfaces. Figure 2 (e) demonstrates the field spreading when light energy in dual core links with energy appears at the interfaces, called coupling or resonance mode because this mode assures a unique resonant peak for a particular analyte.

Figure 3 displays the dispersion relations and variation in loss profile. The blue solid line signifies the RI profile of dual core mode that changes with wavelength. The dotted purple and orange lines denote the fluctuation in effective RI profiles of SPP mode 1 and SPP mode 2. Further, it is noticeable that the dotted purple and orange lines intersect with the solid blue line at point P and K that ensures the wave vector

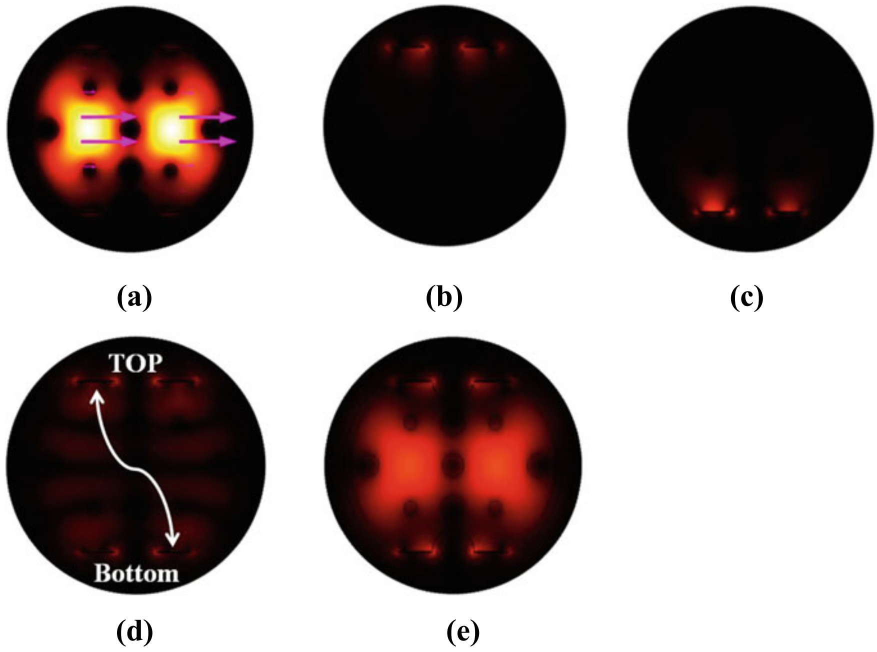


Fig. 2 Electric field expansion for **a** even x-polarized core mode **b** top, **c** bottom and, **d** common SPP mode **e** Coupling mode

matching condition. After fulfilling this condition, two resonance peaks appear at wavelengths 1.45 μm and 2.33 μm, respectively.

Figure 4 (a) shows the loss curve profiles for the analyte whose refractive index varies from 1.30 to 1.34. In Fig. 4 (a), it is noticeable that two resonant peaks appear for a particular analyte RI. Among them, one resonant peak appears due to one side polished while another one appears due to both side polished fiber. While analyzing the resonant peaks for each and every analyte RI, the maximum relative shift in resonant peak is obtained as 340 nm for two side polished fiber when the analyte RI varies from 1.33 to 1.34. Therefore, with the help of obtained resonant shift, one can calculate the wavelength sensitivity as follows [21].

$$S_\lambda = \frac{\Delta\lambda_r}{\Delta n_a}\ \text{nm/RIU} \tag{5}$$

where, Δn_a and $\Delta\lambda_r$ denote the variation in analyte RI and resonant position.

The above expression gives the highest wavelength sensitivity of 34,000 nm/RIU for the change in n_a from1.33 to 1.34.

The RI resolution (R) using the resonant shift can be estimated as follows [21].

$$R = \frac{\Delta n_a \lambda_{min}}{\Delta\lambda_r}\ \text{RIU} \tag{6}$$

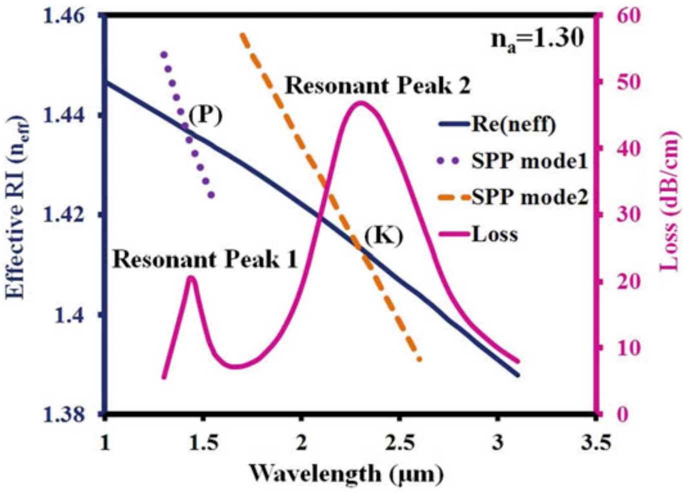


Fig. 3 Loss curve profile and dispersion relation between core and SPP mode for $n_a = 1.30$

Here, λ_{min} has been considered to be 0.1 nm.

Figure 4 (b) demonstrates the amplitude profile with wavelength. The amplitude sensitivity varies while varying n_a from 1.30 to 1.33. The expression to calculate amplitude sensitivity is given as follows [22].

$$A_\lambda = -\frac{1}{CL(\lambda, n_a)} \frac{\partial CL(\lambda, n_a)}{\partial n_a} \ RIU^{-1} \tag{7}$$

The above expression ensures the maximum amplitude sensitivity of 335 RIU^{-1} for two side polished fiber.

Table 1 clarifies that how performance parameter gets affected with the change in analyte RI. The obtained wavelength sensitivities are 5000, 20,000, 22,000 and 34,000 between analyte RI 1.30–1.34 whereas, the amplitude sensitivities are 118,122,101 and 335, respectively.

3.1 Effect of Varying Material Parameters on sensor's Performance

Figure 5(a-b) shows how resonance position changes with the change in thickness of the used materials i.e. gold and TiO_2. In Fig. 5 (a), the utmost loss value is found to be 48.5 dB/cm for gold (Au) thickness equals to 40 nm. The loss value decreases and resonant peak shifts to lower wavelength values while gold thickness increases from 40 to 50 nm. It is noted that increases in gold thickness causes damping loss that

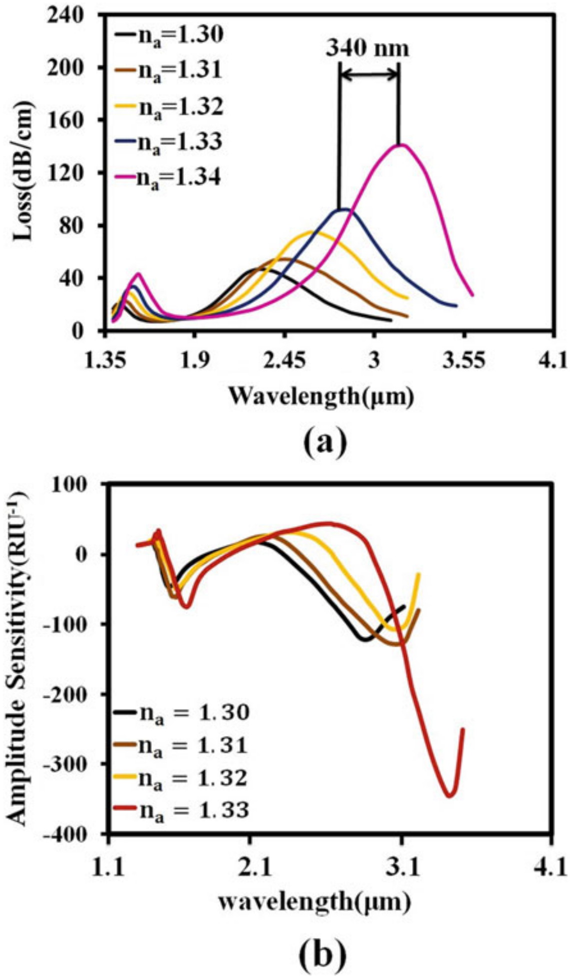


Fig. 4 **a** Variation in resonant peak with the change in n_a from 1.30 to 1.34 **b** Amplitude sensitivity response

Table 1 Performance evaluation of the propounded sensor

n_a	λ_r(nm)	$\Delta\lambda_r$ (nm)	S_λ(nm/RIU)	A_λ (RIU^{-1})	R (RIU)
1.30	2350	50	5000	118	2×10^{-5}
1.31	2400	200	20,000	122	5×10^{-6}
1.32	2600	220	22,000	101	4.5×10^{-6}
1.33	2820	340	34,000	335	2.9×10^{-6}
1.34	3160	–	–	–	–

may reduce the evanescent field penetration depth responsible for analyte detection. Hence, in our analysis, the optimum thickness of gold is taken as 40 nm.

The change in resonant peak position while varying TiO_2 thickness is shown in Fig. 5 (b). The resonant peak shifts to higher wavelength values whileTiO_2 thickness increases to maximum 10 nm. The coupling loss is higher (64.8 dB/cm) at Ti = 10 nm. Therefore, considering the moderate coupling loss, thickness of TiO_2 is optimized to 5 nm.

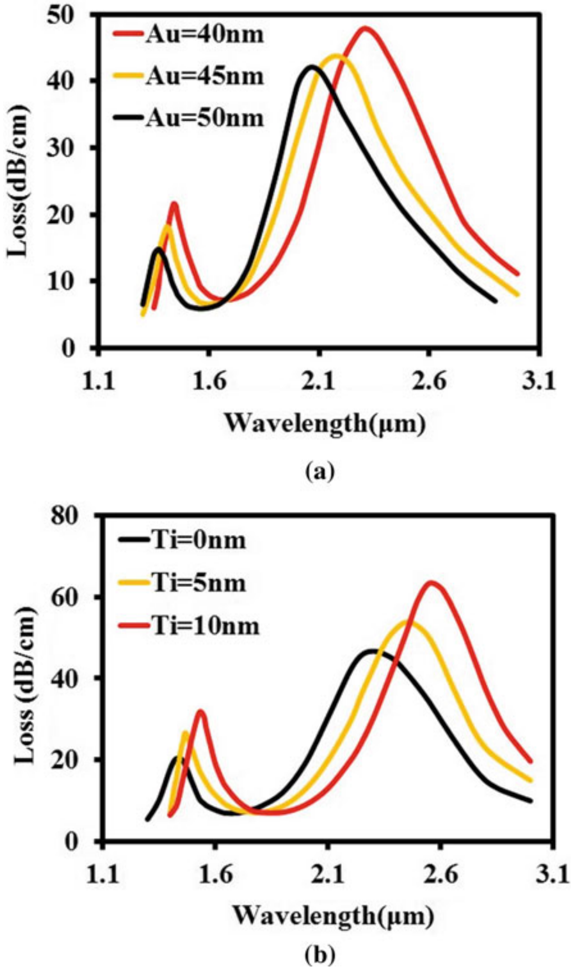
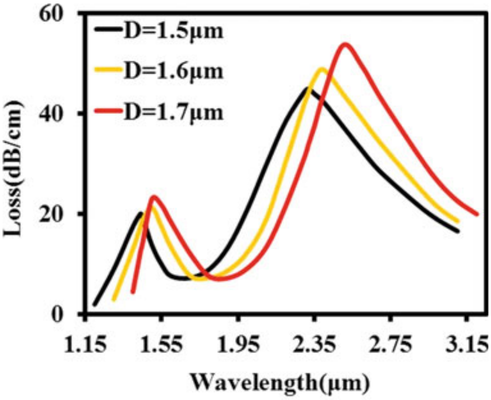


Fig. 5 Variation in resonance wavelength with **a** gold and, **b** TiO_2 layer thickness

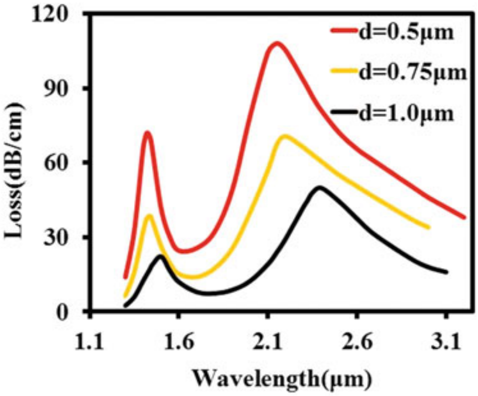
3.2 Effect of Altering Geometrical Parameters

The fluctuation in resonant peak while altering the large air hole diameter (D) is observed in Fig. 6 (a). The resonant peak shifts towards right with respect to wavelength axis when diameter increases from 1.5 μm to 1.7 μm. The right most resonant peak obtained at $D = 1.7$ μm has the highest coupling loss of 53.8 dB/cm while rest of the two peaks have loss values 43.5 dB/cm and 47.3 dB/cm at $D = 1.5$ μm and 1.6 μm, respectively. Further, a wide variation in coupling is clearly visible in Fig. 6 (b) when diameter (d) of small air hole varies from 0.5 to 1.0 μm. It is seen in Fig. 6 (b) that the coupling loss significantly increases with decrease in diameter value. Therefore, the highest loss value of 108 dB/cm is found at $d = 0.5$ μm. It is observed in Fig. 6(c) that varying pitch size (Λ) does not impact much on resonant position. It is the coupling loss that changes with pitch size. Lowering the pitch size causes increment in coupling loss. Hence, the maximum loss value (50.18 dB/cm) is found at Λ=3.0 μm.

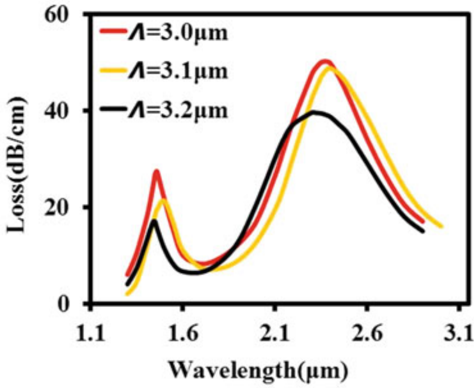
The significance of the proposed sensor can be better understood by comparing it with other already existing sensors. Table 2 compares its major performance parameters and ensures that the wavelength and amplitude sensitivities are quite higher than the reported PCF based sensors.



(a)



(b)



(c)

Fig. 6 Variation in resonance wavelength with diameter of **a** large and, **b** small air holes **c** effect of pitch size on resonant peak

Table 2 A brief comparison of the propounded senor with existing PCF models

Ref.	Geometrical configuration	RI range	S_λ(nm/RIU)	A_λ(RIU^{-1})	R (RIU)
[12]	Symmetrically polished PCF	1.27–1.32	13,500	–	7.4×10^{-6}
[13]	Dual side polished	1.3–1.42	8000	–	1.3×10^{-5}
[23]	PCF with microchannel	1.32–1.34	5000	167	2×10^{-5}
[24]	Dual side polished	1.395–1.415	10,650	252	9.3×10^{-6}
Proposed	**Two side polished**	**1.30–1.34**	**34,000**	**335**	$\mathbf{3 \times 10^{-6}}$

4 Conclusion

Gold-TiO_2 incorporated twin core plasmonic biosensor is numerically analyzed. Two micro openings present on top and bottom side have created four sensing channels that significantly enhanced the coupling between SPP and core mode. For optimum performance, the thickness of gold and TiO_2 layer is kept at 40 nm and 5 nm. With such optimization, we could manage to achieve very high sensitivities of 335 RIU^{-1} and 34,000 nm/RIU in terms of amplitude and wavelength between 1.30 and 1.34. Hence, our propounded model suits well for biological sensing applications.

References

1. Homola, J., Yee, S.S., Gauglitz, G.: Surface plasmon resonance sensors: review. Sens. Actuators B **54**, 3–15 (1999)
2. Yan, X., Li, B., Cheng, T., Li, S.: Analysis of high sensitivity photonic crystal fiber sensor based on surface plasmon resonance of refractive indexes of liquids. Sensors **18**(09), 2922 (2018)
3. Liu, C., Su, W., Liu, Q., Lu, X., Wang, F., Sun, T., Chu, P.K.: Symmetrical dual D-shape photonic crystal fibers for surface plasmon resonance sensing. Opt. Express **26**, 9039–9049 (2018)
4. Islam, M.S., Sultana, J., Rifat, A.A., Ahmed, R., Dinovitser, A., Ng, B.W.H., Ebendor-Heidepriem, H., Abbott, D.: Dual-polarized highly sensitive plasmonic sensor in the visible to near-IR spectrum. Opt. Express **26**, 30347–30361 (2018)
5. Saiful Islam, M.D., Islam, M.R., Sultana, J., Dinovitser, A., Ng, B.W.H., Abbott, D.: Exposed-core localized surface plasmon resonance biosensor. J. Opt. Soc. Am. B **36**, 2306–2311 (2019)
6. Rifat, A.A., Ahmed, R., Mahdiraji, G.A., Adikan, F.R.M.: Highly sensitive D-shaped photonic crystal fiber based plasmonic biosensor in visible to near-IR. IEEE Sens. J. **17**(9), 2776–2783 (2017)
7. Luan, N., Wang, R., Lv, W., Yao, J.: Surface plasmon resonance sensor based on D-shaped microstructured optical fiber with hollow core. Opt. Express **23**(7), 8576 (2015)
8. Qin, W., Li, S., Yao, Y., Xin, X., Xue, J.: Analyte filled core self-calibration microstructured optical fiber based plasmonic sensor for detecting high refractive index aqueous analyte. Opt. Lasers Eng. **58**, 1–8 (2014)
9. Wang, G., Li, S., An, G., Wang, X., Zhao, Y.: Highly sensitive D-shaped photonic crystal fiber biological sensors based on surface plasmon resonance. Opt. Quant. Electron. **48**, 46 (2016)
10. Gangwar, R.K., Singh, V.K.: High sensitive surface plasmon resonance based D-shaped photonic crystal fiber refractive index sensor. Plasmonics **12**, 1367–1372 (2017)

11. Gangwar, R.K., Amorim Vítor, A., Marques, P.V.S.: High performance titanium oxide coated D-shaped optical fiber plasmonic sensor. IEEE Sens. J. **19**, 9244–9248 (2019). https://doi.org/10.1109/JSEN.2019.2927728
12. Wang, F., Liu, C., Sun, Z., Sun, T., Liu, B., Chu, P.: A highly sensitive SPR sensors based on two parallel PCFs for low refractive index detection. IEEE Photonics J. **10**, 1–10 (2018). https://doi.org/10.1109/JPHOT.2018.2856273
13. Wang, S., Li, S.: Surface plasmon resonance sensor based on symmetrical side-polished dual core photonic crystal fiber. Opt. Fiber Technol. **51**, 96–100 (2019)
14. Upadhyay, A., Singh, S., Sharma, D., Taya, S.A.: An ultra-high birefringent and nonlinear decahedron photonic crystal fiber employing molybdenum disulphide (MoS2): a numerical analysis. Mater. Sci. Eng.: B **270**, 115236 (2021)
15. Chen, Y., Xie, Q., Li, X., Zhou, H., Hong, X.: Experimental realization of D-shaped photonic crystal fiber SPR sensor. J. Phys. D **50**, 025101 (2017)
16. Singh, S., Srivastava, A., Pandey, S.K., Prajapati, Y.K.: Single-side-polished gold-coated SPR-based PCF RI sensor. In: Dhawan, A., Tripathi, V.S., Arya, K.V., Naik, K. (eds.) Recent Trends in Electronics and Communication. Lecture Notes in Electrical Engineering, vol. 777. Springer, Singapore. https://doi.org/10.1007/978-981-16-2761-3_28
17. Upadhyay, A., Singh, S., Sharma, D., Taya, S.A.: Analysis of proposed PCF with square air hole for revolutionary high birefringence and nonlinearity. Photonics Nanostruct. – Fundam. Appl. **43**, 100896 (2021)
18. Singh, S., Prajapati, Y.K.: TiO_2/gold-graphene hybrid solid core SPR based PCF RI sensor for sensitivity enhancement. Optik **224**, 165525 (2020)
19. An, G., Hao, X., Li, S., Yan, X., Zhang, X.: D-shaped photonic crystal fiber refractive index sensor based on surface plasmon resonance. Appl. Opt. **56**(24), 6988–6992 (2017)
20. Pandey, S.K., Singh, S., Prajapati, Y.K.: A novel PCF design with an ultra-flattened dispersion and low confinement loss by varying tiny air-hole concentration at core and cladding. Opt. Rev. **28**, 304–313 (2021)
21. Singh, S., Prajapati, Y.K.: Antimonene-gold based twin-core SPR sensor with a side-polished semi-arc groove dual sensing channel: an investigation with 2D material. Opt. Quant. Electron. **52**(2), 1–14 (2022)
22. Rifat, A.A., Mahdhiraji, G.A., Sua, Y.M., et al.: Surface plasmon resonance photonic crystal fiber biosensor: a practical sensing approach. IEEE Photon. Technol. Lett. **27**(15), 1628–1631 (2015)
23. Dash, J.N., Das, R., Jha, R.: AZO coated microchannel incorporated PCF-based SPR sensor: a numerical analysis. IEEE Photon. Technol. Lett. **30**(11), 1032–1035 (2018)
24. Chen, N., Chang, M., Zhang, X., Zhou, J., Lu, X., Zhuang, S.: Highly sensitive plasmonic sensor based on a dual-side polished photonic crystal fiber for component content sensing applications. Nanomaterials **9**, 1587 (2019)

Parametric Analysis of Circular Patch Wideband Antenna for IoT Application

Kuldeep Pandey, A. K. Shankhwar, and Ritesh Sadiwala

Abstract In this paper, a novel topology microstrip dual-side printed antenna is presented for IoT-enabled applications. The novel structured antenna comprises a circular patch by cutting off the ground plane. The ground plane along with a small rectangular edge is exactly below the feed line to increase the antenna bandwidth and a circular notch on the surface of the ground plane is used to increase the antenna properties. Therewith, multiple rectangular notches in the ground plane have increased the effective area of the antenna. The proposed antenna covers the wideband frequency range of 1.15 to 4.83 GHz and shows maximum S11 (return loss) of –44.419, –35.323, and –33.706 dB with 2.99, 3.98, and 5.25 dB gain at 1.9, 3.5, and 4.01 GHz band respectively. Thus proposed antenna can operate all IoT sensor applications between the covered bands such as Wireless local network, global system for mobile (GSM) and RFID modules, etc..

Keywords Internet of Things (IoT) · Microstrip patch antenna · RFID · Sensor · WLAN

1 Introduction

Internet of Things means that the increasing number of connected devices and small sensors are integrated into a huge network with permanent access to the user. One of its real applications is in the Smart home idea, permitting more comfort, effectiveness (at different viewpoints), and security. With an ever-increasing number of gadgets,

K. Pandey (✉)
Department of Engineering and Technology, RKDF University, Bhopal, India
e-mail: kuldeeppandey23@gmail.com

A. K. Shankhwar
Department of Electronics Engineering, HBTU, Kanpur, India
e-mail: akshankhwar@hbtu.ac.in

R. Sadiwala
Department of Electronics Engineering and Technology, RKDF University, Bhopal, India
e-mail: ritesh.sadiwala@rkdf.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_9

nowadays these devices are compulsory, short-power, and more efficient and effective at the same time. Presently, there is a huge demand for IoT-enabled communication technology which attracts the attention of more than one wireless technology [17]. Internet of things shows importance towards interaction among people, different methods, and new technology with an increasing number of connectivity of sensors and devices. It can provide communication among the virtual and real worlds [16]. IoT-enabled devices can be associated with every wireless communication module similar to Wi-Fi, WLAN, Bluetooth, WiMAX, etc. The reality is that IoT is created by the sensor in any wireless communication technology. According to the sensor application, can embed with IoT devices [6]. Such as that device incorporated with the antenna which works for the resonant band according to need. Accordingly, and that antenna works in multiple frequencies and able to operate IoT-enabled devices. IoT makes life much easier and integrated towards multiple wireless applications. It has lots of vast applications such as smart building, automated health care monitoring, and smart home, etc. With the recent advancement of IoTs, an increasing number of IoT applications for short-distance communication has attracted more attention. This has the shortest frequency range communication such as 2.4, 5.2 GHz antenna, usually suitable for mobile devices that require size constraints antenna, broadband spectrum, and easy to integrate with the operational device. Here, microstrip antenna has all these characteristics such as small size, low profile, less complexity, and easy integration with other communicating devices. This antenna makes the system less complex and more efficient.

Microstrip patch antenna having many drawbacks such as low profit, low efficiency, lower power handling capability, narrowband and lower gain so on, but when it is implemented in much practical application it shows good agreement. During the innovation of new technologies, patch antenna gained a lot of attention also due to their furious advantages, so broadband technology, high gain, high impedance, etc. has been the isotropic research topic in microstrip patch antenna [7, 11, 15]. Based on the existing research, the design of compact, efficient, broadband, and high gain antennas has become an essential research trend. At present in this work, a novel design has been proposed and analyzed with a CST environment which can reduce antenna losses and provides high gain at lower bands. This is more efficient to operate IoT applications with wideband characteristics. To identify the significance of the proposed antenna parametric analysis has been done because the structure of the introduced antenna is not in regular size therefore direct analysis of that is not possible. In [5] authors introduced the planar structure of the circular patch and analyze the results on different resonant bands. To get the wideband characteristics response defected ground structure has been introduced in the ground plane. This produces either single or multiple slots on the ground to change the characteristics of the current distribution thus according to the shape and size of the slot, the behavior of the transmission line will change. This shows the changes like inductance and capacitance [12].

The performance of the antenna will be changed according to the shape and size of the patch, substrate material, properties, and feedline applied in the antenna structure [13]. With the circular shape structure, the available mode is supported in a circular motion on a substrate. The height of the substrate is very less compared to the wavelength. According to the dimensions of the circular patches, only one degree of freedom controls the radius of the patch is applied. This will not change the order of the mode but it changes the absolute value of the resonance [3]. Fundamentally circular patch antenna can be analyzed only by the cavity model and full-wave analysis [4]. The cavity model allows the concept of an accurate generalized dielectric substrate that does not radiate any power. In modern wireless communication, it is considered that all handheld communicating device has to be covering all the main frequency bands of IEEE 802.11 and IEEE 802.15.4 with the acceptable gain and far-field pattern [2]. The associated bands of these spectrums are 2.4–2.48, 2.5–2.69, 3.4–3.69 and 5.25–5.85 GHz etc. Additionally, modern antenna design should be flexible to analyze impedance bandwidth for multiple frequency bands independently [9]. Several well known techniques has been introduced previously such as slots in the radiating patch, truncated ground structure (DGS), strips on the antenna and the notches to satisfy all the previous mentioned characteristics [9, 10].

To improve the bandwidth and to diminish the size of the antenna, an antenna with a defected ground structure and two circular notches in the ground plane are presented. The proposed antenna is simulated through a parametric analysis where multiple parameters are analyzed to improve the return loss and gain. Considered two antenna structures in this work which shows a comparative analysis of the antenna concerning antenna design, performance results, etc. In previous work, different antenna structures where radiated patch above and below the overlapped slot with two strips has been presented. compacted size of ground plane length concerning the overall length of antenna has been analyzed to attain 50 impedance matching and to adjust the microstrip width, slit among the feedline and sides of the ground plane [1, 8, 14]. The paper is organized as follows: the proposed structure of the antenna is briefed in Sect. 2. Parametric analysis analyzed results and discussion are explained in Sect. 3 and Sect. 4 concludes the paper.

2 Antenna Design Consideration

The circular shape patch antenna has been introduced here, in which a substrate material has been deposited on the top of the defected ground plane. After that, the circular patches on the substrate are framed. The circular patch considers circular polarization to reduce antenna losses. The antenna structure comprises three layers, the ground plane, the substrate, and the patch which is circular in this design.

Figure 1(a) shows the proposed structure of the wideband antenna. The introduced wideband Microstrip patch antenna is simulated in CST microwave studio 2017 environment, using dielectric constant 4.4 of FR4 substrate and loss tangent 0.02, whose dimension is $80 \times 90 \times 1.6$ mm.

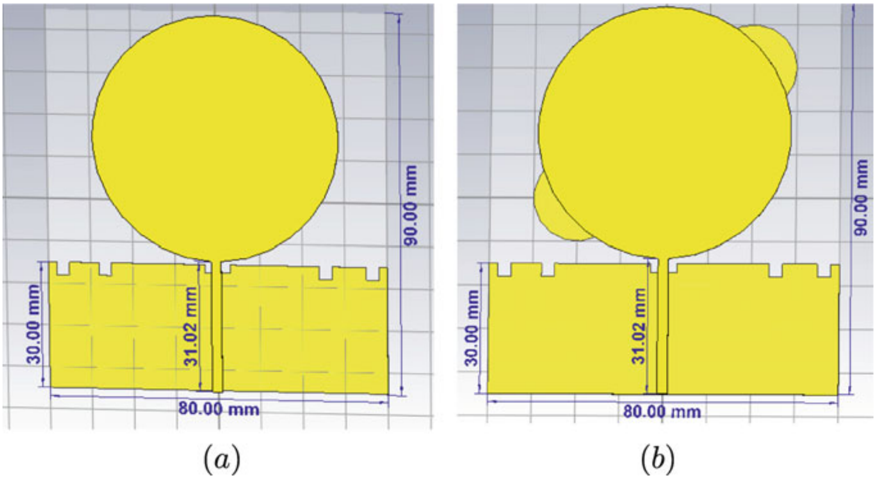


Fig. 1 Antenna design structure **a** reference antenna **b** proposed antenna

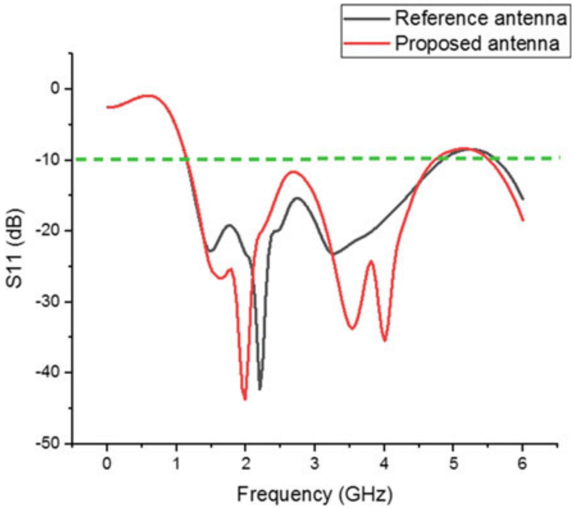


Fig. 2 Simulation results of reflection coefficient **a** Proposed antenna **b** Reference antenna

50 ω microstrip feedline has been linked to the circular radiated patch of 29 mm. Backside on the patch; Modified ground plane is designed to get broad bandwidth along with small rectangular slots on the bottom of the feedline and near the left and right corner.

To increase the gain and bandwidth of the proposed antenna, two circular notches with 10 mm radius are added to the ground plane of the reference antenna, as illustrate in Fig. 1(a). Table 1 show the design parameter of proposed and reference antenna.

Table 1 Dimension of the proposed antenna and reference antenna

Antenna	l_{sub}	W_{sub}	l_{GND}	Patch radius	Ground radius	Feed length	Feed width	Rectangular notch length (Feed line)	Rectangular Notch Width (Feed line)
Reference Antenna	90	80	30	29	10,10 (for both circle)	31.02 31.02	2.40 2.40	2 2	6 6
Proposed Antenna	90	80	30.04	29	–	31.02	2.40	2	6

At the top of the ground plane, additional metallic circular patches are added to increase the effective area of the radiating element and increase the bandwidth along with the gain. It has been observed that the upper cut-off frequency stays together due to the patches because it is managed by the small circle slot in the ground plane. To analyze these differences, the proposed antenna has been analyzed by the extra slots in the ground plane and without slots. Figure 2 reflects the reflection coefficient S11 calculation of antennas. Before the help of the reflection coefficient of both antennas, the proposed antenna S11 results can be seen by showing improvement in the increasing number of bands for many IoT applications in the context of the antenna.

Though, via introducing a patch in the ground plane proposed antenna is able to increase number of bands rather than reference antenna along with reflection coefficient as shown in Fig. 2. Due to various parameters the structure of the antenna may be influence. Such as effects of ground plane, radius of patch and multiple notches on ground plane affects the antenna performance.

3 Result and Discussion

The proposed antenna can reduce the defects of the antenna by introducing modification in-ground and radiating patch geometry. The first considered circular patch with the truncated ground and simulate that results from antenna 1, though modification has been done in antenna 2 with two circular notches in the ground layer. The microstrip feed line is used as a feeding system of antennas center point and this feeding concept is to provide the advantages of maximum efficiency and lesser losses. The patch antenna is typically having circular polarization and that polarization is responsible for a small number of losses in the antenna. This feed line makes the system more compact and easy to fabricate than coaxial feed. With the help of this structure, could increase the radiation efficiency that shows the improved antenna characteristics. The circular patch antenna is perfectly matched for IoT applications due to its circularly polarized characteristics.

4 S11 Measurement and VSWR

Simulated result of proposed antenna with respect to reference antenna of reflection coefficient S11 = –10 dB is shown in Table 2. From the Table 2, it has been analyzed proposed antenna shows better radiation gain and reflection coefficient with geometry modification in ground layer. Gain efficiency of antenna is improved rather than reference antenna. Figure 3 shows the comparative analysis of proposed and reference antenna.

The VSWR plot with the value of 1.012 at 1.986 GHz and smith chart plot are as shown in Fig. 4 and 5 respectively. The far field parameter (radiation pattern) can meet the requirements and the maximum gain of 5.25 dB is quite satisfactory as shown in Fig. 6.

Table 3 shows the comparative analysis of existing work with the proposed work. As of the evaluation table, it is distinguished that for the existing model, defected ground structure, number of resonant, and reflection coefficient S11 is less as compared to the proposed antenna. Two circular notches on the ground structure enhance the antenna properties concerning existing work. This shows the wideband characteristics of the antenna to cover the broadband spectrum for IoT applications. From the Table 3 following consideration have been analyzed:

Table 2 Simulated result of proposed antenna regarding reference antenna

Parameters	Proposed antenna						Reference antenna				
Frequency (GHz)	1.15	1.54	1.98	3.53	4.01	4.71	1.14	1.44	2.2	3.2	4.8
Gain (dB)	1.94	2.42	3.12	3.98	5.25	4.55	1.93	2.27	3.38	4.47	4.85
S11 (dB)	–10.064	–26.132	–44.419	–33.706	–35.323	–10.12	–10.042	–22.628	–42.27	–23.165	–10.044

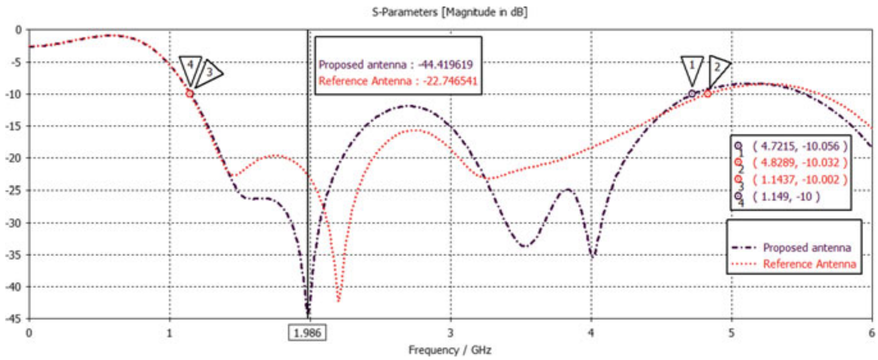


Fig. 3 Comparative S11 parameters of proposed and reference antenna

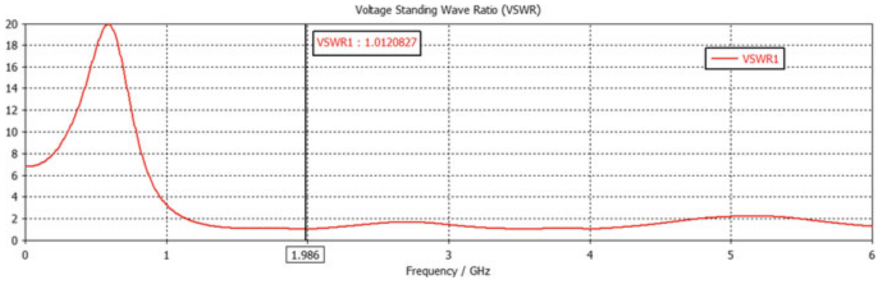


Fig. 4 VSWR plot of proposed antenna

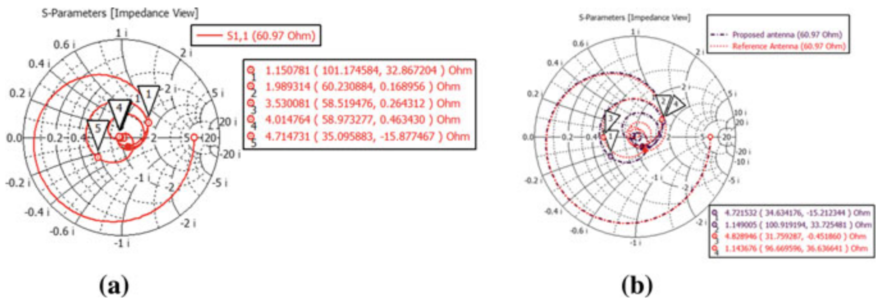
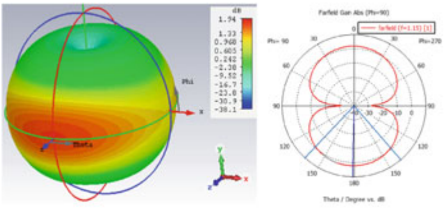


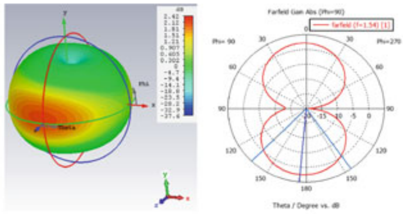
Fig. 5 Smith chart **a** proposed antenna **b** comparison of proposed and reference antenna

Table 3 Comparative Analysis of proposed antenna regarding reference antenna

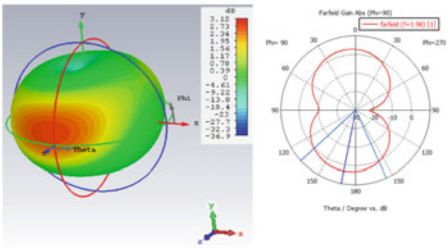
Reference	12 [2018]	1 [2017]	2 [2016]	14 [2016]	16 [2012]	Proposed work
Antenna size	25 × 35 mm	50 × 50 mm	NA	30 × 30 mm	26 × 26 mm	80 × 90 mm
Frequency (GHz)	1.1–2.7 & 3.15–3.65	4.69–5, 5.64–6.41 & 8.1–9.73	2.45	WLAN bands	3.4–5.9	1.15–4.71
Bands	Dual bands	Triple bands	Single band	Dual bands	UWB	Wideband
Gain (dB)	8.9	3.9	6.5	5.2	NA	1.94–5.25
S11 (dB)	NA	∼ –23.122	–40.521	NA	–29.01 @3.4	–44.416, –35.323 & –33.706



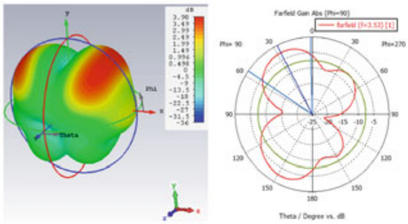
(a) Radiation pattern at 1.15 GHz



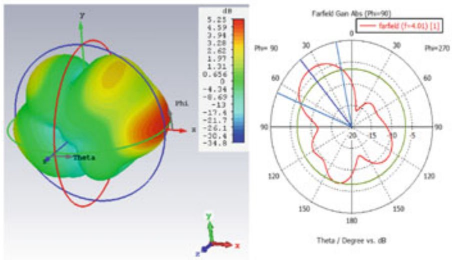
(b) Radiation pattern at 1.54 GHz



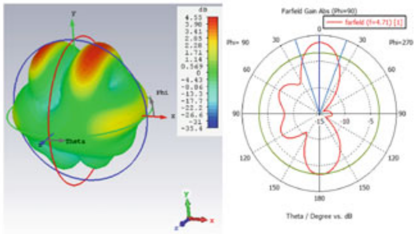
(c) Radiation pattern at 1.98 GHz



(d) Radiation pattern at 3.53 GHz



(e) Radiation pattern at 4.01GHz



(f) Radiation pattern at 4.71 GHz

Fig. 6 Radiation Pattern of proposed antenna

5 Conclusion

In this paper, a wideband antenna with a truncated ground structure has been presented. The simulated results show that the proposed antenna achieves higher bandwidth over the existing results for IoT-enabled applications. The antenna covers the broadband frequency range of 1.15 GHz up to 4.81 GHz and shows maximum return loss of –44.416, –35.323, and –33.706 dB with 3.12, 3.98, and 5.25 dB gains at 1.9, 3.5, and 4.01 GHz band respectively. Enhanced bandwidth over the existing antenna makes the antenna suitable for IoT applications in a specific broad range. Additionally, the proposed antenna is applicable in GPS, Radiolocation, fixed-mobile communications, etc. Far-field analysis of the proposed antennas is simulated and compared with existing work.

References

1. Ahmed, S., Tahir, F.A., Shamim, A., Cheema, H.M.: A compact kapton-based inkjet-printed multiband antenna for flexible wireless devices. IEEE Antennas Wirel. Propag. Lett. **14**, 1802–1805 (2015). https://doi.org/10.1109/LAWP.2015.2424681
2. Awais, Q., Chattha, H.T., Jamil, M., Jin, Y., Tahir, F.A., Rehman, M.U.: A novel dual ultra-wideband CPW-fed printed antenna for internet of things IoT applications. Wirel. Commun. Mob. Comput. **2018**, 1–9 (2018). https://doi.org/10.1155/2018/2179571
3. Balanis, C.A.: Antenna Theory: Analysis and Design, 4 edn. Wiley, Hoboken (2016)
4. K. F. Lee, W.C.: Modern Antennas. Wiley Series in Microwave and Optical Engineering, Hoboken (1997)
5. Kumar, A., Shanmuganantham, T.: A CPW fed octagonal patch UWB antenna with WiMAX band-notched characteristics. In: International Conference on Information Communication and Embedded Systems (ICICES2014), pp. 1–5, February 2014. https://doi.org/10.1109/ICICES.2014.7034023
6. Lazarescu, M.T.: Design of a WSN platform for long-term environmental monitoring for IoT applications. IEEE J. Emerg. Sel. Top. Circuits Syst. **3**(1), 45–54 (2013). https://doi.org/10.1109/JETCAS.2013.2243032
7. Lech, M., Mitchell, A., Waterhouse, R.: Optimization of broadband microstrip patch antennas. In: 2000 Asia-Pacific Microwave Conference. Proceedings (Cat. No.00TH8522), pp. 711–714, December 2000. https://doi.org/10.1109/APMC.2000.925931
8. Madany, Y.M., Elkamchouchi, H.: Analysis of high gain multiband rounded corners dashed rectangular spiral microstrip patch antenna. In: 2005 IEEE Antennas and Propagation Society International Symposium, vol. 1A, pp. 313–316, July 2005. https://doi.org/10.1109/APS.2005.1551312
9. Manouare, A.Z., Ibnyaich, S., Idrissi, A.E., Ghammaz, A.: Miniaturized triple wideband CPW-fed patch antenna with a defected ground structure for WLAN/WiMAX applications. J. Microw. Optoelectron. Electromagn. Appl. **15**, 157–169 (2016). https://doi.org/10.1590/2179-10742016v15i3497
10. Naeem, T., Tahir, F.A., Najam, A.I.: UWB antenna with notches at WLAN bands. In: 2013 11th International Conference on Frontiers of Information Technology, pp. 172–176, December 2013. https://doi.org/10.1109/FIT.2013.39
11. Nishiyama, E., Aikawa, M.: Wide-band and high-gain microstrip antenna with thick parasitic patch substrate. In: IEEE Antennas and Propagation Society Symposium, 2004, vol. 1, pp. 273–276, June 2004. https://doi.org/10.1109/APS.2004.1329624
12. Purushothaman, S., Raghavan, S., Kumar, V.S.: A design of compact metamaterial encumbered monopole antenna with defected ground structure for navigation (L/S-band) applications. In: 2016 IEEE Annual India Conference (INDICON), pp. 1–4, December 2016. https://doi.org/10.1109/INDICON.2016.7839124
13. Drabowitch, S., Papiernik, A., Griffiths, H., Encinas, J., Smith, B.L.: Modern Antennas, 2nd edn. Springer, Heidelberg (2005). https://doi.org/10.1007/978-0-387-26231-4
14. Shameena, V., Mridula, S., Pradeep, A., Jacob, S., Lindo, A., Mohanan, P.: A compact CPW fed slot antenna for ultra wide band applications. AEU Int. J. Electron. Commun. **66**(3), 189–194 (2012). https://doi.org/10.1016/j.aeue.2011.03.015
15. Shaoqiu Xiao, Zhenhai Shao, Bing-Zhong Wang, Ming-Tuo Zhou, Fujise, M.: Design of low-profile microstrip antenna with enhanced bandwidth and reduced size. IEEE Trans. Antennas Propag. **54**(5), 1594–1599 (2006). https://doi.org/10.1109/TAP.2006.874362
16. Vikram, N., Kashwan, K.R.: Design of ISM band RFID reader antenna for IoT applications. In: 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), pp. 1818–1821, March 2016. https://doi.org/10.1109/WiSPNET.2016.7566454
17. Vinodha, E., Raghavan, S.: Double stub microstrip fed two element rectangular dielectric resonator antenna for multiband operation. AEU Int. J. Electron. Commun. **78**, 46–53 (2017). https://doi.org/10.1016/j.aeue.2017.05.020

Dual Band, Dual Port Dielectric Resonator Based MIMO Antenna with Bi-directional Pattern Diversity Using PRS

Komal Srivastava, Ajay Kumar Dwivedi, Nagesh Kallollu Narayanaswamy, and Anand Sharma

Abstract This work deals with pattern diversity where a frequency selective surface (FSS) is utilized along with a cylindrical dielectric resonator MIMO antenna. Its salient feature is bi-directional pattern diversity characteristics. To achieve wider impedance bandwidth and better gain, the dielectric type, a resonator-based antenna has been used. With these features, able to achieve the reflection coefficient $|S_{11}|$ as low as −22.5 dB, and the isolation characteristics are also considerably low. The main lobe is tilted by 50° on either ports and a gain of 5.4 dB has been achieved. The antenna has the operating range in the sub-6 frequency band. It can be used for the applications of vehicular communication in an Intelligent Transport System (ITS).

Keywords Frequency Selective Surface (FSS) · Cylindrical Dielectric Resonator Antenna (CDRA) · Multiple Inputs Multiple Outputs (MIMO) · Sub-6 band · Vehicular communication · Intelligent Transport System (ITS)

1 Introduction

Within the limited spectrum, the fast advancement of wireless communication and technology necessitates higher data speeds and reliability. MIMO technology emerged as an emerging technology to meet these objectives, and it is now used in various wireless standards. Multiple antennas are used on both the transmitter and receiver sides to improve link reliability, data throughput, and channel capacity without requiring additional power or spectrum [1]. DRA researchers have been

K. Srivastava (✉) · A. Sharma
Department of Electronics & Communication Engineering, Motilal Nehru National Institute of Technology, Prayagraj, India
e-mail: srikom.mnnit@gmail.com

A. Sharma
e-mail: anandsharma@mnnit.ac.in

A. K. Dwivedi · N. K. Narayanaswamy
Department of Electronics & Communication Engineering, Nagarjuna College of Engineering and Technology, Bangalore, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_10

working on MIMO antennas for microwave and millimeter-wave frequency bands in recent years because they have several advantages over microstrip patch antennas, such as broader bandwidth, high radiation efficiency, high gain, big power handling capability, and so on [2].

The stringent requirements of vehicular communication in this current scenario of ultra-high-speed smart vehicles, densely populated cities and highways, and high density of vehicles/others mobile devices on road have increased the demand for MIMO antennas. The major issues arising here, solvable by MIMO antennas, are crosstalk and multipath losses. MIMO antennas increase the channel capacity and resolve the fading issues during the propagation of information between the Tx/Rx units [3].

In the open literature, researchers have worked overachieving beam tilting using different antennas and different methods of beam tilting. Dwivedi et al. have proposed a novel configuration of ring dielectric resonator-based multiple input multiple output antenna (MIMO) is proposed with a beam tilting characteristic by using conducting plates in between the antenna elements [4]. Also, Das et al. present a frequency selective surface (FSS) based four-port multiple-input multiple-output (MIMO) cylindrical dielectric resonator antenna (CDRA) system which supports multi-directional pattern diversity characteristics [5]. Another work by Hassan et al. suggested that the port correlation can be reduced satisfactorily by using a superstrate; keeping an FSS above the two ports MIMO patch antenna in a Fabry-Perrot (FP) cavity configuration [6] use of metasurface to achieve LP to CP and better mutual coupling [7].

In this work, a two-port MIMO antenna is proposed in this paper where antenna beams are separated spatially to achieve high isolation and low correlation. To achieve this, one phase gradient FSS has been used and placed on top of the DRA structure as a superstrate. This feature has helped to achieve a good level of beam tilting, along with the advantages of DRA structure.

2 3D Geometrical Arrangement of the Proposed Antenna

Below is the layout of the two-port MIMO structure. It has used a dielectric resonator with a cylindrical shape and has a square-shaped patch. The antenna is sourced using a microstrip line of optimized length and thickness.

On top of the substrate, is the superstrate placed which is a Partially Reflecting Surface (PRS). In order to achieve tilted beams in complementary directions, one phase gradient FSS which act as partially reflecting surfaces (PRS) are used. The PRS is placed at a specific distance from the antenna element to imitate a Fabry–Perot (FP) cavity [8]. The PRS is a combination of several unit cells. A PRS consists of several unit cells and the beam tilting can be done by changing the reflection phase of the PRS. Below is the schematic diagram of the DRA structure with and without PRS.

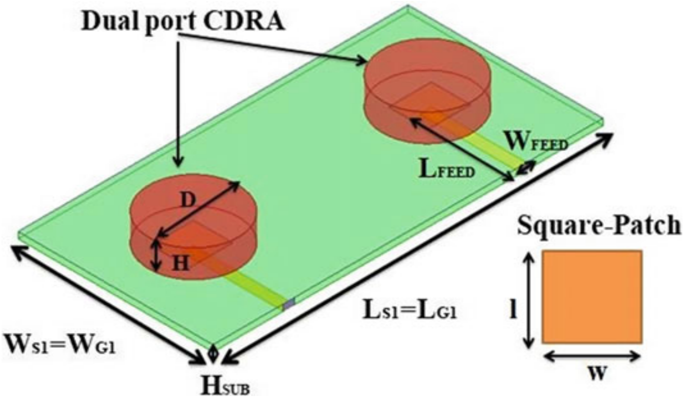


Fig. 1 Shows the dielectric resonator antenna (DRA) with cylindrical shape and square-shaped aperture

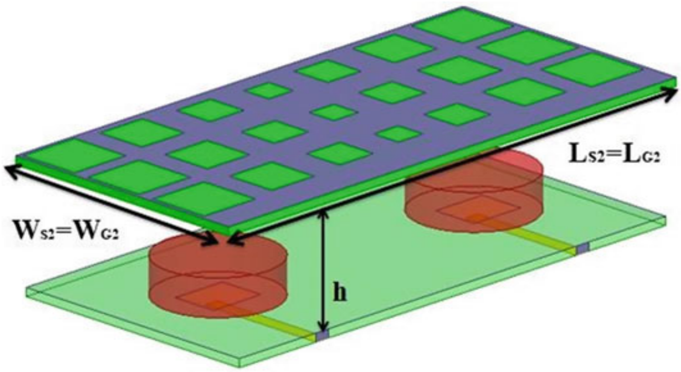


Fig. 2 Shows the DRA structure with the PRS structure as a superstrate

Below are the diagrams of PRS structure, showing both sides of inductive (apertures, top) and capacitive (patches, bottom). The apertures on top are of varying shape causing phase variation. Continuous phase variation can be performed in two ways: by altering the capacitive grid dimension or by altering the inductive grid dimension [9]. The size of the patch is kept constant. The diagram of PRS is as shown in Fig. 3 for top and bottom faces.

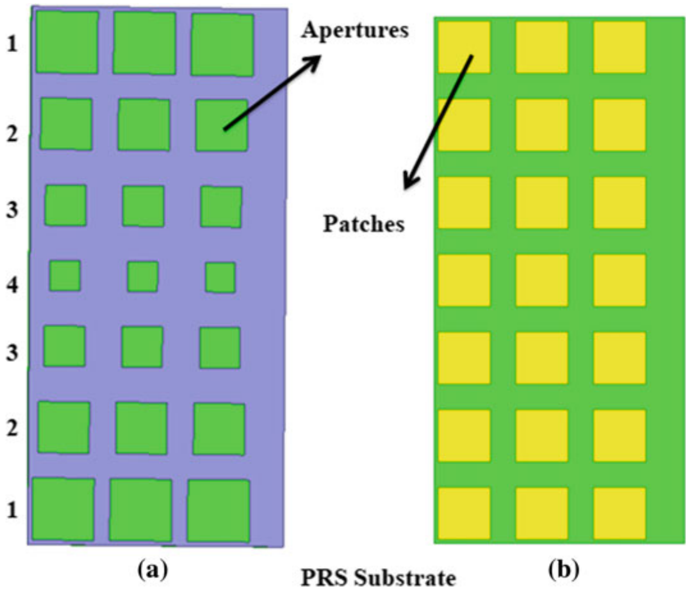


Fig. 3 Shows the PRS substrate **a** Top and **b** Bottom face of the structure

Table I Shows the proposed parametric values for the design of DRA with PRS

Symbols	Dimensions (mm)	Symbols	Dimensions (mm)
$W_{S1} = W_{G1}$	50	w	11
$L_{S1} = L_{G1}$	100	L_{FEED}	26
D	23	W_{FEED}	3
H	7	$W_{S2} = W_{G2}$	50
H	22.2	$L_{S2} = L_{G2}$	100
L	9	H_{SUB}	1.57

Table II Shows the unit cell no. and corresponding dimension of the aperture (Top of FSS). Whereas the patch size is kept constant at 10.75 m

Unit Cell No	Dimensions (mm)
1	12.1
2	10
3	7.9
4	5.8

3 Analysis of the Antenna Structure with and Without PRS

Figs. 4 and 5 show the reflection coefficient graphs of an antenna structure with and without PRS. Fig. 4 shows the $|S_{11}|$ and $|S_{22}|$ graphs where it shows that with PRS

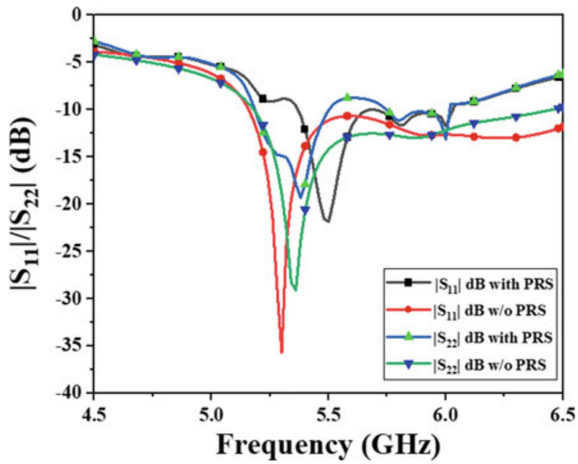


Fig. 4 Shows the $|S_{11}|/|S_{22}|$ graphs of structure with and without PRS

the reflection coefficient has reduced but a considerable value of −22.5 and −20 dB has been achieved at ports 1 and 2. In Fig. 5(a) and (b), it is shown that the $|S_{12}|$ and $|S_{21}|$ graphs of the proposed structure have been slightly improved in comparison to the structure without PRS. Two bands have been achieved in the graph from 5.18 to 5.64 GHz and 5.76 to 5.98 GHz. A considerable gain of 5.35 dB is achieved in the proposed structure. The $|S_{12}|$ and $|S_{21}|$ graphs show dips at 5.5 and 5.38 GHz, respectively.

The length and the width of the rectangular patch, used on the substrate for coupling the microstrip line with the dielectric resonator antenna, show the variation of $|S_{11}|$ graph. The design of the patch has been optimized significantly, as shown in Figs. 5 and 6.

Further analysis of the radiation pattern reflects the beam tilting feature of the proposed structure. Here the major lobe of the pattern has been tilted by 50° on Port 1 at 5.24 GHz and by 40° on Port 2 at 5.26 GHz. Below are the radiation pattern diagrams from Port 1 and Port 2, respectively.

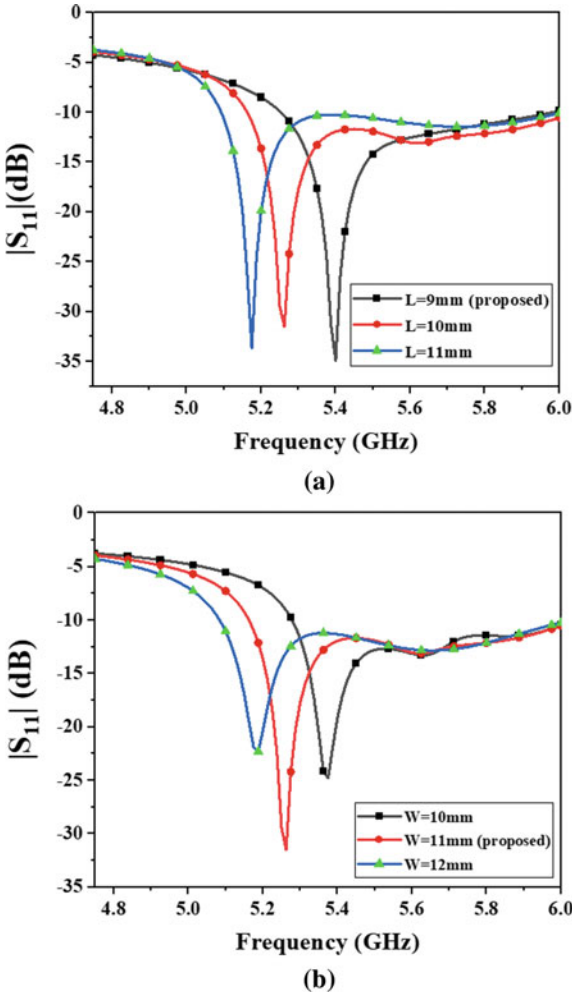
Here, Fig. 9 shows the E-field distribution at 5.4 GHz. A single mode of HE_{11} is achieved when the DRA is placed on a square-shaped aperture.

The resonant frequency for the mode $HE_{11\delta}$ of a cylindrical DRA can be computed mathematically. Initially, here we calculated the effective permittivity of the DRA structure and the effective Height of the CDRA from Eqs. (2) and (3) and then substitute the value in the first equation [10, 11].

$$\mathrm{f_{r,HEM_{11\delta}}} = \frac{6.324\mathrm{c}}{2\pi \mathrm{D1}\sqrt{\varepsilon_r}}\left\{0.27 + 0.36\left(\frac{\mathrm{D}}{2\mathrm{H}}\right) + 0.02\left(\frac{\mathrm{D}}{2\mathrm{H}}\right)^2\right\} \tag{1}$$

$$\varepsilon_{dra,eff} = \frac{H_{eff}}{\frac{H}{\varepsilon_{Al_2o_3}} + \frac{H_{sub}}{\varepsilon_{sub(FR-4)}}} \tag{2}$$

Fig. 5 Shows the variation in $|S_{11}|$ graph **a** with the change in length **b** with the change in width, of the rectangular patch



And

$$H_{eff} = H + H_{sub} \tag{3}$$

By exercising the above calculation, the resonant frequency so obtained as 5.4 GHz from Eq. (1).

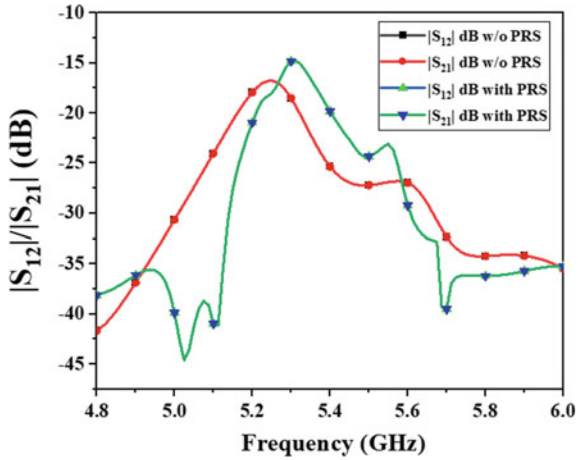


Fig. 6 Shows the $|S_{12}|/|S_{21}|$ graphs of the structure with and without PRS

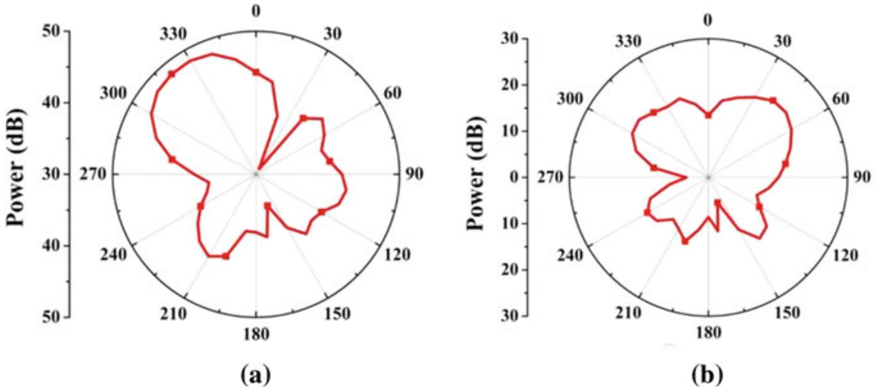
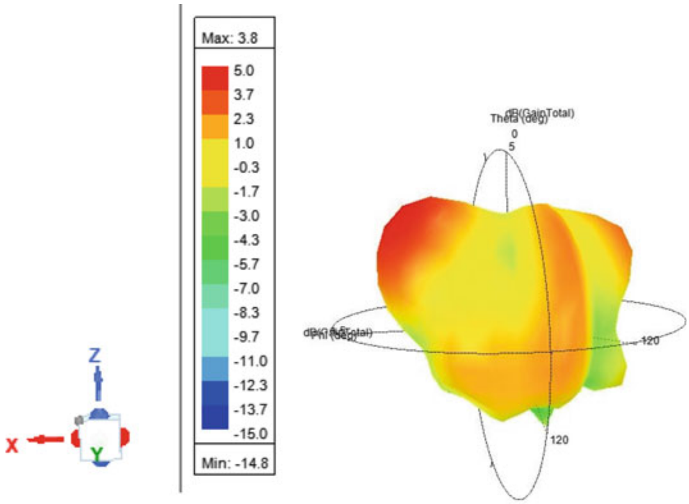


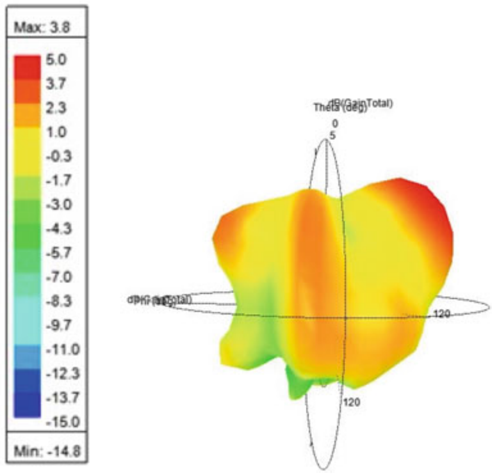
Fig. 7 Shows the beam tilting feature of the PRS on the antenna **a** at 5.24 GHz and **b** at 5.26 GHz

4 Antenna Analysis Using Diversity Parameters

One of the essential diversity parameters that illustrate the performance of the constructed antenna is the ECC and MEG. In MIMO applications, the Envelope channel coefficient and diversity gain are two critical diversity factors. In realistic MIMO antenna systems, ECC is critical for evaluating the impact of one antenna's radiation pattern on the others. When placed in the same unit cell, for better performance ECC (<0.5) is in the acceptable range. Lower the ECC, better the isolation. From Fig. 10 ECC is extremely close to zero, indicating that it has a favorable radiation pattern in MIMO wireless applications. The Scattering parameter can be used to compute ECC using a given equation [12, 13].



(a)



(b)

Fig. 8 Shows the 3-D Polar Plot of **a** Port 1 at 5.24 GHz and **b** Port 2 at 5.26 GHz

$$\rho_{enm} = \left| \frac{\left| S_{nn}^{*} S_{nm} + S_{mn}^{*} S_{mm} \right|}{\left| \left(1 - |S_{nn}|^2 - |S_{mn}|^2\right)\left(1 - |S_{mm}|^2 - |S_{nm}|^2\right) \right|^{1/2}} \right| \tag{4}$$

The value Diversity gain (DG) should ~10, in the above Fig. 10 it is observed to distinguish that the value of DG is closer to 10 at the operating frequency, and also it can be found by the given formula [14].

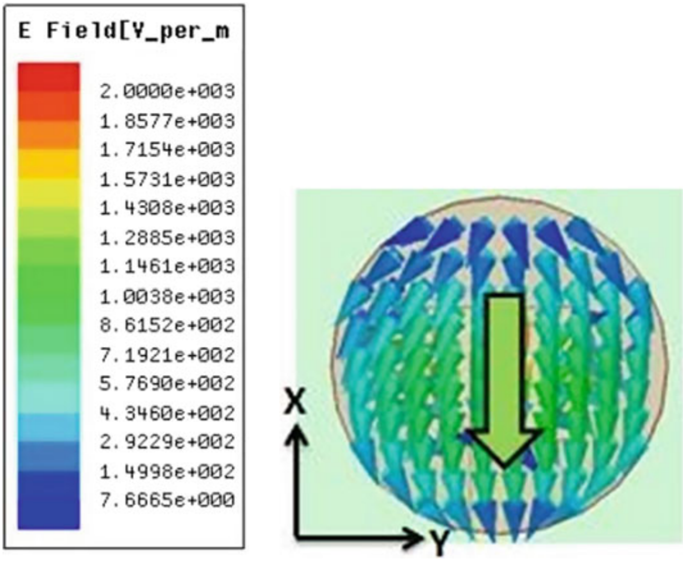


Fig. 9 Shows the E-field distribution at 5.4 GHz (top view)

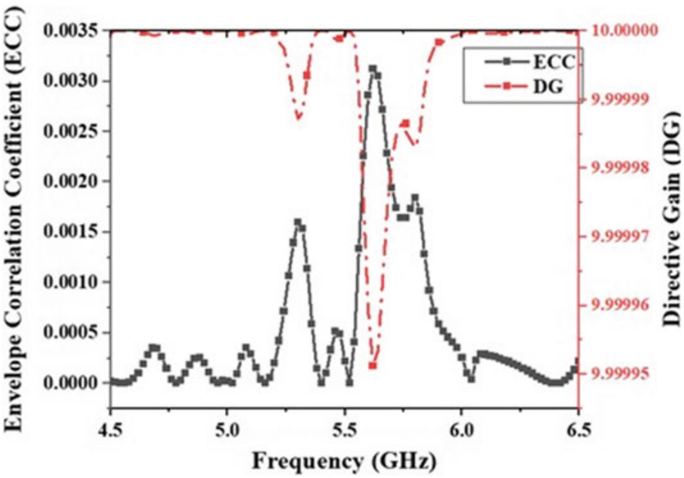


Fig. 10 Shows graph of ECC and DG vs Frequency (GHz) for the proposed MIMO antenna

$$DG = 10\sqrt{1 - ECC^2} \tag{5}$$

Above shows the graph of ECC and DG for the proposed structure with the PRS sheet.

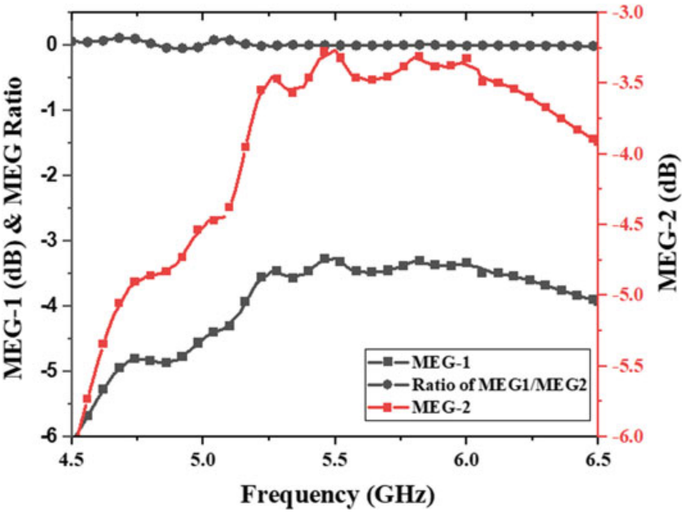


Fig. 11 Shows the graphs of MEG 1, MEG 2, and the Ratio of MEG |MEG1/MEG2 | of the proposed MIMO antenna

MEG is also used to determine whether the antenna is suitable for isotropic or Gaussian media in both indoor and outdoor applications. Figure 11 shows that the antenna is only suitable for the Gaussian medium since its value is less than −3.0 dB in the 2:1 VSWR frequency region for the suggested design [11].

5 Conclusion

Thus in the featured work, with the help of DR-type structure, the authors have been able to achieve a considerable gain of 5.35 dB and dual-band operating range, which makes it suitable for sub 6 GHz range of frequencies. Also, the PRS structure helps to achieve a considerable level of tilting on both the ports providing bi-directional diversity due to beam tilting. The feature of beam-tilting makes it suitable for vehicular communication.

References

1. Zhou, Q., Dai, H.: Joint antenna selection and link adaptation for MIMO systems. IEEE Trans. Veh. Technol. **55**, 243–255 (2006). https://doi.org/10.1109/TVT.2005.861211
2. Petosa, A.: Dielectric Resonator Antenna Handbook. Artech House, Norwood, MA, USA (2007)

3. Mao, C.-X., Gao, S., Wang, Y.: Dual-band full-duplex Tx/Rx antennas for vehicular communications. IEEE Trans. Veh. Technol. **67**, 4059–4070 (2018). https://doi.org/10.1109/TVT.2017.2789250
4. Dwivedi, A.K., Sharma, A., Singh, A.K., Singh, V.: Circularly polarized quad-port MIMO dielectric resonator antenna with beam tilting feature for vehicular communication. IETE Tech. Rev. **39**, 389–401 (2021). https://doi.org/10.1080/02564602.2020.1862714
5. Das, G., Sahu, N.K., Sharma, A., Gangwar, R.K., Sharawi, M.S.: FSS-based spatially decoupled back-to-back four-port MIMO DRA with multidirectional pattern diversity. IEEE Antennas Wirel. Propag. Lett. **18**, 1552–1556 (2019). https://doi.org/10.1109/LAWP.2019.2922276
6. Hassan, T., Khan, M.U., Attia, H., Sharawi, M.S.: An FSS based correlation reduction TECHNIQUE FOR MIMO antennas. IEEE Trans. Antennas Propag. **66**, 4900–4905 (2018). https://doi.org/10.1109/TAP.2018.2842256
7. Dwivedi, A.K., Sharma, A., Singh, A.K., Singh, V.: Metamaterial inspired dielectric resonator MIMO antenna for isolation enhancement and linear to circular polarization of waves. Measurement **182**, 109681 (2021). https://doi.org/10.1016/j.measurement.2021.109681
8. Ratni, B., Merzouk, W.A., de Lustrac, A., Villers, S., Piau, G.-P., Burokur, S.N.: Design of phase-modulated metasurfaces for beam steering in Fabry-Perot cavity antennas. IEEE Antennas Wirel. Propag. Lett. **16**, 1401–1404 (2017). https://doi.org/10.1109/LAWP.2016.2639463
9. Qin, F., Gao, S., Mao, C., Wei, G., Xu, J., Li, J.: Low-profile high gain tilted-beam Fabry-Perot antenna. In: 9th European Conference on Antennas and Propagation (EuCAP). pp. 1–5 (2015)
10. Mongia, R.K., Bhartia, P.: Dielectric resonator antennas—a review and general design relations for resonant frequency and bandwidth. Int. J. Microw. Millimeter-Wave Comput. Eng. **4**, 230–247 (1994). https://doi.org/10.1002/mmce.4570040304
11. Srivastava, K., Dwivedi, A.K., Sharma, A.: Circularly polarized dielectric resonator-based multiple input multiple output antenna antenna with pattern and polarization diversity for vehicular applications. Int. J. Circuit Theory Appl. **49**, 3421–3433 (2021). https://doi.org/10.1002/cta.3078
12. Sharawi, M.S.: Current misuses and future prospects for printed multiple-input, multiple-output antenna systems [wireless corner]. IEEE Antennas Propag. Mag. **59**, 162–170 (2017). https://doi.org/10.1109/MAP.2017.2658346
13. Dwivedi, A.K., Sharma, A., Singh, A.K., Singh, V.: Quad-port ring dielectric resonator based MIMO radiator with polarization and space diversity. Microw. Opt. Technol. Lett. **62**, 2316–2327 (2020). https://doi.org/10.1002/mop.32329
14. Malviya, L., Parmar, A., Solanki, D., Gupta, P., Malviya, P.: Highly isolated inset-feed 28 GHz MIMO-antenna array for 5G wireless application. Procedia Comput. Sci. **171**, 1286–1292 (2020). https://doi.org/10.1016/j.procs.2020.04.137

New Idea for Characterizing of Waves Using Indian MST Radar

Rajendra Prasad Rao, Prakhar Yadav, and R. S. Yadav

Abstract The scientific idea of dynamic process in the earth's atmosphere are performed the past two decades using a 53 MHz Indian MST radar. The MST radar is located at Gadanki India. The MST radar consists of 1024 Yagi-Uda antennas. The antenna calibration is an important aspect for accurate measurements of atmospheric parameters MST radars. In paper we describe two new ideas to calibrate the large antenna system of the Indian MST radar. The "total power radiometer" idea has employed for the first time to adjust the antenna pattern in the mode. The new idea has enabled us to describe for the first time of the antenna pattern. The beam point accuracy in the E-plane. The antenna pattern is characterized with a 3 dB full beam width. The beam width and beam point accuracy are found to be close to the design values. The first side lobe level is 2 dB. The idea of the way antenna beam active radar experiments are conducted. The beam point accuracy has estimated to be using Indian radar waves from the moon. The development of signal-to-noise-ratio is observing the new develop antenna calibration method that used receiver detector.

Keywords MST radar · Yagi antennas · NARL · ADP · SNR

1 Introduction

The very high frequency (53 MHz) radar at Gadanki is a coherent mono-static pulse-coded Doppler radar with broadside phased array antenna. The average power aperture product of MST radar is 7×10^8 Wm^2. The MST radar is a unique system for scientific research of Earth atmosphere. The MST radar has the height regions

---

The original version of this chapter was revised: Affiliations of the authors have been corrected. The correction to this chapter is available at https://doi.org/10.1007/978-981-19-2631-0_66

R. P. Rao (✉) · P. Yadav
JK Institute of Applied Physics and Technology, University of Allahabad, Allahabad, India
e-mail: rajendra.rao434@gmail.com

R. S. Yadav
Electronics and Communication Engineering, JK Institute of Applied Physics and Technology, University of Allahabad, Allahabad, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_11

of Mesosphere (50–100 km), Stratosphere (17–50 km), Troposphere (up to 17 km). The MST radar studies extend up to earth's Mesosphere, a height up to 100 km.

The very high frequency (MST) radar sub systems are phased array with 1024 antenna elements, 32 number of triode-based transmitters. The generating power of MST radar is 2.5 MW. The control and instrumentation room consists of a single channel receiver with −165 dB sensitivity. The MST radar is a controller for automated experiments.

The MST radar is effectively operational for an average of 2500 h each year since installation during 1987–93. They produced many review publications on atmospheric science research. The planning and execution of continuous test maintenance activities, phase calibrations, upgrades of the radar electronic subsystems. The radar challenge is to sustain the quality of the scientific data product. The paper we provide the details of the radar subsystems for 25 years period for its continued successful operation.

The operating frequency of MST radar is 53 MHz. The distribute area in radar control and instrumentation room, transmitters rooms and antenna array that are about 20,000 m^2. The MST radar is a 1024-element array of Yagi-Uda antennas. The arrangement of antenna is a square grid of 32 rows and 32 columns. The Yagi-Uda antenna consists of a dipole, director and reflector. The dimensions are about 3 m length and inter-element spacing of 4 m. The control and instrumentation room consists of exciter, radar controller, large computer, back-end receiver, signal processor and offline data processing system. The 32 number of transmitters are distributed in four transmitter room. The MST radar of transmitter has two rooms each at north and south sides of the array. The transmitter room consists of 8 transmitters, signal distribution network and local processor computer. The use of a transmitter room is antenna phase shift loading and antenna beam navigation.

2 Preparation of Your Paper

2.1 A Brief Description of the Indian MST Radar

The Indian MST radar is a monostatic pulsed Doppler radar operating at 53 MHz. Important specifications of the radar system are given in Table 1. A functional block diagram of the radar system is given in Fig. 1. The Indian MST radar consists of two independent planar antenna each consisting of 1024 Yagi-Uda antennas in linear horizontal polarization. The MST radar is a linear array of 32 arrays, each consist of 32 antennas connected in phase.

Table 1 Specifications of MST radar

SN	Aspect	Specification
1	Location	Gadanki (13.5°N, 79.2°E)
2	Frequency	53 MHz
3	Average power aperture product	7×10^8 Wm^2
4	Peak power	2.5 MW
5	Maximum duty ratio	2.5%
6	Number of Yagi antennas	1024
7	Beam width	3º
8	Pulse width	1–32 μs
9	Pulse repetition frequency	Up to 8 kHz
10	Maximum number of range bins	256
11	Radar controller	Workstation with radar controller software (ADP)
12	Data acquisition	Two channel PCI-card based data acquisition system with 14-bit ADCs

2.2 *Antenna Phase Calibration*

The radar catered about 50 satellite launch campaigns with seamless support with 100% achievement for 25 years. The MST radar providing dynamic wind estimation on launch vehicle up to 20 km height. The performance of the Indian MST radar is measure wind vector for recent launch applications and scientific experiments. The recent measurement of antenna pattern by radio source and active probing of moon showed the antenna pattern pointing accuracy to be within 1%.

The MST radar has been successfully operational by adapting the technological developments. The MST radar has beam steering capability expanded from 5 beam mode to automated 82 beam modes of experiments with beam tilt of 1 degree resolution. The improvement of turbulence due to crystal oscillators in microprocessor systems has solved maximum amount of turbulence problems. The MST radar capability is for wind measurement up to tropopause height.

The development of Atmospheric Data Processor (ADP) software use in processing of the wind measurements. The offline data processing system was based on UNIX program. A magnetic tape data storage system used to record the digital data. The continuous update of the data archival system form tape drive to floppy disks and CD based.

The MST radar control and instrumentation room was initially consisted with 8086 microprocessors. The 8086 microprocessor is a 16-bit. The radar has exciter and coder system, a super heterodyne receiver, local oscillator and a quadrature detector operating at intermediate frequency. The two-channel signal processor for time domain signal processing of the in-phase and quadrature-phase channels is

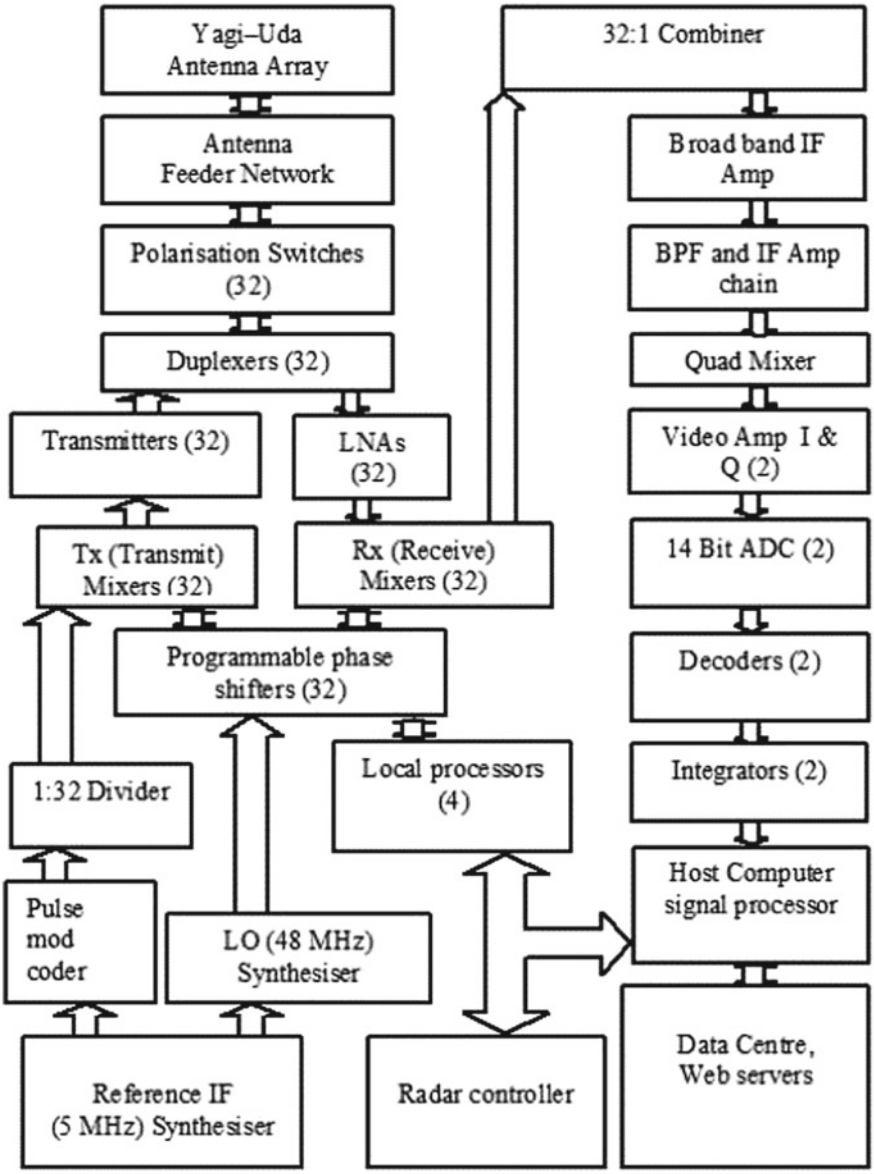


Fig. 1 A Functional block diagram of the Indian MST radar

based on 8086 microprocessors for analog to digital conversion. The radar system has been upgraded multiple times, to catch up the unique changes in technology.

A PCI-bus based data acquisition card developed with two channels of 14-bit ADCs (Analog to digital conversion). The output of each ADC is fed to achieve the decoding operation for coded transmission. The phase shift data of MST radar is loaded to digital phase shift by the 8085 microprocessors with RS-232 serial link. The 8085 is a 8-bit microprocessor. The interface between radar control and instrumentation room and the transmitter rooms are 4 local processor rooms. RS-232 connectivity is upgraded with a fiber optic Ethernet connection to the radar controller. The coaxial cable-based timing signals network is replaced by the optical fiber-based instrumentation network, to overcome the problems during the lightning strikes. The radio frequency sub systems are also tested every year for proper power levels. In radar all the external radio frequency cables between the modules were replaced to overcome the connectivity problems during test and maintenance activities.

The transmitters are maintained to about 85–95% availability for each scientific experiment. The failures of active components of transmitter amplifiers triodes and duplexer pin diodes are attributed to the frequent power outing when the system is operational. The solid-state amplifier module has been upgraded by MRF 141 MOSFET based amplifier. The radar operating in class A, AB the provide an output power of 100 W with a bandwidth of about 8 MHz for an input signal of 0 dBm. The major problems in the transmitter high power modules are failure of Triode output RF coupling capacitors.

The challenge in the failure of active components like triodes and pin diodes is overcome by the installation of the uninterrupted power supply for transmitters. The failure of the high voltage capacitors by the transmitter room air-conditioning. The high-power pin diode based solid state balanced duplexers and triode based high power transmitters demanded daily attention. The lumped element branch line duplexers upgrade and replacement were performed. The annual maintenance activities of transmitters are performed by refurbishing high voltage and RF cavities. The verification of electrical terminals connectivity, testing and adjusting the electrical power supplies of triodes' anode, cathode, filament to within 5% tolerance levels. Fine tuning of transmitter radio frequency cavities are performed for required gain and bandwidth specifications for maintaining the transmitter output RF power. This activity used to take almost one month period with a team of about 10 technicians. Once in three months RF tank circuits' tuning is performed. Total radar preventive maintenance performed for about a month period during December/January. The annually to test each subsystem and the expected problems are solved immediately. RF pulse shape and power level is monitored fortnightly. Daily monitoring of the radar health and data quality, radar performance is monitored.

2.3 Characterization of Antenna Pattern

The Yagi-Uda antenna array phase equalization activities are called phase calibration. They are performed with monthly periodicity. The antenna voltage standing wave ratio (VSWR) is being monitored regularly in the front panel meters of transmitters.

The weather exposed three-element Yagi-Uda antennas and feeder network in the array are maintained with minimum number of spares. The RF feeder network connectivity issues are obviated by yearly application of the RTV sealants' to arrest water entry into electronic modules.

The recent development of active array system with the MST radar antenna array is extending the latest technological advancements. Wind profile up to 20 km is obtained with half the rated power, eliminating RF power loss in the 100 m length cables. The Transmit/ Receive modules are placed near the antenna elements. The MST radar is operational for all the scientific operations along with the parallel execution of the massive development activity in the antenna array for the TR modules, underground electrical, RF and optical fiber cables installation.

Acknowledgements This activity has been carried out by the dedicated technical team for years together at NARL, an autonomous organization under the Department of Space, with the funding of Government of India. Author is grateful to the technicians working with MST radar for their support in carrying out the activities presented in this paper. Discussions and guidance by generation of NARL during the course of the maintenance and upgrade activities are highly respected.

References

1. Balsley, B.B., Gage, K.S.: The MST radar technique: potential for middle atmospheric studies. Pageoph **18**, 453–493 (1980)
2. Blood, S.P., et al.: Studies of high latitude mesospheric turbulence by radar and rocket 2: measurements of small scale turbulence. J. Atmos. Terrest. Phys. **50**(10–11), 963–976 (1988)
3. Chakrabarty, D.K., Beig, G., Sidhu, J.S., Das, S.R.: Fine scale structure and turbulence parameters in the equatorial middle atmosphere. J. Atmos. Terr. Phys. **51**, 19–27 (1989)
4. Avery, S.K., Avery, J.P., Valentic, T.A., Palo, S.E., Leary, M.J., Obert, R.L.: A new meteor echo detection and collection system: Christmas Island mesospheric measurements. Radio Sci. **25**, 657 (1990)
5. Balsely, B.B., Gage, K.S.: The MST radar technique: potential for middle atmospheric studies. J. Pure Appl. Geophys. **118**, 452 (1980)
6. Balsley, B.B.: Poker flat MST radar measurements of winds and wind variability in the mesosphere, stratosphere and troposphere. Radio Sci. **18**, 1011 (1983)
7. Barnett, J.J., Corney, M.: Temperature data from satellites. Handbook for MAP **18**, 3 (1985)
8. Rao, P.B., Jain, A.R., Kishore, K., Balamuralidhar, P., Damle, S.H., Viswanathan, G.: Indian MST radar system 1, system description and sample wind measurements in ST mode. Radio Sci. **30**, 1125–1138 (1995)
9. Prasad, T.R., Patra, A.K., Anandan, V.K., Satyanarayana, P.: Employment of new techniques for characterizing Indian MST radar phased array. IETE Tech. Rev. **33**(6), 584–595 (2016)
10. Hocking, W.K.: A review of Mesosphere–Stratosphere–Troposphere (MST) radar developments and studies, circa 1997–2008. J. Atmos. Solar-Terrest. Phys. **73**(9), 848–882 (2011). https://doi.org/10.1016/j.jastp.2010.12.009

11. Renkwitz, T., Stober, G., Latteck, R., Singer, W., Rapp, M.: New experiments to validate the radiation pattern of the Middle Atmosphere Alomar Radar System(MAARSY). Adv. Radio Sci. **11**, 283–289 (2013)
12. Chau, J.L., Renkwitz, T., Stober, G., Latteck, R.: MAARSY multiple receiver phase calibration using radio sources. J. Atmos. Solar-Terrest. Phys. **118**, 55–63 (2014). https://doi.org/10.1016/j.jastp.2013.04.004

Face Mask Detection and Counting Using Deep Learning and Embedded Systems

Monil Jethva, Stavan Ruparelia, and Ruchi Gajjar

Abstract The pandemic of the coronavirus COVID-19 is having an effect on public health around the globe. The WHO has advised a number of precautions to control the spread of the virus, one of which is the proper usage of a face mask. However, many people are violating this simple measure in public places, which could increase the risk of virus transmission. This inspires us to develop a face mask detection system using deep learning to identify whether individuals are wearing masks and how well they are wearing them in real-time. The system is proposed with two cutting-edge object detection models, YOLOv3 Tiny and YOLOv4 Tiny. This system is capable of detecting and classifying face masks as well as their wearing condition and provides class-wise counting in real-time. For real-time inference, these models were deployed on NVIDIA Jetson Nano hardware. The detection models obtained mAP of 84.2% and 88.4% and FPS of 23.94 and 23.58 for YOLOv3 Tiny and YOLOv4 Tiny respectively. In terms of mAP, YOLOv4 Tiny performed better than YOLOv3 Tiny while both have nearly the same FPS, thereby giving an efficient mask detection in real-time, making it a deployable system for the public or crowded places.

Keywords COVID-19 · Deep Learning · Face Mask Classification · Face Mask Detection

1 Introduction

The global crisis has been brought on by the breakout of coronavirus, which has a major impact on global public health and the economy. Globally, over 178 million COVID-19 cases and 3.86 million deaths had been confirmed as of 20th June 2021 as per WHO records [1]. This pandemic forced us to initiate new rules to control

M. Jethva · S. Ruparelia (✉) · R. Gajjar (✉)
Department of Electronics and Communication Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India
e-mail: stavanrupareliya7878@gmail.com

R. Gajjar
e-mail: ruchi.gajjar@nirmauni.ac.in

M. Jethva
e-mail: jethvamonil99@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_12

its deadly spread. WHO and Governments have provided guidelines that include wearing face masks [2], maintaining social distance [3], and much more in order to stop the transmission of this contagious virus. Researchers have found vaccines for this virus but their production at a large scale and access across the globe is still a challenge. Hence, until each and everyone is fully vaccinated and the virus is gone, precautions like wearing a face mask and keeping a safe distance from others are essential. To ensure that mandatory precautions are taken, surveillance is required at public places, corporate places, and especially crowded areas. Manual monitoring of preventive measures with manpower in a crowded place is less accurate and still practically not possible. This could be implemented using the surveillance systems integrated with artificial intelligence.

This paper's contributions are summarised as follows:

- A robust system is proposed which can detect human face mask efficiently and classify them based on their wearing condition i.e. With Mask, With Improper Mak and Without Mask in real-time and also generates class-wise counting of the same.
- For real-time inference, the proposed face mask detection system is deployed on the NVIDIA Jetson Nano.
- The proposed system can detect, classify and count human face mask and their wearing conditions very accurately in real-time under various lighting conditions, dimension, color and crowd.

The following is an overview of the rest of the work mentioned in the paper. The related work is summarised in Sect. 2. The proposed system methodology is discussed in Sect. 3 along with a description of the dataset and hardware used. The experimental outcomes are described in depth in Sect. 4. At last, Sect. 5 discusses the proposed work's conclusion.

2 Related Work

Deep learning is used significantly in a variety of real-world applications. Since the outbreak of the COVID-19 virus, substantial advances in the field of machine learning has been made in the identification of face masks. A model SSDMNV2 is suggested to identify the mask on the face in work presented by [4]. This model classified whether the individual was wearing a mask or not. They used TensorFlow, Keras, and OpenCV to achieve a higher detection accuracy. The authors of [5] proposed a technique for automatically detecting the wearing conditions of the face mask for smart city real-world application. For the detection of the face mask, they proposed a deep learning-based architecture. Also, the proposed architecture was deployed on CCTV cameras for real-time testing.

In [6], the authors have proposed an approach for face mask detection as well as classification in real-time. They trained YOLOv3 and Faster R-CNN object detection models using a dataset of 7500 images. Faster R-CNN provides higher precision while YOLOv3 detects efficiently with less inference time comparatively. Also, it concludes that the YOLOv3 algorithm is reliable for real world applications and it is easily deployable in surveillance cameras. In the work reported by J. Zhang et al. [7], the research detects human faces with the condition of the wearing mask, i.e. properly or improperly worn mask, to prevent the spread of COVID-19 disease. Here they proposed a practical dataset and new detection framework. This system effectively detected face mask wearing conditions, giving a mAP of 84.1%.

In [8], the authors suggested a new concept for mask detection using various YOLO models as well as a new architecture. They used a variety of datasets to train deep learning models, as well as the dataset proposed. Moreover, the mAP of existing deep learning models was compared to the proposed models. The comparison revealed that the proposed models performed better than existing models. In [9], to detect medical face masks, the authors proposed a new architecture combining YOLOv2 with ResNet-50. The authors used a combination of two available datasets, Medical Masks Detection (MMD) and Face Mask Detection (FMD), to train the proposed architecture. The proposed architecture was trained using two optimizer, SGDM and Adam. The authors were able to achieve better average precision using Adam than with SGDM, according to the results. To summarise the outcomes of these research, deep learning seems to be a very effective method for detecting and classifying masks on face recognition.

Recently, machine learning models have been deployed on low computing devices. In [10], many such applications have been discussed along with the latest optimization techniques. With the help of different optimization techniques and the latest frameworks, deployment of computational intensive machine learning models has been made possible with negligible loss in the accuracy. However, training of models are still memory and computational intensive hence require high computing devices or GPU servers for training of the models.

3 Proposed Methodology for Mask Detection and Classification

3.1 System Overview

In this proposed work, we have implemented and compared different Convolutional Neural Network architectures which can detect and count the face mask as well as classify into three classes based on its wearing condition i.e. With Mask, With Improper Mask, or Without Mask in real-time. For this work, we have manually collected the data from different sources, pre-processed it and annotated the images. After creating a processed dataset, we used pre-trained Models, YOLOv3 Tiny and YOLOv4 Tiny, and trained them on our own dataset. After training each model, we deployed these models on NVIDIA Jetson Nano hardware and attached a camera to Jetson Nano for taking in real-time video feed. Figure 1 depicts the proposed system's

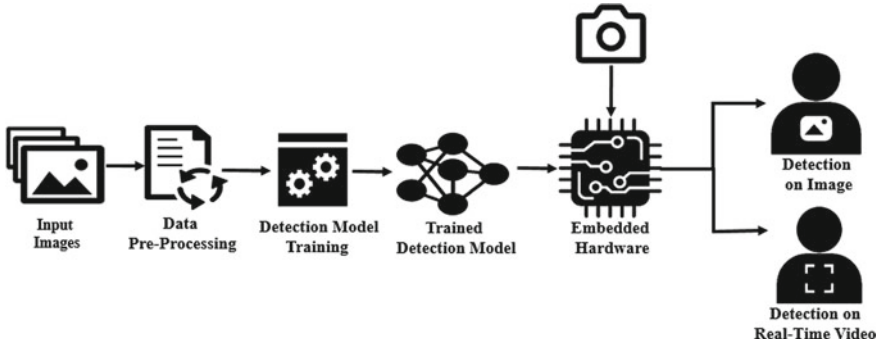


Fig. 1 Diagram of the proposed system

block diagram. The proposed embedded system can efficiently detect face masks and wearing conditions i.e. With Mask, With Improper Mask, or Without Mask by making a bounding box around the face, in real-time along with the class-wise counting.

3.2 Dataset Preparation

3.2.1 Data Collection

In this paper, a deep learning based approach has been proposed for detecting and classifying face masks based on wearing conditions i.e. With Mask, Without Mask, and With Improper Mask, and display class-wise counting in real-time. To train YOLOv3 Tiny and YOLOv4 Tiny models, we needed a dataset that consists of images from each class but most of the available datasets were not suitable for our model training. Hence we decided to collect images and create our own dataset. We collected images from different sources available online as well as captured images with a camera. In total, we collected 2600 images out of which 1082 images were of With Mask Face (assigned a label of Class 0), 1000 images were of Without Mask (assigned a label of Class 1) and 518 images were of With Improper Mask Face (assigned a label of Class 2). Figure 2 shows some representative images from our newly prepared dataset.

3.2.2 Data Pre-processing

Collected raw images were having different height and width. For training the models, we required the uniform width and height of all images. Hence we resized all images into 512 × 512 size using Python script. All raw images were labeled manually using a labelImg [11] tool. The PASCAL VOC [12] format was used as it is a standard annotation format for object detection. An example of manual annotation of the image

Fig. 2 Examples of our face mask detection dataset

Fig. 3 Image annotation for dataset preparation

datasets is shown in Fig. 3. After manual labelling, the labelImg tool generated an XML file in which annotations are stored. Later, the dataset was splitted into two categories i.e. Training and Testing. We splitted the dataset with the ratio of 4:1 resulting into 2080 training images and 520 test images.

3.3 Detection Models

Object detection, which is a combination of object localization and classification, has been really challenging task for researchers. Achieving good accuracy in object detection usually involves a very complex and deep convolutional neural network architecture which increases the accuracy but comes with a trade-off of decrease in FPS as the processing power required increases. Another challenge was detecting objects of different size in a image using a single object detection model method.

YOLO is an object detection algorithm that balances between accuracy and FPS along with the advantage of detecting objects at different scales. In 2016, the first paper on YOLO [13] was published, after which many different variants of YOLO [14–17] were published. To detect objects, YOLO processes an image in two stages. The first stage extracts features from the input image that will aid the detection of objects, in the input image, by the second stage. The second stage is made up of detection layers. These detection layers take extracted features from the first stage as an input and generate the bounding boxes, class detected and a confidence score for the objects detected in an image.

In this paper, we have used two different versions of YOLO; YOLOv3 Tiny and YOLOv4 Tiny. The reason we chose these Tiny versions is because the processing power required for YOLOv3 and YOLOv4 is pretty high for the NVIDIA Jetson Nano. To detect objects, these models use only convolution layers hence they are also known as fully deep convolutional neural networks. For feature extraction, Convolution layers are used in YOLOv3 Tiny, along with Batch Normalization [18] and Leaky ReLU activation function and also uses Max Pooling layers. Whereas YOLOv4 Tiny uses Batch Normalization layer and Leaky ReLU, an activation function following Convolutional layer along with Cross Stage Partial (CSP) [19] architecture. After feature extraction, YOLOv3 Tiny and YOLOv4 Tiny, both use Batch Normalization layer and Leaky ReLU, an activation function following Convolutional layer as a detection layers to detect objects at various scales from an input image.

3.4 Hardware Description

To train detection models, we require high processing power. Hence we decided to train both models using Google Colab with GPU acceleration. Google Colab allotted Tesla T4 GPU of NVIDIA having CUDA version 11.2. NVIDIA Jetson Nano hardware with a camera module was used to implement the trained models in real-time on the embedded system. It has a 128 core NVIDIA Maxwell GPU with a maximum speed of 472 GFLOPs for hardware acceleration. Additionally, it has 4 GB RAM and other peripherals [20].

4 Experimental Results

Model training and validation loss vs. epochs graphs are discussed in this section. A brief discussion of model evaluation and performance comparison is provided. Lastly, real-time face mask detection and counting inference results are shown.

4.1 Model Training

To detect the face mask and classify its condition i.e. With Proper Mask, Improper Mask, or Without Mask in real-time, we trained YOLOv3 Tiny and YOLOv4 Tiny on Google Colab. Google Colab allotted Tesla T4 GPU of NVIDIA having CUDA version 11.2.

With a ratio of 4:1, our dataset was divided into training and test datasets, generating the training dataset having 2080 images and the test dataset having 520 images. We trained both models for 10 epochs with the step size of 520 and batch size of 4. Adam optimizer and decaying learning rate ranging from 1e−4 to 1e−5 was used for training YOLOv3 Tiny and YOLOv4 Tiny models. To observe the training process and avoid overfitting, we generated the graphs of GIoU Validation loss [21] Vs. Epochs, Probability Validation loss Vs. Epochs, Confidence Validation loss Vs. Epochs, and Total Validation loss Vs. Epochs for both models. In Fig. 4, graphs (a) and (b) are respective graphs of YOLOv3 Tiny and YOLOv4 Tiny. From graphs, we can track that confidence validation loss was much higher than other losses resulting in a higher value of total validation loss. Over time, the learnable parameters of the model improved by backpropagation resulting in a decrease in loss and an increase in model performance.

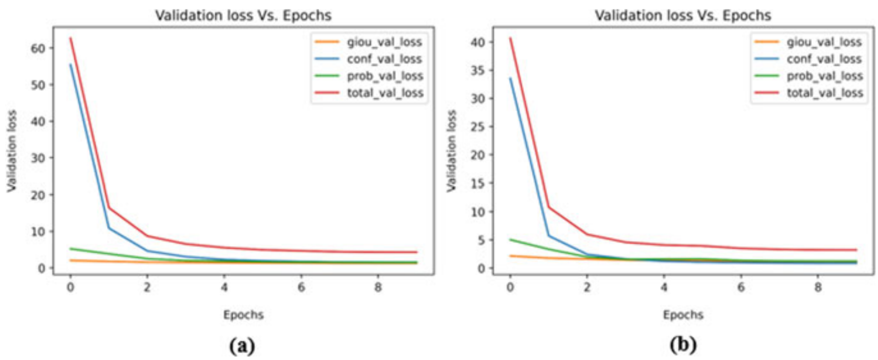


Fig. 4 Validation results **a** YOLOv3 tiny **b** YOLOv4 tiny

4.2 Model Evaluation

To assess the performance of YOLOv3 Tiny and YOLOv4 Tiny, we generated inferences on 520 test images and compared them with ground truth labels. If the predicted and ground truth bounding boxes' intersection over union (IoU) values are greater than the threshold, it is termed as True Positive (TP). If the IoU value is less than the threshold, it is termed as False Positive (FP). False Negative (FN) occur when ground truth bounding boxes exist but the model fails to predict them. These data can be used to calculate Precision, Recall, and f1 Score using Eqs. (1), (2), and (3), respectively. Table 1 presents the class-wise values of Precision, Recall, and f1 Score for both models. To assess the performance of the YOLOv3 Tiny and YOLOv4 Tiny models in terms of real-time face mask detection and its various conditions, such as With Mask, Without Mask, and Improper Mask, we calculated frames per second (FPS). We achieved FPS of 23.94 and 23.58 respectively for YOLOv3 Tiny and YOLOv4 Tiny.

$$Precision = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \tag{1}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{2}$$

$$f1\ Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{3}$$

Table 1 Model evaluation

Detection models	Class	Precision	Recall	f1 score
YOLOv3 Tiny	**With mask**	91.96	52.86	67.13
	With improper mask	94.35	45.09	61.02
	Without mask	94.87	51.95	67.14
YOLOv4 Tiny	**With mask**	94.24	56.18	70.45
	With Improper mask	93.88	48.83	64.24
	Without mask	95.15	55.04	69.74

4.3 Performance Comparison

To evaluate our system's performance, our results are compared with other models. Table 2 represents the comparative analysis among our models, YOLOv3 Tiny and YOLOv4 Tiny, with two other models, Faster R-CNN and SSD300. We have compared the class-wise accuracy and mean Average Precision (mAP) among these four models. As it can be observed from this table, the model YOLOv4 Tiny achieved the highest accuracy of 91.6% for class With Mask whereas the other three models have very similar accuracy. If we compare the accuracy of a class With Improper Mask, YOLOv4 Tiny and SSD300 achieved almost the same accuracy of 80%. For class Without Mask, YOLOv4 Tiny achieved the highest accuracy of 92.9% whereas YOLOv3 Tiny also achieved 89.1% which is significantly higher than Faster R-CNN and SSD300. Comparing the mAP of all these models, YOLOv4 Tiny achieved the highest mAP of 88.4% whereas YOLOv3 Tiny achieved mAP of 84.2% which is better than the Faster R-CNN and SSD300 model. Overall, the YOLOv4 Tiny model performed best as compared to other models.

4.4 Real-Time Face Mask Detection System Results

The detection models were evaluated in real-time using NVIDIA Jeston Nano hardware. YOLOv3 Tiny and YOLOv4 Tiny real-time inference test results are presented in Fig. 5. The proposed system can be seen to be capable of detecting and classifying human face masks, as well as accurately counting them by class. In Fig. 5, (a) and (c) is the inference results generated using the YOLOv3 Tiny model and (b) and (d) are the inference results generated using YOLOv4 Tiny. In Fig. 5, it can be observed that YOLOv3 Tiny and YOLOv4 Tiny, both models were able to detect and classify human face masks and their wearing condition efficiently in-crowd. As well as the counting of the face mask has been displayed class-wise on the top-left corner of the image. In the image (a), two with mask class and one with improper mask have been detected. As a result, its count has been displayed class-by-class in the image's top-left corner. When comparing images (c) and (d), YOLOv4 Tiny model was good at detecting without mask person from blurred region of the image. If we compare the results, both models detected almost the same bounding boxes for each type

Table 2 Comparative accuracy analysis of face mask detection models

Detection models	With Mask (%)	With Improper Mask (%)	Without Mask (%)	mAP
Faster R-CNN [22]	85.8	72.2	73.8	77.3
SSD300 [23]	88.0	80.9	74.9	81.2
YOLOv3 Tiny	85.5	77.9	89.1	84.2
YOLOv4 Tiny	91.6	80.7	92.9	**88.4**

Fig. 5 Real-time testing results: **a** and **c** from YOLOv3 Tiny, and **b** and **d** from YOLOv4 Tiny

of image. But if we compare the confidence score of each bounding box between YOLOv3 Tiny and YOLOv4 Tiny inference results, YOLOv4 Tiny was more confident than YOLOv3 Tiny. In terms of real-time inference performance, YOLOv4 Tiny outperformed YOLOv3 Tiny, making it an efficient choice for deployment on hardware for real-time mask detection in public or crowded places.

5 Conclusion

This paper proposed a real-time face mask detection, classification, and counting system to automate the surveillance of preventive measures to prevent the spread of COVID-19. This system identifies a human face mask and its wearing condition with bounding boxes of different colours. Popular models, YOLOv3 Tiny and YOLOv4 Tiny were trained on our collected dataset for this system. After evaluating trained models on multiple images, YOLOv3 Tiny achieved the mAP of 84.2% and FPS of 23.94 whereas YOLOv4 Tiny achieved the mAP of 88.4%

and FPS of 23.58. The real-time experimental results shows that YOLOv4 Tiny is more accurate than YOLOv3 Tiny. Numeric results also revealed that YOLOv4 Tiny has better accuracy than YOLOv3 Tiny while having almost the same FPS. The proposed models, YOLOv3 Tiny and YOLOv4 Tiny were deployed on embedded hardware to classify and count face masks in real-time. The proposed models can be integrated with surveillance systems at public and office areas to monitor if people are wearing masks properly or not.

References

1. Who coronavirus (COVID-19) dashboard. https://covid19.who.int/
2. When and how to use masks. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/when-and-how-to-use-masks
3. Coronavirus disease (COVID-19): Masks. https://www.who.int/news-room/q-a-detail/coronavirus-disease-covid-19-masks
4. Nagrath, P., Jain, R., Madan, A., Arora, R., Kataria, P., Hemanth, J.: SSDMNV2: a real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2. Sustain. Cities Soc. **66**, 102692 (2021)
5. Rahman, M.M., Manik, M.M.H., Islam, M.M., Mahmud, S., Kim, J.H.: An automated system to limit COVID-19 using facial mask detection in smart city network. In: 2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), pp. 1–5. IEEE (2020)
6. Singh, S., Ahuja, U., Kumar, M., Kumar, K., Sachdeva, M.: Face mask detection using YOLOv3 and faster R-CNN models: COVID-19 environment. Multimedia Tools Appl. **80**(13), 19753–19768 (2021)
7. Zhang, J., Han, F., Chun, Y., Chen, W.: A novel detection framework about conditions of wearing face mask for helping control the spread of COVID-19. IEEE Access **9**, 42975–42984 (2021)
8. Kumar, A., Kalia, A., Verma, K., Sharma, A., Kaushal, M.: Scaling up face masks detection with YOLO on a novel dataset. Optik **239**, 166744 (2021)
9. Loey, M., Manogaran, G., Taha, M.H.N., Khalifa, N.E.M.: Fighting against COVID-19: a novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. Sustain. Cities Soc. **65**, 102600 (2021)
10. Ajani, T.S., Imoize, A.L., Atayero, A.A.: An overview of machine learning within embedded and mobile devices-optimizations and applications. Sensors **21**(13), 4412 (2021)
11. Tzutalin: Labeling. Free Software: MIT License (2015). https://github.com/tzutalin/labelImg
12. Everingham, M., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A.: The pascal visual object classes (VOC) challenge. Int. J. Comput. Vis. **88**(2), 303–338 (2010)
13. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788 (2016)
14. Redmon, J., Farhadi, A.: YOLO9000: better, faster, stronger. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7263–7271 (2017)
15. Li, X., Wang, J., Xu, F., Song, J.: Improvement of YOLOv3 algorithm in workpiece detection. In: 2019 IEEE 9th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER) (2019). https://doi.org/10.1109/cyber46603.2019.9066490
16. Adarsh, P., Rathi, P., Kumar, M.: YOLOv3-tiny: object detection and recognition using one stage improved model. In: 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 687–694. IEEE (2020)

17. Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M.: YOLOv4: optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934 (2020)
18. Xu, B., Wang, N., Chen, T., Li, M.: Empirical evaluation of rectified activations in convolutional network. arXiv preprint arXiv:1505.00853 (2015)
19. Wang, C.Y., Liao, H.Y.M., Wu, Y.H., Chen, P.Y., Hsieh, J.W., Yeh, I.H.: CSPNet: a new backbone that can enhance learning capability of CNN. In: Proceedings of the IEEE/CVF Conference On Computer Vision and Pattern Recognition Workshops, pp. 390–391 (2020)
20. NVIDIA Developer: Nvidia Jetson Nano developer kit (2019). https://developer.nvidia.com/embedded/jetson-nano-developer-kit
21. Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., Savarese, S.: Generalized intersection over union: a metric and a loss for bounding box regression. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 658–666 (2019)
22. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. arXiv preprint arXiv:1506.01497 (2015)
23. Liu, W., et al.: SSD: single shot multibox detector. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) Computer Vision – ECCV 2016. ECCV 2016. LNCS, vol. 9905, pp. 21–37. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46448-0_2

Faith Value Based On-Demand Approach to Improve QoS in MANETs

Nishant Nandwani, Vishal Anand, and Surjeet Balhara

Abstract MANETs are made up of group nodes which are mobile and are connected by wireless manner in a self-building, self-mending network, and don't require an established infrastructure. Because the network topology of MANETs changes regularly, nodes are free to move around at will. Each node behaves as a router as well source and destination node. Major issue of MANETs is to require each and every device to regularly store the data necessary to correctly route traffic. In this paper, a problem related to broadcasting behavior of Ad hoc On-demand Distance Vector (AODV) routing protocol is addressed. The observed problem affects performance of Ant Colony Optimization (ACO) algorithm based AODV routing. The solution for observed problem is using the faith values, wherein it is assigned to every node in the network which describes the faithfulness of the node. Thus an improvement of ACO using faith values, protocol named as FVAODV is proposed, described, verified and simulated using NS-3 software tool in this paper.

Keywords Ad hoc on-demand Distance Vector (AODV) · Ant Colony Optimization (ACO) · Congestion · Mobile Ad-Hoc Networks (MANETs) · Quality of Service · Routing protocols · Throughput

1 Introduction

Mobile Adhoc Networks (MANETs) are wireless networks in which no infrastructure is required for nodes to communicate information in order to join the entire network [1]. While sending and receiving data, nodes may be able to shift from one point to another. When desired path for data transmission is found through message broadcasting, the delay is increased. As a result, effective routing methods are required to keep data flowing when nodes relocate. This includes, when necessary,

N. Nandwani (✉) · V. Anand · S. Balhara
Bharati Vidyapeeth's College of Engineering, New Delhi, India
e-mail: nishantnandwani.ece2@bvp.edu.in

V. Anand
e-mail: vishalanand.ece2@bvp.edu.in

S. Balhara
e-mail: surjeet.Balhara@bharatividyapeeth.edu

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_13

upgrading routing table to find better data transmission paths. Vehicle Area Networks (VANETs), Wireless Sensor Networks (WSNs), and Mesh Networks are the three main types of networks that a MANET can be categorized as [2]. In this paper, MANETs are considered, due to the challenges involved in using AODV protocol for infrastructure constrained networks. MANETs can either function alone or as part of a bigger internet [3]. With the inclusion of one or more distinct transceivers between nodes, they establish a highly dynamic autonomous topology.

The unscheduled real-time business can benefit from network architecture. The ad hoc networks can be used in a variety of ways. MANETs is a classification based on its application. MANETs can be used in road safety, home, healthcare, disaster rescue operations, air/land/navy defence, etc. [4].

MANETs are a collection of mobiles, laptops, and other devices which can on the spur of the moment link and allocate resources over wireless medium without the need for a fixed network infrastructure or centralized management [5]. Without any backbone support from fixed infrastructure network, MANETs can be put together quickly and at a low cost. Because of infrastructure less nature of MANETs, central monitoring is not possible. Mobile devices comprising MANETs must be compatible to communicate with each other without considering a type of device. As shown in Fig. 1, they can include laptops, mobiles, printers, routers, and, phones among other devices. The nodes are endowed with antennas that serve as both wireless transmitters as well as receivers, depending on their configuration. As per the application demand, omnidirectional, highly directional, or a combination of antennas may be used. Mobile nodes are limited in respect to bandwidth, network range and battery capacity due to limited resources available with them. In MANETs, various reactive routing protocols such as AODV, Temporally Ordered Routing Algorithm (TORA) protocol, Dynamic Source Routing (DSR) protocol, as proactive routing protocol is not efficient considering such ever-changing scenarios. When mobility is limited proactive protocol may be used with reactive protocol, like in Classless Inter-Domain Routing (CIDAR) protocol or Zone Routing Protocol (ZRP).

AODV protocols routing nature is that it broadcasts the route discovery packets throughout the network during route discovery process. Although Ant Colony Optimization (ACO) algorithm-based scheme improved the AODV routing protocol, it did not consider any QoS parameter and continued for route discovery by broadcasting as per the standard protocol [6]. To overcome this inefficient use of resources by simply broadcasting Route Request (RREQ) packet in network, a faith value-based approach is proposed. Faith Value is defined as number of received packets to the number of lost packets. Instead of broadcasting the RREQ packet, it is sent directly to the node with highest faith value in the network. In video streaming throughput is needed while End-to-End (E-2-E) delay can be compromised. Taking this into consideration, in this paper, throughput is considered as the Quality of Service (QoS) parameter and some constraint are applied on AODV protocol to improve network performance.

In this paper, Sect. 2 discusses the work related to throughput improvement by modifying AODV or any other standard protocol which has already been proposed by

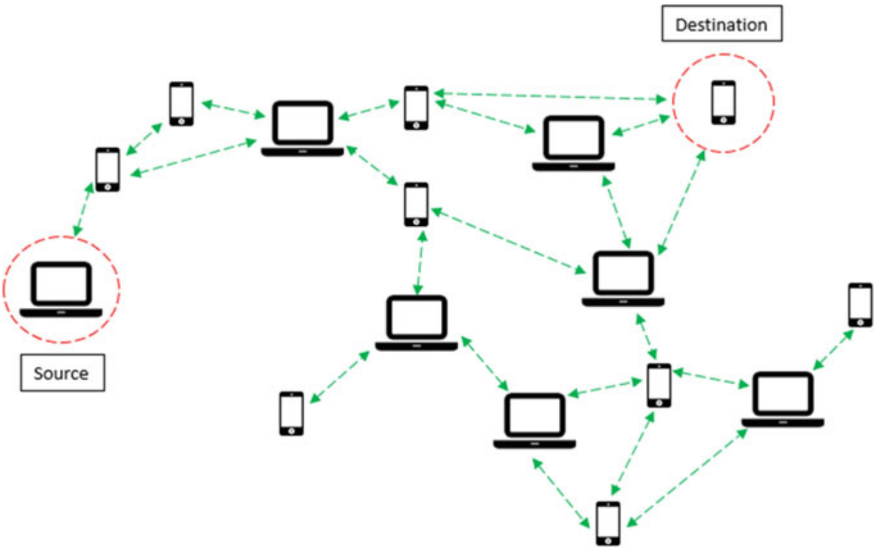


Fig. 1 Mobile Ad hoc network nodes

various authors. Section 3 explains the proposed work, which is built upon ACO algorithm. Section 4 provides simulation results and discussion. In the end, conclusion and future scope is discussed in Sect. 5.

2 Related Work

In their paper, S. Radha and S. Shanmugavel proposed that the source node is responsible for the formation of the power matrix [7]. The power level of each node in a matrix is recorded. After a period of time, the matrix is arranged in increasing order of powers, resulting in the formation of an index matrix. Afterwards, a frequency matrix is constructed by determining which node has responded with the same power for the greatest number of times. After that, a matrix is created by multiplying the frequency of occurrences of each node in the index matrix together. For each node in the table, the first entry represents the finest neighbor to a particular node, and the following entry represents the successive finest neighbor, and trend continue. MAC and AODV protocols were proposed by M. Adil and colleagues, as a hybrid protocol [8]. Both the MAC Address and the IP Address are used to identify a machine on the internet in a unique manner. The MAC Address is used to ensure that the physical address of the computer is different from other computers. IP Address is a logical address assigned to a computer and is used to uniquely identify and locate computers connected to a network system. In the concerned database, the MAC address of each device is stored in a mac table, which is unique to that database (offline). A. A. Aouiz et al conveyed that multipath mechanisms can diminish blockages in the network [9]. E. O. Ochola et al. proposed adding an extra authentication packet to the process of route discovery [10]. When the source node receives an RREP packet, it saves the packet in table maintained for routing proposes and subsequently passes on an instantaneous authentication packet over the backward route of the received RREP packet. It is made up of a random number generated by the source node and sent to the destination node. When more than two authentication packets arrive at the destination node, each arriving via a different route, the destination node keeps them in its routing table and verifies to see if they store the same arbitrary numbers as the preceding ones. If the authentication packets contain the same arbitrary numbers even though they are sent via a dissimilar pathways, the authentication confirm packet is transmitted from the destination node to the source node by the destination node. Among the contents of the confirm packet is an arbitrary number produced by the destination node. If each authentication confirm packet stores a dissimilar set of arbitrary numbers, the source node will wait till more than two authentication confirm packets contain the same set of arbitrary numbers before proceeding with the authentication. When the source node receives two or more authentication confirm packets with the same arbitrary numbers, it will choose the shortest route to forward the data to the destination node in order to avoid wasting resources. The mechanism for security in this approach is that a malicious node posing as the destination node will not dispatch the correct authentication confirm packet to the source node if the malicious node is impersonating the destination node. Ahmed M. Abd Elmoniem et al. suggested an AODV and Ant Colony Optimization (ACO)-based multi-path routing protocol [6]. This protocol is referred to Multi-Route AODV Ant routing (MRAA). Antra Bhardwaj and Hosam El-Ocla proposed a protocol for routing titled as Ad hoc On-demand Multipath Distance Vector with Fitness Function (AOMDV-FFn) [11]. The authors integrated the AOMDV mechanism on top of the genetic algorithm. Even if some loss of packets occurs, the protocol enables an enhanced mechanism to choose the most effective routes with the highest fitness values, taking the shortest path, maximum surplus energy, and minimal traffic of data. They proposed a mechanism in which the TCP Congestion Control Enhancement for Random Loss (TCP CERL) in the FFn might be used to optimize optimal route. The goal of protocol proposed by Ibrahim Kacem et al. was to find the minimal-value stake in nominal capacities that provides small traffic routing and tenacity in the case of any node collapse [12]. When taking in account the routing and also the decision functions which are replicated using a synchronized fuzzy transition proposal, fuzzy synchronized Petri net is heavily used, and the ant system is used to solve issue of unknown scenarios in ad-hoc networks. The acquired results of synchronized Fuzzy Ant System (SynFAnt) exhibits the efficacy of SynFAnt approach. SynFAnt protocol for routing enhances the packet delivery ratio, throughput and end-to-end delay by reducing their jitter and increasing their acceptance rate. The non-cooperative behavior of a mobile node, according to Prasannavenkatesan Theerthagiri and Menakadevi T, leads to a major issue in routing and networking [13]. Non-cooperativeness is due to resource limitations like non-centralized nodes, battery capacity, or harmful nodes, among other things. Effective routing paths can only be provided by network nodes that work together as a group. Chunfeng Liu et al. proposed a protocol based on kalman filter theory. In this, an estimated state of previous transmission is used to predict current state of a node for better neighbour discovery. This approach, though applied in VANETs in the said paper, can be used in future along with working of FV in the same network [14]. M. Anand proposed intelligent routing AODV (IRAODV) to reduce energy consumption at nodes in MANETs. It works by finding nodes in a particular region and letting one node out of those in that region to transmit the packet. Same is done for acknowledgement of these packets. This approach resulted in better performance of the proposed approach to that of AODV [15].

3 Proposed Work

AODV is intended for starting on itself in an environment with mobile nodes, and to withstand a mixture of network behaviors such as node mobility, link failures, and packet losses, among other things. A brief description of the AODV protocol is provided in the first part of this section. AODV maintains a routing table at each node in the network at the time of route of discovery phase. Nodes discover routes in AODV through the use of route request (RREQ) cycles. RREQ messages are broadcast from a node's neighbor nodes in order to request a route to a destination. AODV protocol finds the shortest path between the source and destination nodes. It does not take into account any QoS parameters in its operation and is thus modified accordingly.

3.1 AODV

AODV routing protocol offers unicast, broadcasts and multi broadcast communication for ad hoc transmission in a network [16]. If route is required by a source node, AODV commences route discovery. Routes are kept as long as they are needed by source node and the routes always exist using sequence numbers. In AODV, nodes keep an information in route table for next hop which is stored till target node (Fig. 2).

In route discovery process, after receiving RREQ message but failing to provide any route to the required destination, a node broadcasts an RREQ message to all other nodes within range of the node that received the message. In addition, it retains a reverse-path to the origin node, which might be utilize to send following responses to this RREQ after the initial request. This process is repeated until it finds a node with correct route to the destination. Now, this node that received RREQ message replies by sending Route Reply (RREP) acknowledgement. A bidirectional route is established between the requesting node and the destination at the conclusion of this request-response cycle. As soon as connectivity is lost by a node to its following hop, then the same node annuls its route by broadcasting RERR to each and every node that may have encountered its RREP.

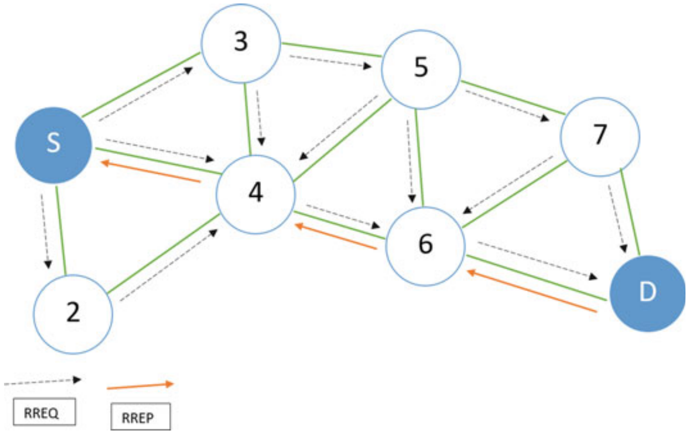


Fig. 2 Mechanism of AODV

3.2 Faith Value AODV (FVAODV)

In this paper, a faith value-based protocol is proposed that is further build on the ACO algorithm. ACO is based on the food searching behavior of ants [6]. Ants leave a chemical, named pheromone, on the path that they travel. The remaining ants after the first one follows the pheromone tracks of first ant. Similarly, authors proposed to use this pheromone value in a network, so that RREQ packets in the network follow that first RREQ packet which went through an efficient route and left behind its pheromone. The faith value (FV) for a node is defined as a ratio of received packets to discarded packets by that node. Higher the FV, more is the faithfulness of node in the network. The issue with ACO is that it broadcasts messages in the AODV protocol. Instead of broadcasting itself from a source node, an ant packet in the network will go directly to a member node with a higher faith value. Framework of the proposed method can be explained as follows (See Fig. 3):

ACO Algorithm with FV

Nodes keep a FV of all one hop neighbors
Check Condition
if There is no route to the destination and no neighbor with highest pheromone
then Take path with highest faith value
else if There is no path to the destination nonetheless an operative neighbor with the greatest pheromone exists
then transmit the packet to this neighbor

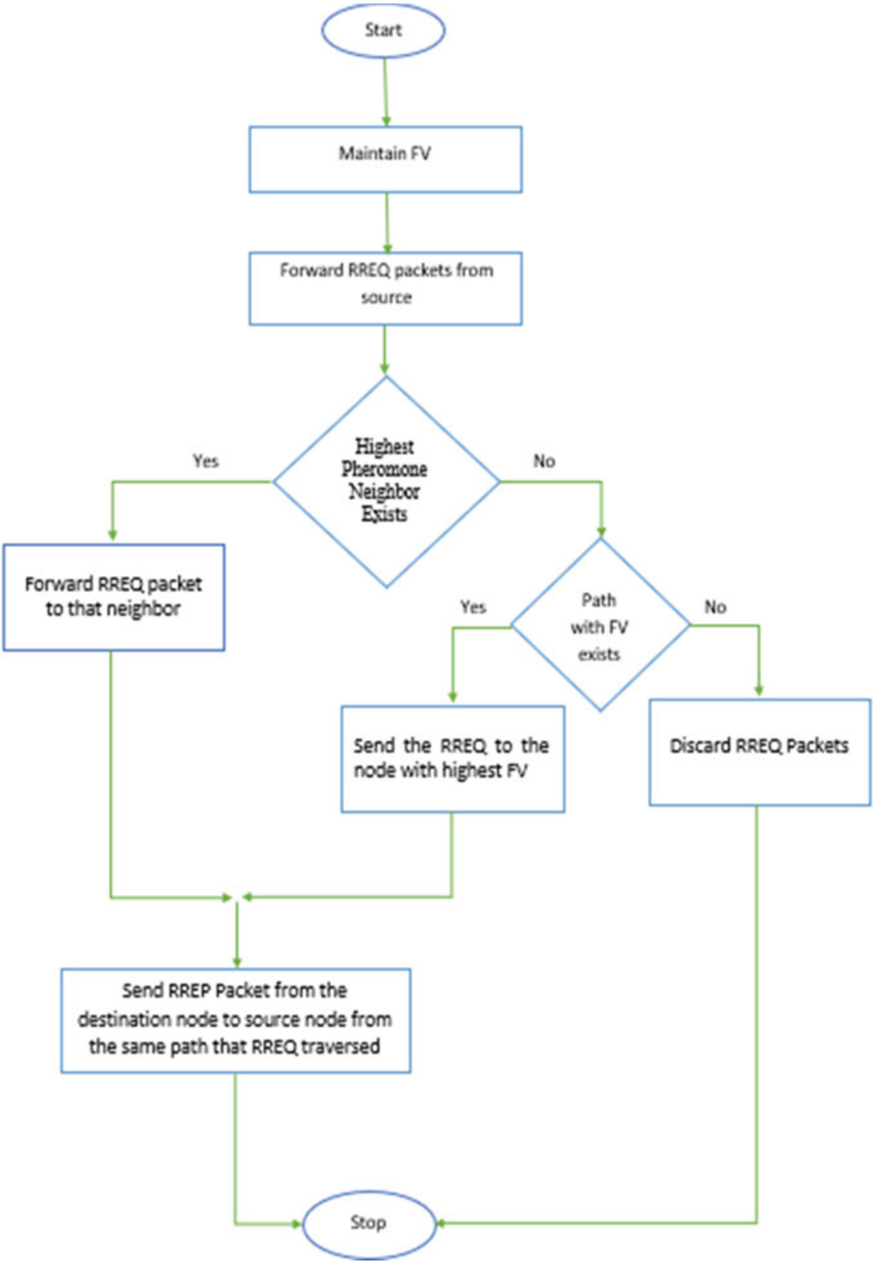


Fig. 3 Flowchart of FVAODV

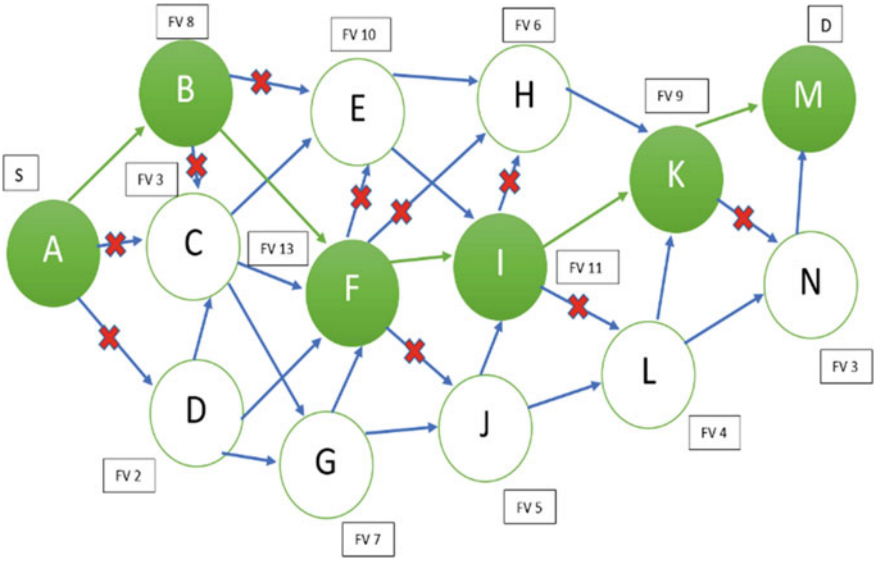


Fig. 4 Working of FVAODV

Working of FVAODV is explained in Fig. 4, with A being the source node and M being the destination node. From A, RREQ packet is sent towards the node with highest faith value that is to a most faithful node amongst the available ones. In Fig. 4 node B has highest FV among B, C and D, thus RREQ packet goes towards B. Similarly, from node B, RREQ goes to node F, then node I and then finally to node K, following the same method for finding route. While the RREQ goes from node to node with highest faith value, it leaves a pheromone packet at each node, so that other route discovery packets may follow the same route until an error has occurred in the pathway. From node K to node M (destination node), the transmission is done as per the basic route discovery of AODV, by seeing minimum number of hop counts to the destination.

4 Simulations

(See Table 1).

4.1 Network Performance Parameters

Throughput refers to the maximum average data rate provided by network during transmission of packets from one location in the network to another.

Table 1 Simulation parameters

Number of nodes	20–100
Simulation time	200 s
Simulation area	300×1500 m^2
Propagation model	Random waypoint
Protocols	AODV, FVAODV
Mobility	5–75 m/s
Start-up time	50 s
Transmitted power	7.5 dBm
MAC layer	802.11 b
Antenna	Omnidirectional

Packet delivery ratio (PDR) is defined as the ratio of the number of packets delivered to the destination to the number of packets sent by the sender.

End-to-End (E-2-E) delay is the time taken by the packets to go from origin node to destination node in the network.

Jitter is a term used to describe variation in delay that occurs during data transfer.

4.2 Network Simulator 3

Network Simulator 3.32 (NS 3.32) which is an open source/discrete event simulator that works on Linux platform is used to carry out simulations.

4.3 Results and Discussion

As you can see in Figs. 5, 6, 7 and 8 throughput and E-2-E delay metrics are observed of AODV and FVAODV and simulated using NS 3.32. The simulation results are obtained by comparing QoS parameters, E-2-E delay and throughput, between the AODV and the FVAODV simulation models.

Figure 5 depicts network throughput achieved as a function of mobility parameter in AODV and FVAODV. Mobility in this environment ranges from 5 to 75 m/s. Throughput increases and reaches a maximum at a mobility of 20 m/s in both cases. Throughput increases from 14.256 kbps at 5 m/s to 22.135 kbps at 15 m/s and reaches a maximum of 24.564 kbps at 20 m/s in FVAODV. It can also be seen from the same figure that it decreases with increasing mobility. The proposed method results in an average improvement of throughput by 12%–13% if compared with AODV.

The network throughput achieved by changing the number of nodes in the AODV and FVAODV protocols is depicted in Fig. 6. In this case, the number of nodes spans from 20 to 100. Throughput of both protocols decreases as the number of nodes

Fig. 5 Throughput VS Mobility in AODV & FVAODV

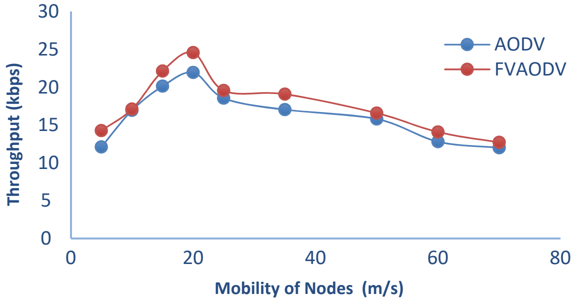


Fig. 6 Throughput VS Number of Nodes in AODV & FVAODV

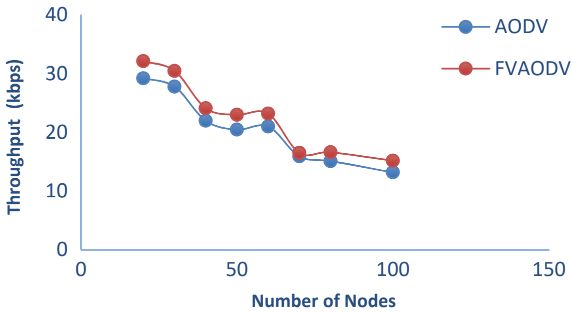


Fig. 7 End-to-End Delay VS Mobility in AODV & FVAODV

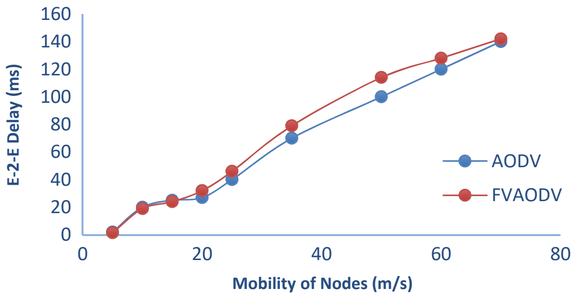
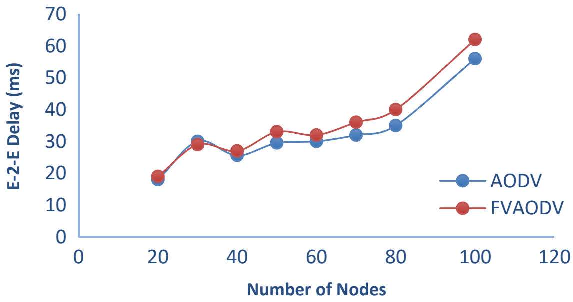


Fig. 8 End-to-End Delay VS Number of Nodes in AODV & FVAODV



increases. Throughput decreases from 32.083 kbps at 20 nodes to 22.964 kbps at 50 nodes and reaches a minimum of 15.14 kbps at 100 nodes in FVAODV. The proposed method results in an average improvement of throughput by 10%-12% as compared with AODV.

The network delay obtained by varying mobility from 5 to 75 m/s while maintaining the number of nodes stagnant at 40 in the AODV and FVAODV protocols is depicted in Fig. 7. As mobility increases, delay is minimally affected in both protocols. FVAODV and AODV both exhibit a comparable delay when changing mobility.

The network delay obtained by changing the number of nodes in the AODV and FVAODV protocols is shown in Fig. 8. The number of nodes in this case ranges from 20 to 100, while mobility is kept constant at 20 m/s. In both protocols, the general trend observed is that the delay appears to increase as the number of nodes increases. However, FVAODV has a marginally longer delay than the AODV protocol.

5 Conclusion and Future Scope

Many Control packets are propagated unnecessarily in MANETs when on-demand route discovery is performed through broadcasting. Furthermore, in the MANETs, the redundancy of information for the routing (i.e., RREP and RREQ) processed is high. With concept of faith values incorporated into this paper, which is an extension of the ACO based ad hoc on-demand routing protocol, in order to decrease the overhead related to message routing and delivery is proposed. FVAODV protocol is straightforward and well-suited for use in a MANETs, as demonstrated in the paper.

Performance of FVAODV protocol against the original AODV routing protocol under a variety of mobility and node configurations for throughput and E-2-E delay is tested. The simulation results show that FVAODV achieves a greater throughput, on average, by 10%-12% if compared with AODV without having a significant impact on the E-2-E delay. Although, the specifics of FVAODV, which is developed on the AODV routing protocol, are discussed, faith value concept developed in this paper can be applied to a wide range of other on-demand routing schemes in general. The usage of FV values can be many and further research can be done in using FV concept with routing algorithms, retransmission strategies [17], scalable networks [18] or using with probability density functions [19] of destination nodes for faster routing, etc.

References

1. Perkins, C.E., Royer, E.M., Das, S.R., Marina, M.K.: Performance comparison of two on-demand routing protocols for ad hoc networks. IEEE Pers. Commun. **16**, 16–28 (2001)

2. Atto, M., Mstafa, R.J., Alkhayyat, A.: Improving AODV routing protocol for image transmission over mobile video sensor networks. IEEE Access **8**, 169396–169407 (2020)
3. Nakayama, H., Kurosawa, S., Jamalipour, A., Nemoto, Y., Kato, N.: A dynamic anomaly detection scheme for AODV-based mobile ad hoc networks. IEEE Trans. Veh. Technol. **58**(5), 2471–2481 (2009)
4. Helen, D., Arivazhagan, D.: Applications, advantages and challenges of ad hoc networks. J. Acad. Ind. Res. **2**(8), 453–457 (2014)
5. Sivaram, M., Porkodi, V., Mohammed, A.S., Manikandan, V., Yuvaraj, N.: Retransmission DBTMA protocol with fast retransmission strategy to improve the performance of MANETs. IEEE Access **7**, 85098–85109 (2019)
6. Abdel-Moniem, A. M., Mohamed, M. H., Hedar, A.R.: An ant colony optimization algorithm for the mobile ad hoc network routing problem based on AODV protocol. In: IEEE Xplore, pp. 1333–1337 (2010)
7. Radha, S., Shanmugavel, S.: Implementation of node transition probability based routing algorithm for MANET and performance analysis using different mobility models. J. Commun. Netw. **5**(3), 202–214 (2003)
8. Adil, M., et al.: MAC-AODV based mutual authentication scheme for constraint oriented networks. IEEE Access **8**, 44459–44469 (2020)
9. Aouiz, A., Hacene, S.B., Lorenz, P.: Channel busyness based multipath load balancing routing protocol for ad hoc networks. IEEE Network **33**, 99–107 (2019)
10. Ochola, E.O., Mejaele, L.F., Eloff, M.M., van der Poll, J.A.: Manet reactive routing protocols node mobility variation effect in analysing the impact of black hole attack. In: South African Institute of Electrical Engineers, pp. 80–91 (2017)
11. Bhardwaj, A., El-ocla, H.: Multipath routing protocol using genetic algorithm in mobile ad hoc networks. IEEE Access **8**, 177534–177548 (2020)
12. Kacem, I., Sait, B., Mekhile, S., Sabeur, N.: A new routing approach for mobile ad hoc systems based on fuzzy petri nets and ant system. IEEE Access **6**, 65705–65720 (2018)
13. Theerthagiri, P., Menakadevi, T.: Resource-based routing protocol for Mobile Adhoc Networks. Songklanakarin J. Sci. Technol. **42**(4), 889–896 (2020)
14. Liu, G. Zhang, W. Guo, and R. He, "Kalman Prediction-Based Neighbor Discovery and Its Effect on Routing Protocol in Vehicular Ad Hoc Networks", IEEE Transactions on Intelligent Transportation Systems, 2019, pp. 1–11
15. Anand, M., Sasikala, T.: Efficient energy optimization in mobile ad hoc network (MANET) using better-quality AODV protocol. Cluster Comput. **22**(27), 12681–12687 (2018)
16. Roger, E.M., Perkins, C. E.: An implementation study of the AODV routing protocol. In: IEEE Wireless Communication and Network Conference, pp 1003–1008 (2000)
17. Baumann, R., Heimlicher, S., Plattner, B.: Routing in large-scale wireless mesh networks using temperature fields. IEEE Netw. **22**(1), 25–31 (2008)
18. Valera, A.C., Seah, W.K.G., Rao, S.V.: Improving protocol robustness in ad hoc networks through cooperative packet caching and shortest multipath routing. IEEE Trans. Mob. Comput. **4**(5), 443–457 (2005). https://doi.org/10.1109/TMC.2005.67
19. Rios, M.: Variable route expiration time based on a fixed probability of failure for ad-hoc networks routing applications. IEEE Lat. Am. Trans. **13**(1), 383–389 (2015)

Sensitivity Enhancement of $Ti_3C_2T_x$ Based Bimetallic SPR Biosensor Using SiO_2

Rajeev Kumar, Sarika Pal, J. P. Saini, and Y. K. Prajapati

Abstract We proposed here a novel surface plasmon resonance (SPR) sensor configuration based on silicon dioxide (SiO_2) and $Ti_3C_2T_x$ (MXene) with bi-metal layers for detection the biomolecules using angular interrogation method. Whereas $Ti_3C_2T_x$ are greater contact area, metallic conductivity and hydrophilic functionalized surface termination for biological molecules present in aqueous solution. $Ti_3C_2T_x$ is used as a bio-recognition (BRE) layer. Through theoretical analysis, we calculated the several performance parameters at optimized thickness of Cu and Ni layer. Our result shows the maximum sensitivity of 290°/RIU at optimized Cu and Ni layer thicknesses. It is observed that the use of SiO_2 layer enhances performance parameters of proposed sensor. Also, the effect of the increasing the Ni and SiO_2 layer on the reflectance curve is analyzed with increasing the RI of sensing medium.

Keywords MXene ($Ti_3C_2T_x$) · Surface plasmon resonance · Sensitivity · Quality factor

1 Introduction

Surface plasmon resonance (SPR) is an emerging technology for fast and improved sensing of biomolecules. It has encouraged the rapid rise in the field of optical sensing

R. Kumar (✉) · S. Pal
National Institute of Technology, Uttarakhand, Uttarakhand 246174, India
e-mail: rajeevkrc@gmail.com; rajeevkumar.phd2020@nituk.ac.in

S. Pal
e-mail: sarikapal@nituk.ac.in

J. P. Saini
Netaji Subhas University of Technology (NSUT), New Delhi 110078, India
e-mail: jps@nsut.ac.in

Y. K. Prajapati
Motilal Nehru National Institute of Technology, Allahabad, Prayagraj, Uttar Pradesh 211002, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_14

devices. The SPR detection technique is a label-free, highly efficient, and real-time analysis and useful for the fields of medical diagnosis, environmental safety and control, health science research, industrial monitoring, and food safety etc. [1–4]. It is extensively used in biosensors because it is reasonable for detecting smaller size biomolecules with higher resonance angle shift. In 1980s, first the SPR phenomenon was demonstrated at the metal-dielectric surface [1] for gas sensing [2]. The SPR sensors rely on the measurement of binding induced RI changes of sensing region on adsorption of analytes. In the SPR phenomenon, the charge carriers oscillate at metal–dielectric interface due to positive real part of permittivity of dielectric material and negative real part of permittivity of metal. This phenomenon generates surface plasmon wave (SPWs) at the metal–dielectric interface. To achieve the resonance condition, incident light needs to be coupled to this SPW wave at proper phase matching condition [3]. Only transverse magnetic or p-polarized waves contributes to the surface plasmon's (SPs) excitation, while s-polarized light (transverse electric waves) will serve as a reference signal. Otto [4] and Kretschmann Raether (K-R) [3] are two well-known prism-based SPR measurement configurations used for light coupling. SPR is typically studied using a Kretschmann-configuration based on attenuated total internal reflection (ATR). In this configuration, a metal film is layered on high RI prism. The prism-based SPR sensors offers various advantage such as simple, easy and flexible design, popular for in-situ monitoring [5]. Au and Ag are commonly used metals for sensor applications. Au and Ag are more conductive material for long term plasmonic and other than Cu is much cheaper plasmonic material. Whereas Cu gets oxidized easily and it has not received more attention as a plasmonic material [6]. Singh et al. discussed the oxidization problem of Cu and which can be avoided by suitable oxide coating. The Cu can be protected from oxidation by using a Ni layer on top of it. Apart from that, Ni is a chemically inactive ferromagnetic metal with a poor near-infrared plasmonic response due to large ohmic losses [7–9]. The adoption of a Cu/Ni bimetallic combination can significantly increase the sensitivity [7]. The combination of Cu/Ni shows a strong response by localizing EM fields allied with the plasmonic response [10]. However, the bimetallic combinations such as Ag-Au, Cu-Ni and Ag–Ni of metal layers may also be used to exhibit trade-off between in different performance parameters such as sensitivity (S), detection accuracy (DA) and quality factor (QF) or figure of merit (FOM) etc. [7, 12–14].

Metal oxide layer silicon dioxide (SiO_2) has been used much used for improved sensing performance [15]. It has low refractive index to other oxide layer and show effective light trapping. In recent years, various 2D materials have been investigated including graphene [16], transition-metal-dichalcogenides (TMDs) [17], black phosphorus [18], which are used on the top of plasmon supportive materials of the SPR sensors. Each 2D material has been recommended as potential sensing materials and has its peculiar merits onto target analytes depending on their bandgap, adsorption energy, and physical properties. A new composition or structure of MXene, among 2D nanomaterial family i.e., $Ti_3C_2T_x$ has been recently investigated for plasmonic applications [18–24]. Due to its excellent physical, chemical, electrical and mechanical properties, it has been used in numerous applications energy-conversation devices, and electrochemical capacitors etc. [20–23]. Considering the unique properties of

$Ti_3C_2T_x$, we propose to enhance the sensing performance of $Ti_3C_2T_x$ based SPR sensors for biosensing, biochemical, gas sensing etc. **Wu et al.** remarkably improved the sensitivity of the proposed SPR biosensor by using Gold and a few-layer $Ti_3C_2T_x$ (MXene). They obtained the largest sensitivity of 160°/RIU by using 4-layer $Ti_3C_2T_{x,}$ when a change in the RI of sensing-medium layer, was fixed as $\Delta n_s = 0.005$ [24]. **Xu et al.** theoretically discussed the performance of MXene based SPR sensors with the configuration of Au/TMDs/Au/MXene under 633 nm wavelength. The highest sensitivity of 198°/RIU was obtained with monolayer MXene and five layers of WS_2 [25]. **Srivastava et al.** [26], demonstrated the Au/MXene/TMDs/BP based SPR biosensor. They obtained the higher sensitivity 190.22°/RIU using single layer. The penetration depth was 1.52 times larger against the convention SPR sensor. **Kumar et al**. [27] proposed SPR biosensor where Ag-silicon-BP-MXene is used and exhibited the highest sensitivity of 264°/RIU at 633 nm. **Raikwar et al.** [28], theoretically presented MXene based SPR sensor with Antimonene. They obtained the maximum sensitivity of 224.22°/RIU.

In this manuscript, we propose a Cu/Ni/SiO_2/$Ti_3C_2T_x$ based SPR refractive index (RI) sensor suitable to operate at fixed wavelength (i.e. 633 nm). A thin SiO_2layer has been used as adhesive layer between Ni and $Ti_3C_2T_x$ layer. The $Ti_3C_2T_x$ is used on the top for biomolecule's recognition. The unique properties of $Ti_3C_2T_x$ such as larger surface area, chemical stability, smaller work function, layered architecture, highly accessible hydrophilic surface terminations, and strong light-matter interaction of proposed senor are utilized to enhance the sensitivity.

2 Theoretical Modeling and Mathematical Expression

2.1 Configuration of the Proposed SPR Sensor

The proposed structure of SPR sensor based on SiO_2 and $Ti_3C_2T_x$, has been shown in Fig. 1. We have chosen prism BK7, metal layer of Cu/Ni, silicon dioxide (SiO_2) layer, 2D material $Ti_3C_2T_x$, and a sensing layer for the proposed sensor. It operates at 633 nm wavelength. The first layer is BK7 prism with the RI 1.5151[11]. The second and third layer are Cu and Ni metal films are RIs can be obtained as per the Drude model as mentioned below [7, 11].

$$\mathbf{n} = (\boldsymbol{\varepsilon}_\mathbf{r} + \boldsymbol{\varepsilon}_\mathbf{i})^{\frac{1}{2}} = \left(1 - \frac{\boldsymbol{\lambda}^2\boldsymbol{\lambda}_\mathbf{c}}{\boldsymbol{\lambda}_\mathbf{c}^2(\boldsymbol{\lambda}_\mathbf{c} + \mathbf{i}\boldsymbol{\lambda})}\right)^{\frac{1}{2}} \quad (1)$$

where, λ_c and λ_p are representing collision and plasma wavelength respectively. The value of λ_c and λ_p for the Cu and Ni metal films are shown in Table 1.

The third layer is of SiO_2 with 1 nm, 3 nm and 5 nm thickness considered for this work and its RI is 1.4570 [17]. It has purely real RI and used as adherence layer on Ni

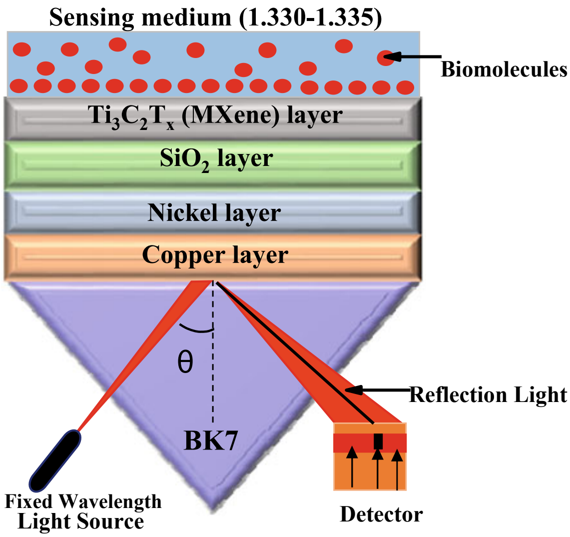


Fig. 1 Proposed SPR sensor

Table 1 λ_c and λ_p of Cu-Ni metals according to the Drude model at $\lambda = 633$ nm

S. No	Metals film	Collision's wavelength (λ_c) (μm)	Plasma wavelength (λ_p) (μm)
1	Copper (Cu)	4.0852×10^{-5}	1.3617×10^{-7}
2	Nickel (Ni)	2.8409×10^{-5}	2.5381×10^{-7}

layer. The fourth layer is $Ti_3C_2T_x$, having the RI $2.38 + i1.33$ [24–28]. The sensing medium is considered as the last layer of the aqueous solution with the RI variation $1.33 + \Delta n_s$, where $\Delta n_s = 0.005$ represents the RI shift of the sensing medium with analyte absorption over the$Ti_3C_2T_x$.

2.2 Mathematical Equations for Reflectivity

This section covers the N-Layer modelling for calculating the reflectance and different performance factors via as a transfer matrix method (TMM) and Fresnel equation for the proposed sensor [26]. The different layers possessing the thickness (d_k), dielectric constant (ε_k), and RI (n_k), are stacked along the z-axis. After applying the boundary conditions for the tangential component of electric (E) and magnetic (H) fields, the tangential field at $Z = Z_1 = 0$ can be related in terms of $Z = Z_{N-1}$ for the first and last layers as given:

$$\begin{bmatrix} X_1 \\ Y_1 \end{bmatrix} = M \begin{bmatrix} X_{N-1} \\ Y_{N-1} \end{bmatrix} \tag{2}$$

where, [X_1, Y_1] and [X_{N-1}, Y_{N-1}] are representing the tangential E and H fields at first and last (N^{th})layer boundary respectively. The characteristics matrix (M) for all assorted multilayer structure of the proposed design is shown as below:

$$M = \prod_{k=2}^{N-1} M_k = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \tag{3}$$

with,

$$M_k = \begin{bmatrix} \cos\beta_k & \frac{-i}{p_k}\sin\beta_k \\ -ip_k\sin\beta_k & \cos\beta_k \end{bmatrix} \tag{4}$$

where, k represents an arbitrary number, β_k represents phase thickness and p_k is represent the transverse refractive indices that can be explained below:

$$p_k = \sqrt{\frac{\mu_k}{\varepsilon_k}}\cos\theta_k = \frac{\sqrt{\varepsilon_k - n_1^2\sin^2\theta_1}}{\varepsilon_k} \tag{5}$$

and

$$\beta_k = \frac{2\pi}{\lambda} n_k \cos\theta_k (z_k - z_{k-1}) = \frac{2\pi d_k}{\lambda}\sqrt{\varepsilon_k - n_1^2\sin^2\theta_1} \tag{6}$$

where, λ and θ_1 is represent the free space wavelength and incident angle.

The reflection coefficient (r) acquired for p-polarized-light may be given as;

$$r = \frac{(m_{11} + m_{12}p_n)p_1 - (m_{21} + m_{22}p_n)}{(m_{11} + m_{12}p_n)p_1 + (m_{21} + m_{22}p_n)} \tag{7}$$

where, [p_1, p_n] acquired from Eq. 6 are relative components of [1^{st}, n^{th}] layer respectively. Lastly, the reflectance (R_p) for the multilayer configuration can be achieved by using expressions shown below;

$$R_p = |r|^2 \tag{8}$$

2.3 *Performance Parameters of the SPR Sensor*

To evaluate the performance of the proposed SPR sensor, following parameters are computed from the SPR characteristics curves; that are simulated using MATLAB.

Sensitivity (S) is defined in terms ratio of difference between two resonance angles ($\Delta\theta_{res} = \theta_2 - \theta_1$) and sensing layer RI of 1.335–1.330 ($\Delta n_s = 0.005$) measured from the reflectance curve, it indicates the sensor's sensing capability.

$$S = \frac{\Delta\theta_{Res}}{\Delta n_s}(^\circ/\text{RIU}) \tag{9}$$

Full Width at Half Maximum (FWHM) [30] is difference of resonance angles at 50% reflection intensity. It also indicates the angular width of the resonance curve. Detection accuracy (DA) is defined as inversely proportional to the FWHM.

$$DA = \frac{1}{FWHM}(1/^\circ) \tag{10}$$

Figure of merit (FoM) indicates the multiplication of the sensitivity and detection accuracy.

$$FoM = S * DA(1/\text{RIU}) \tag{11}$$

3 Discussions and Simulation Results

3.1 Optimization of Copper, Nickel and Silicon Dioxide Layer

After defining the RIs of each layer, the angular interrogation technique is used to analyze the five-layer configuration (Prism BK7/Cu/Ni/SiO_2/MXene/Sensing medium) of the proposed SPR sensor. MATLAB software has been used to analyses the performance parameter. First, the constituent layers of proposed sensor configuration are optimized to get better performance. The sensitivity and minimum reflectance are the two well suited parameters for optimization; as minimum reflectance position shows the maximum SPs generation and complete energy transfer from incident light to sensor configuration. The Fig. 2(a–c) shows thickness optimization for Cu, Ni and SiO_2 layers in terms of minimum reflectance. It shows the Rmin. variation vs. Cu thicknesses (20–60 nm), at Ni thicknesses (1, 3, 5, 7, 10, 15 nm), for 1, 3, 5 nm of SiO_2 layers. The Rmin. decrease first near to zero and then increase with increase of the Cu thickness at all Ni thickness layer. In Fig. 2, it is clearly observed that, as the Ni layer thickness is increased, the Rmin. dip shift towards the smaller thicknesses of Cu. This is due to use of Ni layer, which enhances the electromagnetic field in the sensing medium [8]. The optimized Cu layer thickness are obtained at (43, 41, 39, 37, 35 and 29 nm) for Ni thicknesses (1, 3, 5, 7, 10, and 15 nm), at 1, 3, and 5 nm of SiO_2 layer respectively.

Sensitivity is an important parameter to analyze the sensor performance. So, the thicknesses of constituent layers of proposed sensor are optimized in term of sensitivity too in Fig. 3(a–c). The results demonstrate higher sensitivity at higher Ni and SiO_2 layer thicknesses. The reason behind the sensitivity increment is use of Ni and

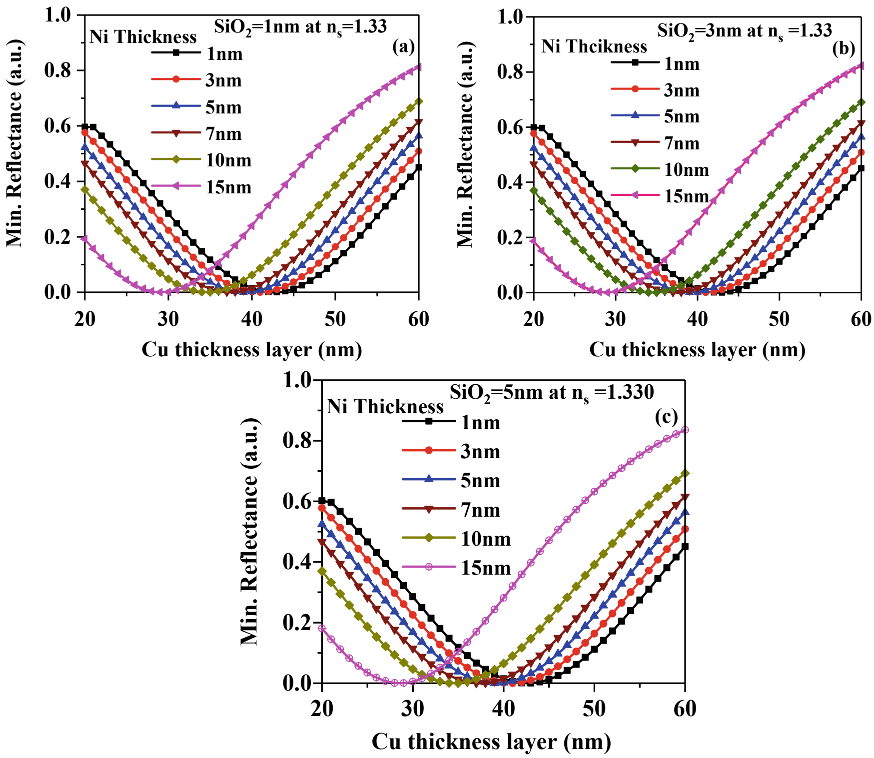


Fig. 2 Cu, Ni and SiO_2 layer thickness optimization in terms of minimum reflectance (Rmin.)

oxide layer SiO_2. The use of SiO_2 layer increases the adsorption energy and Ni layer magnifies the electromagnetic field in sensing medium [7, 27]. From Fig. 3(c), it is clearly observed that the maximum sensitivity 290°/RIU achieved at the thicknesses 38, 15, and 5 nm of Cu, Ni, and SiO_2, respectively. The Table 2 tells the sensitivity and Rmin. obtained for different thickness combinations of Cu, Ni and SiO_2 layers.

The $Ti_3C_2T_x$ is used at the top of the proposed sensor as the biomolecular recognition elements (BRE) material that enhances the sensitivity by providing more absorption sites for the selective biomolecules present in the sensing medium due to its unique sensing and optical properties. The Fig, 4(a) demonstrates the maximum sensitivity, DA, and FoM are reached to 290°/RIU, 0.16/° and 46.4/RIU, respectively, for proposed configuration of the SPR sensor. The metal layer plays a wide roll in the SPR sensor as it is responsible for SPs generation. The effect of Cu thickness, with the fixed thicknesses of other layers, is shown in Fig. 4(b). It worth noting here that the reflectance curve is shifted in upward direction after increasing the thickness of Cu layer with fixed thickness of Ni and SiO_2 layer.

Figure 5 indicates the reflectance curve for 1, 3 and 5 nm thickness of SiO_2 layer, respectively. The solid line of reflectance curve indicates the RI of 1.330 sensing medium, whereas the dotted line indicates the RI of 1.335 sensing medium. The

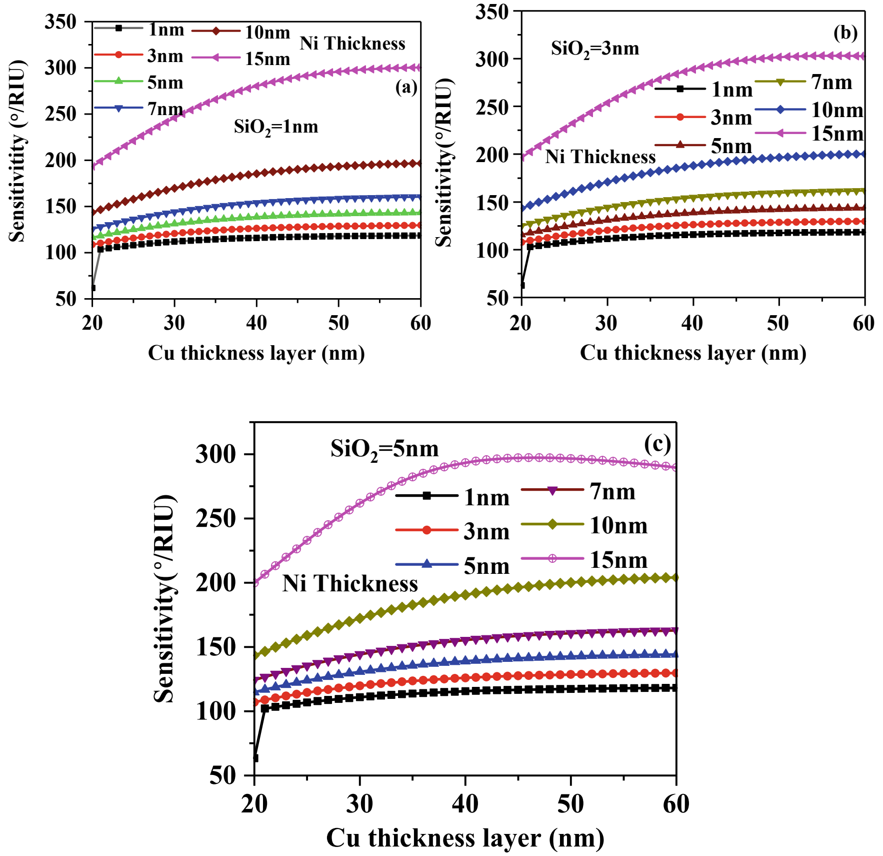


Fig. 3 Cu, Ni and SiO_2 layer thickness optimization in terms of sensitivity

Table 2 Shows the Sensitivity and Rmin at optimized thickness of Cu, Ni and SiO_2 layer with monolayer $Ti_3C_2T_x$ for the sensing layer RI 1.33

Thickness		$SiO_2 = 1$ nm		$SiO_2 = 3$ nm		$SiO_2 = 5$ nm	
Cu	Ni	S	Rmin	S	Rmin	S	Rmin
43	1	116.88	3.8×10^{-4}	116.53	3.8×10^{-4}	116.19	3.6×10^{-4}
41	3	126.62	4.4×10^{-5}	126.50	3.9×10^{-5}	126.27	3.2×10^{-5}
39	5	138.08	3.9×10^{-5}	138.19	4.5×10^{-5}	138.31	5.2×10^{-5}
37	7	151.83	3.2×10^{-4}	152.29	3.2×10^{-4}	152.75	3.3×10^{-4}
35	10	178.87	3.1×10^{-4}	180.71	3.8×10^{-4}	182.77	4.9×10^{-4}
29	15	241.32	1.4×10^{-4}	248.66	7.9×10^{-6}	256.57	4.4×10^{-4}

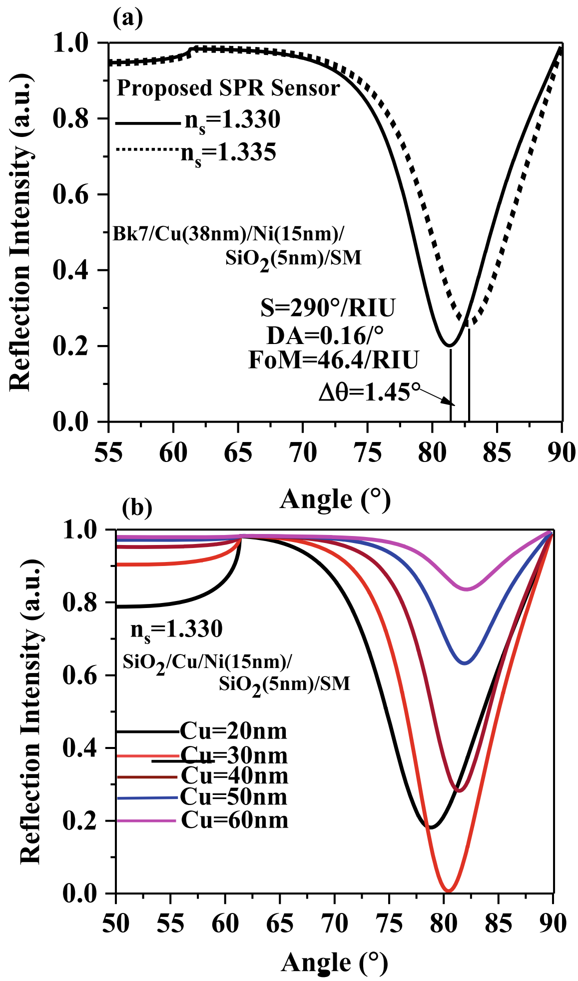


Fig. 4 Reflectance Curve **a** Proposed SPR sensor at optimized constituent layers **b** At various thicknesses of Cu

sensitivity and detection accuracy obtained from reflectance curve shown in Fig. 5 for 1, 3, 5 nm of SiO_2 layer thickness are 241.32°/RIU, 248.66°/RIU, 256.57°/RIU and 0.12/°, 0.11/°,0.12/°respectively. On increasing the Ni and SiO_2 layer thickness, the resonance angle is shifted to the larger side by enhancing the field at the sensing medium interface with reasonable FoM of 29.53/RIU,29.56/RIU, and 31.13/RIU at minimum reflectance. The SiO_2 layer is used as adhesion layer in the proposed sensor, the plasmonic effect occurs near the SiO_2-$Ti_3C_2T_x$ interface and the light trapping is enhanced [29]. Because of enhanced light trapping, more SPs are generated and resulted in improved sensitivity.

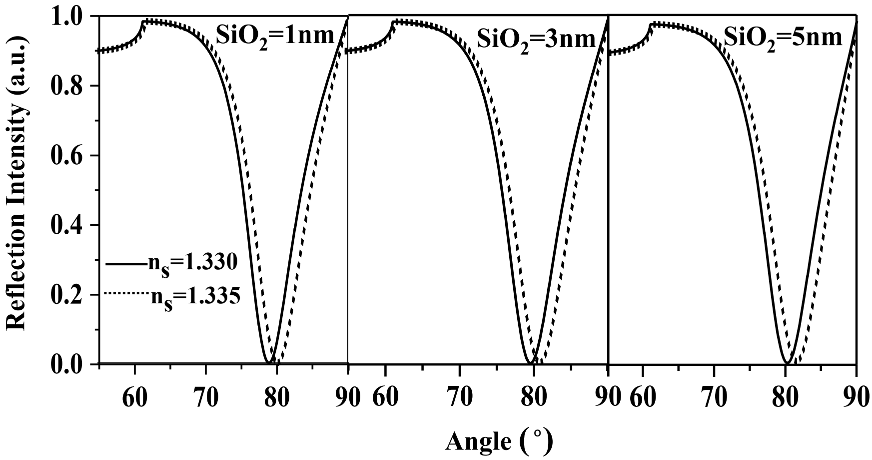


Fig. 5 Reflectance curve for the optimized thickness of SiO_2 layers

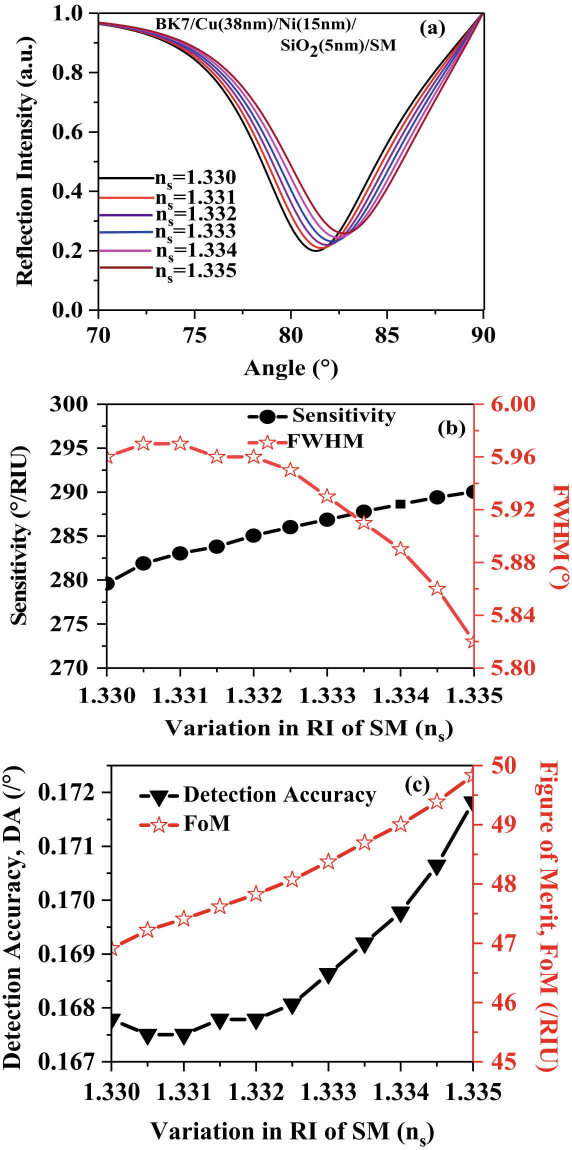
3.2 *Effect of Change in RI of Sensing Layer*

It is observed from Fig. 6(a) that the reflectance curve is shifted towards the higher angle with increasing RI of sensing medium as it changes the propagation constant (wave vector) of the SP waves. Figure 6(b) shows the variation of sensitivity and FWHM with respect to RI of sensing medium. The sensitivity and FWHM variations are from 279.60°/RIU to 290°/RIU and 5.92 to 5.82°/RIU respectively, corresponding to sensing medium RI variation from 1.330 to 1.335. It is clear that the FWHM decreases with increasing RI of sensing medium. The angles corresponding to the half maxima of the reflectance curve are obtained to calculate FWHM and DA, and FoM are calculated from the Eqs. (10) and (11). One can visualize from Fig. 6(c) that the D.A and FoM are increases with the RI of sensing medium [20].

3.3 *Electric Field Intensity Plot*

Figure 7 represents the variation of electric field distribution for 6-layer proposed SPR sensor. The electric field intensity of p-polarized light is plotted in Fig. 7 (a & b) using COMSOL Multiphysics software. The electric field intensity enhancement plays an important role in the performance of the SPR sensor. Figure 7(a) and (b) represents TM field distribution and SPs propagation respectively at different interface of the proposed structure [BK7-Cu(38 nm)-Ni(15 nm)-SiO_2(5 nm)-$Ti_3C_2T_x$(0.993 nm)] with respect to the normal distance form prism to sensing medium at resonance condition. Similarly, Fig. 7(c) represents the 1D TM field distribution for the proposed

Fig. 6 Variation in RI of sensing medium **a** Reflectance Curve **b** Sensitivity vs FWHM **c** Detection Accuracy vs FoM



sensor. It clearly indicates the rise in TM field at Cu/Ni interface due to field enhancement of SPs generated at Cu layer on use of Ni layer, then it again rises to peak value at $Ti_3C_2T_x$–sensing medium interface. Then, this TM field intensity falls exponentially and slowly within the sensing medium indicating the maximum interaction volume of the evanescent field with the biomolecules and in turn maximizing the sensitivity. Here, the propagation depth (PD), which is the normal distance along the sensing medium at which the evanescent field decays to 1/e of its maximum field intensity is also calculated. The PD evaluated for the proposed sensor is 142 nm.

Table 3 Shows the comparison of the proposed sensor with existing sensors in literature for the detection of the biomolecule in terms of the sensitivity. It is clear that the maximum sensitivity is achieved for the proposed sensor as compared to existing sensor.

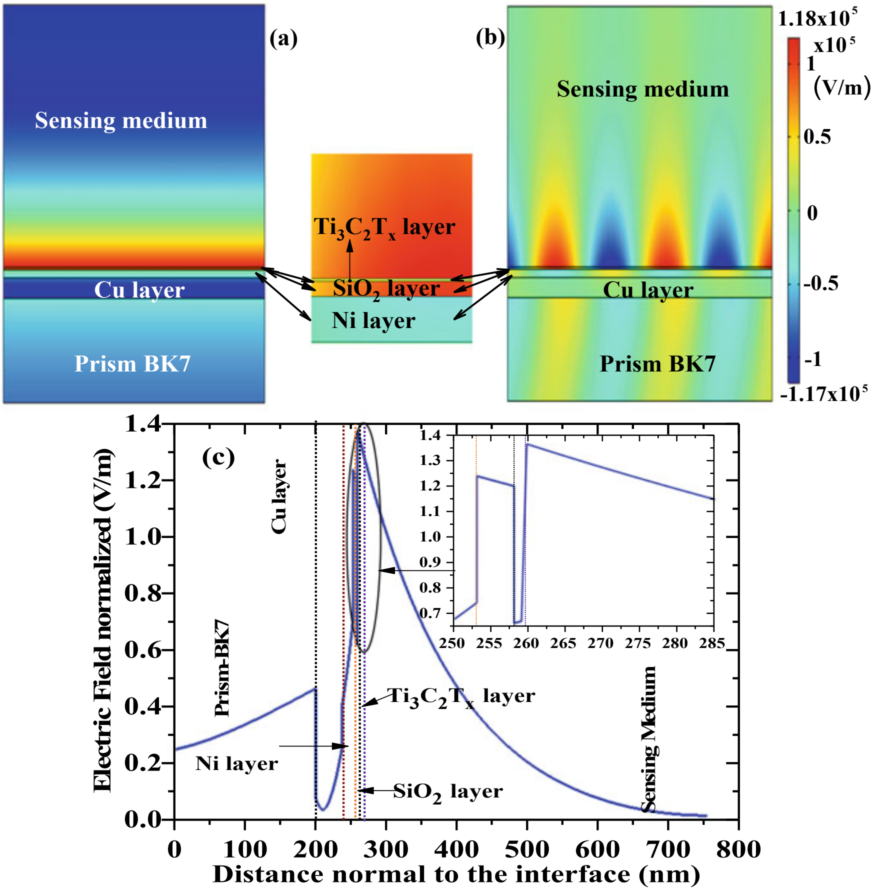


Fig. 7 Simulated TM field intensity at 633 nm for the proposed SPR sensor **a** Distribution of normalized electric field **b** Distribution of Ey component of the electric field **c** Electric field distribution as a function of distance from prism to sensing medium

Table 3 Comparison Table of the proposed sensor and existing sensor in terms of sensitivity

Structure & Reference	Sensitivity
BK7-Metal-MXene-SM [24]	160
BK7-Au-MXene(1L)-Au-TMDs-SM [25]	198
Bk7-Au-MXene-WS_2-BP-SM [26]	190.22
BK7-Ag-Silicon-BP-MXene-SM [27]	264
BK7-TiO_2-Au-MXene-Antimonene-SM [28]	224
Proposed SPR sensor	290

4 Conclusion

In this manuscript, a theoretically designed the SPR sensor based on Cu-Ni-SiO_2-$Ti_3C_2T_x$ layers is presented. Our numerical analysis shows that after adding the SiO_2 layer, the maximum sensitivity of 290°/RIU is obtained with remarkable FoM. Moreover, it is observed that the FWHM is decreased due to which the DA and FoM enhanced after increasing the RI of sensing medium. It is further analyzed that on using SiO_2 layer over $Ti_3C_2T_x$, the sensitivity is enhanced after increasing its thickness. As a result, it is expected that the proposed biosensor can be used for detection the biomolecules with remarkable performance.

References

1. Gordon Ii, J.G., Ernst, S.: Surface plasmons as a probe of the electrochemical interface. Surf. Sci. **101**(1–3), 499–506 (1980)
2. Nylander, C., Liedberg, B., Lind, T.: Gas detection by means of surface plasmon resonance. Sens. Actuators **3**, 79–88 (1982)
3. Raether, Heinz: Surface Plasmons on Smooth and Rough Surfaces and on Gratings. Springer, Heidelberg (1988)
4. Otto, A.: Excitation of surface plasma waves in silver by the method of frustrated total reflection. Z Physik **216**, 398–410 (1968)
5. Raether, H., Kretschmann, E.: Radiative decay of nonradiative surface plasmons excited by light. Z Naturforsch. TEIL A **23A**, 2135–2136 (1968)
6. Singh, S., Mishra, S.K., Gupta, B.D.: Sensitivity enhancement of a surface plasmon resonance based fibre optic refractive index sensor utilizing an additional layer of oxides. Sens. Actuators A: Phys. **193**, 136–140 (2013)
7. Vibisha, G.A., Nayak, J.K., Maheswari, P., Priyadharsini, N., Nisha, A., Jaroszewicz, Z., Rajesh, K.B., Jha, R.: Sensitivity enhancement of surface plasmon resonance sensor using hybrid configuration of 2D materials over bimetallic layer of Cu–Ni. Opt. Commun. **463**, 125337 (2020)
8. Maharana, P.K., Jha, R., Palei, S.: Sensitivity enhancement by air mediated graphene multilayer-based surface plasmon resonance biosensor for near infrared. Sens. Actuators B Chem. **190**, 494–501 (2014)
9. Shukla, S., Sharma, N.K., Sajal, V.: Theoretical study of surface plasmon resonance-based fiber optic sensor utilizing cobalt and nickel films. Braz. J. Phys. **46**(3), 288–293 (2016)
10. Atmatzakis, E., Papasimakis, N., Fedotov, V., Vienne, G., Zheludev, N.I.: Magneto-optical response in bimetallic metamaterials. Nanophotonics **7**(1), 199–206 (2018)

11. Kumar, R., Pal, S., Pal, N., Mishra, V., Prajapati, Y.K.: High-performance bimetallic surface plasmon resonance biochemical sensor using a black phosphorus–MXene hybrid structure. Appl. Phys. A **127**(4), 1–12 (2021)
12. Pal, S., Verma, A., Prajapati, Y.K., Saini, J.P.: Influence of black phosphorous on performance of surface plasmon resonance biosensor. Opt. Quant. Electron. **49**(12), 403 (2017)
13. Jha, R., Sharma, A.K.: Chalcogenide glass prism based SPR sensor with Ag–Au bimetallic nanoparticle alloy in infrared wavelength region. J. Opt. A: Pure Appl. Opt. **11**(4), 045502 (2009)
14. Wang, Q., Jiang, X., Niu, L.Y., Fan, X.C.: Enhanced sensitivity of bimetallic optical fiber SPR sensor based on MoS_2 nanosheets. Opt. Lasers Eng. **128**, 105997 (2020)
15. Chhowalla, M., Shin, H.S., Eda, G., Li, L.J., Loh, K.P., Zhang, H.: The chemistry of two-dimensional layered transition metal dichalcogenide nanosheets. Nat. Chem. **5**(4), 263 (2013)
16. Maurya, J.B., Prajapati, Y.K., Singh, V., Saini, J.P.: Sensitivity enhancement of surface plasmon resonance sensor based on graphene–MoS_2 hybrid structure with TiO_2–SiO_2 composite layer. Appl. Phys. A **121**(2), 525–533 (2015)
17. Pal, S., Pal, N., Prajapati, Y.K., Saini, J.P.: Performance evaluation of SPR biosensor using metamaterial over conventional SPR and graphene based SPR Biosensor. In: 2018 5th International Conference on Signal Processing and Integrated Networks (SPIN), pp. 696–701. IEEE (2018)
18. Pal, S., Verma, A., Saini, J.P., Prajapati, Y.K.: Sensitivity enhancement using silicon-black phosphorus-TDMC coated surface plasmon resonance biosensor. IET Optoelectron. **13**(4), 196–201 (2019)
19. Wang, X., Mathis, T.S., Li, K., et al.: Influences from solvents on charge storage in titanium carbide MXenes. Nat. Energy. **4**(3), 241–248 (2019)
20. Huang, K., Li, Z., Lin, J., Han, G., Huang, P.: Two-dimensional transition metal carbides and nitrides (MXenes) for biomedical applications. Chem. Soc. Rev. **47**(14), 5109–5124 (2018)
21. Wu, L., Lu, X., Wu, Z.S., Dong, Y., Wang, X., Zheng, S., Chen, J.: 2D transition metal carbide MXene as a robust biosensing platform for enzyme immobilization and ultrasensitive detection of phenol. Biosens. Bioelectron. **107**, 69–75 (2018)
22. Pal, S., Verma, A., Prajapati, Y.K., Saini, J.P.: Sensitive detection using heterostructure of black phosphorus, transition metal di-chalcogenides and MXene in SPR sensor. Appl. Phys. A **126**(10), 1–10 (2020)
23. Wu, Q., et al.: A 2D transition metal carbide MXene-based SPR biosensor for ultrasensitive carcinoembryonic antigen detection. Biosens. Bioelectron. **144**, 111697 (2019)
24. Wu, L., You, Q., Shan, Y., Gan, S., Zhao, Y., Dai, X., Xiang, Y.: Few-layer $Ti_3C_2T_x$ MXene: a promising surface plasmon resonance biosensing material to enhance the sensitivity. Sens. Actuators B: Chem. **277**, 210–215 (2018)
25. Xu, Y., Ang, Y.S., Wu, L., Ang, L.K.: High sensitivity surface plasmon resonance sensor based on two-dimensional MXene and transition metal dichalcogenide: a theoretical study. Nanomaterials **9**(2), 165 (2019)
26. Srivastava, A., Verma, A., Das, R., Prajapati, Y.K.: A theoretical approach to improve the performance of SPR biosensor using MXene and black phosphorus. Optik **203**, 163430 (2020)
27. Kumar, R., Pal, S., Verma, A., Prajapati, Y.K., Saini, J.P.: Effect of silicon on sensitivity of SPR biosensor using hybrid nanostructure of black phosphorus and MXene. Superlattices Microstruct. **145**, 106591 (2020)
28. Raikwar, S., Srivastava, D.K., Saini, J.P., Prajapati, Y.K.: 2D-antimonene-based surface plasmon resonance sensor for improvement of sensitivity. Appl. Phys. A **127**(2), 1–8 (2021)
29. Xu, J., Xiao, X., Stepanov, A.L., Ren, F., Wu, W., Cai, G., Zhang, S., Dai, Z., Mei, F., Jiang, C.: Efficiency enhancements in Ag nanoparticles-SiO_2–TiO_2 sandwiched structure via plasmonic effect-enhanced light capturing. Nanoscale Res. Lett. **8**, 73 (2013)
30. Maurya, J.B., Prajapati, Y.K.: A novel method to calculate beam width of SPR reflectance curve: a comparative analysis. IEEE Sens. Lett. **1**(4), 1–4 (2017)

Design and Simulation of 1Tbps Coherent Detection Dual Polarization Free Space Optics Based OFDM System

Priyanshu Kumar Pandey, Shiksha Jain, and Parimal Tiwari

Abstract The requirement for high-speed communication has become significant due to technological growth. In Industry 4.0 everything around us is going to be interconnected. This defines a new challenge for internet systems in terms of speed and latency requirements. Introduction of Optical Wireless Communication/Free Space Optics will not only address the challenges of speed and latency requirements but it will also address the issue of interconnectivity in the remotest area as well using satellite internet. OFDM (Orthogonal Frequency Division Multiplexing) is one of the most widely used technologies in wireless communication. It has gained popularity due to advancements in DSP (Digital Signal Processing). The use of OFDM in Free Space Optics (FSO) leads to the prohibition of various signal impairments and therefore increases data rate and transmission distance. The capacity can be further be increased by employing polarization diversity i.e. Dual Polarization (DP) with OFDM. The proposed system has been designed and simulated to achieve a 1Tbps data rate for a distance up to 3000 km with BER less than 1×10^{-6}. The results indicate that the proposed system be used for 4000 km with BER 8×10^{-4}. and 5000 km with BER 9.4×10^{-3}.

Keywords Free space optics · OFDM · Dual polarization · DSP

1 Introduction

Optical Wireless Communication (OWC) or Free Space Optics (FSO) refers to the transmission of optical signals wirelessly into free space. The potential advantages offered by OWC/FSO are in terms of speed and bandwidth. When the first laser was introduced in 1960, the first optical communication system using a laser optical source was demonstrated. Optical communications initially reported significant attenuation of 100 dB/km [23]. Subsequently, the attenuation has been reduced

P. K. Pandey (✉) · S. Jain · P. Tiwari
Institute of Engineering and Technology, Dr Rammanohar Lohia Awadh University, Ayodhya, Uttar Pradesh, India
e-mail: hipriyanshu@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_15

to a minimum value. The modern internet infrastructure is heavily dependent upon the Optical Fiber backbone. The other alternative to that is the Satellite Network backbone. Inter-satellite links (ISLs) are considered important in providing global coverage for the transmission of information [2]. Three geostationary satellites 120° apart in free space can cover the entire globe for communication purposes [3, 4]. Optical wireless communication links (OWCs) are widely used today due to many legacy advantages over traditional radio frequency-based wireless links which are heavily congested [5–8]. There are various modulation techniques available for OWC/FSO but due to its inherent advantage in terms of prohibiting signal impairments, OFDM is a promising choice. OFDM scheme is widely applied to wireless local area networks (WLANs), which are wireless systems for digital audio and video transmission, according to Worldwide Interoperability for Microwave Access (WiMAX) IEEE 802.11g over Asynchronous Digital Subscriber Line (ADSL) and Wired Access System [9]. OFDM scheme was first introduced to the optical fiber transmission systems with direct detection [10]. Shortly thereafter, COOFDM was introduced, which was made possible by coherent digital detection [11]. New optical OFDM methods such as guard interval transmission [12] and multimode fiber optic transmission [13] have also been reported. In this paper the requirement of high data rate and better noise performance requirement have been addressed using the integration of OFDM with FSO. To the author's knowledge, the proposed design offers higher data rates, longer transmission distances, and better noise performance than its predecessors. The differences from the prior work and performance comparison have been tabulated in Table 2.

2 System Design and Operation

The design and operation of optical OFDM has been given in the following subsection.

2.1 *Coherent Optical OFDM*

OFDM has been employed in many different applications because of its benefits. The OFDM optical system can be divided into five functional blocks, including (i) OFDM RF Transmitter (ii) (RTO) RF to Optical converter (iii) Optical Wireless Channel (iv) (OTR) Optical to RF down-converter and (v) OFDM RF Receiver.

The block diagram of typical OFDM transmitter and receiver has been given in Figs. 1 and 2 respectively.

The data bits are first converted from serial to parallel bitstreams. As OFDM is multi-carrier modulation (MCM) scheme each of these bits is mapped with the subcarriers.

The transmitted $y(t)$ signal using MCM can be given by Eq. (1)

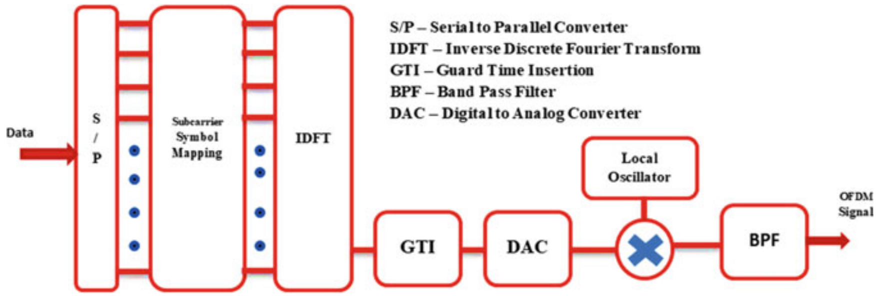


Fig. 1 Block diagram of OFDM transmitter

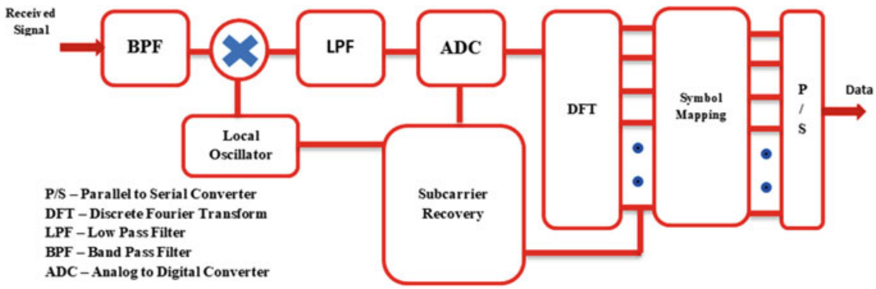


Fig. 2 Block diagram of OFDM receiver

$$y(t) = \sum_{l=-\infty}^{\infty} \sum_{m=1}^{N} c_{ml} y_m(t - lT) \tag{1}$$

Here c_{ml} is the l th information symbol at m th sub-carrier, y_m is m th sub-carrier waveform, T is symbol period, N is number of subcarrier and f_m is frequency of m th sub-carrier. Also

$$y_m(t) = G(t)e^{j2\pi f_m t} \tag{2}$$

and

$$G(t) = \begin{cases} 1, 0 \le t \le T \\ 0, Otherwise \end{cases} \tag{3}$$

At the receiving side correlator can be used to detect c'_{ml}(information symbol) for optimum detection. If $r(t)$ is the received signal c'_{ml} can be given as

$$c'_{ml} = \int_0^T r(t - lT)y_m^* dt = \int_0^T r(t - lT)e^{-j2\pi f_m t} dt \tag{4}$$

The major drawback of MCM is in terms of Bandwidth (BW) requirements due to non-overlapping signals. OFDM uses orthogonal yet overlapped signals. For orthogonal signals correlation

$$r_{mp} = \frac{1}{T}\int_0^T y_m y_p^* dt = \frac{1}{T}\int_0^T e^{j2\pi(f_m - f_p)t} dt \tag{5}$$

$$\text{or, } r_m p = \frac{1}{T}\frac{[e^{(j2\pi(f_m - f_p)T)} - 1]}{j2\pi(f_m - f_p)} = e^{(j\pi(f_m - f_p)T)}\frac{\sin[\pi(f_m - f_p)T]}{\pi(f_m - f_p)T} \tag{6}$$

For orthogonal signals, Eq. (6) can be satisfied by Eq. (7) as

$$(f_m - f_p) = \frac{n}{T} \tag{7}$$

If the subcarriers (narrow sub bands) satisfy Eq. (7) it can be recovered using correlator given in Eq. (4) despite the overlapping.

Modulation is achieved by performing IDFT (Inverse Discrete Fourier Transform) using IFFT (Inverse Fast Fourier Transform) algorithm on the input symbol c_{ml}. After the insertion of guard bits digital to analog conversion is done with the help of DAC. RF signals to optical signals conversion can be achieved by biasing the (MZM) Mach-Zehnder Modulator at zero point [14, 15].

The OFDM IQ modulator is used to generate QAM signals. For the design of a dual-polarization system, two such arrangements of OFDM transmitters are needed. The basic idea for employing (DP) Dual Polarization is to increase the capacity.

At the receiving end, IQ Demodulator is used to convert the optical signal to an RF signal. The demodulation process is based on coherent detection using a PIN photodiode. After converting it to digital signals Discrete Fourier Transform is performed and by the subcarrier recovery signals are again mapped into bitstreams. To improve the computational efficiency DFT is computed using FFT (Fast Fourier Transform) algorithm. Finally parallel to serial conversion is done to reconstruct the original bit pattern. As the system employs a Dual Polarization scheme, two such subsystems are needed.

2.2 Spectral Efficiency of OFDM with 4 QAM and Dual Polarization

For OFDM, let the symbol period is T and the number of subcarriers is N then the Symbol Rate R can be given as

$$R = \frac{N}{T} \tag{8}$$

Normally to denote the boundary of each wavelength channel bandwidth of the first null is used [16]. The OFDM BW (Bandwidth) can be expressed as

$$BW = \frac{N-1}{t} + \frac{2}{T} \tag{9}$$

here "t" is the observation period.

Bandwidth Efficiency (η) is defined by ratio of symbol rate (R) and Bandwidth (BW). However, for dual-polarization, the factor of 2 is multiplied i.e.

$$\eta = 2\frac{R}{BW} \tag{10}$$

By adopting higher-order QAM modulation, spectral efficiency η can be further improved [17].

2.3 Free Space Optical Wireless Channel (OWC)

As FSO/OWC utilizes free space as a channel and Line of Sight (LOS) propagation, modelling the effects of the channel is very crucial. For the LOS system, the received optical signal power is given by [18].

$$P_R = \eta_T \eta_R P_T G_T G_R \left(\frac{\lambda}{4\pi Z}\right)^2 L_T L_R \tag{11}$$

where, P_R denotes received power,η_T denotes optical efficiency of the transmitter, η_R denotes optical efficiency of the receiver, P_T denotes transmitted power, G_T denotes transmitter gain, G_R denotes receiver gain, λ denotes operating wavelength, Z denotes the distance between transmitter and receiver, L_T denotes pointing loss factor of the transmitter, L_R denotes pointing loss factor of the receiver.

The transmitter gain G_T and receiver gain G_R are given by [19].

$$G_T \approx \left(\frac{\pi D_T}{\lambda}\right)^2 \tag{12}$$

and

$$G_R \approx \left(\frac{\pi D_R}{\lambda}\right)^2 \tag{13}$$

where D_T denotes transmitter telescope diameter and D_R denotes receiver telescope diameter.

Here, the transmitter pointing loss factor and receiving pointing loss factor are given by [4]

$$L_T = e^{-G_T \theta_T^2} \tag{14}$$

and

$$L_R = e^{-G_R \theta_R^2} \tag{15}$$

where θ_T denotes transmitter pointing error and θ_R denotes receiver pointing error.

3 Simulation Setup

Simulation setup parameters are given in Table 1 for the proposed system. The simulation has been done using the OptiSystem simulation tool. On the transmitter side, Laser Beam is split into two polarized beams. These two beams are given to two different optical OFDM systems. The data bit sequence is generated through PRBS (Pseudo Random Bit Sequence) Generator. Two such optical OFDM signals with different polarization are combined together with the help of Polarization Combiner and after the amplification, it is launched into space through a transmitter telescope.

The Dual Polarized Optical OFDM signal travels through the free space and reaches the receiver telescope which is in Line of Sight (LOS) with the transmitter telescope. It is then split into two polarizations with the help of a Polarization Splitter. The polarized light from receiving end laser is used for coherent detection. OFDM demodulator demodulates the signal and feeds it to the QAM sequence decoder.

Table 1 Simulation parameters

Parameters	Value
No. of subcarriers	104
No. of FFT points	128
No. of prefix points	14
No. of input/output port	1/1
Modulation (Mapper)	4 QAM
Bit rate	1 Tbps
Tx and LO laser power	30 dBm at 1550 nm
Line width	0.1 MHz
Tx/Rx pointing error	1.1 urad
Tx/Rx optical efficiency	0.8

Table 2 Performance Comparison with proposed design

Parameters	Ref. [20]	Ref. [21]	Ref. [22]	Ref. [1]	Proposed
Technique	256 QAM	256 QAM	256 QAM	256 QAM	OFDM
Bit rate	160 Gbps	80 Gbps	256 Gbps	640 Gbps	1 Tbps
No. of channel	1	1	1	1	1
BER	2×10^{-3}	2×10^{-3}	2×10^{-2}	2×10^{-3}	9.4×10^{-3}
Link distance	160 km	150 km	20 km	4 km (FSOL)	5000 km

4 Simulation Results and Discussion

The optical spectrum and the time domain representation of the signal are shown in Fig. 3. The spectrum Fig. 3(a) is mainly centered at 1550 nm of wavelength. The time-domain representation Fig. 3(b) can be visualized for the first 50 bits with the help of Optical Time Domain Visualizer.

At the receiving end after coherent detection and OFDM signal demodulation, the signal constellation diagram can be analyzed. The setup shows very promising results in terms of electrical signal constellation at distance up to 1000 km. As the distances increases, a scattered constellation is observed. In Fig. 4(a) and (b) the constellation diagram is shown for the distance of 3000 km.

As we further increase the distance, the scattering of the constellation is observed which affects the decision-making at the receiving end. Fig. 4(c) and (d) show such an issue and therefore it affect the BER as well. In Fig. 4(e) and (f) the constellation shows the interference with the adjacent bits and therefore it affects the decision-making process and BER.

The minimum BER has been almost zero or less than 1×10^{-6} for the distance up to 3000 km however as the distance increases the BER starts increasing as it is clearly visible in the constellation diagram itself. As shown in Fig. 5(c) for the distance of 4000 km the BER is 8×10^{-4} and for the distance of 5000 km, BER is 9.4×10^{-3}. The BER Analyzer X and Y show the eye diagram for both the polarization X and

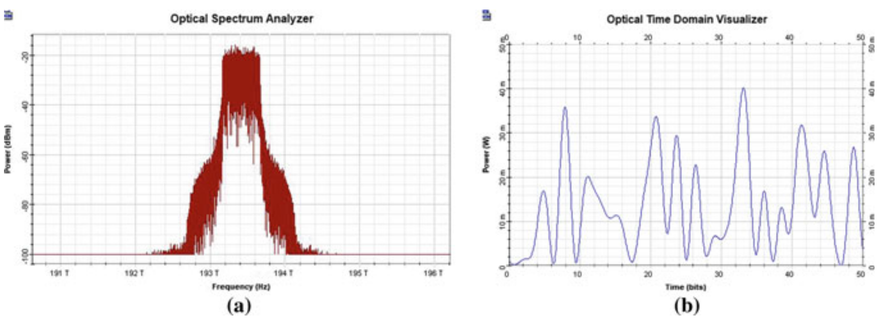


Fig. 3 **a** Optical spectrum **b** Time Domain representation at the output of transmitter

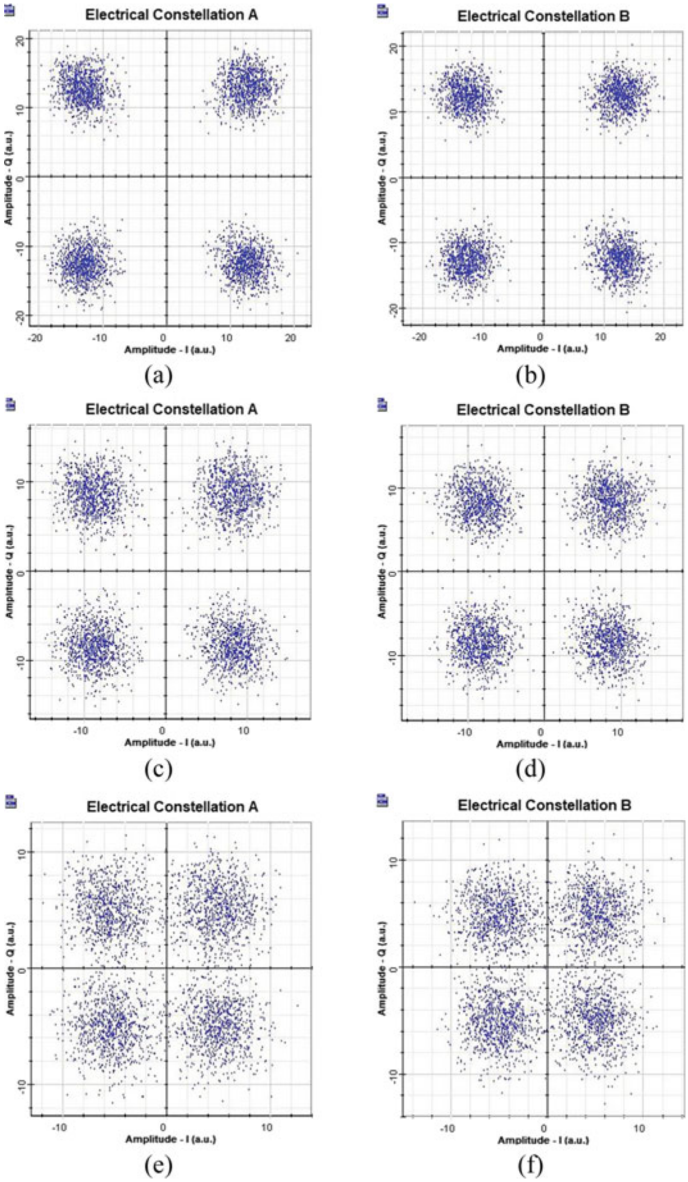


Fig. 4 Signal Constellation diagram for dual polarization **a** and **b** for 3000 km, **c** and **d** for 4000 km, **e** and **f** for 5000 km transmission distance

Y up to 3000 km as shown in FigS. 5(a) and 5(b). The opening of the eye in the eye diagram is desirable with eye height of 0.865 (a.u.) and the total jitter is 0.03 (bit period) which indicate the signal quality with minimal distortion.

As shown in Table 2 it depicts the performance comparison of the proposed system with others. The proposed system offers a significant advantage in terms of Bit rate and link distance. The wireless link impairments are generally taken care of by the use of OFDM. Due to the multicarrier modulation feature of OFDM, the data is carried out by many subcarrier tones. For the frequency selective fading channel, the use of OFDM has led to performance efficient and very high data rate systems.

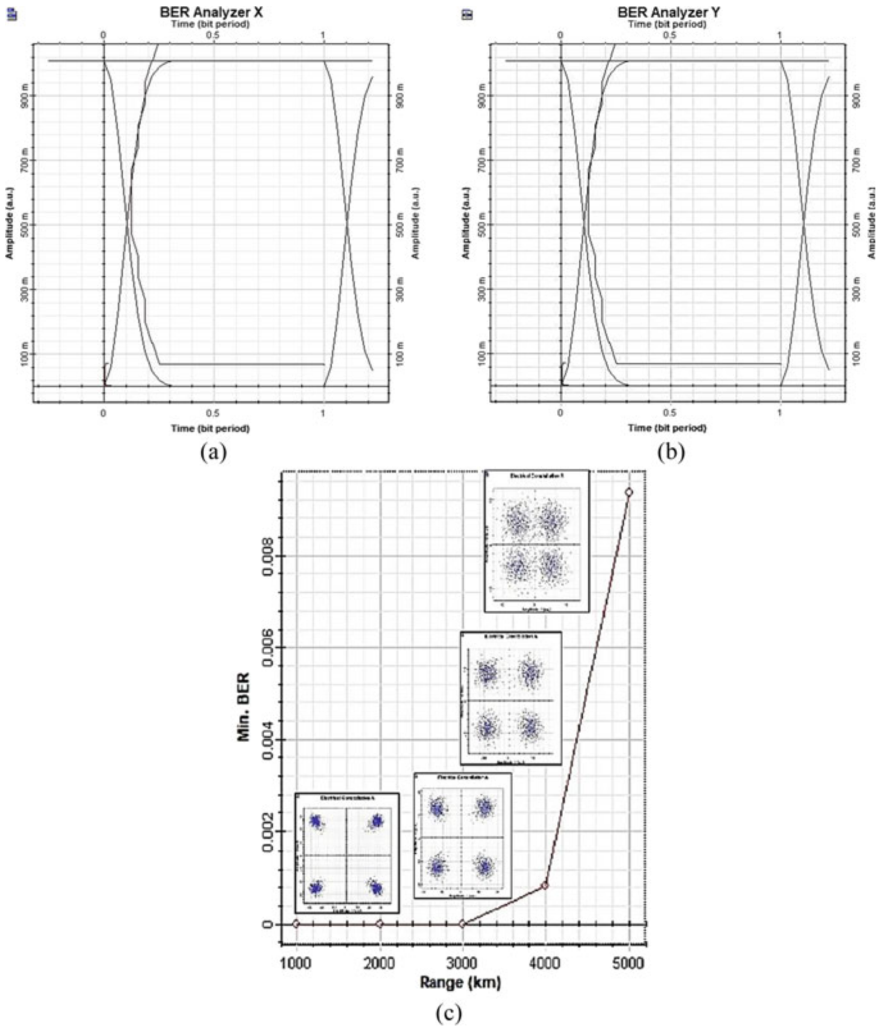


Fig. 5 BER Analysis **a** and **b** Dual Polarization (X and Y) and BER for different ranges **c**

While coding with OFDM generally subcarriers carry most of the information in the less affected band of the channel and least information into the band which is most affected. The center frequencies of such bands are decided by the frequency of the subcarrier. This approach offers a significant reduction in the effects of noise on the signal.

5 Conclusion

The simulation results show that the proposed system is suitable for providing 1Tbps of bit rate up to 5000 km with BER 9.4×10^{-3}. As there is a significant development in the field of Digital Signal Processing it is easier to implement and design OFDM systems. To increase the capacity Dual Polarization scheme has been adopted as it increases Bandwidth/Spectral Efficiency by the factor of 2. Enabling technology leads to improved data rates and capacity enhancement. The inherent advantage of OFDM can be integrated with the optical systems to model very high-speed communication systems. The proposed system can be very useful for the purpose of high-speed wireless optical links for satellites. As satellite-based internet systems are gaining popularity demand for high-speed interlinking of these satellites is inevitable.

References

1. Kakati, D., Arya, S.C.: A 640-Gbps, 15.2344-b/s/Hz full-duplex optical fiber/wireless single-channel coherent communication system using IQM-based DP-256-QAM and DSP techniques. Photonic Netw. Commun. **39**(1), 26–38 (2020). https://doi.org/10.1007/s11107-019-00875-7
2. Patnaik, B., Sahu, P.K.: Inter-satellite optical wireless communication system design and simulation. IET Commun. **6**, 2561–2567 (2012). https://doi.org/10.1049/iet-com.2012.0044
3. Sharma, V., Kumar, N.: Modeling of 2.5 Gbps-intersatellite link (ISL) in inter-satellite optical wireless communication (IsOWC) system. Optik **124**(23), 6182–6185 (2013). https://doi.org/10.1016/j.ijleo.2013.04.094
4. Sharma, V., Kumar, N.: Improved analysis of 2.5 Gbps-inter-satellite link (ISL) in inter-satellite optical-wireless communication (IsOWC) system. Opt. Commun. **286**, 99–102 (2013). https://doi.org/10.1016/j.optcom.2012.08.055
5. Ahriz, I., Douin, J.-M., Lemoine, F., Wei, A.: Performance evaluations in optical and wireless networks for CONDOR project. Procedia Comput. Sci. **130**, 73–81 (2018). https://doi.org/10.1016/j.procs.2018.04.014
6. Fazili, Y., Nafarieh, A., Raza, M.H., Robertson, B., Philips, W.: The effect of adopting green SLA on key parameters of optical WDM networks. J. Ambient Intell. Humaniz. Comput. **7**, 371–384 (2016). https://doi.org/10.1007/s12652-016-0364-5
7. Nafarieh, A., Raza, M., Robertson, W.: A comprehensive analysis of QoS-based routing mechanisms over shared mesh protected optical infrastructures. J. Ambient Intell. Humaniz. Comput. **6**, 463–472 (2015). https://doi.org/10.1007/s12652-015-0291-x
8. Siano, P., Graditi, G., Atrigna, M., Piccolo, A.: Designing and testing decision support and energy management systems for smart homes. J. Ambient Intell. Humaniz. Comput. **4**, 651–661 (2013). https://doi.org/10.1007/s12652-013-0176-9

9. Buchali, F., Dischler, R., Liu, X.: Optical OFDM: a promising high-speed optical transport technology. Bell Labs Tech. J. **14**(1), 125–146 (2009). https://doi.org/10.1002/bltj.20358
10. Lowery, A.J., Du, L.B., Armstrong, J.: Performance of optical OFDM in ultralong-haul WDM lightwave systems. J. Light. Technol. **25**, 131–138 (2007). https://doi.org/10.1109/JLT.2006.888161
11. Shieh, W., Chen, W., Tucker, R.S.: Polarisation mode dispersion mitigation in coherent optical orthogonal frequency division multiplexed systems. Electron. Lett. **42**(17), 996 (2006). https://doi.org/10.1049/el:20061639
12. Sano, A., et al.: No-guard-interval coherent optical OFDM for 100-Gb/s long-haul WDM transmission. J. Light. Technol. **27**, 3705–3713 (2009). https://doi.org/10.1109/JLT.2009.2023369
13. Lowery, A.J., Armstrong, J.: 10 Gb/s multimode fiber link using power-efficient orthogonal-frequency-division multiplexing. Opt. Express **13**, 25 (2005)
14. Shieh, W., Bao, H., Tang, Y.: Coherent optical OFDM: theory and design. Opt. Express **16**, 841 (2008). https://doi.org/10.1364/oe.16.000841
15. Tang, Y., Yi, X., Shieh, W., Evans, R.: Optimum design for coherent optical OFDM transmitter. Opt. InfoBase Conf. Pap. **19**, 483–485 (2007)
16. Hara, S., Prasad, R.: Multicarrier Techniques for 4G Mobile Communications. Artech House, Boston (2003)
17. Li, C., Yang, Q.: Optical OFDM/OQAM for the future fiber-optics communications. Procedia Eng. **140**, 99–106 (2016). https://doi.org/10.1016/j.proeng.2015.09.238
18. Arnon, S.: Performance of a laser µsatellite network with an optical preamplifier. J. Opt. Soc. Am. A. **22**, 708 (2005). https://doi.org/10.1364/josaa.22.000708
19. Polishuk, A., Arnon, S.: Optimization of a laser satellite communication system with an optical preamplifier. J. Opt. Soc. Am. A. **21**, 1307 (2004). https://doi.org/10.1364/josaa.21.001307
20. Toyoda, K., Koizumi, Y., Omiya, T., Yoshida, M., Hirooka, T., Nakazawa, M.: Marked performance improvement of 256 QAM transmission using a digital back-propagation method. Opt. Express **20**, 19815 (2012). https://doi.org/10.1364/oe.20.019815
21. Kasai, K., Wang, Y., Beppu, S., Yoshida, M., Nakazawa, M.: 80 Gbit/s, 256 QAM coherent transmission over 150 km with an injection-locked homodyne receiver. Opt. Express **23**, 29174 (2015). https://doi.org/10.1364/oe.23.029174
22. Chien, H.C., Jia, Z., Yu, J.: 256-Gb/s single-carrier PM-256QAM implementation using coordinated DD-LMS and CMA equalization. In: European Conference on Optics and Communication, ECOC, November 2015, pp. 2–4 (2015). https://doi.org/10.1109/ECOC.2015.7341619.
23. Kao, K.C., Hockham, G.A.: Dielectric-fibre surface waveguides for optical frequencies. Elektron **14**, 11–12 (1997). https://doi.org/10.1049/piee.1966.0189
24. Pandey, P.K., Jain, S., Tiwari, P.: Realization of 1Tbps FSO / OWC based inter satellite link using DP-QPSK for next generation LEO satellite internet system. In: 2021 IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), pp. 1–5 (2021). https://doi.org/10.1109/UPCON52273.2021.9667581

FBG-Based Large Scale WDM-PON Optical Broadcast Communication System

Divya Sisodiya and Gurjit Kaur

Abstract First-class quality and minimum signal errors are the basic requirements of this era, and these requirements are constantly expanding according to the prerequisites of different applications such as real-time communication, network video, and video conferencing. In this research, the FBG-based WDM broadcast communication system is designed to transmit different types of information, such as sound, video, design, text, etc., within the frequency range of 1550 nm. The basic idea behind the framework comes from the passive optical network and adjusts it. The framework is planned in such a way that customers can only identify information of one frequency data at a time. The presentation of the frame is concentrated between 2.5 to 10 gigabits per second, and then the correlation between the Q factor and the bit error rate (BER) is completed based on the results obtained in the bit error analyzer. The framework is simulated with uniform and non-uniform FBG, it can be seen that from obtained results, non-uniform FBG will reduce the dispersion by itself and provide a better- Quality factor of 60. If there is a uniform FBG, then pumping by using a coupler at the same frequency can reduce the effect of dispersion and produce an ideal quality factor of more than 11 for the signal. EDFA amplifiers are associated in the channel to enhance different optical signals at the same time, followed by single-mode fiber (SMF). The framework has been simulated in Opti system's Opti Wave system software.

Keywords Passive optical networks · WDM · Single mode fibers

1 Introduction

The development of fiber access such as fiber to the home and fiber to the building. FTTX has outstanding advantages such as high exchange speed and long data transmission capacity and is currently circulating globally. Fiber-to-the- home (FTTH)

D. Sisodiya (✉) · G. Kaur
Delhi Technological University, Delhi, India
e-mail: divsisnit27@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_16

networks are becoming more familiar with point-to-point (P2P) plans and point-to-point multiplexing (P2MP) passive time-division multiplexing (PON) optical networks [1]. Due to the Internet and network improvement, voice and text authorization management has evolved into image and data-based management. In addition, it has been seen that the traffic plan gradually becomes symmetrical. This difference of opinion will require new access networks to help enable high-speed information transactions (2.5, 10 Gbps or higher), better bandwidth, and ensure superior television quality (HDTV) for future video management. After the start of the fiber to the home field, passive optical networks have become popular [1, 7].

The transmission strategy adopted in the study is based on the GPON network. The network is connected to the WDM pair and can be used as a metropolitan area network. GPON supports high-speed data transmission (2.5 to 1.25 Gbps). GPON can support various services, including voice, Ethernet, ATM, and remote location. In addition, it manages different baud rates. At present, the data transmission rate of uplink communication is 0.24 Gbps, and the data transmission rate of downlink communication is 2.44 Gbps [7]. The effective wavelength range used by GPON is 1480 to 1500 nm for downstream communications.

Wavelength division multiplexing (WDM)-based PON is the most advanced PON, dedicated to achieving huge electronic transmission speeds at electronic speeds. Unlike GPON or any TDM PON (Time Division Multiplexed PON) that uses a single fiber to transmit information, WDM-PON uses different frequencies in the contrast fiber between the ONT and OLT according to the needs of the end customer [2]. The basic problem associated with high data rate optical communication networks is dispersion. A uniform FBG provides a non-dispersive signal with the help of a coupler associated with a similar connection path, while a non- uniform FBG generates a dispersion compensated signal due to the chirped grating. For these two FBGs, the framework is implemented as efficiently as possible at 2.5 Gbps or higher data rates. In practical applications, the framework uses components such as FBG to make it non-dispersive. The coupler is used to enhance the signal. The two are used together as a switch to provide ideal information according to the customer's preconditions.

The research paper is divided into sections describing the complexity of the framework. Further explanation is given in the following sections. Section 2 describes the subsystem model of WDM and FBG frameworks. Section 3 describes the concept of FBG. Section 4 described the simulation and modeling of the WDM PON broadcast communication system in the Opti system. The results and observations are presented in Sect. 5. Section 6 includes the conclusion of research.

2 Coupler and FBG-Based WDM Broadcast Transmission System

Figure 1 below is the subsystem module of the WDM-PON transmission system. The simulation settings are created using Opti-system software, where we use the switching mechanism produced by the combination of pump coupler and FBG. These FBGs can be uniform or non-uniform. The data signal is an optical signal with a frequency of 1550 nm. Dispersion is considered to be one of the basic problems in high data rate optical communication frames. Dispersion obviously affects the speed and quality of data transmission, leading to the disappearance of the bitstream. Diffusion repair fibers are widely used to compensate for this effect. Fiber Bragg grating is a fully mature development, especially suitable for compensating the dispersion in the optical communication frame. FBG-based development can be seen as an economically wise choice rather than other competitive advancements, providing huge advantages such as smaller size, no nonlinear effects, and low loss. Dispersion compensation using FBG is based on the wavelength-specific time delay used by chirp FBG. The ultra-long FBG adopts an adaptive suppression granulation strategy, which can provide us with unprecedented possible results.

The signal is then directed towards forward pump. In forward pumping, the information signal and the pump signal cause comparison of routes within the fiber. Inside the pump coupler, the energy of the pump signal is transferred to the information signal and is enhanced with the performance of the optical fiber [3]. After passing kilometers of single-mode fiber, the signal passes through the EDFA, where the signal is combined with Er3 + particles and enhanced. In terms of performance, we have obtained improved data from the 1550 nm information signal. WDM DEMUX demultiplexes the original data of a single user at the output of the receiver. The purpose of the demonstration is to develop an operational simulation model that can be successfully transmitted on an FBG-based wavelength division multiplexing (WDM-PON) passive optical network.

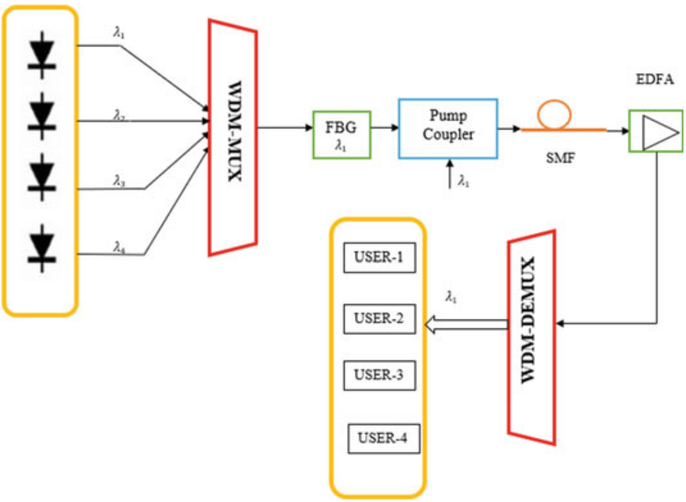


Fig. 1 Subsystem of FBG-based WDM broadcasting system

3 Concept of FBG

Bragg's fiber grating is one of the most flexible and advanced technologies for dispersion compensation. In most cases, it can be used as a single-channel or multi-channel dispersion compensation module (DCM) and has been proven to be the basic parameter of wavelength division multiplexing (WDM) and optical time-division multiplexing (OTDM) [4, 5]. Due to the better optical characteristics of chirped or non-uniform FBG, it is more suitable for WDM systems.

The grating period of FBG is not a fixed value but linearly changes as the length of the grating changes, so that the grating period at the beginning of the grating becomes shorter. The grating period of FBG can be defined as the distance between the two closest extreme refractive index estimates. FBG reflects a narrow spectral wavelength focused on λ_B and allows all other wavelengths to pass.

The basic principle of the FBG process is Fresnel reflection, which causes light emitted between supports of different refractive indices to be reflected and refracted at the interface. To obtain the reflected wavelength, the Bragg condition in Eq. (1) should be satisfied.

$$\lambda_B = 2\Lambda n_g \tag{1}$$

λ_B is the reflected bragg wavelength.
Λ is symbol of grating period.
n_g is the refractive index (effective).

The wavelength spacing between the first minima, or the bandwidth is given by Eq. (2).

$$\Delta\lambda = \left[\frac{2\delta n_0 \eta}{\pi}\right]\lambda_B \tag{2}$$

where, δn_0 is variation in refractive index, and η is the fraction power in the core.

4 Simulation Model and Work Explanation

The created model appears in Fig. 2 and was additionally used for the perception of single wavelength information according to the customer's interest, which prompted the use of different channels (four channels in one subsystem) on the transmission side and all are associated with wavelength division multiplexers after going through couplers and FBGs (both act as switches). As for the data transmission system, the model we created uses a modulation plan of non-return to zero NRZ coding. The transmitters used in optical line terminals with a system speed of 10 Gbit/s are on the service provider's side. From a basic point of view, we have considered several component configurations, where each transmitter includes a 10 Gbit/s pseudorandom binary sequence (PRBS) information source, an NRZ driver, a Mach–Zehnder optical modulator (MZM), and a continuous wave (CW) laser. The four information signals shown in Fig. 2 are set as locations A, B, C, D with wavelengths 193.1, 193.2, 193.3, 193.4 THz. After setting the wavelengths, signals are combined together to one subsystem, then three additional transmitters' subsystems are created, after that signal is passed through optical switches and amplifiers, and then coupled through Wavelength division multiplexers (MUX) for further transmission in the optical distribution network.

Through these center frequencies, different types of data can be sent. The FBG and the coupler act as switching devices, because by changing the frequency range of the FBG and sending data at the same frequency together with the pump coupler, we can access a specific range of data on the receiver side. After the switching operation, the data is sent through a channel that includes an erbium-doped fiber amplifier (EDFA) and single-mode fiber. The length of single-mode fiber can vary from 10 to 40 km to obtain error-free results. The use of EDFA amplifiers ensures that dispersion is eliminated and the bit error rate is reduced. At the receiving end, the demultiplexer is used to separate the transmission channel data by WDM DMUX. The optical receiver network unit uses a PIN photodiode with a responsivity of 1 A/W, and the output

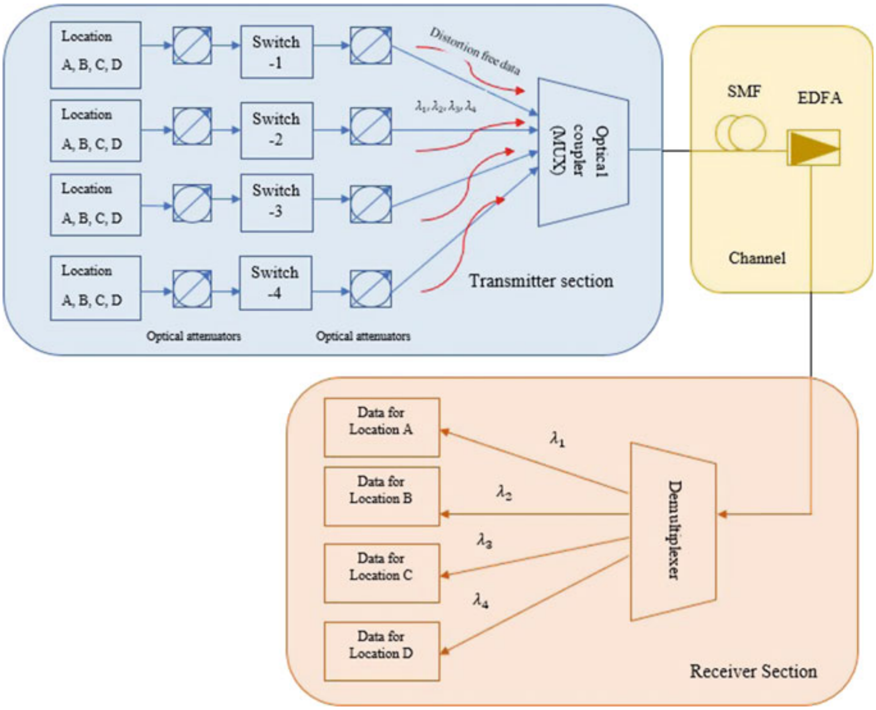


Fig. 2 Simulation model of FBG-based WDM-PON broadcast communication system

Table 1 System parameters used for simulation in Opti system

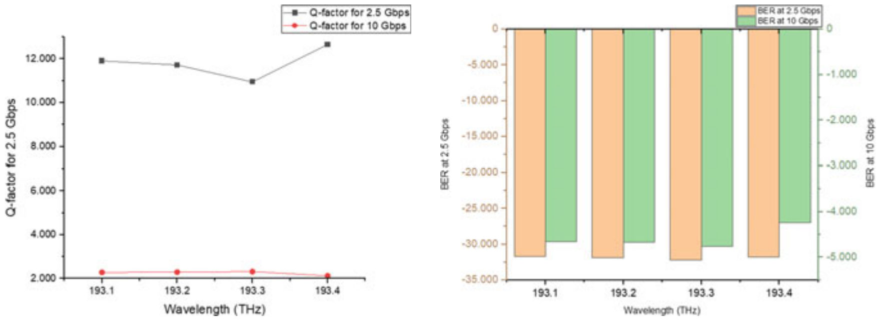
System parameters	Values used
Wavelength	193.1,193.2,193.3,193.4 THz
System bit-rate	2.5, 10 Gbps
Optical fiber length	10 to 40 km
Sensitivity	−100 dBm
Resolution	0.1 nm

of the photodiode is connected to an electrical Bessel low pass filter of bandwidth 1.875 GHz if the bit rate is 2.5 Gbps, and bandwidth changes to 7.5 GHz if the bit rate is 10 Gbps with a cutoff frequency of 3 dB, and then connected to the BER estimator. The system parameters established in the simulation process are shown in Table 1.

5 Results and Discussions

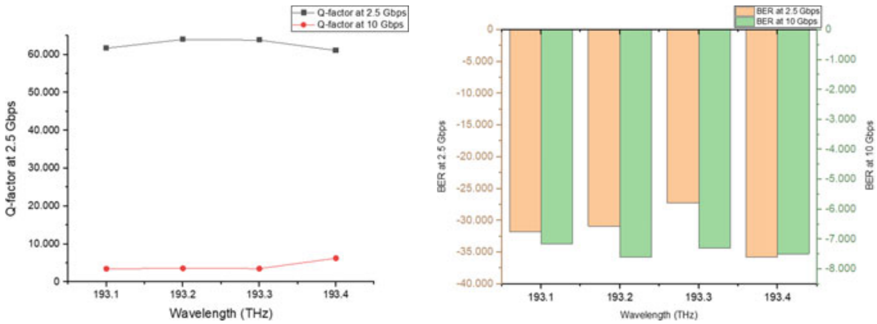
The results were obtained according to each of the four sub- systems used at the transmitter and receiver sides. Frequencies are changed on each FBG and pump coupler one by one. Moreover, types of FBGs are changed as uniform and non-uniform. It has been observed that Uniform FBGs reflect the matched wavelength signal very strongly and that's why at the receiver, the signal of Bragg wavelength is distorted. If the same signal that is passing through Uniform FBG is passed through a pump coupler with the pumping of the same wavelength, then the distortion can be reduced. While, in the case of Non-uniform FBGs, due to their non-uniform structure distortion is getting reduced and it does not need any other pumping technique in the circuit to get a good-Quality factor. Basically, the quality of the system specifically for uniform FBGs has been tried to increase by connecting pump couplers, and then results are compared by using chirped FBGs also so that performance of the model can be checked for our research. This system is simulated at the rate of gigabits per second which is from 2.5 to 10Gbit/s. The comparative results of FBGs (uniform and non-uniform) are taken in terms of Q-factor and Bit error rate (BER).

For a uniform FBG, the Q factor is 12, which corresponds to 2.5 Gbit/s. As the bit rate increases to 10 Gbit/s, the signal quality decreases. It can be seen from Fig. 3(a) that the value of Q is constant. The quality factor Q of non-uniform FBG is about 60, which is excellent at 2.5 Gbps, but it has dropped to 9 to 10 Gbit/s, as shown in Fig. 4(a). Figure 3(b) shows the bit error in the exponential factor of uniform FBG performance. Looking at the exponential factor, we can say that the bit error is small at 2.5 Gbps, but large at 10 Gbps. When the FBG is uneven, no error will occur at 2.5 Gbit/s. When the bit rate becomes 10 Gbit/s, errors are observed, as shown in Fig. 4(b).



(a) (b)

Fig. 3 **a** Q-factor performance for uniform FBG, **b** Bit errors performance for uniform FBG



(a) (b)

Fig. 4 **a** Q-factor performance of Non-Uniform FBG, **b** Bit error performance for Non-uniform FBG

Table 2 Observation at 2.5 Gbps for uniform and non-uniform FBGs

Type of FBG	Average Q-factor at 2.5 Gbps	Average BER at 2.5 Gbps
Uniform FBG	11.807	−31.902
Non-uniform FBG	62.649	−31.401

Table 2 is showing the average value of Q -factor and BER at 2.5 Gbps. Values are calculated in the range of 193 THz.

Table 3 is showing the average value of Q -factor and BER at 10 Gbps. From Tables 2 and 3, it is observed that system model is working efficiently for 2.5 Gbps as compared to 10 Gbps.

Figure 5 shows the observation of the eye diagram of uniform and non-uniform FBG.

It can be clearly seen from the figure that as the bit error rate increases, the signal will be distorted. For a distortion-free and noise-free fit, the standard BER rate should

Table 3 Observation at 10 Gbps for uniform and non-uniform FBGs

Type of FBG	Average Q-factor at 10 Gbps	Average BER at 10 Gbps
Uniform FBG	2.257	−4.583
Non-uniform FBG	4.227	−7.383

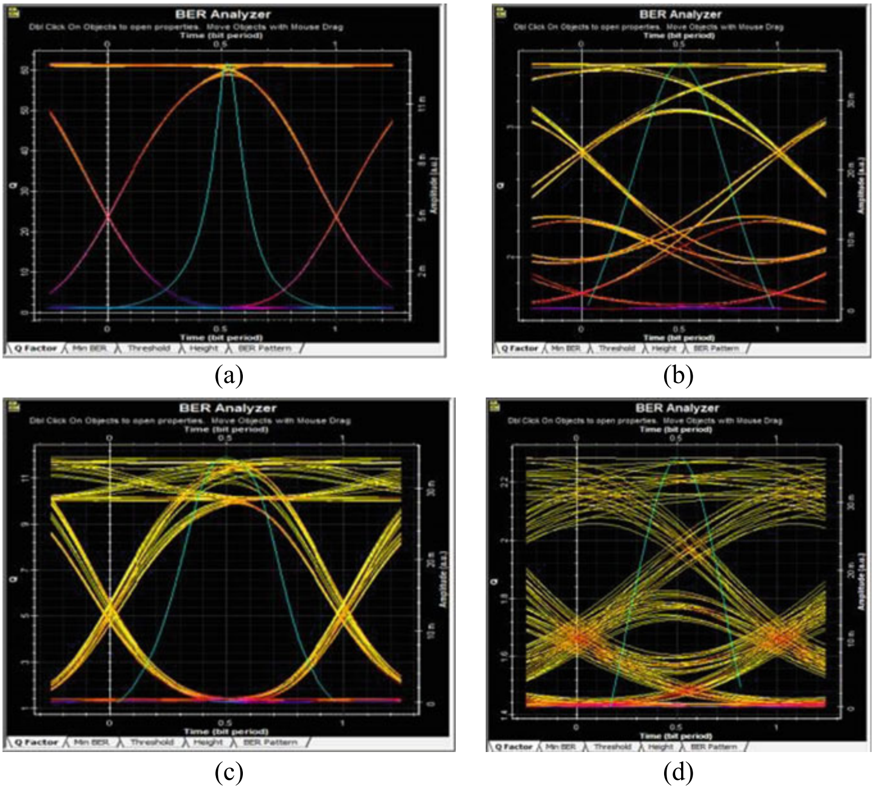


Fig. 5 **a** Non-Uniform FBG Curve for 2.5 Gbps, **b** Non-Uniform FBG Curve for 10, **c** Uniform FBG Curve for 2.5, **d** Uniform FBG Curve for 10

be less than 10 ^ (−12) [1, 6]. According to the BER guidelines for the transmission or communication of feasible signals in remote optical organizations, the best re-creation results are obtained at 2.5 Gb/s. The system can be used by organizations to simultaneously transmit different data and receive specific data according to needs. The system can be used to connect multiple cameras from different areas of the city or different rooms in a building to multiple PCs, and the screen can display the data of one area/room at a time. The system can also be created as an optical map, which can effectively detect data of one location at a time. Since the system displays reliable results at a high bit rate, it can be used in metropolitan area networks to receive high-quality signals.

6 Conclusion

Dispersion is an important factor that must be minimized in order to receive high data rates in the transmission network in order to take full advantage of the optical fiber communication indicated by the customer's interest. Therefore, the project work is mainly focused on the maximum Q factor and minimum bit error rate when EDFA amplifier and pump technology are introduced into the system. The maximum Q factor is obtained for the 2.5 Gbps system data rate. As the data rate further increases and reaches 10 Gbps, the Q factor decreases at the same time. The system is implemented on a large scale for broadcast transmission. Due to its minimal distortion, high-quality factor, and negligible distance correlation, this system will bring gratifying results to metropolitan area networks.

References

1. Verma, S., Kakati, A., Saurabh, L., Bhulania, P.: Performance analysis of Q-Factor and polarization for GPON network using Optisystem. In: 2016 International Conference on Information Technology (InCITe)- The Next Generation IT summit (2016)
2. Senkans, Ugis, Braunfelds, Janis, Lyashuk, Ilya, Porins, Jurgis, Spolitis, Sandis, Bobrovs, Vjaceslavs: Research on FBG-based sensor networks and their coexistence with fiber optical transmission systems. J. Sens. **2019**, 1–13 (2019). https://doi.org/10.1155/2019/6459387
3. Tangonan, G.L.: Fiber optic couplers. Ateneo de Manila University (1991)
4. Kocher, D., Kaler, R.S., Randhawa, R.: 50 km bidirectional FTTH transmission comparing different PON standards. Electronics & Communication Engineering Department, Thapar University, Patiala, India. Elsevier GmbH (2013)
5. Lee, C.-H, Sorin, W.V., Kim, B.Y.: Fiber to the home using a PON infrastructure. J. Lightwave Technol. **24**(12), 4568–4583 (2006)
6. Kim, Hyoungsoo, et al.: An electronic dispersion compensator (EDC) with an analog eye-opening monitor (EOM) for 1.25-Gb/s gigabit passive optical network (GPON) upstream links. IEEE Trans. Microw. Theory Tech. **55**(12), 2942–2950 (2007). https://doi.org/10.1109/TMTT.2007.909887
7. ITU-T G.984.1 Telecommunication Standardization Sector of ITU (03/2008). Series G: Transmission Systems and Media, Digital Systems and Networks, Digital sections and digital line system – Optical line systems for local and access networks. Printed in Switzerland Geneva (2009)

Design of a Dispersion Compensating Hexagonal Photonic Crystal Fiber (DC-HPCF) for High Nonlinearity and Birefringence

Sanat Kumar Pandey, Priyanka Dwivedi, Shivam Singh, and Y. K. Prajapati

Abstract This manuscript deals with a novel hexagonal lattice structure based PCF having circular shaped air holes in the cladding region. A rectangular slot filled with Gallium phosphide (GaP) is introduced inside the core region to confirm a photonic crystal fiber with the characteristics of ultra-high nonlinearity and high birefringence. The complete analysis of the proposed structure is performed on the COMSOL multiphysics platform using Finite element method (FEM). Different optical parameters of the proposed DC-HPCF like chromatic dispersion (D), numerical aperture (NA), birefringence, effective refractive index, v-number, effective area (EA), nonlinearity (NL) and effective material loss (EML) are analysed with different geometrical variables. In the proposed structure of DC-HPCF the optimisation of all the above optical characteristics is achieved with the pitch value of 0.80 μm. With the help of this optimised proposed structure of DC-HPCF the ultra-high nonlinearity coefficient value of 28,978 W^{-1} km^{-1} at birefringence value 0.121 are confirmed with high negative dispersion and low effective area of −2885 ps/(nm.km) and 9.4×10^{-13} m^2, respectively at wavelength 1.55 μm. Therefore, with above confirmed optical characteristics the proposed structure of DC-HPCF is best suited to perform as a four wave mixer, polarization splitter, polarization-maintaining fiber and also as a dispersion compensating fiber.

S. K. Pandey (✉)
Department of Electronics Engineering, Government Polytechnic Lucknow, Lucknow 226016, Uttar Pradesh, India
e-mail: spandey@mnnit.ac.in

P. Dwivedi
Department of Information Technology, Government Girls Polytechnic Lucknow, Lucknow 226016,
Uttar Pradesh, India

S. Singh
Department of Electronics and Communication Engineering, Aditya Engineering College, Surampalem 533437, Andhra Pradesh, India

Y. K. Prajapati
Department of Electronics and Communication Engineering, Motilal Nehru National Institute of Technology Allahabad, Allahabad 211004, Uttar Pradesh, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_17

Keywords Photonic crystal fiber (PCF) · Nonlinearity · Chromatic dispersion · Numerical aperture · Effective material loss

1 Introduction

PCF is very popular technology to handle the light wave. It has major improvement over the traditional fiber such as negative dispersion with tuneable birefringence [1], high effective area [2], polarization-maintaining capability [3], and high nonlinearity etc. Conventional optical fibers have limitation in its versatile uses like in sensing applications [4], high power mechanism [5], nonlinear fiber optics [6], spectroscopy [7], metrology, super-continuum generation [8], optical coherence tomography [9], quantum dots [10], etc. PCF technology offers a high degree of freedom towards structural parameters due to which it becomes possible to design different structures with admirable design flexibility, which is not possible in conventional fibers. The compact and robust nature, less environmental effects, light-weight, and minimum cost makes PCF based devices more popular to work even in enormous environmental conditions [11]. PCF possess special characteristics such as high birefringence and nonlinearity which can be tuned with structure design parameters. High value of nonlinearity are needed for many applications like sensing, laser technology and supercontinuum generation [12]. To achieving high nonlinearity in PCFs, it is highly recommended to maintain the small effective mode area. Besides, it can be also improved by using high refractive index materials such as GaP, and carbon disulphide (CS_2) [13].

Nowadays, researches have been effectuated to reduce the dispersion and flattened the curve for dispersion compensation. Razzak et al. [14] suggested a distinguish HNL-PCF for the telecommunication window in which finite difference method with PML is used for simulation. They have obtained high value of nonlinearity coefficient as 27 W^{-1} km^{-1} at 1.55 μm and within wavelength range of 1.46 to 1.66 μm achieved flattened value of dispersion as 0 ± 0.5 ps/(nm.km) with very low confinement loss in the overall band of interest. In the year 2014 and in year 2016, Hasan et al. suggested an octagonal shaped structure PCF with elliptically structured core as well as the equiangular spiral structure [15]. In their manuscript, they showed high negative dispersions of −544.7 and −526.99 ps/(nm.km) respectively for both the structures. They have also reported another single mode octagonal PCF which has an ultra-flattened dispersion with negative sign of −608.93 ps/(nm.km) with variation of −12.7 ps/(nm.km) in the wavelength of 1.46–1.625 μm [16]. In 2018, Y.K. Prajapati et al. proposed a hexagonal PCF design which is doped with germanium for wideband supercontinuum generation and achieved nonlinearity coefficient value of 0.0166 W^{-1} m^{-1} with the chromatic dispersion smaller from − 11.8 ps/(nm.km) at 1.55 μm [17].

The photonic crystal fibers which have smaller values of nonlinearity coefficient are designed with pure silica [14]. To compensate this problem SF-57, Bismuth and, GaP etc. materials with high nonlinear can be used.

The main objective of the proposed DC-HPCF is to analyse the presence of GaP filled rectangular hole present at the core region to high negative dispersion, nonlinearity as well as high birefringence. In our proposed work, we analyse the performance of DC-HPCF which have five hexagonal rings. Among these, four outer rings of cladding region have uniform air holes distribution while a core rectangular slot filled with GaP is surrounded by the innermost ring. To see the effect of different pitch sizes (i.e. p = 0.80, 0.82 and 0.84 μm) on the performance of DC-HPCF, the wavelength is varied from 0.8 to 2.8 μm.

2 Design and Modal Realization

2.1 Geometry Modelling

The optical characteristics of PCF are described by the geometrical parametric quantities such as pitch value (the distance between consecutive air-holes, diameter of small air-holes, position of the air-holes and air filling fraction. Figure 1 represent the 2D-anatomical of the proposed DC-HPCF comprises of total five rings in hexagonal pattern of air holes (circular) are present in cladding section of diameter value d (0.96 μm). Core section of the crystal is bored with rectangular shaped GaP filled hole for high value of negative dispersion and nonlinearity with high birefringence. The circular perfectly matched layer (PML) of size 1 μm is applied across the boundary of DC-HPCF. The FEM based COMSOL Multiphysics (5.1) as a simulation software is used to analyzed the entire model. By selecting the proper pitch value (i.e. 0.80, 0.82 or 0.84 μm), the value of effective refractive index can be controlled which efficaciously enhance the negative dispersion. Here, the proposed PCF has two degrees of freedom i.e. pitch (p) and wavelength. By optimizing these parameters, we can achieve the required dispersion and nonlinearity with high birefringence. To measure the performance parameters of DC-HPCF for communication applications, we have thoroughly done the modal analysis.

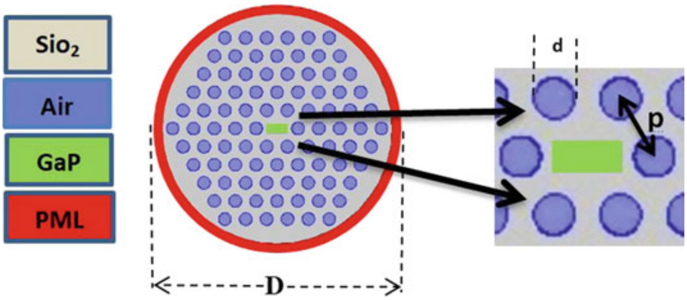


Fig. 1 2D schematic of the DC-HPCF with its enlarged core view

2.2 The Proposed Structure Synthesis and Analysis of Various Parameters

2.2.1 Dispersion

This is a commonly known fact that an optical source can transmit a range of frequencies along with a particular fixed frequency component, however the broadening in transmitting modes occurs due to the occurrence of propagation delay between the various component of light rays, which is known as intermodal dispersion. Which is analysed by the mathematical expression equation [18].

$$D(\lambda) = -\frac{\lambda}{c}\frac{d^2 R_{re}(n_{eff})}{d\lambda^2}\,\mathrm{ps}/\mathrm{nm.km} \tag{1}$$

Here, $D(\lambda)$ represents total effective dispersion of DC-HPCF, the combined effect of optical waveguide and material dispersion, represented with the help of the following mathematical expression as [19].

$$D(\lambda) = D_m(\lambda) + D_w(\lambda) \tag{2}$$

$D_m(\lambda)$ is analysed with the help of Sellmeier equation for a silica glass waveguide and $D_w(\lambda)$ in the terms of geometrical optical model index of HDC-PCF [19].

2.2.2 Numerical Aperture (NA)

Numerical aperture is an optical geometrical term with the help of we can find out of geometrical structure the effect of an optical fiber, it is measure of light power gathering ability of the core an fiber. It is an unit less parameter. The value of NA is depends on effective area A_{eff} as given below [20].

$$\text{Numerical Aperture (NA)} = \left[1 + \frac{\pi A_{eff}}{\lambda^2}\right]^{-1/2} \tag{3}$$

2.2.3 Birefringence (B)

Birefringence value of an optical fiber is a parameter which shows the effect of geometry and material used on polarisation of an optical fiber. The real part modulus value of difference between the *x*-polarized effective RI and y-polarized effective RI at a particular wavelength known as Birefringence which can be expressed as given [20]:

$$B = \left| Re\left\{ n_{eff}^{x} - n_{eff}^{y} \right\} \right| \tag{4}$$

where, n_{eff}^{x} and n_{eff}^{y} shows the net amount of refractive index values in both x and y direction respectively.

2.2.4 Effective Refractive Index (n_{eff})

The effective refractive index of an optical fiber is a property which also govern the speed of light inside the core of an fiber. The effective value of refractive index of pure silica can be estimated with the Sellmeier expression which is given as [21]:

$$n_s^2 = \left(1 + \frac{0.692\lambda^2}{\lambda^2 - 0.0047} + \frac{0.408\lambda^2}{\lambda^2 - 0.014} + \frac{0.897\lambda^2}{\lambda^2 - 97.934} \right) \tag{5}$$

where, n_s is the net value of refractive index of silica material in the fiber core and λ is the value of wavelength under consideration.

2.2.5 V-Number (V_{eff})

The V parameter is an optical parameter of an fiber which shows the amount of optical power associated with a particular mode of an fiber core, which is represented by V-number which can be expressed as [22].

$$V_{eff} = \frac{2\pi p}{\lambda} \sqrt{n_{core}^2 - n_{eff}^2} \tag{6}$$

where, p stands for pitch of the PCF, n_{core} and n_{eff} express the net core and cladding material effective RI.

2.2.6 Effective Area (A_{eff})

Effective area is basically a parameter of an fiber which shows the energy handling capacity. The ability of a particular fiber which measure how much amount of optical intensity of light or how much energy a fiber can handle without any nonlinear effects is known as effective area. It can be limiting by designing proper pitch value(p) and filling fraction of air inside the fiber core. It can be expressed as [2].

$$A_{eff} = \frac{\left(\int \int_{-\infty}^{+\infty} |E|^2 dxdy \right)^2}{\int \int_{-\infty}^{+\infty} |E|^4 dxdy} \mu m^2 \tag{7}$$

A_{eff} value is totally depends on type of fiber under consideration, the input wavelength and also on fiber's RI profile.

2.2.7 Nonlinearity Coefficient (γ)

The nonlinearity of a photonic crystal fiber is totally depends on the core diameter, the high value of nonlinearity is preferred over the conventional fiber to perform various nonlinear effects. To achieve high nonlinear effect with a optical fiber, the fibers are designed with very small core radius. The nonlinear coefficient (γ) can be calculated as [2].

$$\gamma = \left(\frac{2\pi}{\lambda}\right)\left(\frac{n_2}{A_{eff}}\right) \tag{8}$$

Here n_2 represents the nonlinear net refractive index of fused silica (2.33×10^{-20} m^2/w) material present inside the core region and λ is the operating wavelength of light.

2.2.8 Effective Material Loss (EML)

The effective material loss of an optical fiber shows the effect of material on loss of the energy fraction traveling from fiber material. It is also a measure of the ability of a fiber core to what fraction of power confined within the core out of total input power is known as effective material loss (EML), it is an important parameter for a PCF. it can be calculated as [23]:

$$\alpha_L = \frac{1}{2}\left(\sqrt{\frac{\varepsilon_0}{\mu_0}}\right)\left(\frac{\int_{A_m} n\alpha_m \left|E^2\right| dA}{\int_{All} S_z dA}\right) \tag{9}$$

Here, α_m is absorption loss of material used to fabricate the fiber. A_m is associated area occupied by fiber material, E is the field vector. The term μ_0 and ε_0 are represents the permeability and permittivity of material used in free space and n represents the refractive index of material.

3 Results and Discussions

Figure 2 represents the distribution of field vector in which Fig. 2(a) shows x- polarized mode indicated with horizontal red arrow while Fig. 2(b) is for y- polarized core mode shown with vertical red arrow at 1.55 μm wavelength.

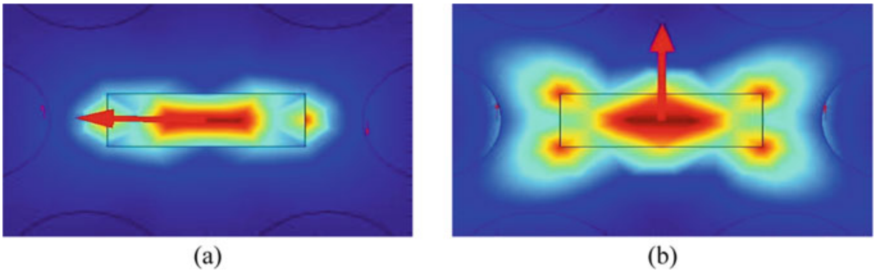


Fig. 2 Field distributions at 1.55 μm wavelength

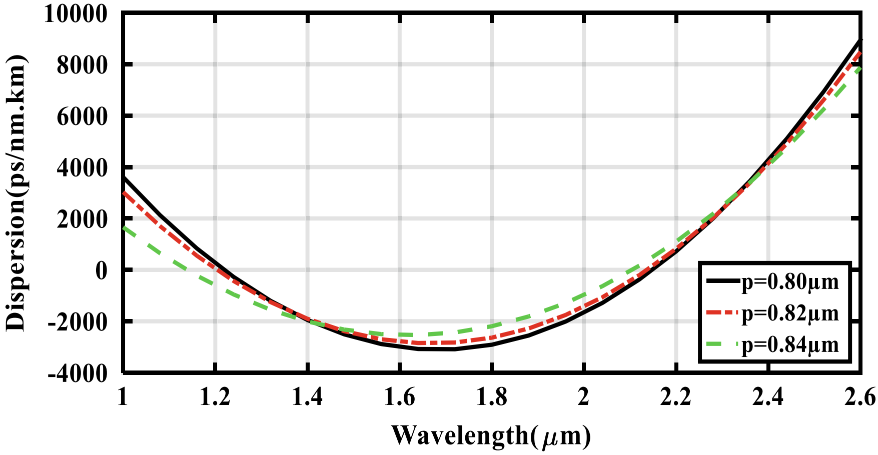


Fig. 3 Chromatic dispersion of DC-HPCF structure in reference of wavelength, which is bored with rectangular GaP filled holes at core region

Figure 3 represents properties of chromatic dispersion for the fiber structure in reference with wavelength whose core is bored with a rectangular shaped GaP hole. Here, we plot the wavelength response for three different pitch values as p = 0.80, p = 0.82 and p = 0.84 μm, we find that the proposed structure with pitch value p = 0.80 μm has the highest negative dispersion of −2885 ps/(nm.km) at 1.55 μm wavelength.

From Fig. 4, we can observe that the value of numerical aperture is increasing gradually with the increase in wavelength value for all three different pitch values under consideration. Therefore, the peack value of numerical aperture for the considered structure is obtained as 0.697 with p = 0.80 μm at wavelength 1.55 μm.

Figure 5 represents the graph of birefringence for the structure with p = 0.80, p = 0.82 and p = 0.84 μm. We find that all the three curves shown in Fig. 5 have increasing nature with wavelength. Here, we have optimized the proposed structure with pitch value p = 0.80 μm. At this value, the highest birefringence value is obtained as 0.121 at wavelength 1.55 μm.

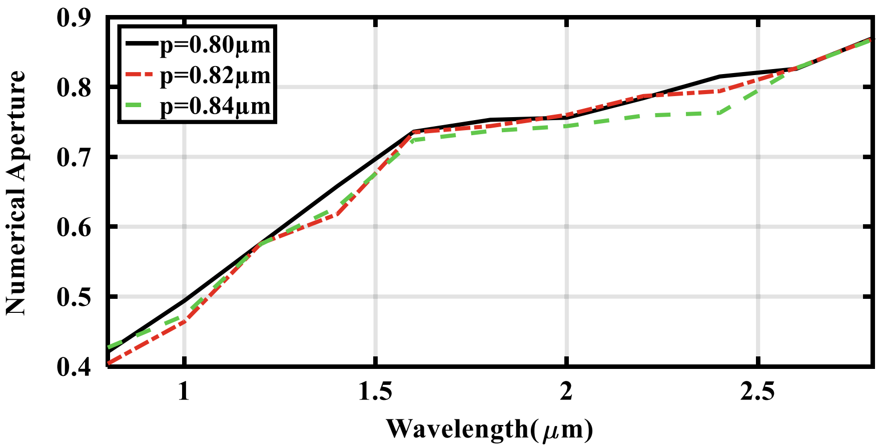


Fig. 4 Response of numerical aperture of DC-HPCF structure with wavelength, which is bored with rectangular GaP filled holes at core region

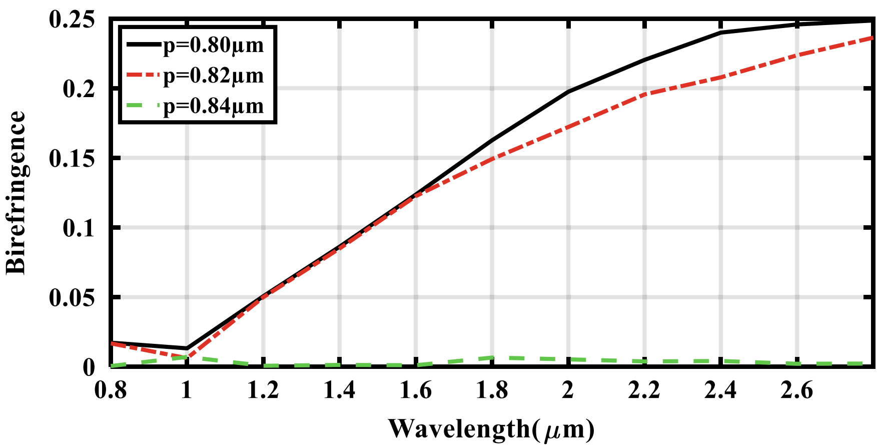


Fig. 5 Rresponse of numerical aperture of DC-HPCF structure in reference with wavelength, which is bored with rectangular GaP filled holes at core region

In Fig. 6, it is seen that the V- parameter value is greater than 2.405 for wavelength 1.55 µm while it is less than 2.405 over 1.2 to 1.55 µm band. Therefore, the proposed DC-HPCF ensures single mode operation for the entire communication window wavelength under consideration.

Figure 7 represent the effective area for three different structures of DC-HPCF with pitch values p = 0.80, p = 0.82 and p = 0.84 µm, we find that the all three are increasing in nature. From Fig. 7, we found that the proposed structure with p = 0.80 µm has the lowest effective area of 0.948(µm^2) at wavelength 1.55 µm.

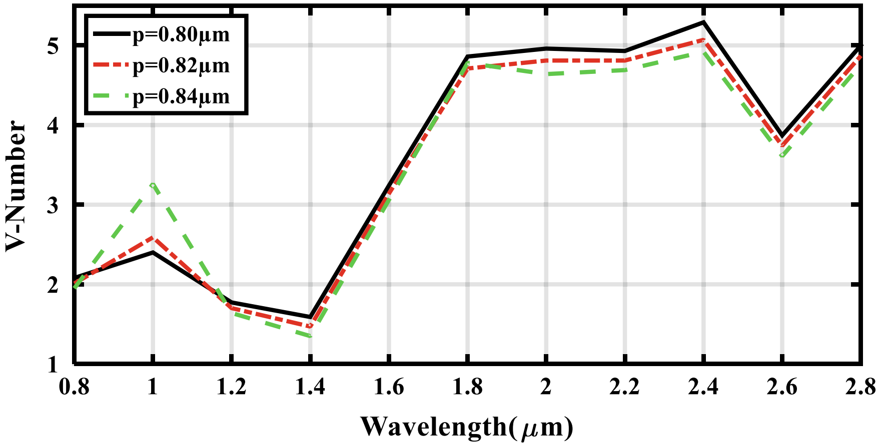


Fig. 6 Response of v-number of DC-HPCF structure reference with wavelength which is bored with rectangular GaP filled holes at core region

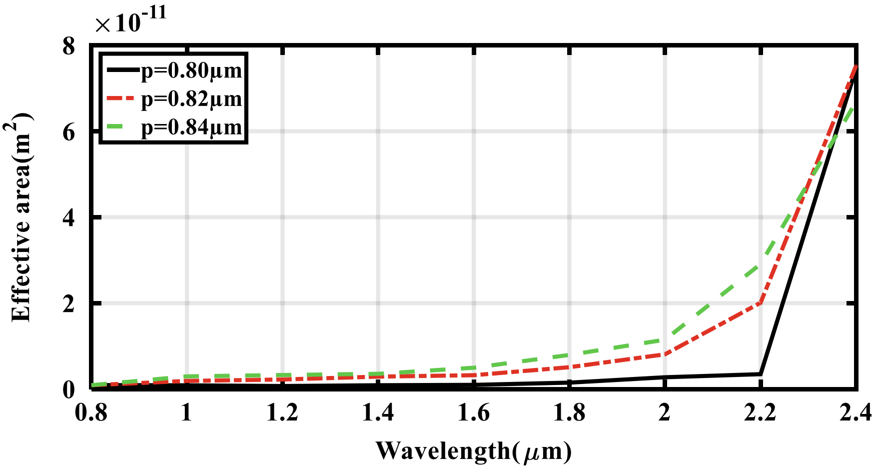


Fig. 7 Response of effective area of proposed DC-HPCF structure reference with wavelength, which is bored with rectangular GaP filled holes at core region

Figure 8 represents the nonlinearity variation for DC-HPCF structure with p = 0.80, p = 0.82 and p = 0.84 µm. In Fig. 8, we noticed that the all three curves have decreasing in nature for complete wavelength under consideration. Here, we found that the proposed structure with p = 0.80 µm has the nonlinear coefficient value of 28978 w^{-1} km^{-1} at 1.55 µm wavelength.

Figure 9 represent the effective material loss (EML) for three different structures of DC-HPCF with pitch values p = 0.80, p = 0.82 and p = 0.84 µm. It is observed that the all three curves show a slow decreasing EML response at the starting wavelength

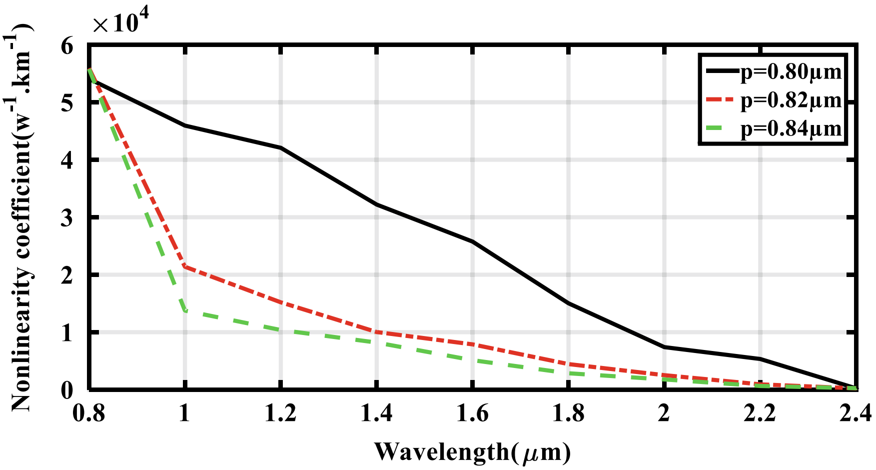


Fig. 8 Response of nonlinearity coefficient of proposed DC-HPCF structure reference with wavelength, which is bored with rectangular GaP filled holes at core region

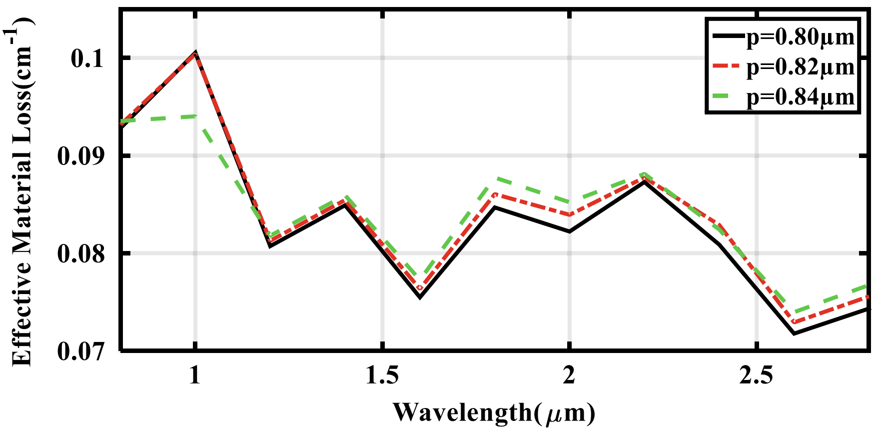


Fig. 9 Response of effective material loss of proposed DC-HPCF structure reference with wavelength, which is bored with rectangular GaP filled holes at core region

then a sudden downward curve can be observed at higher wavelength. Therefore, we can observe that the proposed structure with pitch value p = 0.80 μm has the EML value of 0.076(cm^{-1}) at wavelength 1.55 μm (Table 1).

Table 1 Comparative study of performance parameter of proposed DC-HPCF with pitch p = 1.8 μm having a rectangular GaP filled hole at the core region as shown in Fig. 1 with other PCFs at wavelength 1.55 μm

Refs.	Year	D [ps/(nm.km)]	γ ($W^{-1}Km^{-1}$)	Birefringence
Ref. [24]	2013	**−457.4**	11.26	10^{-7}
Ref. [25]	2013	**−588**	31.85	0.0181
Ref. [26]	2014	**−562.52**	130.20	0.021
Ref. [27]	2014	**−723.1**	–	0.0379
Ref. [16]	2016	**−608**	–	0.0181
Ref. [17]	2018	**−118**	0.0166	–
Ref. [28]	2019	**−722.48**	21.23	0.245
Ref. [29]	2019	**–**	7.94×10^3	–
Ref. [21]	2019	**–**	2.2×10^2	5.7×10^{-3}
Ref. [30]	2019	**–**	6.0×10^2	0.018
Proposed	**2021**	**−2885**	**28,978**	**0.121**

4 Conclusion

The high speed signal transmission with optical fiber requires a dispersion compensating (i.e. PCF with high negative dispersion) as well as polarization maintaining fibers (i.e., PCF with high nonlinearity). Therefore, taking all the above facts under consideration, it is observed that the proposed DC-HPCF whose core is bored with a rectangular shaped GaP filled hole having pitch value 0.80 μm shows the high negative dispersion and nonlinearity as 2885 ps/(nm.km) and 28,978 $w^{-1}.km^{-1}$ at 1.55 μm wavelength. Besides, the birefringence and effective area are also found good as 0.121 and 0.948 μm^2, respectively. Therefore, the proposed DC-HPCF is a suitable option for high speed optical communication and sensing purpose.

Finally, we observe that a DC-HPCF with pitch p = 0.80 μm has a drastic reduction in dispersion response. The main beauty of this proposed structure is its simple arrangement of GaP filled rectangular hole at the core region in silica atmosphere to design a highly nonlinear PCF.

References

1. Knight, J.C., Birks, T.A.: All silica single mode optical fiber with photonic crystal (1996)
2. Arif, M.F.H., Biddut, M.J.H.: Enhancement of relative sensitivity of photonic crystal fiber with high birefringence and low confinement loss. Optik Int. J. Light Electron Optics **131**, 697–704 (2017)
3. Cordeiro, C., et al.: Microstructured-core optical fiber for evanescent sensing applications. Opt. Express **14**, 13056–13066 (2006)
4. Shivam, S., Prajapati, Y.K.: Highly sensitive refractive index sensor based on D-shaped PCF with gold-graphene layers on the polished surface. Appl. Phys. A **125**, 1–7 (2019)

5. Feehan, J.S., Price, J.H.V., et al.: Efficient high-harmonic generation from a stable and compact ultrafast Yb-fiber laser producing 100 μJ, 350 fs pulses based on bendable photonic crystal fiber. Appl. Phys. B **123**, 1–12 (2017)
6. Humbert, G., et al.: Supercontinuum generation system for optical coherence tomography based on tapered photonic crystal fibre. Opt. Express **14**, 1596–1603 (2006)
7. Holzwarth, R., Udem, T., Hänsch, T.W., Knight, J., Wadsworth, W., Russell, P.S.J.: Optical frequency synthesizer for precision spectroscopy. Phys. Rev. Lett. **85**, 2264–2267 (2000)
8. Zhao, T., et al.: Highly nonlinear polarization maintaining As2Se3 based photonic quasi crystal fiber for supercontinuum generation. Opt. Mater. **73**, 343–349 (2017)
9. Shivam, S., Prajapati, Y.K.: "Dual-polarized ultrahigh sensitive Gold/MoS_2/Graphene based D-shaped PCF refractive index sensor in visible to near- IR region. Opt. Quant. Electron. **52**(17), 1–15 (2019)
10. Chibrova, A.A., et al.: The red shift of the semiconductor quantum dots luminescence maximum in the hollow core photonic crystal fibers. Opt. Mater. **73**, 423–427 (2017)
11. Knight, J.C.: Photonic crystal fiber. Nature **424**, 847–851 (2003)
12. Renversez, G., Kuhlmey, B.: Dispersion management with micro-structured optical fbers: ultra-flattened chromatic dispersion with low losses. Opt. Lett. **28**, 989–991 (2003)
13. Anas, M.T., Asaduzzaman, S., Ahmed, K., Bhuiyan, T.: Investigation of highly birefringent and highly nonlinear Hexa Sectored PCF with low confinement loss. Results Phys. **11**, 1039–1043 (2018)
14. Razzak, S.M.A., Namihira, Y.: Proposal for highly nonlinear dispersion-flattened octagonal photonic crystal fibers. IEEE Photon. Technol. Lett. **20**(4), 249–251 (2008)
15. Hasan, M.I., Mahmud, R.R., Morshed, M., Hasan, M.R.: Ultra-flattened negative dispersion for residual dispersion compensation using soft glass equiangular spiral photonic crystal fiber. Journal of Mod. Opt. **63**(17), 1681–1687 (2016)
16. Hasan, M.R., Anower, M.S., Hasan, M.I.: a polarization maintaining single-mode photonic crystal fiber for residual dispersion compensation. IEEE Photon. Technol. Lett. **28**(16), 1782–1785 (2016)
17. Prajapati, Y.K., Srivastava, V.K., Singh, V., Saini, J.P.: Effect of germanium doping on the performance of silica based photonic crystal fiber. Optik **155**, 149–156 (2018)
18. Krishna, G.D., Pillai, V.M., Gopchandran, K.G.: Design of low dispersion and low loss photonic crystal fiber: Defected core circular-octagon hybrid lattices. Opt. Fiber Technol. **51**, 17–24 (2019)
19. Brückner, V.: To the Use of Sellmeier formula. Senior Experten Service (SES) Bonn and HfT Leipzig, Germany (2011)
20. Paul, B.K., Khalek, M.A., Chakma, S., Ahmed, K.: Chalcogenide embedded quasi photonic crystal fiber for nonlinear optical applications. Ceramics Int. **44**(15), 18955–18959 (2018)
21. Arif, M.F.H., Hossain, M.M., Islam, N., Khaled, S.M.: A nonlinear photonic crystal fiber for liquid sensing application with high birefringence and low confinement loss. Sens. Bio-Sens. Res. **22**, 100252 (2019)
22. Das, S., De, M., Singh, V.K.: Single mode dispersion shifted photonic crystal fiber with liquid core for optofluidic applications. Opt. Fiber Technol. **53**, 102012 (2019)
23. Sultana, J., Islam, M.S., Faisal, M., et al.: Highly birefringent elliptical core photonic crystal fiber for terahertz application. Opt. Commun. **407**, 92–96 (2018)
24. Tee, D.C., Bakar, M.A., Tamchek, N., Adikan, F.M.: Photonic crystal fiber in photonic crystal fiber for residual dispersion compensation over E + S + C + L + U wavelength bands. IEEE Photon. J. **5**(3), 7200607 (2013)
25. Habib, M.S., Habib, M.S., Razzak, S.M.A., Hossain, M.A.: Proposal for highly bi-refringent broadband dispersion compensating octagonal photonic crystal fiber. Opt. Fiber Technol. **19**, 461–467 (2013)
26. Hasan, M.I., Razzak, S.M.A., Habib, M.S.: Design and characterization of highly birefringent residual dispersion compensating photonic crystal fiber. J. Lightw. Technol. **32**(23), 4578–4584 (2014)

27. Hasan, M.I., Habib, M.S., Habib, M.S., Razzak, S.M.A.: Highly nonlinear and highly birefringent dispersion compensating photonic crystal fiber. Opt. Fiber Technol. **20**, 32–38 (2014)
28. Prajapati, Y.K., Rahul, R., Singh, V.: Design of a photonic crystal fiber for dispersion compensation and sensing applications using modified air holes of the cladding. Braz. J. Phys. **49**(5), 745–751 (2019)
29. Monfared, Y.E., Ponomarenko, S.A.: Extremely nonlinear carbon-disulphide-filled photonic crystal fiber with controllable dispersion. Opt. Mater. **88**, 406–411 (2019)
30. Prabua, K., Malavikab, R.: Highly birefringent photonic crystal fiber with hybrid cladding. Opt. Fiber Technol. **47**, 21–26 (2019)

Survey of Supervised Machine Learning Techniques in Wireless Sensor Network

Kalyan Krishna Awasthi and Manish Kumar

Abstract Wireless Sensor Networks has fascinated a number of research scholars because of its wide applications and potential scope of improvement. In a Wireless Sensor Network (WSN), the data is sensed through the sensing nodes, processed according to some rule and then sent back to the base station. The major design concern in any WSN is the efficient utilization of the limited energy and improvement in the network lifetime. This is achieved by using a proper routing technique. A number of hierarchical routing protocols are introduced in this regard. Large Scale WSNs are finding vast applications with the recent technological advancements. Such Large scale WSNs require a lot of data computation. This data needs to be periodically sent to the base station. There is always a constraint of data rate in WSNs. The large amount of data that has to be sent at a low data rate can create an overhead at nodes which already have an energy constraint. Thus, there is a scope of introducing Machine Learning (ML) algorithms in WSNs which can make a network self-oriented. Machine Learning is a branch of Artificial Intelligence that is based on the core idea that machines can learn from the data fed to them and can take accurate and futuristic decisions. This review paper demonstrates that Supervised ML is a practical approach that can address a number of complex problems in WSNs.

Keywords Wireless sensor network · Sensor nodes · Machine learning · Supervised machine learning

1 Introduction

Wireless Sensor Network (WSN) is a collection of large number of autonomous, miniatured, low power and low-cost sensing units called sensor nodes [1, 2]. These nodes accumulate the sensed data from their environment and coordinate amongst themselves to send this sensed data to a centralized facility called base station or sink

K. K. Awasthi (✉) · M. Kumar
Department of Electronics and Communication Engineering, School of Engineering Babu, Banarasi Das University, Lucknow, India
e-mail: awasthi.kalyan@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_18

for further computation [3, 4]. A variety of sensors like thermal, chemical, acoustic etc. can be used for the purpose of sensing the data. The WSN for the same reason has vast applications [5]. A number of challenges in terms of network security threat, network architecture, deployment, data collection and network coverage are a major threat [6]. Some applications of WSN have to cater considerably very large number of sensor nodes. To manage such large WSNs we require scalable and efficient algorithms. The highly dynamic nature of WSNs may further require redesigning of the network. The conventional approaches for WSN does not adapt with the dynamic environment requirements. This creates a need of a self-organized and adaptable approach.

Machine Learning (ML) is a method that automatically lends itself to the changing environment [7, 8]. ML is a technique that learns from its past experiences and acts with almost nil programming requirement [9]. ML was first introduced as a technique for artificial intelligence [10]. Over the years, ML evolved and proved to be a computationally viable technique for WSNs [11]. ML algorithms are iteration dependent which helps them to attain experience-based knowledge that ultimately gives better results. The application of ML techniques in WSN not only limits the human intervention in the network but also improves the performance of the network in terms of energy efficiency. The common ML algorithms used in WSNs are fuzzy logic [12], neural network [13], reinforcement learning [14], evolutionary algorithm [15] and swarm intelligence [16]. Machine Learning can be supervised, unsupervised, semi-supervised or reinforcement learning. The detailed classification of ML algorithms used in WSN is discussed in Sect. 2. This paper highlights the supervised ML techniques. The incorporation of these algorithms has addressed the WSN issues to a great extent [17]. The important abbreviations used in the paper are listed in Table 1.

The rest of the paper is organized as follows: Necessary Background of ML in WSNs is discussed in Sect. 2. The role of Supervised ML in WSN is summarized in Sect. 3. The Research directions and challenges are discussed in Sect. 4. The future Scope and Conclusion are discussed in Sect. 5.

Table 1 List of abbreviations used

Acronym	Description
WSN	Wireless Sensor Network
ML	Machine Learning
ANN	Artificial Neural Network
SVM	State Vector Machine
SVD	Singular Value Decomposition
PCA	Principle Component Analysis
ICA	Independent Component Analysis
DT	Decision Tree
RF	Random Forest
k-NN	K Nearest Neighbor
DL	Deep Learning

2 Machine Learning in Wireless Sensor Networks

Machine Learning since its conception has been conceived as a combination of tools and algorithms to generate a predictive model. However, in current scenario ML is much more than just a predictive tool because of the collection of extensive patterns. These patterns can be very helpful for the researchers who wish to apply ML in WSN applications.ML techniques, based on the learning style, can be cdivided into supervised learning, unsupervised learning, semi-supervised learning and reinforcement learning [18]. The taxonomy of machine learning is shown in Fig. 1.

2.1 *Supervised Learning*

In Supervised learning a model is developed on the basis of some input and output set of values. These values are predefined and based on these values (datasets with labels) the system is trained to find a relation between them. At the end of the training a function is mapped between input and output values. Supervised Learning addressed various WSN challenges like localization, coverage problem, fault detection, routing, energy harvesting [19]. Supervised Learning is further divided into classification and regression. Classification can be logic based (decision tree and random forest), perception based (ANN and deep learning), statistical learning based (Bayesian and SVM) or instance based (k-NN) algorithms.

Regression is the simplest supervised learning approach in which certain values are predicted based on a given set of values. This method gives minimum error. The mathematical model for regression is explained in [20]. The WSN problems addressed through regression are localization [21], connectivity problem [22], data aggregation [23] and energy harvesting [24]. A simple regression model is depicted in Fig. 2. In the mentioned model the input is ‘x’ which is an independent variable

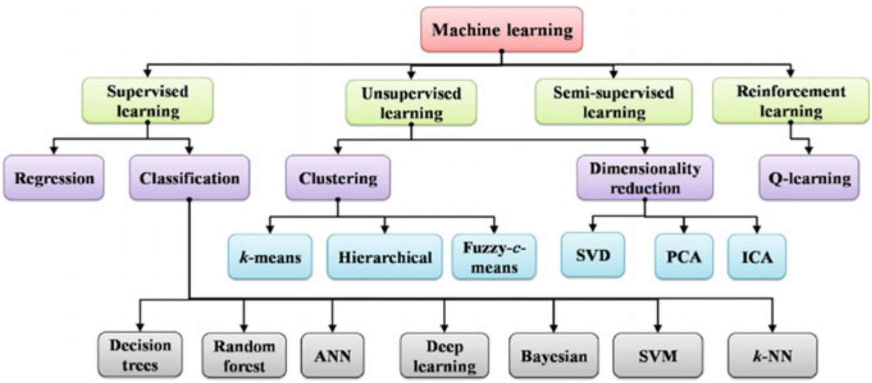


Fig. 1 Machine learning taxonomy

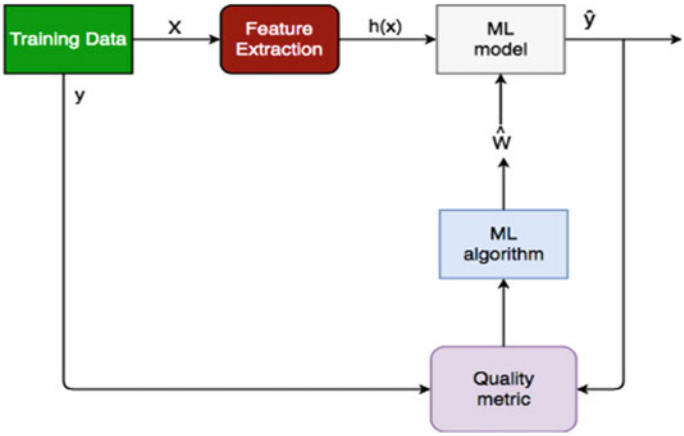


Fig. 2 Simple linear regression model

and the output is ‘y’ which is a dependent variable. ‘w’ here is functions that relates ‘y’ to ‘x’.

In order to improve the reliability, a supervised learning approach called **Decision Tree (DT)** is developed on the basis of if–then rules. The constituent nodes of a decision tree are leaf nodes (for final result) and decision nodes (for selecting available choices). A simple structure of DT is discussed in [25]. The various issues of WSN resolved by DT method are connectivity [26], anomaly detection [27], data aggregation [28] and mobile sink path selection [29]. A general decision tree is depicted in Fig. 3.

Random Forest (RF) algorithm is another supervised ML technique in which we take a collection of trees**.** Every tree in the forest tells us something about the basic classification. The RF algorithm functions in two basic stages. In first stage we create a random forest classifier and in the second stage the results are predicted [30]. RF approach is advantageous when we are dealing with large datasets in a heterogenous environment. The algorithm randomly selects a subset from the training sample and separates the variables at each node of the tree. This generates a very large number of decision trees. These algorithms can readily address the coverage [31] and MAC protocol [32] issues.

Artificial Neural Networks (ANN) technique exploits the human neuron-based model for classification of data [33]**.** ANN can be conceptualized as a combination of large number of processing units (like neurons in a human brain) that process the information to give accurate results. The basic ANN layered architecture is depicted in Fig. 4. Every ANN mainly consists of three layers- input layer, hidden layers (can be one or more) and output layers. The main advantage of ANN is its ability to classify the complex computational datasets easily. There is no restriction of the data input in ANN which makes it superior to other supervised ML techniques. However, the major issue with ANN is the higher computational requirement which in turn can be a drawback in terms of efficient energy management. The various issues of WSN

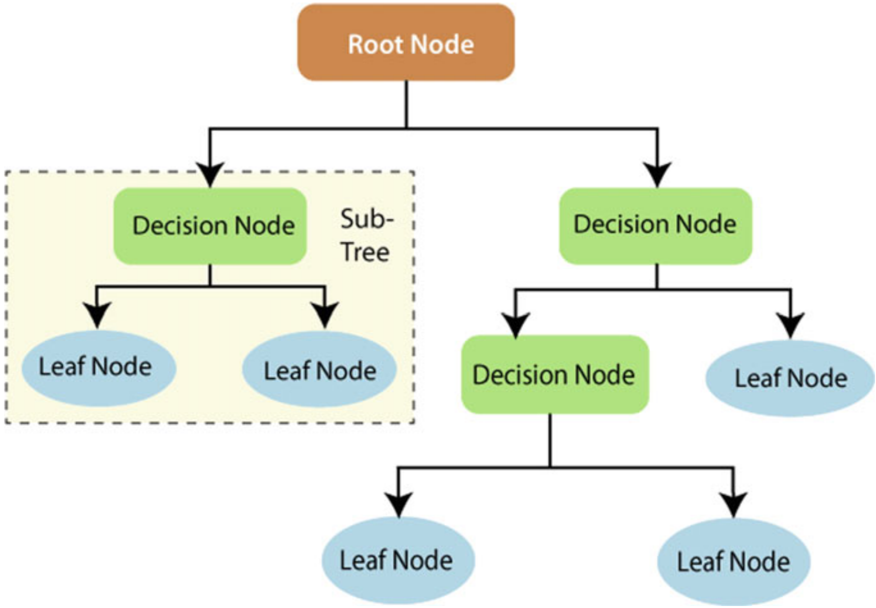
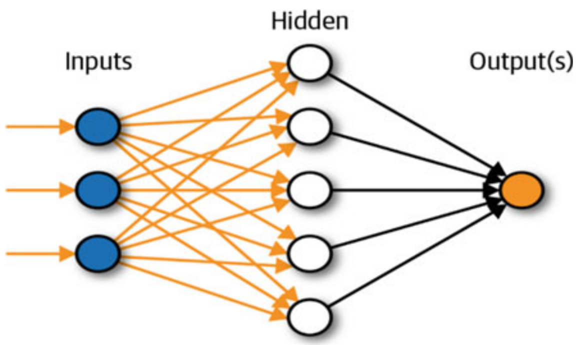


Fig. 3 Graphical representation of a decision tree

addressed through ANN are localization [34], fault detection [35], routing efficiency [36], data aggregation [37] and congestion control [38].

Deep Learning (DL) is a subclass of ANN**.** In deep learning a multiple layered representation between the input and output is considered. These multiple level representations are obtained through simple yet non-linear modules. Each module transforms the representation from one level into a more abstract higher level. When a number of such modules are generated, complex functions can be decoded and learnt which enables a researcher to address multiple objectives [39]. The WSN

Fig. 4 Layered architecture of ANN



issues addressed through ANN are routing [36], data quality estimation [40], energy harvesting [41] and fault detection [42].

Bayesian Learning approach makes use of probabilistic models where the entire analysis is carried out through Bayesian inference [43]. The core concept of Bayesian Model is based on Bayes theorem. Equation 1 represents the Bayes theorem.

$$P(A/B) = P(B/A).\,P(A)/P(B) \tag{1}$$

P(A/B) is the probability of occurrence of A when B has already occurred and it is called posterior probability that we wish to calculate. P(A) is the prior probability and P(B/A) is the probability of occurrence of B when A has already occurred. Bayesian approach is a specialized supervised ML technique which is quite handy in solving the regression and classification problems. This approach required relatively less training samples. An application of Bayesian learning is a statistical learning algorithm names Gaussian Process regression model [44].

Support Vector Machine (SVM) is another supervised ML approach which is used to predict a continuous response. SVM uses hyperplane and coordinate individual observation [45]. If we are able to identify the boundary conditions for a problem, most of the data becomes redundant as we have to focus on a specific set of data only. This approach uses support vectors to identify the boundary points. This approach gives the best classification for a dataset. This method is suitable for high dimensional data where we have a very large number of variables present for the detection purpose. SVM in WSNs can be used as a regression and sensor data forecasting [46]. An example of SVM is depicted in Fig. 5.

k-Nearest Neighbor (k-NN) is one of the simplest supervised machine learning techniques. In this algorithm we find the similarities between the new data and the existing data. Based on the similarity quotient the new data is assigned to either of the pre-existing data sets. It is a non-parametric algorithm often referred to as lazy learner algorithm as it does not perform the learning task immediately rather it stores

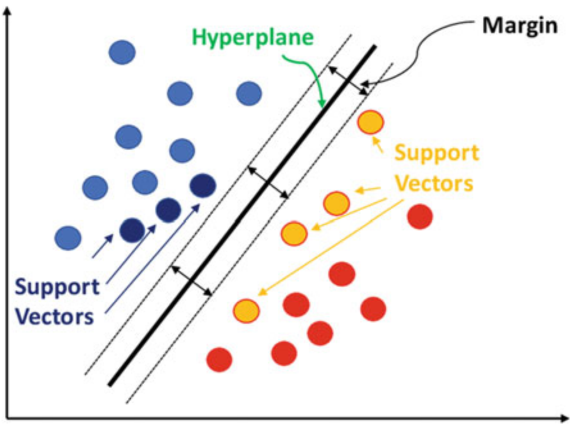


Fig. 5 An example of SVM

the data and at the time of classification it performs the necessary action on the data sets. The classification is mainly done on the basis of training samples and the test sample distances. The basic concept and application of k-NN in ML is explained well in [47].

The comparison of different supervised learning ML techniques on the basis of important WSN parameters is shown in Table 1 [48].

2.2 Unsupervised Learning

In unsupervised learning output and input are non-associated. This approach makes use of unlabeled datasets without having an output vector. The similarity between the groups is investigated and based on those similarities the classification of the samples into different sets is done. The major application of this technique is in data aggregation problems and node clustering [49, 50]. The unsupervised ML is further classified into clustering (hierarchical, k-means and fuzzy-c- means) and dimensionality reduction (Principle component analysis, independent component analysis and singular value decomposition). A detailed analysis of these methods is discussed in [48].

2.3 Semi-supervised Learning

Most practical applications require a blend of labeled and unlabeled data. For labeled data we have supervised learning and for unlabeled data we have unsupervised learning. When a combination of these two is required, we use semi-supervised learning. This method generally uses a very small labeled data and a large unlabeled data. In first step the similar data is clustered using unsupervised ML approach and then the existing labeled data is used to label the unlabeled data. These objectives of a semi-supervised ML technique can be sub-categorized as -*transductive learning* and *inductive learning*. The basic problems of WSN addressed through semi-supervised learning are localization [51] and fault detection [52].

2.4 Reinforcement Learning

In reinforcement learning [53] a sensor node interacts within its environment and gathers information for future actions. The optimal result is derived from the interacting environment over a long-term interaction combined with the sensors own experience. The functionality of reinforcement learning is depicted in Fig. 6. Few of the important reinforcement techniques are Q-learning [54], SARSA Algorithms, Temporal Differences and Deep Q Networks.

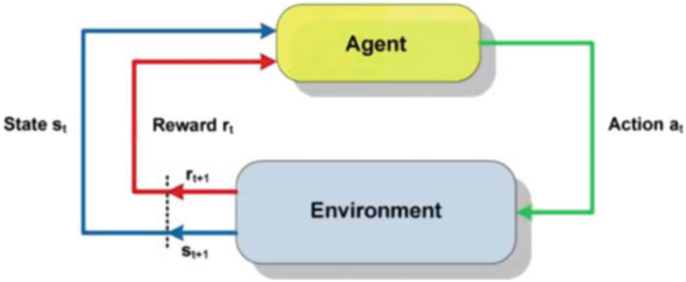


Fig. 6 Functional diagram of reinforcement learning

2.5 Evolutionary Computation

Evolutionary computational model is an advanced problem-solving technique which incorporates the models that are related to natural and biological evolution. This can be considered as a special class of artificial intelligence. In this method we first develop a set of solutions and then recursively use it through iteration [55]. This method is basically used for solving the optimization problem. The major WSN issues handled through this approach are localization, coverage, routing, target detection and mobile sink. A survey of different evolutionary approaches followed for various WSN applications is given in [56].

3 Role of Supervised Machine Learning in Wireless Sensor Networks

The ML techniques used for resolving WSN problems lend a solution mainly through Reinforcement learning, supervised learning and unsupervised learning. Of all these three methods the supervised ML technique is most widely used. In WSN the main focus is on sensing the data and ensuring that the sensed data is successfully dumped to base station or sink node. If a data is traversed through an inappropriate path the decrease in the energy quotient of the nodes may ultimately lead to network failure. So, it is very important to ensure that the correct functionality of a node is defined and the sensing and data processing is then carried with appropriate techniques. Supervised ML follows the basic rule that the machine automatically trains itself according to the reference data (datasets) provided to it to give accurate results. This inherent property of supervised ML makes it more appropriate for resolving WSN issues. In order to understand the role of Supervised Machine Learning in solving the WSN problem, it is very important to understand the research issues related to WSN and the potential WSN threats [6, 19]. The major issues that are addressed through supervised ML are Node localization, Node coverage and Connectivity, Routing Layer Issues, MAC Layer Issues, Event Monitoring and Data Aggregation issues, Energy harvesting and Node Query Procession. The solution of these problems provided by supervised ML approach are already mentioned in Sect. 2 with every individual supervised ML method. Other ML approaches like Reinforcement, Q-learning, k-means have also attracted the researchers to a great extent but this paper mainly focusses on the supervised learning approach. A summary of the issues related to WSN along with their solutions/remarks through supervised learning is given in Table 2 [48, 57] (Table 3).

The major ML algorithms used in WSN issue resolution are depicted in Fig. 7.

A summarized study shows that supervised ML techniques are the best and most suited for addressing WSN issuesQuery. The same can be conclusively verified through Fig. 8 [48].

Table 2 Comparison of supervised ML techniques based on important WSN parameters

WSN parameters	DT	RF	ANN	DL	Bayesian	SVM	k-NN
Speed classification	A	A	A	A	A	A	D
Learning speed	B	C	D	D	A	D	A
Accuracy	C	C	B	B	D	A	C
Parameter handling	B	B	D	C	A	D	B

The performance parameters are: A = Best Performance, B = Good Performance, C = Satisfactory Performance, D = Poor Performance

Table 3 WSN issues with supervised ML-based solutions

WSN issues	Supervised ML technique	Solutions/Remarks
Node localization	k-NN	Distance Estimation in efficient manner
Node coverage & connectivity	Decision Tree, ANN	Segregation of Connected and failed node in a network
Routing layer issues	Decision Tree, Random Forest	Optimal routing path prediction
MAC layer issues	SVM, Decision Tree, ANN	Efficient channel assignment
Event monitoring & target detection	Deep Learning, Bayesian Learning	Efficient monitoring and multiple target detection
Energy harvesting	SVM, Deep Learning	Energy management for improving network lifetime
Node query processing	k-NN	Handshake for data transfer and Node Beacon Sending

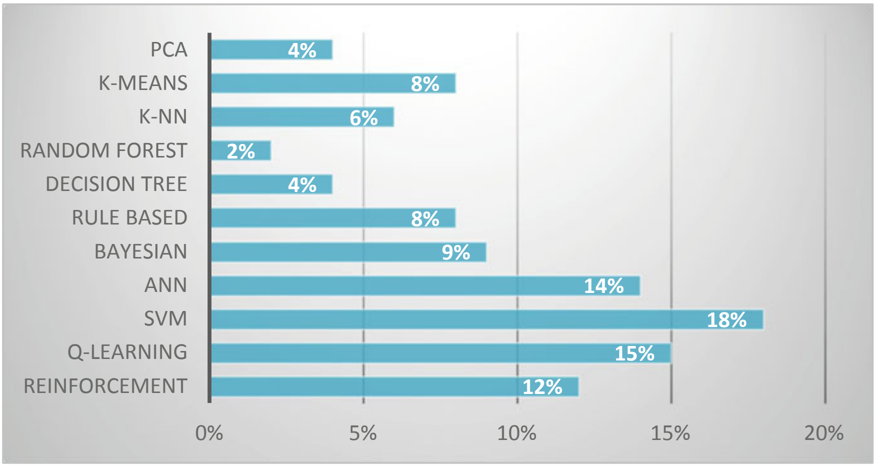


Fig. 7 Chart depicting the use of ML algorithms in WSN

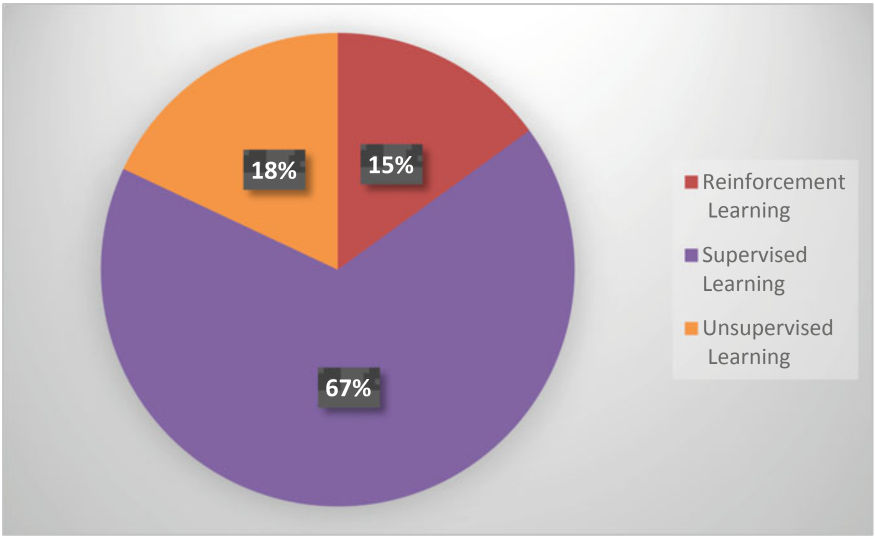


Fig. 8 Classification of ML algorithms for WSN

4 Challenges and Research Directions

There are various issues that can still be addressed through supervised ML techniques. The major issues are listed below:

4.1. **WSN Anomaly Detection:** Any anomaly in a sensor network can lead to transmission delay, communication overhead etc. An anomaly is further an application driven issue that can have varying effects on the WSN lifetime. Use of supervised ML techniques to cater a specific anomaly and deal with it has a great research scope.

4.2. **WSN Routing:** In a more practical and realistic application of WSN there may be a requirement of collecting the data from multiple sources and routing them to multiple destinations. This obviously has a very high probability of packet collisions. There is indeed a requirement of cooperative WSN routing protocol that can avoid collision. Further a more dynamic routing protocol can be designed for the networks where the nodes change their positions.

4.3. **WSN Coverage and Connectivity:** It is always desirable to have minimum number of nodes covering the maximum region. The random deployment of nodes which is preferrable done in WSN environment sometimes create a coverage hole. If through supervised ML approaches such situations can be predicted then it can be very helpful to improve the network lifetime.

4.4. **WSN Data Aggregation:** In certain applications of WSN like a smart city application, a number of sensor nodes sense the same data and send to the sink. At gateways, it becomes very difficult to handle this data. So, there is a requirement of summarizing the useful data from multiple sources. The aggregation of data further removes the data redundancy which in turn improves the network lifetime. A more advanced data aggregation-based ML approach may be developed.

4.5. **WSN Energy Harvesting:** For a large scale WSN, energy is a major constraint. Energy harvesting means to derive energy from alternate sources (environmental) and use it for the network. Most of the WSN protocols used today are energy efficient, they can be further revisited through supervised ML techniques to improve or enhance their performance in terms of energy.

4.6. **WSN Node Deployment:** In most of the WSN applications the nodes are structured in a three-dimensional space. The random structuring of nodes in three-dimension cannot give good results with the existing two-dimensional based protocols. Further, path learning approach can be very handy for efficient path planning. This has a good research scope.

4.7. **QoS in WSN:** The QoS standards in WSN are application based. They must meet the requirement of the user and application. It is very difficult to define the QoS for a specific requirement and application. Supervised ML techniques may be used to design cross-layer protocols that can very well address this issue specially in a heterogenous WSN environment.

4.8. **WSN Event Monitoring ad Target Detection:** The recent smart city application of WSN relies heavily on event monitoring and target detection [57]. The major issues of concern are intrusion detection and traffic monitoring. Supervised ML techniques like Bayesian method can be very handy in this regard.

4.9. **WSN Congestion Control and Avoidance:** The conventional congestion control and avoidance protocols for WSN have a limitation in case of a dynamically changing environment. Congestion control techniques should be self-adapting to respond to these dynamic changes. Supervised ML techniques can be used to develop enhanced protocol that can collect the sensed data directly.

5 Conclusion and Future Scope

In this review paper we have highlighted the basic ML techniques that can be used for addressing various issues in WSN. The paper justifies that ML algorithms are versatile and a single ML algorithm can address multiple issues in a WSN. The supervised ML algorithms are more frequently used in WSNs because of their natural ability to adapt in dynamic environment and relatively easy self-learning capability. We have also discussed the challenges and open research issues in WSN that can be addressed through supervised ML algorithms. The supervised ML techniques can very well be the solution to these problems. This may lead to a significant use of supervised ML techniques in recent advanced applications. With the Industry 4.0 [58] flourishing at a rapid pace, supervised ML techniques have immense scope in this regard.

References

1. Singh, P., Gupta, O.P., Saini, S.: Adv. Comput. Sci. Technol. **10**(5), 733–739 (2017). ISSN 0973–6107
2. Rawat, P., Singh, K.D., Chaouchi, H., Bonnin, J.M.: Wireless sensor networks: a survey on recent developments and potential synergies. J. Supercomput. **68**(1), 1–48 (2014)
3. Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E.: Wireless sensor networks: a survey. Comput. Netw. **38**(4), 393–422 (2002)
4. Yick, J., Mukherjee, B., Ghosal, D.: Wireless sensor network survey. Comput. Netw. **52**(12), 2292–2330 (2008)
5. Wang, X., Ding, L., Wang, S.: Trust evaluation sensing for wireless sensor networks. IEEE Trans. Instrum. Meas. **60**(6), 2088–2095 (2011)
6. Pinar, Y., Zuhair, A., Hamad, A., Resit, A., Shiva, K., Omar, A.: Wireless Sensor Networks (WSNs). In: 2016 IEEE Long Island Systems, Applications and Technology Conference (LISAT), pp. 1–8 (2016). https://doi.org/10.1109/LISAT.2016.7494144
7. Duffy, A.H.: The "what" and "how" of learning in design. IEEE Expert **12**(3), 71–76 (1997)
8. Langley, P., Simon, H.A.: Applications of machine learning and rule induction. Commun. ACM **38**(11), 54–64 (1995)
9. Sarker, I.H.: Machine learning: algorithms, real-world applications and research directions. SN Comput. Sci. **2**, 160 (2021)
10. Ayodele, T.O.: Introduction to machine learning. In: New Advances in Machine Learning. InTech (2010)
11. Alsheikh, M.A., Lin, S., Niyato, D., Tan, H.-P.: Machine learning in wireless sensor networks: algorithms, strategies, and applications. IEEE Commun. Surv. Tutor. **16**(4), 1996–2018 (2014)
12. Maksimovic, M., Vujovic, V., Milosevi, V.: Fuzzy logic and wireless sensor networks – a survey. J. Intell. Fuzzy Syst. **27**, 877–890 (2014). https://doi.org/10.3233/IFS-131046

13. Aliyu, F., Umar, S., Al-Duwaish, H.: A survey of applications of artificial neural networks in wireless sensor networks. In: 2019 8th International Conference on Modeling Simulation and Applied Optimization (ICMSAO), pp. 1–5 (2019). https://doi.org/10.1109/ICMSAO.2019.8880364
14. Arya, A., Ram, D., Malik, A., Garg, R.: Reinforcement learning based routing protocols in WSNs: a survey. Int. J. Res. Appl. Sci. Eng. Technol. **6**, 3523–3529 (2018)
15. Nan, G., Li, M.: Evolutionary based approaches in wireless sensor networks: a survey. In: 2008 Fourth International Conference on Natural Computation, pp. 217–222 (2008). https://doi.org/10.1109/ICNC.2008.98
16. Cao, L., Cai, Y., Yue, Y.: Swarm intelligence-based performance optimization for mobile wireless sensor networks: survey, challenges, and future directions. IEEE Access **7**, 161524–161553 (2019). https://doi.org/10.1109/ACCESS.2019.2951370
17. Kulkarni, R.V., Förster, A., Venayagamoorthy, G.: Computational intelligence in wireless sensor networks: a survey. IEEE Commun. Surv. Tutor. **13**(1), 68–96 (2011)
18. Abu-Mostafa, Y.S., Magdon-Ismail, M., Lin, H.-T.: Learning From Data. AMLBook (2012)
19. Das, S., Abraham, A., Panigrahi, B.K.: Computational Intelligence: Foundations, Perspectives, and Recent Trends, pp. 1–37. John Wiley & Sons Inc., Hoboken (2010)
20. Montgomery, D.C., Peck, E.A., Vining, G.G.: Introduction to Linear Regression Analysis, vol. 821. John Wiley & Sons, Hoboken (2012)
21. Zhao, W., Su, S., Shao, F.: Improved DV-hop algorithm using locally weighted linear regression in anisotropic wireless sensor networks. Wirel. Pers. Commun. **98**(4), 3335–3353 (2018)
22. Chang, X., Huang, J., Liu, S., Xing, G., Zhang, H., Wang, J., Huang, L., Zhuang, Y.: Accuracy-aware interference modeling and measurement in wireless sensor networks. IEEE Trans. Mob. Comput. **15**(2), 278–291 (2016)
23. Gispan, L., Leshem, A., Be'ery, Y.: Decentralized estimation of regression coefficients in sensor networks. Dig. Signal Process. **68**, 16–23 (2017)
24. Tan, W.M., Sullivan, P., Watson, H., Slota-Newson, J., Jarvis, S.A.: An indoor test methodology for solar-powered wireless sensor networks. ACM Trans. Embedded Comput. Syst. **16**(3), 82.1-82.25 (2017)
25. Quinlan, J.R.: Induction of decision trees. Mach. Learn. **1**(1), 81–106 (1986)
26. Shu, J., Liu, S., Liu, L., Zhan, L., Hu, G.: Research on link quality estimation mechanism for wireless sensor networks based on support vector machine. Chin. J. Electron. **26**(2), 377–384 (2017)
27. Garofalo, A., Di Sarno, C., Formicola, V.: Enhancing intrusion detection in wireless sensor networks through decision trees. In: Dependable Computing, pp. 1–15 (2013)
28. He, H., Zhu, Z., Mäkinen, E.: Task-oriented distributed data fusion in autonomous wireless sensor networks. Soft Comput. **19**(8), 2305–2319 (2015)
29. Kim, S., Kim, D.-Y.: Efficient data-forwarding method in delay-tolerant P2P networking for IoT services. Peer-to-Peer Netw. Appl. **11**, 1176–1185 (2017)
30. Breiman, L.: Random forests. Mach. Learn. **45**(1), 5–32 (2001)
31. Elghazel, W., et al.: Random forests for industrial device functioning diagnostics using wireless sensor networks. In: Aerospace Conference, pp. 1–9. IEEE (2015)
32. Alotaibi, B., Elleithy, K.: A new MAC address spoofing detection technique based on random forests. Sensors **16**(3), 1–14 (2016)
33. White, H.: Learning in artificial neural networks: a statistical perspective. Neural Comput. **1**(4), 425–464 (1989)
34. Banihashemian, S.S., Adibnia, F., Sarram, M.A.: A new range-free and storage-efficient localization algorithm using neural networks in wireless sensor networks. Wirel. Pers. Commun. **98**(1), 1547–1568 (2018)
35. Chanak, P., Banerjee, I.: Fuzzy rule-based faulty node classification and management scheme for large scale wireless sensor networks. Expert Syst. Appl. **45**, 307–321 (2016)
36. Lee, Y.: Classification of node degree based on deep learning and routing method applied for virtual route assignment. Ad Hoc Netw. **58**, 70–85 (2017)

37. Habib, C., Makhoul, A., Darazi, R., Salim, C.: Self-adaptive data collection and fusion for health monitoring based on body sensor networks. IEEE Trans. Ind. Inf. **12**(6), 2342–2352 (2016)
38. Rezaee, A.A., Pasandideh, F.: A fuzzy congestion control protocol based on active queue management in wireless sensor networks with medical applications. Wirel Pers. Commun. **98**(1), 815–842 (2018)
39. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436–444 (2015)
40. Wang, Y., Yang, A., Chen, X., Wang, P., Wang, Y., Yang, H.: A deep learning approach for blind drift calibration of sensor networks. IEEE Sens. J. **17**(13), 4158–4171 (2017)
41. Chen, F., Fu, Z., Yang, Z.: Wind power generation fault diagnosis based on deep learning model in internet of things (IoT) with clusters. Cluster Comput. **22**, 1–13 (2019)
42. Li, C., Xie, X., Huang, Y., Wang, H., Niu, C.: Distributed data mining based on deep neural network for wireless sensor network. Int. J. Distrib. Sens. Netw. **11**(7), 1–7 (2015)
43. Barber, D.: Bayesian Reasoning and Machine Learning. Cambridge University Press, Cambridge (2012)
44. Rasmussen, C.E.: Gaussian processes for machine learning. In: Adaptive Computation and Machine Learning. Citeseer (2006)
45. Vapnik, V.N.: An overview of statistical learning theory. IEEE Trans. Neural Netw. **10**(5), 988–999 (1999)
46. Ruß, G.: Data mining of agricultural yield data: a comparison of regression models. In: Industrial Conference on Data Mining, pp. 24–37. Springer, Heidelberg (2009)
47. Wang, L.: Research and implementation of machine learning classifier based on KNN. In: IOP Conference Series: Materials Science and Engineering, vol. 677, p. 052038 (2019). https://doi.org/10.1088/1757-899X/677/5/052038
48. Kumar, D.P., Amgoth, T., Annavarapu, C.S.R.: Machine learning algorithms for wireless sensor networks: a survey. Inf. Fusion **49**, 1–25 (2019)
49. Lee, S., Chung, T.: Data aggregation for wireless sensor networks using self-organizing map. In: Artificial Intelligence and Simulation, Series Lecture Notes in Computer Science, vol. 3397, pp. 508–517. Springer, Heidelberg (2005)
50. Macua, S., Belanovic, P., Zazo, S.: Consensus-based distributed principal component analysis in wireless sensor networks. In: 11th International Workshop on Signal Processing Advances in Wireless Communications, pp. 1–5 (2010)
51. Yoo, J., Kim, H.J.: Target localization in wireless sensor networks using online semi-supervised support vector regression. Sensors **15**(6), 12539–12559 (2015)
52. Zhao, M., Chow, T.W.: Wireless sensor network fault detection via semi-supervised local kernel density estimation. In: 2015 IEEE International Conference on Industrial Technology (ICIT), pp. 1495–1500. IEEE (2015)
53. Kaelbling, L.P., Littman, M.L., Moore, A.W.: Reinforcement learning: a survey. J. Artif. Intell. Res. **4**, 237–285 (1996)
54. Watkins, C.J., Dayan, P.: Q-learning. Mach. Learn. **8**(3–4), 279–292 (1992)
55. Hashim, H.A., Ayinde, B.O., Abido, M.A.: Optimal placement of relay nodes in wireless sensor network using artificial bee colony algorithm. J. Netw. Comput. Appl. **64**, 239–248 (2016)
56. Kulkarni, R.V., Forster, A., Venayagamoorthy, G.K.: Computational intelligence in wireless sensor networks: a survey. IEEE Commun. Surv. Tutor. **13**(1), 68–96 (2011)
57. Sharma, H., Haque, A., Blaabjerg, F.: Machine learning in wireless sensor networks for smart cities: a survey. Electronics **10**, 1012 (2021). https://doi.org/10.3390/electronics10091012
58. Kumar, S., Narkhede, B., Jain, K.: Industry 4.0: literature review and future research directions (2018)

Performance Analysis of Hetero-Dielectric Stacked Buried Oxide on Modified Source-Drain FDSOI MOS Transistor

Sudhir Rai and R. K. Chauhan

Abstract In this paper, a proposed structure of hetero-dielectric stacked buried oxide on the modified source-drain fully depleted silicon on insulator (HB MS-MD FDSOI). It presents improved electrical characteristics over conventional MS-MD FDSOI. The buried oxide (BOX) layer of this structure uses the combination of high-k dielectric (Si_3N_4_HfO_2) to minimize the value of OFF-state current, subthreshold slope, drain induced barrier lowering, and the highest value of ON to OFF current ratio. The impact of buried oxide (Si_3N_4_HfO_2) represents the improvement by 13.77% and 24.63%in subthreshold slope, DIBL, and static power dissipation respectively as compared to conventional MS-MD FDSOI. The proposed structure is simulated using ATLAS 2-D device simulator.

Keywords FDSOI · ON current to OFF current ratio · Short channel effects (SCEs)

1 Introduction

Continual scaling down of MOS transistors have invited a lot of problems like short channel effects (SCEs), gate induced drain leakage (GIDL), drain induced barrier lowering (DIBL) hot electron effects, threshold voltage roll-off, OFF-state leakage, [1]. These effects cause to weaken the control of the gate over the channel region. The subthreshold slope is raised at room temperature because of the scaling effect and it can lead to reduced threshold voltage [2, 3].

SOI-based CMOS technology is effective in reducing SCEs and OFF-state leakage [4]. The fully depleted silicon on insulator (FDSOI) device is widely used in the low-power application. It has been studied in detail by Cheng and Khakifirooz et al. [5]. The reliability of the FDSOI MOSFET often declines due to short-channel effects the device sizes are scaled down to the sub-100-nm region. The challenges associated with modern SOI-based CMOS Technology limits its electricalefficiency [6]. The subthreshold slope can be improved in the case of the long channel SOI MOSFET

S. Rai (✉) · R. K. Chauhan
Madan Mohan Malaviya University of Technology, Gorakhpur, India
e-mail: sdece@mmmut.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_19

by increasing the thickness of the buried oxide (BOX) [7]. On the other side, in a shorter gate length SOI MOSFET, because of the electric field distribution into the buried oxide, a thicker BOX creates a greater DIBL. As a consequence, with a rise in the thickness of the BOX, the subthreshold slope worsens [8].

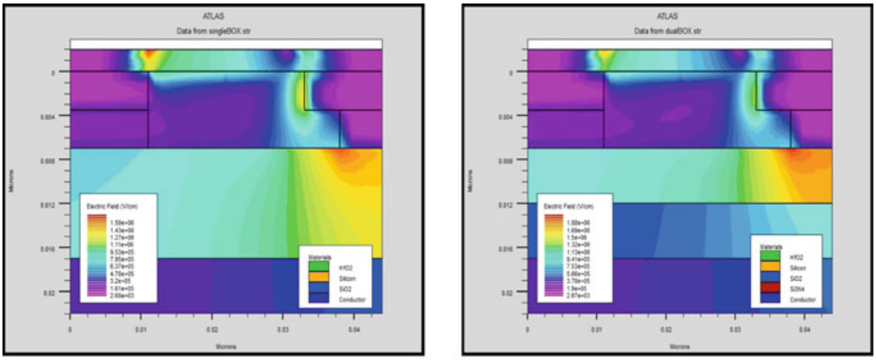
Most of the researchers has studied to boost the electrical performance of FDSOI MOS transistor [9]. The influence of the floating body in FDSOI MOSFET [10] is discussed by Tseng et al., Ohtou et al. and Chouksey et al., have researched to maximize the FD-SOI MOSFET's electrical efficiency at a very thin BOX thickness [11, 12]. The introduction of the ground plane in the BOX results in reduced DIBL effects. Upon increasing BOX thickness, BOX capacitance decreases, thus improving subthreshold slope [13]. All these works involved engineering on single layer BOX. But the impact of Hetro dielectric stacked buried oxide in FDSOI MOSFET M. Lebby presented a multilayered Buried oxide structure fabrication methodology and suggested that this structure can be effective in curbing leakage problems and short channel effects [14].

In this paper, a proposed structure using hetero-dielectric stacked buried oxide on the modified source-drain fully depleted silicon on insulator (HB MS-MD FDSOI) transistor is proposed. It presents improved electrical characteristics over conventional MS-MD FDSOI. This structure used Si_3N_4_HfO_2 combination to minimize the value of OFF-state current, subthreshold slope, drain induced barrier lowering, and the highest value of ON to OFF current ratio. The Silvaco Atlas 2D device simulator was used to carry out all the designs and simulations [15].

2 Device Structure and Specification

The conventional and proposed hetero-dielectric stacked buried oxide (HB MS-MD FDSOI) structure is shown in Fig. 1(a) and (b), respectively. The source region is split vertically into two areas where the upper part is doped with a high concentration value while the bottom region is doped with a low concentration value [3]. In the proposed structure, the buried oxide layer is divided into two segments, the BOX includes a bottom layer (layer 1) of material with a high dielectric constant and thickness (= 5 nm) equal to the top layer (layer 2) of material with low dielectric constant. For a better understanding of the effect of the hetero-dielectric BOX structure, different hetero-dielectric material combinations have been considered in Table 2, for which equivalent dielectric constant will be given by Eq. 1, where $\in_1$ and $\in_2$ are dielectric constants of hetero dielectric BOX($\in_{eq}$) materials. The device specifications of conventional and proposed structures have been tabulated in Table 1. The physical models area drift–diffusion model, Shockley-Reed-Hall,conmob, fldmob and bgn, used to simulate the electrical characteristics.

$$\in_{eq} = \frac{\in_1 \in_2}{5(\in_1 + \in_2)} \tag{1}$$



(a) (b)

Fig. 1 The **a** Conventional and **b** Proposed hetero-dielectric stacked buried oxide (HB MS-MD FDSOI) structure

Table 1 Device parameters of conventional MS-MD FDSOI MOSFET [3] and Proposed HB MS-MD FDSOI MOSFET

Device parameters	Conventional devices [3]	Proposed devices	
Gate length	22 nm	22 nm	
Buried oxide thickness (BOX)	10 nm	Layer 1	Layer 2
		5 nm	5 nm
Gate oxide thickness (T_{ox})	2 nm	2 nm	
Silicon thickness (T_{si})	7 nm	7 nm	
Substrate doping	1e16	1e16	
Drain/high doping N^+ region	1e20	1e20	
Source/ high doping N^+ region	1e20	1e20	
Low doping N^-	1e17	1e17	
The work function of gate material	4.52	4.52	
Gate oxide material	HfO_2	HfO_2	

Table 2 Different buried combination and equivalent dielectric constant ($\in_{eq}$) (calculated in Eq. 1)

Hetero-dielectric material combination	Equivalent dielectric constant ($\in_{eq}$)
SiO_2	3.9
SiO_2_Si_3N_4	5.13
SiO_2_HfO_2	6.625
Si_3N_4_HfO_2	11.186

3 Results and Discussions

The electric field distribution throughout the device is shown in Fig. 1. The stacked buried oxide layer helps in strengthening the electric field in the channel region to a larger extent. Hence it can be concluded that the controllability of the gate in the proposed structure has increased owing to the effect produced by hetero-dielectric box arrangement.

The transfer characteristic is plotted for the hetero-dielectric stacked buried oxide (HB MS-MD FDSOI) and conventional MS-MD FDSOI structure at fixed V_{DS} = 0.7 V, are shown in Fig. 2. To establish an incremental peak at the channel/drain interfaces, the channel region was expanded into the drain region. The proposed device's I_{ON}/I_{OFF} ratiois higher than the conventional FDSOI device. The I_{ON}/I_{OFF} ratio is an essential parameter for device switching, the term switching power may be defined as relative switching power (RSP) [16] to compare the switching capacity.

$$\mathrm{RSP} = \left[\frac{\frac{I_{ON}}{I_{OFF}} \text{ of Proposed HB MS - MD FDSOI MOSFET}}{\frac{I_{ON}}{I_{OFF}} \text{ of Conventional MS - MD FDSOI MOSFET}}\right] \quad (2)$$

The value is 8.47, which is much higher than 1, using Eq. 1 for RSP measurement, so it indicates that the suggested structure has a greater switching performance than the conventional FDSOI.The off-condition current of the proposed structure is still closely matching that suggested in the ITRS for the FDSOI MOSFET for low power digital applications [17].

In Fig. 2 the transfer characteristic is plotted for the proposed HB MS-MD FDSOI and conventional MS-MD FDSOI [3] structure at fixed V_{DS} = 0.7 V. The

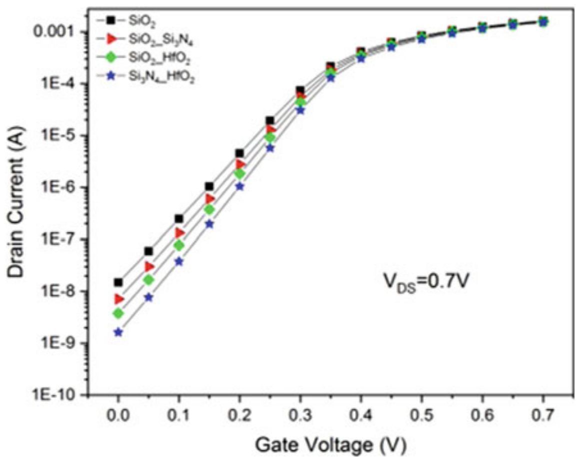


Fig. 2 The transfer characteristic of HB MS-MD FDSOI with hetero-dielectric buried oxide material arrangement

Si_3N_4_HfO_2 stacked BOX gives out minimum off-state current. Beside this, SiO_2 BOX produces maximum leakage current. The buried oxide reduces the parasitic capacitance between the source, drain, and substrate. A low K_{BOX} decreases the vertical electric field [18].

Figure 3 shows the variation in I_{OFF} and I_{ON}/I_{OFF} ratio for different BOX material arrangement sat $V_{DS} = 0.7$ V. If we increase the dielectric constant, the leakage current will be decreased. The Si_3N_4_HfO_2 has the highest equivalent dielectric constant among all combinations so, Si_3N_4_HfO_2 has the least leakage. In the case of SiO_2, ON current observed is, i.e. 1.62 mA higher as compared to all other combinations, Si_3N_4_HfO_2 has the least ON current value, i.e. 1.52 mA. The device has improved the I_{ON}/I_{OFF} ratio 8.46 times as conventional structure, Si_3N_4_HfO_2 has a higher value obtained, i.e. 9.242×10^5, and the least value of SiO_2 is 1.091×10^5. The variation in intrinsic gate delay against buried oxide material arrangement is shown in Fig. 4, the intrinsic gate delay depends on the I_{ON} current. From Fig. 4 the highest value of Si_3N_4_HfO_2 arrangement but this arrangement gives the least ON current. Another essential factoris threshold voltage, it is defined as the gate voltage below which no current flows in the transistor. The variation in threshold voltage (V_{th}) value for a different buried oxide material arrangement is shown in Table 3.

Another essential factor is the Sub-threshold slope and DIBL for low power digital application. Figure 5 shows the variation of sub-threshold slope and DIBL against buried oxide material arrangement. The sub-threshold slope depends upon the oxide capacitance and also depends upon depletion capacitance in the channel. It also depends upon trapped-charge density at the channel-oxide interface. The preliminary

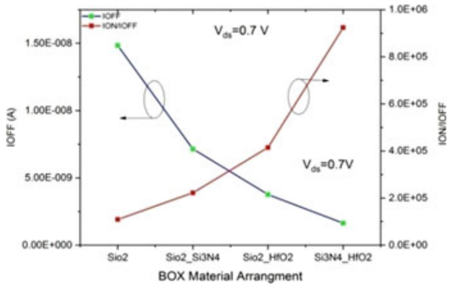


Fig. 3 Variation in I_{OFF} and I_{ON}/I_{OFF} ratio for different BOX material arrangement

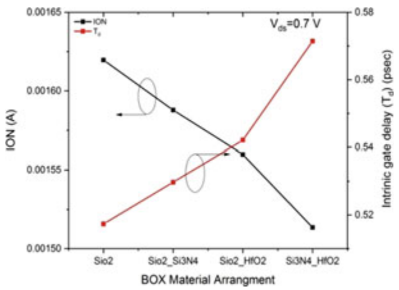


Fig. 4 Variation in I_{ON} and Intrinsic gate delay different BOX material arrangement

Table 3 Electrical performance comparison between MS-MD FDSOI and HB MS-MD FDSOI based on various parameters

Structure	OFF current (A)	ON current (A)	I_{ON}/I_{OFF}	Threshold voltage (V)	Sub-Threshold slope (mV/decade)	DIBL (mV/V)	Static power dissipation (nW)
SiO2 [3]	1.483×10^{-8}	1.62×10^{-3}	1.091×10^{5}	0.327	78.4022	86.62	10.38744
SiO_2_Si_3N_4	7.147×10^{-9}	1.59×10^{-3}	2.224×10^{5}	0.326	74.8149	75.72	5.00308
SiO_2_HfO_2	3.764×10^{-9}	1.56×10^{-3}	4.144×10^{5}	0.326	71.4062	68.99	2.63494
Si_3N_4_HfO_2	1.633×10^{-9}	1.51×10^{-3}	9.242×10^{5}	0.329	67.599	65.28	1.14364

result obtained for the sub-threshold slope is the least for the proposed structure with Si_3N_4_HfO_2 BOX, i.e.around 67 mV/decade which is 13.77%. Improvement over conventional SiO_2 BOX layer. The sub-threshold slope of the SiO_2_Si_3N_4 and SiO_2_HfO_2 is 74 mV/decade and 71 mV/decade respectively.

Subthreshold slope denotes all device effectively go into off state when Vgs is less than threshold voltage, which shows reduces in subthreshold leakage current, concequently I_{ON} to I_{OFF} increases.

The threshold voltage is lowered in the short-channel MOSFET due to the rise in drain voltage. This occurrence is known as DIBL, the electric field is strong within the drain region [19]. Figure 5 shows the reduction in DIBL by varying the buried oxide stacking material. The DIBL is calculated by given formula.

$$\mathrm{DIBL} = \left[\frac{Vth_{\mathrm{VDS}=0.7\mathrm{V}} - Vth_{\mathrm{VDS}=0.05\mathrm{V}}}{V_{\mathrm{DS}\,=\,0.7\mathrm{V}} - V_{\mathrm{DS}\,=\,0.05\mathrm{V}}}\right] \tag{3}$$

Single-layer (SiO_2) BOX has the highest value obtained around 87 mV/V. In multilayer stacked (Si_3N_4_HfO_2), BOX has the least value obtained around 65 mV/V. DIBL for SiO_2_Si_3N_4 & SiO_2_HfO_2 is 75 & 70 mV/V respectively.

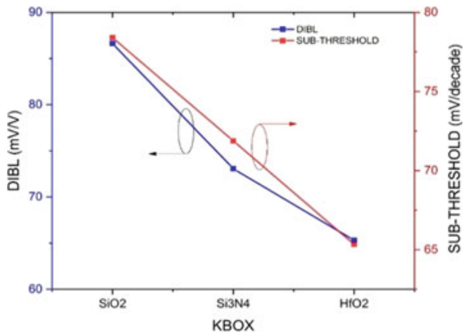


Fig. 5 Variation in DIBL and subthreshold slope for buried oxide material arrangement

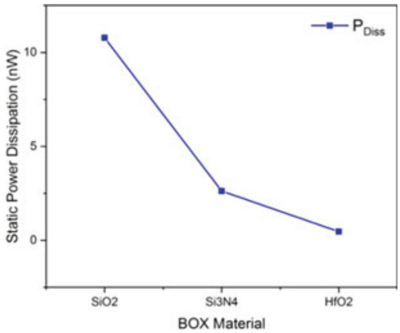


Fig. 6 Variation in static power dissipation for buried oxide material arrangement

Hence, Si_3N_4_HfO_2 BOX will produces minimum short channel effects among all the structure discussed.

Static power dissipation is the multiplication of small leakage current and applies voltage. As OFF current increases, static power dissipation also increases [20], which is calculated by using Eq. 4. Figure 6 shows the static power dissipation variation against the BOX material arrangement. The least value is obtained in the arrangement Si_3N_4_HfO_2 i.e. 1.143 nW, and the highest value is obtained in the arrangement of a single layer (SiO_2) is 10.387 nW.

$$P_{Diss.} = (I_{OFF}.V_{DD}) \tag{4}$$

4 Conclusion

The proposed hetero-dielectric stacked buried oxide-based FDSOI structure has been designed using various box material arrangements. It is observed that the Si_3N_4_HfO_2 box arrangement has performed better in terms of DIBL, SS, I_{OFF}, I_{ON}/I_{OFF}, P_{DISS} as compare to other arrangements. It has shown that improvement in DIBL, SS, I_{OFF}, I_{ON}/I_{OFF}, and SPD by 24.63, 13.77%, 9 times, 8 times, and 88% correspondingly over Single Sio_2 Box Layer. Hence it is suggested that the proposed device can be suitable for low-power digital circuit applications.

References

1. Doris, B.: Device design considerations for ultra-thin SOI MOSFETs. In: IEEE International Electron Devices Meeting, Washington, DC, USA, pp. 27.3.1–27.3.4 (2003)
2. Anvarifard, M.K., Orouji, A.A.: Proper electrostatic modulation of electric field in a reliable nano-SOI with a developed channel. IEEE Trans. Electron Dev. **65**(4), 1653–1657 (2018)

3. Mishra, V.K., Rao, N.: Electrostatic investigation of intended source drain ultra thin body FD-SOI MOSFET. SILICON **12**, 2819–2827 (2020)
4. Ranka, D., Rana, A.K., Yadav, R.K., Giri, D.: Performance analysis of FD-SOI MOSFET with different gate spacer dielectric. Int. J. Comput. Appl. **18**(5), 22–27 (2011)
5. Cheng, K., Khakifirooz, A.: Fully depleted SOI (FDSOI) technology. Sci. China Inf. Sci. **59**(6), 1 (2016)
6. Xiong, W., Ramkumar, K., Jang, S.J., Park, J.T., Colinge, J.: Self-aligned ground-plane FDSOI MOSFET. In: 2002 IEEE international SOI conference, Williamsburg, VA, USA, 2002, pp. 23–24 (2002)
7. Numata, T., Takagi, S.: Device design for subthreshold slope and threshold voltage control in sub-100-nm fully depleted SOI MOSFETs. IEEE Trans. Electron Dev. **51**(12), 2161–2167 (2004)
8. Xiong, W., Colinge, J.P.: Self-aligned implanted ground-plane fully depleted SOI MOSFET. Electron. Lett. **35**(23), 2059 (1999)
9. Mishra, V., Chauhan, R.K.: Performance analysis of fully depleted soi tapered body reduced source (FD-SOI TBRS) MOSFET for low power digital applications. Springer Nature, Singapore (2017)
10. Tseng, Y.C., Huang, W.M., Diaz, D.C., Ford, J.M., Woo, J.C.S.: AC floating-body effects in submicron fully depleted (FD) SOI nMOSFETs and the impact on analog applications. IEEE Electron Dev. Lett. **19**(9), 351–353 (1998)
11. Chouksey, S., Fossum, J.G., Agrawal, S.: Insights on design and scalability of thin-BOX FD/SOI CMOS. IEEE Trans. Electron Dev. **57**(9), 2073–2079 (2010)
12. Ohtou, T., Yokoyama, K., Shimizu, K., Nagumo, T., Hiramoto, T.: Threshold-voltage control of AC performance degradation-free FD SOI MOSFET with extremely thin BOX using variable body-factor scheme. IEEE Trans. Electron Dev. **54**(2), 301–307 (2007)
13. Kumar, M.J., Siva, M.: The ground plane in buried oxide for controlling short-channel effects in nanoscale SOI MOSFETs. IEEE Trans. Electron Dev. **55**(6), 1554–1557 (2008)
14. M.L.V.: Multilayered box in fdsoi mosfets (7,821,066 B2 L). U.S patent (2010)
15. ATLAS User's manual, SILVACO, Santa Clara, CA, USA (2016)
16. Anvarifard, M.K.: A nanoscale-modified junctionless with considerable progress on the electrical and thermal issue. Int. J. Numer. Model. Electron. Netw. Dev. Fields **32**(3), e2537 (2018)
17. Neisser, M., Wurm, S.: ITRS lithography roadmap: 2015 challenges. Adv. Opt. Technol. **4**(4), 1 (2015)
18. Paul, J., Mondal, C., Biswas, A.: Enhancing digital performance of nanoscale GeOI MOSFETs through optimization of buried oxide properties and channel thickness. Microsyst. Technol. **28**, 645–651 (2018)
19. Wagaj, S.C., Patil, S.C.: Comparison study of dual material gate silicon on insulator junctionless transistor and with junction transistor for analog performance. Int. J. Mater. Mech. Manuf. **7**(3), 144–149 (2019)
20. Mishra, V.K., Chauhan, R.K.: Area efficient layout design of CMOS circuit for high-density ICs. Int. J. Electron. **105**(1), 73–87 (2017)

Simulation Engineering of Heterojunction Colloidal Quantum Dot- Solar Cell Using Tungsten Trioxide (WO_3) as an Electron Transport Layer

Priyanka Dubey, Sadanand, Shambhavi Rai, B. K. Pandey, and D. K. Dwivedi

Abstract Quantum dot solar cell is one of the most potential third generation solar cell due to some peculiar properties of quantum dots like size and composition tunability. Despite of these properties, the colloidal quantum dot solar cell (CQDSCs) performance is still lagging behind conventional solar cells because of the phenomenon of recombination of carriers in quasi-neutral region (QNR). The solar cell device performance is highly dependent on the choice of electron transport layer (ETL) and hole transport layer (HTL). In the present work, WO_3 is used as ETL and Cu_2O is used as HTL, tetrabutylammonium iodide treated PbS (PbS-TBAI) is used as absorber layer. All simulation work is done by using SCAPS-1D simulator program. The obtained power conversion efficiency (PCE) of the present solar cell device structure is 16.32%. This device is then optimized by varying the doping density of ETL and HTL and 5×10^{17} cm^{-3} is finalized for both ETL and HTL. Further, the investigation of series and shunt resistance is performed. The optimized values for series and shunt resistance are 1 and 1×10^{6} Ω cm^2 respectively. The optimized value of the PCE is obtained as 18.07%.

Keywords Colloidal quantum dot solar cells (CQDSCs) · Transparent conductive electrode (TCE) · Electron transport layer (ETL) · Hole transport layer (HTL) · Power conversion efficiency (PCE) · Indium doped tin oxide (ITO)

1 Introduction

In the present era of modernization, the demand for energy consumption is increasing rapidly. Due to the limitation on availability of the commercial resources like fossil fuels, the focus is being shifted towards renewable resources of energy. Among all the resources of renewable energy, solar energy can be a good alternative for fossil

P. Dubey · Sadanand · S. Rai · B. K. Pandey · D. K. Dwivedi (✉)
Amorphous Semiconductor Research Lab, Department of Physics and Material Science, Madan Mohan Malaviya University of Technology, Gorakhpur 273010, India
e-mail: todkdwivedi@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_20

fuels in near future because sun is a source of enormous amount of energy which is solely capable of fulfilling the electricity requirement of the entire world [1, 2].

The third-generation colloidal quantum dot solar cell (CQDSCs) is seemed to be one of the most promising candidates for renewable energy technologies. CQDSCs have very high efficiency, stability, low cost, and solution processibility [3]. In CQDSCs, quantum dots are employed as absorbing material for conversion of sunlight into electrical energy. Quantum dots are very small sized particles usually having dimensions between 1 and 10 nm and show some peculiar properties like multi-exciton generation, tunable bandgap and size dependent optical, electrical and mechanical properties because of Quantum Confinement effect [4]. The structure of quantum dot solar cell consists of ETL, HTL and quantum dots as absorber layer. There are several kinds of quantum dots having narrow bandgap which are used as absorber layer for instance, CdTe (Cadmium Telluride) [5], Cadmium Sulphide (CdS) [6], Lead Selenide (PbSe) [7], Cadmium Selenide (CdSe) [8] etc. but these materials have very low efficiencies. The best reported material for absorber layer is PbS QDs [9].

In the present work, a device structure having different layers- Au/Cu_2O/PbS-TBAI/WO_3/ITO has been designed. Cu_2O is used for hole transportation (HTL) as it is environment friendly and is an abundant material. Further, the Cu_2O has high absorption coefficient and cost-effective fabrication which makes it one of the best candidates for hole transport material. For electron transport layer WO_3 is used, it has a transmittance of 80% in the visible region and has a good conductivity of nearly $10^{-3}/\Omega$-cm. WO_3 has a good electron mobility and during its fabrication and processing, by controlling the oxygen content it's density of states and Fermi level can be tuned. Also, WO_3 is easily fabricable and cost efficient. The absorber layer is treated with tetra-butyl ammonium iodide PbS-TBAI, where the electron-hole pair generation takes place. For this device structure, the obtained efficiency is 16.32%. The similar device structure with ETL TiO_2 has been reported earlier with efficiency of 13.72%. Further, the overall optimization of the device stricture has been performed. The enhanced efficiency of optimized device structure is obtained.

2 Simulation Methodology and Device Structure

In this proposed device structure, the indium doped tin oxide (ITO) coated on glass substrate act as transparent conductive electrode (TCE). The coating of ITO makes the glass low resistance and highly transparent. The role of TCE is to allow the incoming light to reach the photoactive layer, where the electron hole pairs are generated. The WO_3 ETL and ITO layer are paired together and act as a front contact. For the formation of back contact, the HTL Cu_2O is paired with Au metal electrode [10]. The sun rays hit the device from the glass side. The absorber layer absorb photons which generate the electron-hole pairs at junction. For separating the oppositely charged carriers, we have ETL and HTL which work as electron and hole transport layers, respectively. In this device structure, window layer is used to avoid the recombination

Table 1 Device parameters of different layers used for analysis [10–12]

Parameters	Cu_2O	PbS-TBAI	WO_3
W (μm)	0.05	0.23	0.08
E_g (eV)	2.17	1.14	2.6
χ (eV)	3.2	4.0	3.8
ε	7.11	20	4.8
N_c (cm^{-3})	2.02×10^{17}	1×10^{19}	2.2×10^{21}
N_v (cm^{-3})	1.1×10^{19}	1×10^{19}	2.2×10^{21}
μ_e (cm^2V^{-1} s^{-1})	2.0×10^2	2.0×10^{-2}	30
μ_h (cm^2V^{-1} s^{-1})	8.0×10^1	2.0×10^{-2}	30
N_D (cm^{-3})	0	1×10^{15}	6.35×10^{17}
N_A (cm^{-3})	1×10^{16}	1×10^{15}	0
N_t (cm^{-3})	–	–	1×10^{15}

Table 2 Defect parameters used for analysis [11]

Parameters	HTL/PbS-TBAI interface	PbS-TBAI/ETL interface
Defect type	Neutral	Neutral
Cross-section area of e^- (cm^2)	1.2×10^{-13}	1.0×10^{-19}
Cross section area of holes (cm^2)	1.2×10^{-13}	1.0×10^{-19}
Distribution of energy	Single	Single
Reference for defect energy level Et	Above the highest E_v	Above the highest E_v
Reference energy (eV)	0.5	0.6
Total density (integrated over all energies) ($1/cm^2$)	1×10^{16}	2×10^{14}

of the carriers at the surface of junction. So that maximum no. of the electron hole pairs generates at junction of the absorber layer. Now with the help of external load, the electrons injected to ETL are transported to HTL. So, the sequence of phenomenon happening in the device are absorption of photon, generation of electron-hole pairs, separation of carriers and movement of generated carriers (Tables 1 and 2).

To study the impact of the different solar cell parameters we have used SCAPS-1D simulator program. One dimensional simulation program is developed by University of Gent Belgium. This simulation program is accessible to all at free of cost. The basic mathematical equations used are Continuity and Poisson's equations which is solved by SCAPS-1D simulator program.

Continuity equations:

$$-\frac{\partial J_n}{\partial x} + G - U_n = \frac{\partial n}{\partial t}$$
$$-\frac{\partial J_p}{\partial x} + G - U_p = \frac{\partial p}{\partial t}$$

$$J_n = -\frac{\mu_n}{q} n \frac{\partial E_{F_n}}{\partial x}$$

$$J_p = +\frac{\mu_p}{q} p \frac{\partial E_{F_p}}{\partial x}$$

where, J_p denotes the current density of holes, J_n represents the current density of electron, the recombination rate for electrons is denoted by U_n whereas U_p represents the rate of recombination of the holes, generation rate of carriers is denoted by, μ_n is the mobility of electrons, μ_p is the mobility of holes, E_{F_n} and E_{F_p} represents the quasi Fermi level of electron hole respectively.

Poisson's equation

$$\frac{\partial}{\partial x}\left[\varepsilon_0 \varepsilon \frac{\partial \Psi}{\partial x}\right] = -q[-n + p + N_D^+ - N_A^- - n_t + p_t]$$

Where, Ψ denotes the electrostatic potential, q is the electronic charge, p and n are the no. of unbounded holes and unbounded electrons respectively, n_t and p_t represents trapped electrons and trapped holes, N_D^+ and N_A^- are ionized donor-like

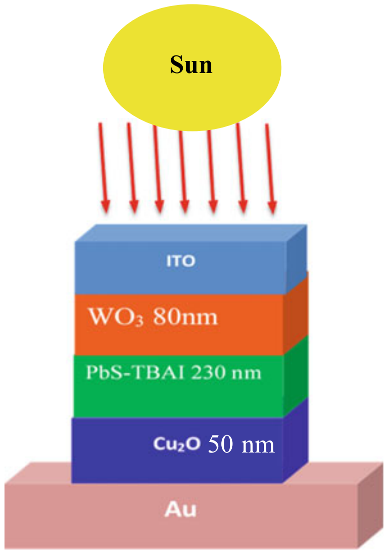


Fig. 1 Schematic diagram

doping and ionized acceptor-like doping concentration respectively, Ɛ represents the permittivity and ε_o represents the permittivity of free space (Fig. 1).

3 Result and Discussion

3.1 Impact of Acceptor Doping Concentration of HTL

In this sub-section, we have studied the impact of variation of acceptor doping density of HTL on the PV parameters Fig. 2(a–b). The acceptor doping density is varied from 1×10^{16} to 1×10^{19} cm^{-3} in by keeping the other device parameters constant. It can be seen from the graph that the device performance has improved significantly with increasing the acceptor doping density of Cu_2O. There is a remarkable increase in the value of V_{oc} and FF with increasing acceptor doping concentration of HTL.

The fill factor is being affected by the series resistance of the device. There is a decrease in the resistivity of the HTL with increase in doping of HTL and hence enhance the flow of holes from PbS-TBAI layer to HTL. The fill factor increases from 75.48 to 78.82% as doping concentration is increased from 1×10^{16} to 1×10^{19} cm^{-3}.

With increasing doping density of HTL, there is a rise in built-in potential (Vb_i) and hence electric field between HTL and absorber layer. This helps in extraction of generated holes and increase in V_{oc} from 762 to 810 mV. The PCE of the proposed device increases from 16.32 to 18.43% as doping density increases from 1×10^{16} to 1×10^{19} cm^{-3}.

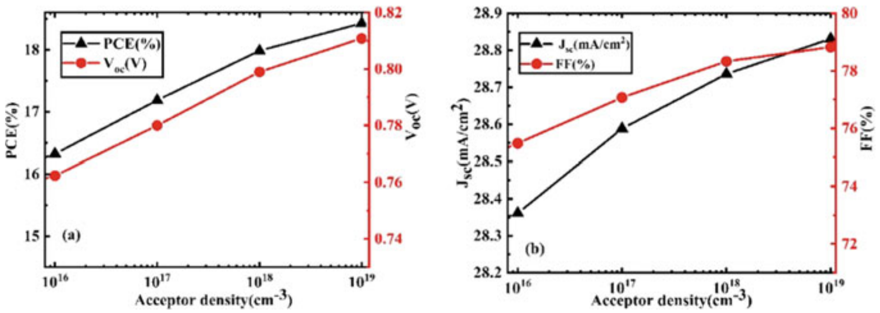


Fig. 2 Impact of acceptor doping density on PCE, V_{oc}, J_{sc}, and FF

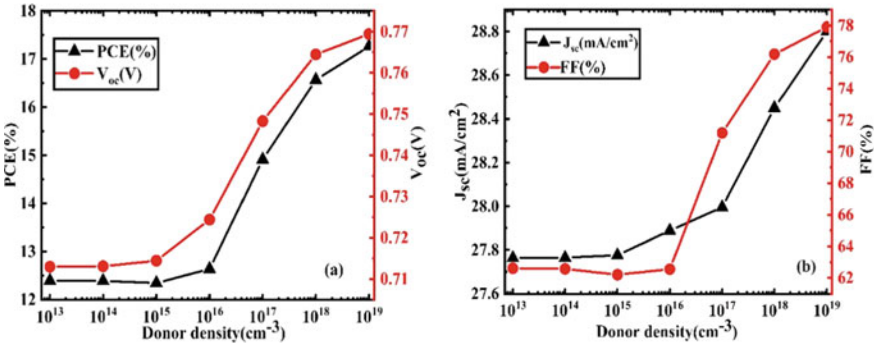


Fig. 3 Impact of donor doping density on PCE, V_{oc}, $J_{sc,}$ FF

3.2 Impact of Donor Doping Concentration of ETL

In the present sub-section, we have examined the effect of donor doping density of WO_3 on the photovoltaic performance of the solar device. The donor doping density is varied from 1×10^{13} to 1×10^{19} cm^{-3} while other parameters kept unchanged.

It is clear from the Fig. 3(a–b) that all the four parameters V_{oc}, J_{sc}, PCE, FF increases with increasing the donor density of ETL. The role of ETL is to extract the electrons from PbS-TBAI absorber layer which reduce the interfacial recombination at PbS-TBAI/ETL interface.

At low doping concentration, the conductivity is lower, and recombination of charge carriers increases, so the V_{oc}, J_{sc}, FF and PCE are reduced. As we increase the concentration of n-type dopants in ETL, the ETL offer high conductivity and mobility to generated electrons which results in the better performance of the device. The maximum V_{oc}, J_{sc}, FF and PCE obtained as 769.4 mV, 28.80 mA/cm^2, 77.90 and 17.28% at donor doping concentration of 1×10^{19} cm^{-3} respectively.

3.3 Impact of Series and Shunt Resistances on the Efficiency of the Device

In this section, we have studied the effect of variation of series and shunt resistance on the PCE of the device. The role of resistances on the device performance is quite significant. Ideally the solar cell device should have zero series resistance and infinite shunt resistance but in actual practice series and shunt resistances have some finite value. Series and shunt resistances represents losses. Series resistance is due to the material resistances which offers hinderance to the path of electrons. The power conversion efficiency of the proposed device is affected by the series and shunt resistances as shown in Fig. 4(a–b). With increase in the value of series resistance Rs, the PCE goes on decreasing. At the Rs value of 0 Ωcm^2, the PCE is maximum

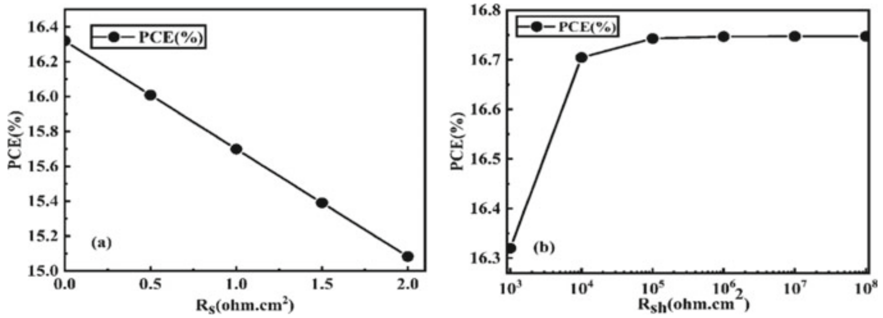


Fig. 4 Impact of series and shunt resistances on the PCE of the solar cell device

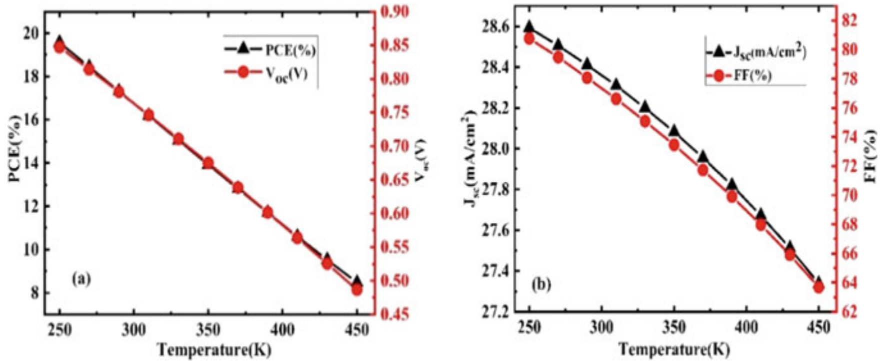


Fig. 5 Impact of increasing temperature on the QDSC performance

16.32% which is shown in Fig. 4(a). As we increase the shunt resistance of the device, we can see from the Fig. 4(b) that PCE increases to a specific value and then get saturated. At small values of shunt resistances there is a power loss in the device and performance of the device which gets degraded.

3.4 Impact of Increasing Temperature

In this section, the variation of four photovoltaic parameters with the increasing temperature is studied. The temperature is varied from 250 to 450K at a step of 20K and the corresponding variations is depicted in the Fig. 5(a–b). With the increase in temperature, there is an increase in the energy of electrons which in turn decreases the bandgap of the material. After gaining energy, the electron recombine with the hole [13]. The increment in the rate of recombination results in the decrease in V_{oc}, PCE, J_{sc} and FF. There is a linear decrement in V_{oc} and PCE with increase in temperature due to reverse saturation current. The FF and J_{sc} also decreases with increase in temperature (Table 3).

Table 3 Comparison of parameters among recently reported device structures

Device	V_{oc} (mV)	J_{sc} (mA/cm^2)	PCE (%)	FF (%)	Reference
ITO/MZO/PbS-TBAI/PbS-EDT/Au	620	24.50	9.41	62.0	[9]
ITO/TiO_2/PbS-TBAI/PbS-EDT/Au	635	24.46	9.87	63.8	[11]
ITO/TiO_2/PbS-TBAI/Cu_2O/Au	755.9	24.60	13.72	73.76	[14]
ITO/WO_3/PbS-TBAI/Cu_2O/Au	762.4	28.36	16.32	75.48	This work

4 Conclusion

In the present study, by varying donor density of electron transport layer, Acceptor density of HTL, series and shunt resistances and working temperature, the performance of the device has been investigated. From this study we found that the results are best at 1×10^{19} cm^{-3} for both ETL and HTL but taking care of fabricating difficulties 5×10^{17} cm^{-3} has been finalized for both ETL and HTL.

It is clear from this study that the performance of the device is best at small series and high shunt resistances. So, the series and shunt resistances are taken to be 1 and 1×10^6 Ωcm^2 respectively. For, optimization of the device, the working temperature is kept at 300K because at temperatures higher than this, performance of the device goes on decreasing.

The comparison among the parameters of proposed device and optimized device is listed below

Device	V_{oc} (mV)	J_{sc} (mA/cm^2)	PCE %	FF %
Before optimization	762.4	28.36	16.32	75.48
After optimization	792.0	28.67	18.07	79.55

References

1. Kannan, N., Vakeesan, D.: Solar energy for future world: a review. Renew. Sustain. Energy Rev. **62**, 1092–1105 (2016). https://doi.org/10.1016/j.rser.2016.05.022
2. Sadanand, S.P.K., Rai, S., Lohia, P., Dwivedi, D.K.: Comparative study of the CZTS, CuSbS2 and CuSbSe2 solar photovoltaic cell with an earth-abundant non-toxic buffer layer. Sol. Energy. **222**, 175–185 (2020). https://doi.org/10.1016/j.solener.2021.05.013
3. Sahu, A., Garg, A., Dixit, A.: A review on quantum dot sensitized solar cells: past, present and future towards carrier multiplication with a possibility for higher efficiency. Sol. Energy. **203**, 210–239 (2020). https://doi.org/10.1016/j.solener.2020.04.044
4. Smith, A.M., Nie, S.: Semiconductor nanocrystals: Structure, properties, and band gap engineering. Acc. Chem. Res. **43**, 190–200 (2010). https://doi.org/10.1021/ar9001069
5. Shen, X., Jia, J., Lin, Y., Zhou, X.: Enhanced performance of CdTe quantum dot sensitized solar cell via anion exchanges. J. Power Sources **277**, 215–221 (2015). https://doi.org/10.1016/j.jpowsour.2014.12.022

6. Li, L., Yang, X., Gao, J., Tian, H., Zhao, J., Hagfeldt, A., Sun, L.: Highly efficient CdS quantum dot-sensitized solar cells based on a modified polysulfide electrolyte. J. Am. Chem. Soc. **133**, 8458–8460 (2011). https://doi.org/10.1021/ja201841p
7. Hu, L., et al.: Synergistic effect of electron transport layer and colloidal quantum dot solid enable PbSe quantum dot solar cell achieving over 10% efficiency. Nano Energy **64**, 103922 (2019). https://doi.org/10.1016/j.nanoen.2019.103922
8. Choi, Y., Seol, M., Kim, W., Yong, K.: Chemical bath deposition of stoichiometric CdSe quantum dots for efficient quantum-dot-sensitized solar cell application. J. Phys. Chem. C. **118**, 5664–5670 (2014). https://doi.org/10.1021/jp411221q
9. Zhang, X., Santra, P.K., Tian, L., Johansson, M.B., Rensmo, H., Johansson, E.M.J.: Highly efficient flexible quantum dot solar cells with improved electron extraction using MgZnO nanocrystals. ACS Nano **11**, 8478–8487 (2017). https://doi.org/10.1021/acsnano.7b04332
10. Hossain, M.I., Alharbi, F.H., Tabet, N.: Copper oxide as inorganic hole transport material for lead halide perovskite based solar cells. Sol. Energy. **120**, 370–380 (2015). https://doi.org/10.1016/j.solener.2015.07.040
11. Pandey, R., Khanna, A., Singh, K., Patel, S.K., Singh, H., Madan, J.: Device simulations: toward the design of >13% efficient PbS colloidal quantum dot solar cell. Sol. Energy. **207**, 893–902 (2020). https://doi.org/10.1016/j.solener.2020.06.099
12. Rai, N., Rai, S., Singh, P.K., Lohia, P., Dwivedi, D.K.: Analysis of various ETL materials for an efficient perovskite solar cell by numerical simulation. J. Mater. Sci. Mater. Electron. **31**, 16269–16280 (2020). https://doi.org/10.1007/s10854-020-04175-z
13. Sadanand, D.D.K.: Numerical modeling for earth-abundant highly efficient solar photovoltaic cell of non-toxic buffer layer. Opt. Mater. (Amst). **109**, 11040 (2020). https://doi.org/10.1016/j.optmat.2020.110409
14. Prasad, S., Sadanand, L.P., Dwivedi, D.K.: Efficient PbS colloidal quantum dot solar cells employing Cu_2O as hole transport layer. Opt. Quant. Electron. **53**, 1–14 (2021). https://doi.org/10.1007//s11082-021-03075-8

Numerical Study of Colloidal Quantum Dot Photovoltaic Solar Cell for Highly Efficient Device

Rashmi Tiwari, Priyanka Dubey, Sadanand, Pooja Lohia, and D. K. Dwivedi

Abstract Nowadays, colloidal quantum dot solar cell is comprehensively used in day-to-day life because of its specific properties such as tunable size. This paper defines a brief detail on the effect of thickness of tin dioxide SnO_2 layer and presents their impact on solar cell performance. The further doping concentration of ETL Material has been varied to study the impact on photovoltaic parameters of the solar cell device. The performance of the solar cell device decreases with the escalation in the thickness of the SnO_2 layer. The influence of the doping concentration of the ETL material has been explored. The optimum efficiency of the solar cell 23.57% has been inspected, which could prove a highly efficient quantum dot solar cell experimentally using a simulated device structure.

Keyword Multi-excitation generation · CuI · PbS-TBAI · SnO2

1 Introduction

An implausible amount of solar energy is received by the earth in just one second. It supplies massive amounts of energy in one minute, which is enough for the world's need for renewable energy for one year. The energy hold on altogether fossil energy sources is concerning up to the radiation hanging the world in three days [1]. Therefore, a photovoltaic system is an alternative way to use this renewable energy source by direct conversion in electricity for consumption. The production of electrical energy is increasing day by day using photovoltaic devices. Solar cells are evolving for the last few years, but still, there are some issues to fulfill the global energy

R. Tiwari · P. Dubey · Sadanand · D. K. Dwivedi (✉)
Amorphous Semiconductor Research Lab, Department of Physics and Material Science, Madan Mohan Malaviya University of Technology, Gorakhpur 273010, India
e-mail: todkdwivedi@gmail.com

P. Lohia
Department of Electronics and Communication Engineering, Madan Mohan Malaviya University of Technology, Gorakhpur 273010, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_21

demand. The solar cell is categorized into three different generations. In the third generation, the quantum dot solar cell is introduced [2–4].

Due to some specific properties like tunable size bandgap, low temperature, and adaptable optoelectronic properties, Colloidal Quantum Dot Solar Cell (CQDSCs) are comprehensively applied in advanced photovoltaic systems [5–7]. So as to match the wide absorption of solar spectra size of quantum dots can be changed by the quantum size effect for solar cell application. Recently, the multi-excitation generation (MEG) impact in colloidal quantum dot-based solar cells has been specified [8]. Within the context of next-generation electrical phenomenon devices for harnessing alternative energy, it additionally offered challenges to grasp basic ideas like multi-exciton generation. The size of quantum dots is tiny, and it has different shapes such as geometrical shapes that square measure confined in three dimensions with their sizes compared to Bohr's exciton radius [9]. In PbS QDSCs, metal oxide such as ZnO, TiO_2 has generally been employed as ETL in CQDSCs attributable to their appropriate or high dielectric constant ($\varepsilon = 7$) for exciton detachment [10–12] and suitable energy levels. The ETL layer limitations due to this tin dioxide (SnO_2) appear as a fresh ETL because of its remarkable properties such as wider bandgap, high mobility, good anti-reflection nature, and lower temperature process [13–15] compatibility with CQD device construction procedure, etc.

The main goal of this work is to supply low value and high potency of QDSCs by changing the (ETL) electron transport layer from the initial device, which is ITO/TiO_2/PbS-TBAI/ PbS-EDT/Au. The device consists of different layers ITO/SnO2/PbS-TBAI/PbS-EDT/Au in the present work. The absorber layer was treated with tetra butyl ammonium iodide PbS-TBAI while HTL layer of PbS-EDT and ETL layer of tin dioxide (SnO_2) has been used for simulation. The efficiency obtained of QDSCs is 9.37% by recollecting SnO_2-Cl to achieve interface passivation in QDSSc [10].

2 Device Structure and Material Parameters

Using 1D-SCAPS software, we have studied the variation of parameters by altering the size and donor density of the SnO_2 layer. The influence of changing doping concentration and thickness on the photovoltaic device is analyzed under AM 1.5G spectrum with intensity 1000 W/m^2 at temperature 300K (Table 1).

The device structure ITO/SnO_2/PbS-TBAI/PbS-EDT/Au with various thicknesses of layers is shown in Fig 1.

SCAPS-1D simulation software was developed at Ghent University, Belgium. This software is freely accessible to all research communities [20]. Basically, in SCAPS-1D software, it governs the equations such as the Poisson equation and therefore the continuity equations for e^- and holes, and these equations are given as follows

Table 1 Device parameters of dissimilar layers [16–19]

Parameter	PbS-EDT	PbS-TBAI	SnO_2
Thickness (μ m)	0.005	00.23	0.05
Band gap (eV)	01.14	1.14	3.5
Electron affinity (eV)	3.9	4.00	4.0
Dielectric Permittivity	20	20	9
CB density (cm^{-3})	1×10^{19}	1×10^{19}	2.2×10^{17}
VB density (cm^{-3})	1×10^{19}	1×10^{19}	2.2×10^{16}
Electron mobility ($cm^2V^{-1}\ s^{-1}$)	2.0×10^{-4}	2.0×10^{-2}	20
Hole mobility ($cm^2V^{-1}\ s^{-1}$)	2.0×10^{-4}	2.0×10^{-2}	10
N_D (cm^{-3})	1×10^{14}	1×10^{15}	1×10^{17}
N_A (cm^{-3})	1×10^{16}	1×10^{15}	0
Defect density (cm^{-3})			1×10^{15}

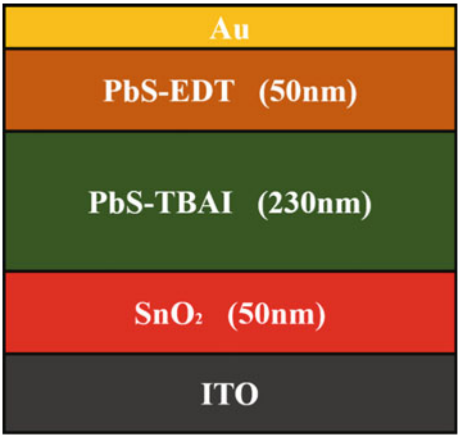


Fig. 1 Schematic representation of the device

$$\frac{\partial}{\partial x}(\varepsilon(x)\frac{\partial \Psi}{\partial x}) = -\frac{q}{\varepsilon_0}[-n + p - N_A^- + N_D^+ + \frac{1}{q}\rho_{def}(n, p)] \tag{1}$$

$$-\frac{\partial j_n}{\partial x} + G - U_n(n, p) = \frac{\partial n}{\partial t} \tag{2}$$

$$-\frac{\partial j_p}{\partial x} + G - U_p(n, p) = \frac{\partial p}{\partial t} \tag{3}$$

For charge carrier transport description, drift and diffusion expression for holes and electrons are as follows

$$j_n = -\frac{\mu_n}{q}\frac{\partial E_{F_n}}{\partial x} \tag{4}$$

$$j_p = -\frac{\mu_p}{q}\frac{\partial E_{F_p}}{\partial x} \tag{5}$$

where ε_o and ε are the permittivity of vacuum and also the permittivity of semiconductor material, Ψ is the electrostatic potential. Where p, n are free carrier concentrations, $N_D^{\pm}$ are the defect distributions, j_n, j_p are the electron and hole current densities, and G and U are the generation rate and recombination rate [21].

3 Result and Discussion

3.1 Influence of Variation in Thickness of ETL

In this part, we have inspected the effects of SnO_2 layer thickness on the solar cell demonstration. The SnO_2 layer thickness varies from 0.005 to 0.10 μm and keeps other parameters constant like thickness and bandgap of absorber and hole transport layer. It has been found that on changing the thickness of ETL all parameters almost decrease. It occurs due to the recombination of carriers and less generation of electron-hole pairs. It means that at the thicker ETL layer, the generation of electron and hole pairs becomes less, and thus a reduction in FF, V_{oc}, J_{sc}, and PCE have been observed. As a result, at a higher thickness, of ETL efficiency of the device is smaller, the photons are absorbed in thicker ETL. Therefore, to obtain high efficiency, the ETL layer should be small (Fig. 2).

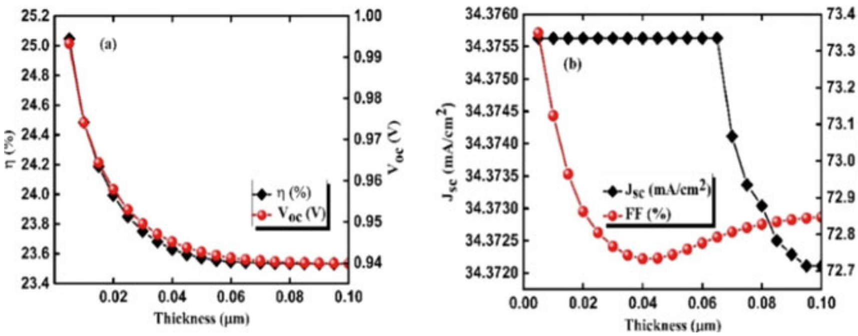


Fig. 2 Impact of ETL thickness on PV parameters **a** PCE, V_{oc} **b** J_{sc}, FF

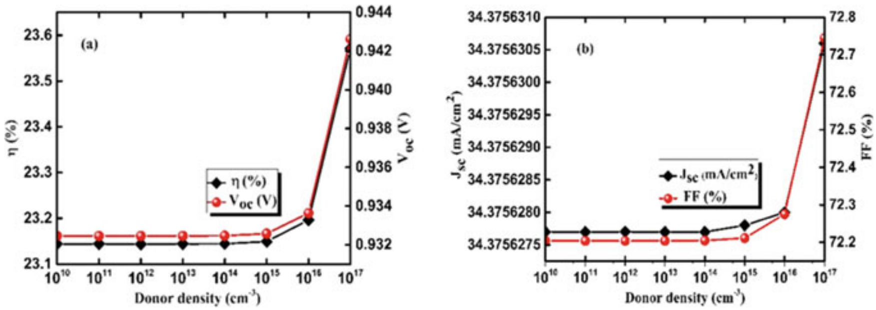


Fig. 3 Influence of doping on PV parameters **a** PCE, V_{oc} **b** J_{sc}, FF

3.2 Impact of Doping Density of ETL

In this simulation, we have varied doping concentrations of (ETL) Electron transport layer (SnO_2) from 1×10^{10} to 1×10^{17} cm^{-3} keeping other parameters unchanged throughout variation of doping concentration. It is seen that on changing the donor density of SnO_2 the photovoltaic cell parameters like open-circuit voltage (V_{oc}), short circuit current density (J_{sc}), Fill Factor (FF), and Power Conversion Efficiency (PCE) are improved, which shown in Fig. 3(a–b). The improvement of donor density of the ETL improved the incorporated electric field of a photovoltaic device, which is virtuous for the open-circuit voltage of the solar cell. In Fig. 3, we have observed that all photovoltaic parameters increase. It means that as photons are incident on the device, it generates a large number of electron-hole pairs. Thus, it results in a large number of current and efficiency of solar cell device.

3.3 J-V and Quantum Efficiency Curve

In this section J-V curve and QE curve have been studied to understand the performance of solar cell devices. Figure 4(a) displays the J-V curve, and Fig. 4(b) shows the QE curve. It is perceived that through the escalation in temperature, the PCE of quantum dot solar cell devices goes on diminishing; therefore, we have simulated this device at room temperature. The efficiency of the device, 23.57% has been observed [22–24]. By forward-facing surface passivation, the photogeneration of charge carriers adjacent to the superficial remains exaggerated—high forward-facing surface recombination due to absorption of blue light nearby the surface. In the greenish part of the spectrum truncated diffusion length distress, the possibility of solar cell accretion and reduction in quantum efficiency is perceived [23–27]. Normally, QE is not prominent in subordinate wavelength because of a smaller amount of power from the spectrum. QE declines through the escalations in

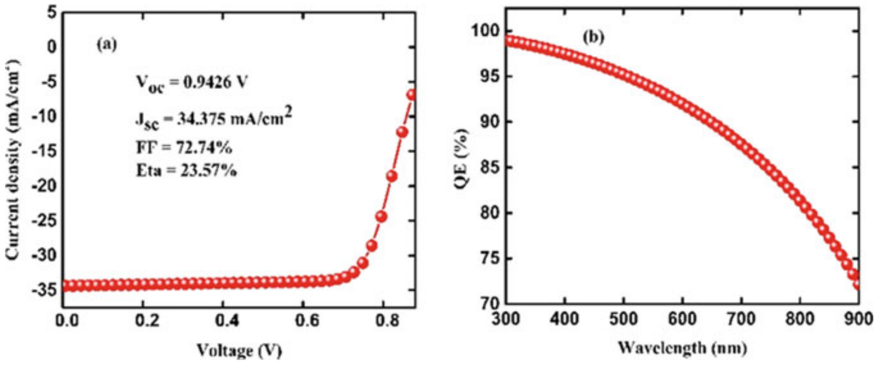


Fig. 4 Short circuit current density and Quantum efficiency curve of device **a** J-V **b** QE

Table 2 Photovoltaic parameters of the experimental and simulated device

Device	V_{oc} (V)	J_{sc} (mA/cm^2)	FF (%)	PCE (%)
Experimental device [10]	0.57	23.79	0.58	7.92
Simulated device (This work)	0.9426	34.375	72.74	23.57

wavelength, shown in Fig. 4(b). In Table 2 comparison between experimental and simulated devices is given.

4 Conclusions

Throughout the simulation, SCAPS-1D software has been used. In this work, we have carried out simulation studies to compare the performance of the initial device by changing ETL from the initial device structure, which consists of TiO_2 as an ETL layer. Increasing the thickness of ETL current density of the device is constant up to a certain value, and after that value, it decreases abruptly. Varying the various parameters of the device structure, we have optimized the performance of the photovoltaic device, and we can see that at higher thickness, the PCE is less than the lower thickness. Variations in doping concentration have been investigated to explore the performance of solar cell devices. Based on doping concentration, the device gives maximum efficiency at a higher doping concentration of ETL. Parameters like doping concentration and thickness of ETL layer have been varied to investigate the impact on solar cell performance. The optimum PCE of the solar cell obtained is 23.57%. The calibrated device's performance is more efficient compared to the device structure of TiO_2 as an ETL. Further, it is also observed that to achieve good performance of the solar device, the thickness of ETL should be small. The solar cell device is a renewable energy source. That's why the worldwide demands are increasing day by day, and it is eco-friendly.

References

1. Kannan, N., Vakeesan, D.: Solar energy for future world: a review. Renew. Sustain. Energy Rev. **62**, 1092–1105 (2016). https://doi.org/10.1016/j.rser.2016.05.022
2. How, G.T.S., Jothivenkatachalam, K., Pandikumar, A., Huang, N.M.: Metal nanoparticle decorated ZnO nanostructure based dye-sensitized solar cells. rational design of solar cells for efficient solar energy conversion, p. 1 (2018)
3. Wang, T., Ren, S., Li, C., Li, W., Liu, C., Zhang, J., Wu, L., Li, B., Zeng, G.: Exploring window buffer layer technology to enhance CdTe solar cell performance. Sol. Energy **164**, 180–186 (2018)
4. Ichikawa, Y., Yoshida, T., Hama, T., Sakai, H., Harashima, K.: Production technology for amorphous silicon-based flexible solar cells. Sol. Energy Mater. Sol. Cells **66**(1–4), 107–115 (2001)
5. Liu, S., Hu, L., Huang, S., Zhang, W., Ma, J., Wang, J., Guan, X., Lin, C.H., Kim, J., Wan, T., Lei, Q.: Enhancing the efficiency and stability of PbS quantum dot solar cells through engineering an Ultrathin NiO nanocrystalline interlayer. ACS Appl. Mater. Interfaces. **12**(41), 46239–46246 (2020)
6. Tang, J., Kemp, K.W., Hoogland, S., Jeong, K.S., Liu, H., Levina, L., Furukawa, M., Wang, X., Debnath, R., Cha, D., Chou, K.W.: Colloidal-quantum-dot photovoltaics using atomic-ligand passivation. Nat. Mater. **10**(10), 765–771 (2011)
7. Liu, Z., Lin, C.H., Hyun, B.R., Sher, C.W., Lv, Z., Luo, B., Jiang, F., Wu, T., Ho, C.H., Kuo, H.C., He, J.H.: Micro-light-emitting diodes with quantum dots in display technology. Light Sci. Appl. **9**(1), 1–23 (2020)
8. Lazer, D., Pentland, A.S., Adamic, L., Aral, S., Barabasi, A.L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T.: Life in the network: the coming age of computational social science. Science (New York, NY) **323**(5915), 721–723 (2009)
9. Memming, R.: Semiconductor Electrochemistry. John Wiley & Sons, Hoboken (2015)
10. Khan, J., Yang, X., Qiao, K., Deng, H., Zhang, J., Liu, Z., Ahmad, W., Zhang, J., Li, D., Liu, H., Song, H.: Low-temperature-processed SnO 2–Cl for efficient PbS quantum-dot solar cells via defect passivation. J. Mater. Chem. A. **5**(33), 17240–17247 (2017)
11. Ning, Z., Voznyy, O., Pan, J., Hoogland, S., Adinolfi, V., Xu, J., Li, M., Kirmani, A.R., Sun, J.P., Minor, J., Kemp, K.W.: Air-stable n-type colloidal quantum dot solids. Nat. Mater. **13**(8), 822–828 (2014)
12. Ding, C., Liu, F., Zhang, Y., Hayase, S., Masuda, T., Wang, R., Zhou, Y., Yao, Y., Zou, Z., Shen, Q.: Passivation strategy of reducing both electron and hole trap states for achieving high-efficiency PbS quantum-dot solar cells with power conversion efficiency over 12%. ACS Energy Lett. **5**(10), 3224–3236 (2020)
13. Guo, Y., Lei, H., Wang, C., Ma, J., Chen, C., Zheng, X., Yang, G., Xiong, L., Tan, Z.: Reconfiguration of interfacial and bulk energy band structure for high-performance organic and thermal-stability enhanced perovskite solar cells. Solar RRL. **4**(4), 1900482 (2020)
14. Liu, Q., Qin, M.C., Ke, W.J., Zheng, X.L., Chen, Z., Qin, P.L., Xiong, L.B., Lei, H.W., Wan, J.W., Wen, J., Yang, G.: Enhanced stability of perovskite solar cells with low-temperature hydrothermally grown SnO2 electron transport layers. Adv. Func. Mater. **26**(33), 6069–6075 (2016)
15. Lan, X., Voznyy, O., Kiani, A., de Armure, F.P.G., Abbas, A.S., Kim, G.H., Liu, M., Yang, Z., Walters, G., Xu, J., Yuan, M.: Passivation using molecular halides increases quantum dot solar cell performance. Adv. Mater. **28**(2), 299–304 (2016)
16. Azri, F., Meftah, A., Sengouga, N., Meftah, A.: Electron and hole transport layers optimization by numerical simulation of a perovskite solar cell. Sol. Energy **181**, 372–378 (2019)
17. Kawano, Y., Chantana, J., Nishimura, T., Minemoto, T.: Influence of halogen content in mixed halide perovskite solar cells on cell performances through device simulation. Sol. Energy Mater. Sol. Cells **205**, 110252 (2020)
18. Dwivedi, D.K.: Numerical modeling for earth-abundant highly efficient solar photovoltaic cell of non-toxic buffer layer. Opt. Mater. **109**, 110409 (2020)

19. Zhang, X., Santra, P.K., Tian, L., Johansson, M.B., Rensmo, H., Johansson, E.M.: Highly efficient flexible quantum dot solar cells with improved electron extraction using MgZnO nanocrystals. ACS Nano **11**(8), 8478–8487 (2017)
20. Burgelman, M., Nollet, P., Degrave, S.: Modelling polycrystalline semiconductor solar cells. Thin Solid Films **361**, 527–532 (2000)
21. Movla, H.: Optimization of the CIGS based thin film solar cells: Numerical simulation and analysis. Optik **125**(1), 67–70 (2014)
22. Simya, O.K., Mahaboobbatcha, A., Balachander, K.: A comparative study on the performance of Kesterite based thin film solar cells using SCAPS simulation program. Superlattices Microstruct. **82**, 248–261 (2015)
23. Dwivedi, D.K.: Modeling of CZTSSe solar photovoltaic cell for window layer optimization. Optik **222**, 165407 (2020)
24. Li, J., Wang, D., Li, X., Zeng, Y., Zhang, Y.: Cation substitution in earth-abundant kesterite photovoltaic materials. Adv. Sci. **5**(4), 1700744 (2019)
25. Dwivedi, D.K.: Theoretical investigation on enhancement of output performance of CZTSSe based solar cell. Solar Energy. **193**, 442–451 (2019)
26. Babu, P.S., Singh, P.K., Thakur, A.K., Dwivedi, D.K.: Optimization of photovoltaic solar cell performance via the earth abundant Zn3P2 back surface field. Optik **229**, 166235 (2021)
27. Dwivedi, D.K., Alharthi, F.A., El Marghany, A.: One-step hydrothermal synthesis of Cu2ZnSn (S, Se) 4 nanoparticles: structural and optical properties. Nanosci. Nanotechnol. Lett. **12**, 338–344 (2020)

Design of Dual-Band Printed Antenna for RF Energy Harvesting Applications

Rashmi Pandey, A. K. Shankhwar, and Ashutosh Singh

Abstract In this work dual-band printed antenna is designed and analyzed the performance of return loss and gain which suitably works for energy harvesting. It works on 2.1 GHz (UMTS) and 2.6 GHz (Wi-Fi) frequency bands. Two different patch structuresare designed. The first structure was designed without the slits and the second one with the slits and observed the response of antenna parameters with both the structure. The proposed antenna with the slits shows improvement in return loss and gain due to radiating patch structure. The proposed antenna is observed the gain of 2.430 and 2.363 dBi at 2.1 and 2.6 GHz respectively. Structure one shows the response at a single frequency band at 2.23 GHz with 2.7 dBi gain and −15.09 dB return loss. The gain of the proposed antenna is quite different but it considers two frequency bands for the proposed work. For RF energy harvesting required dual-band antenna, hence the proposed antenna fulfills the requirement of the RF energy harvesting.

Keywords Dual band · Gain · Printed patch antenna · RF energy harvesting · UMTS · Wi-Fi

1 Introduction

At present, self-sufficient electronic devices are in great demand in our daily life activity, for instance, health trackers with sensor-enabled, domestic electronic appliances, RFID, and other wireless devices with IoT [1]. Most of these devices are powered with the use of batteries and this also requires less power to operate.

R. Pandey (✉) · A. K. Shankhwar · A. Singh
Department of Electronics Engineering, Harcourt Butler Technical University, Kanpur, India
e-mail: rashmi821@gmail.com

A. K. Shankhwar
e-mail: akshankhwar@hbtu.ac.in

A. Singh
e-mail: ashuhbti@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_22

The life span of the battery is limited to years and in most cases, maintenance and replacements of the batteries are also tough nowadays [2]. Thus, many researchers are working on a substitute to power up these devices. The concept of RF energy harvesting has invented in 1970 as an alternative method to empower the IoT, sensors, and battery-enabled devices by exploiting RF signals from an environment. The device used for the purpose is known as rectenna. Rectenna is a combined part of antenna and rectifier, this receives RF signal by the antenna and converts that signal into DC power by the rectifier. It is a green energy concept and this is very useful to power low-power devices. Some dual-band antennas for RF energy harvesting have been reported in the literature. A differentially fed antenna with dual polarization has been reported in [3, 4], for the bands operated at 2.4 and 5.5 GHz. Dual-band dielectric resonator antenna has been proposed in [5], dual-band stacked patch antenna has presented in [6]. A circularly polarized patch antenna at 2.4 GHz has been reported in the literature [7].

Few works related to dual-band antennas have been reported in the literature [4, 8–10]. A Dual-band patch antenna to harvest energy at 1.95, and 2.45 GHz has been introduced in [8]. In [8] author uses a circular patch to analyze the performance of the antenna at the targeted bands. Dual-band folded dipole antenna for RF energy harvesting has been reported in work [9] at the operating bands 0.9 GHz and 2.45 GHz. The authors presented their work on a dual-band antenna for RF energy harvesting application [4, 10] has satisfied the requirement of available power at the targeted bands. A triple-band antenna has also been presented in the [11] to receiving the radio frequency signals for UMTS, WLAN, and WiMax. The works represent the differential antenna for the harvesting energy from the targeted frequency bands. Another dual-band differentially fed antenna has also been reported in the literature [12] to receive RF signals for the base station. Smart IoT devices application from the RF energy harvesting has shown in [13]. All these reported work shows enough antenna parameters and targeted frequency bands, and few of them has size constraints, gain difference, and complexity issue. In the presented work author fulfill the basic requirements of harvesting RF energy from nearby sources.

The primary objectives of the work are reported in the paper:

- A dual-band antenna is designed and analyzed to operate at 2.1 and 2.6 GHz frequency bands.
- To achieve the appropriate response of the proposed antenna rectangle slits are introduces in the radiating patch.
- The proposed structure has fulfilled the requirement of UMTS and Wi-Fi/WiMaxbands power requirement for energy harvesting.
- Therefore, it increases the chance to optimize the designed antenna that collects more energy from the targeted band with a quite considerable gain value.

The proposed antenna geometry of the radiating patch is introduced in Sect. 2. Section 3 discusses the result response of the antenna in terms of return loss $S_{11,}$ Radiation properties, and surface current distribution along with the gain. Section 4 concludes the presented work.

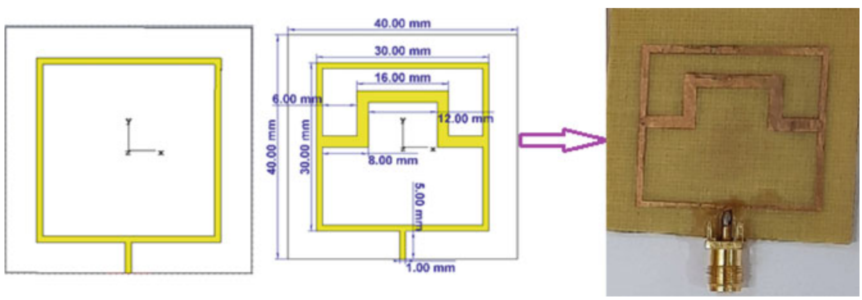
2 Antenna Design Specifications

In the beginning, a microstrip patch antenna is adopted for its outstanding features such as being easy to integrate, low cost, and light profile. To design the antennaFR4 substrate $\varepsilon_r = 4.3$ with 1.6 mm thickness is used. Low-cost substrate thickness can be increased but the increment in the thickness shows an increment in the antenna weight. This is responsible for system complexity. The dimension of the substrate is 40×40 mm^2 (which is $0.28\,\lambda_0\,x\,0.28\,\lambda_0$ at 2.1 GHz, $0.34\,\lambda_0\,x\,0.34\,\lambda_0$ at 2.6 GHz) is considered to achieve the dual-band characteristics.The wavelength of the proposed antenna is calculated according to Eq. 1. The partial ground structure is introduced in the ground to reduce the losses due to material and it also maintains the return loss and radiation properties of the proposed antenna.

$$\lambda = \frac{C}{f} \tag{1}$$

where λ represents the antenna wavelength, C is the speed of light (3×10^9), and f represents the resonating frequency.

Figure 1 represents the geometry and prototype of the proposed antenna structure. The antenna in step 1 is designed with simple rectangular slots and step 2 is designed with bend rectangular slits. Slits are created in the main radiating patch to enhance the impedance bandwidth and it is also responsible to cover the different frequency bands. The dimension of the antenna is calculated with the standard equation of the antenna [14]. A 50Ω microstrip transmission line as a feed is used to transmit the electrical signal towards the antenna. The design process of the proposed antenna is initializing with the simple rectangular geometry in the radiating patch. After that slit introduces in the patch to observe the response at 2.1 and 2.6 GHz. To improve the performance of the proposed antenna, step-by-step slits are introduced in the structure. Through the analysis of radiating patch final dual-band antenna for 2.1 GHz (UMTS) and 2.6 GHz (Wi-Fi) is achieved.



(a) Step 1 (b) Step 2 (c) Final Structure

Fig. 1 Antenna geometry **a** Antenna without slits **b** Antenna with slits **c** Proposed antenna

3 Result Discussion

From the analysis of the proposed antenna, observed the response of different parameters such as return loss S_{11}, Radiation characteristics of the antenna, gain, and surface current distribution. The proposed antenna is designed with two different structures;the basic structure at step 1 analyzes the single operating band at 2.23 GHz with 2.7 dBi gain and −15.09 dB return loss. Apart from this, the antenna structure at step 2 contributes the dual-band radiation at 2.1 and 2.6 GHz with 2.43dBi and 2.36dBi gain. The dual-band antenna shows remarkable advantages rather than single-band and wideband antennas. A Single-band antenna collects RF energy from a single RF source this limits the application at any specific bands. In the case of wideband antennas, it collects more RF power, which is also the trade-off. Therefore, a dual-band antenna performs better for RF energy harvesting from the targeted bands. The proposed analysis of the dual-band antenna is appropriate to harvest the RF power. The return loss of the proposed antenna is analyzed by using the CST simulation tools. It shows the losses during the radiation of energy in the free space. Figure 2 shows the characteristics of the antenna return loss at the targeted frequency bands. Three bands are considered to evaluate the performance of both the designed structures. The observed S_{11} response at 2.1, 2.6, and 2.23 GHz is −11.06, −24.16, and −15.12 dB respectively.

Figure 3 shows the details about the power transmitted, absorbed, stimulated, and radiated at all ports. It is a part of the excitation during the initialization of the feed point. It displays the energy radiation characteristics of the proposed antenna.

Figure 4 demonstrates the behavior of the surface current distribution at 2.1 and 2.6 GHz. The surface current is mostly scattered on the radiating slits of the antenna, this shows the maximum energy is emitted by the designed antenna. Hence, the desired performance of the proposed antenna has been achieved during the simulation. Thus, it fulfills the main objective of the work and is efficient for RF energy harvesting applications. Figure 5 shows the radiation characteristics of the proposed

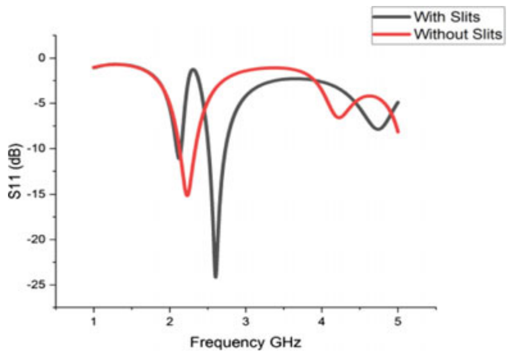


Fig. 2 Return loss (S_{11}) analysis of the proposed antenna

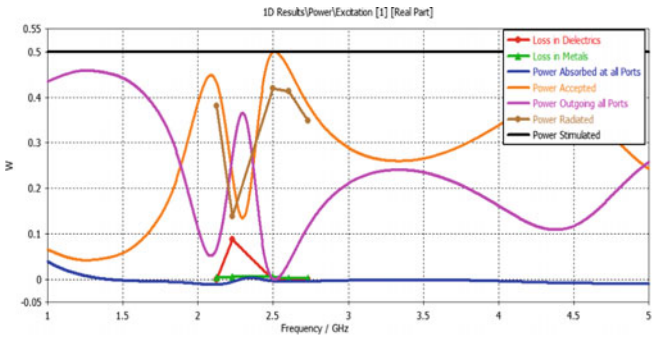


Fig. 3 Excited power of the proposed antenna

antenna. Which means how much RF power is received by the antenna. It is an important characteristic of any designed antenna.

From the above analysis, it has been observed that the proposed antenna radiates efficiently in targeted frequency bands. It fulfills the necessity of energy harvesting.

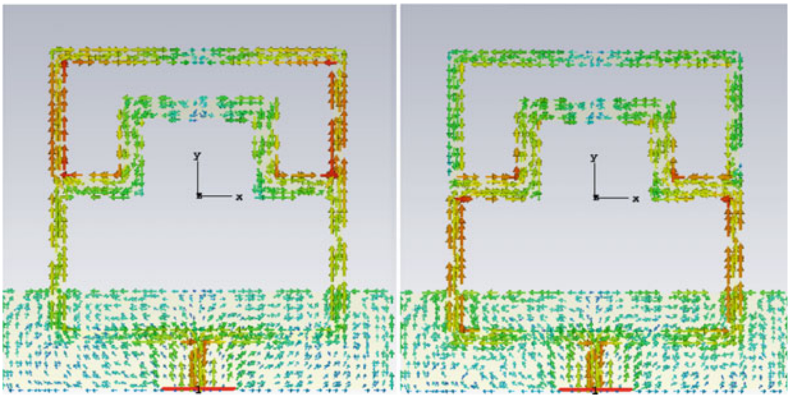


Fig. 4 Surface current distributions at 2.1 and 2.6 GHz

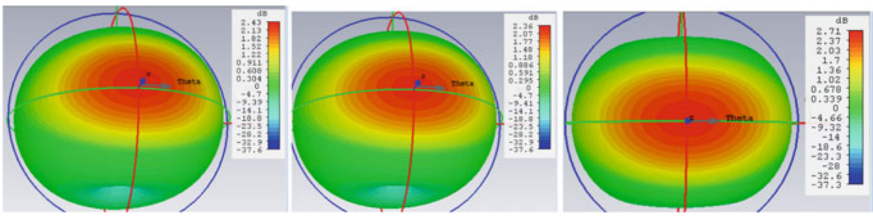


Fig. 5 Energy radiation of the proposed antenna at 2.1, 2.6, and 2.23 GHz

Analyzed parameters are quite enough to harvest RF power to power up the low power devices.

4 Conclusions

Dual-band printed antenna is designed and analyzed for RF energy harvesting. It works on 2.1 GHz (UMTS) and 2.6 GHz (Wi-Fi) frequency bands. Both the targeted bands are suitable to provide RF power from the UMTS and Wi-Fi devices. The proposed antenna with the slits shows improvement in return loss and gain. The proposed antenna has observed the gain of 2.430 and 2.363 dBi at 2.1 and 2.6 GHz respectively. A Dual-band antenna is preferred for radiofrequency energy harvesting. Therefore, an antenna with bend rectangle slits has been considered for the final work. The proposed antenna suitably works with the targeted application bands. It shows good candidature to harvest the RF power from the proposed antenna.

References

1. Kuhn, V., Lahuec, C., Seguin, F., Person, C.: A multi-band stacked RF energy harvester with RF-to-DC efficiency up to 84%. IEEE Trans. Microwave Theory Techniques **63**(5), 1768–1778 (2015)
2. Palazzi, V., et al.: A novel ultra-lightweight multiband rectenna on paper for RF energy harvesting in the next generation LTE bands. IEEE Trans. Microwave Theory Tech. **66**(1), 366–379 (2017)
3. Tang, Z., Liu, J., Yin, Y.: Enhanced cross-polarization discrimination of wideband differentially fed dual-polarized antenna via a shorting loop. IEEE Ant. Wirel. Propagat. Lett. **17**(8), 1454–1458 (2018)
4. Mattsson, M., Kolitsidas, C.I., Jonsson, B.L.G.: Dual-band dual-polarized full-wave rectenna based on differential field sampling. IEEE Ant. Wirel. Propag. Lett. **17**(6), 956–959 (2018)
5. Tang, H., Tong, C., Chen, J.-X.: Differential dual-polarized filtering dielectric resonator antenna. IEEE Trans. Ant Propag. **66**(8), 4298–4302 (2018)
6. Tang, Z., Liu, J., Cai, Y.W., Wang, J., Yin, Y.: A wideband differentially fed dual-polarized stacked patch antenna with tuned slot excitations. IEEE Trans. Ant. Propag. **66**(4), 2055–2060 (2018)
7. Wu, J., Yin, Y., Wang, Z., Lian, R.: Dual-band circularly polarized antenna with differential feeding. Progr. Electromagn. Res. **49**, 11–17 (2014)
8. Aboualalaa, M., Abdel-Rahman, A.B., Allam, A., Elsadek, H., Pokharel, R.K.: Design of a dual-band microstrip antenna with enhanced gain for energy harvesting applications. IEEE Ant. Wirel. Propag. Lett. **16**, 1622–1626 (2017)
9. Niotaki, K., Kim, S., Jeong, S., Collado, A., Georgiadis, A., Tentzeris, M.M.: A compact dual-band rectenna using slot-loaded dual band folded dipole antenna. IEEE Ant. Wirel. Propag. Lett. **12**, 1634–1637 (2013)
10. Wang, M., Fan, Y., Yang, L., Li, Y., Feng, J., Shi, Y.: Compact dual-band rectenna for RF energy harvest based on a tree-like antenna. IET Microwaves Ant. Propag. **13**(9), 1350–1357 (2019)
11. Chandravanshi, S., Sarma, S.S., Akhtar, M.J.: Design of triple band differential rectenna for RF energy harvesting. IEEE Trans. Ant. Propag. **66**(6), 2716–2726 (2018)

12. Eltresy, N.A., Dardeer, O.M., Al-Habal, A., Elhariri, E., Abotaleb, A.M., Elsheakh, D.N., Khattab, A., Taie, S.A., Mostafa, H., Elsadek, H.A., Abdallah, E.A.: Smart home IoT system by using RF Energy Harvesting. Preprints 2020070280 (2020)
13. Boursianis, A.D., Papadopoulou, M.S., Koulouridis, S., Rocca, P., Georgiadis, A., Tentzeris, M.M., Goudos, S.K.: Triple-band single-layer rectenna for outdoor rf energy harvesting applications. Sensors **21**, 3460 (2021). https://doi.org/10.3390/s21103460
14. Girase, N., Tiwari, R., Sharma, A., Singh, H.: Design and simulation of slotted rectangular microstrip patch antenna. Int. J. Comput. Appl. **103**, 19–23 (2014). https://doi.org/10.5120/18294-9264

Sensitivity Enhancement of Platinum Diselenide Based SPR Sensor Using Titanium Dioxide as Adhesion Layer

Rashmi Tiwari, Sachin Singh, Pooja Lohia, and D. K. Dwivedi

Abstract This paper reports a theoretical representation of SPR biosensor consisting of a two-dimensional transition dichalcogenides material, gold as a plasmonic metal, calcium difluoride (CaF_2) glass prism and an adhesive material titanium dioxide (TiO_2). The performance of proposed biosensor is assessed in terms of its performance parameters; sensitivity (S), Figure of merit (FOM), Limit of detection (LOD), FWHM (full width half maximum) and DA (detection accuracy). The highest sensitivity noted in this work is $218^0 RIU^{-1}$ which is 4.80% greater than the sensitivity noted for SPR biosensor without adhesive material i.e., without TiO_2. The sensitivity of proposed structure without titanium dioxide is $208^0 RIU^{-1}$. Figure of merit, FWHM, limit of detection and detection accuracy of the proposed structure is 53.322 RIU^{-1}, 4.087^0, 4.587×10^{-6} and 0.2446 deg^{-1} respectively.

Keywords Plasmons · Sensitivity · Limit of detection · Resonance · Adhesion layer

1 Introduction

For the applications in the field of biosensing, the method widely known is surface plasmon resonance (SPR) technique. Since the last four decades SPR based biosensors are used for disease diagnostics, food safety and for environmental monitoring [1, 2]. The features such as fast, accurate, label-free, reliable, and cost-effective processes has made SPR based biosensors most favourable worldwide [2]. The standard structure of surface plasmon based biosensors contains a metal, a prism, and analyte. The SPR technique involves excitation of surface plasmons (SPs) or surface

R. Tiwari · S. Singh · D. K. Dwivedi (✉)
Department of Physics and Material Science, Amorphous Semiconductor Research Lab, Madan Mohan Malaviya University of Technology, Gorakhpur 273010, India
e-mail: todkdwivedi@gmail.com

P. Lohia
Department of Electronics and Communication Engineering, Madan Mohan Malaviya University of Technology, Gorakhpur 273010, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_23

plasmon waves (SPWs). Resonating free electrons inside the metal are called surface plasmons, it is excited when p-polarized light falls at the metal–dielectric interface [3]. Single metal based SPR biosensors were unable to enhance the performance of sensors, this disadvantage gave researchers an opportunity to demonstrate that the biosensor performance can be enhanced by integrating different 2D materials over the standard structure of sensor such as TMDs, black phosphorous, blue phosphorene, MXene, and antimony [2] etc. Platinum diselenide is new emerging group-10 TMDC material. Oxide materials like, titanium dioxide (TiO2) zinc oxide (ZnO), and silicon dioxide (SiO2) have high refractive index, the thin layer of these oxides can give a strong SPR effect when combine with the metal of biosensor forming bimetallic layer. The bimetallic layer is responsible for strong SPR effect [4–8], this layer helps in the enhancement of sensitivity because more surface plasmons are generated at the metal–dielectric interface [7]. In this paper, TiO2 is used as an adhesion layer fixed between metal and prism with platinum diselenide directly in contact with metal.

2 Design and Mathematical Formulation of SPR Sensor

2.1 Theoretical Structure Overview

The proposed SPR sensor operates at the wavelength of 633 nm (λ). The proposed SPR biosensor consists of CaF_2 prism, TiO_2, Au, and $PtSe_2$. In the standard biosensor the metal layer is directly attached to base of the prism while in the proposed SPR biosensor we have introduced an adhesion layer named TiO_2 (Titanium dioxide) in between the metal deposit and the prism. The use of adhesive layer on the prism can overcome the problem of low adhesiveness of silicon based SPR biosensor which is responsible for the degradation of sensitivity of sensor, TiO_2 have real and high refractive index which improves the plasmonic effect [6, 7]. The considered SPR biosensor is shown in Fig. 1 while Table 1 summarizes the associated SPR structures. $PtSe_2$ has a higher optical absorption coefficient, charge transfer, work function, and mobility than conventional 2D nanomaterials. These properties aid in the prevention of oxidation and provide advantageous electrical and optical properties for sensing applications.

Further, a 2D nanomaterial PtSe2 is coated over the Au film surface to enhance the sensitivity of the biosensor which is directly connected to the analyte or sensing medium. Sensing medium is the final layer of the proposed biosensor, it is an aqueous solution consist of biomolecules which must be sensed. RI of the analyte for the biosensor is 1.330, the interaction between the biomolecules and PtSe2 leads to a change in RI of sensing medium, i.e., 1.330 $+\Delta$ns, where Δns $= 0.005$ is the modification in refractive index of the sensing analyte. The details regarding the layers of proposed sensor are given in Table 2.

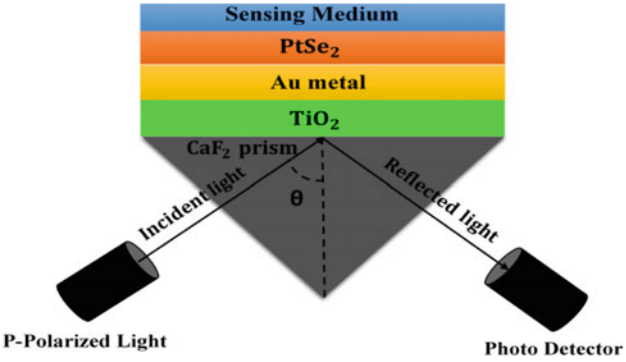


Fig. 1 Proposed SPR sensor structure

Table 1 Layer arrangement of the SPR sensor structure

No. of structure	Sensor type	Arrangement of multi layers
Structure-1	Conventional SPR	$CaF_2/Ag/PtSe_2/SM$
Structure-2	Proposed SPR	$CaF_2/TiO_2/Ag/PtSe_2/SM$

Table 2 Parameters of SPR sensor structure

No. of layers	Materials	RI for λ = 633 nm	d (nm)	References
Layer-I	CaF_2 Prism	1.4329	–	[9]
Layer-II	TiO_2	2.5837	10	[2]
Layer-III	Gold	0.1726 + 3.4220i	50	[10]
Layer-IV	$PtSe_2$	2.9189 + 0.9593i	2	[11]
Layer-V	SM	1.330 to 1.335	–	

2.2 *Theoretical Structure Overview*

For the calculation of reflectance, TMM (transfer matrix method) [12] is considered. Tangential components of EM (electromagnetic) fields are:

$$\left[\frac{P_1}{H_1}\right] = B_{ij}\left[\frac{P_{N-1}}{H_{N-1}}\right] \tag{1}$$

where, P_1 & $P_{N\text{-}1}$, and H_1 & $H_{N\text{-}1}$ are the tangential components of electric and magnetic fields at 1^{st} and N^{th} boundary, respectively. The characteristic matrix B for joined structure is specified as:

$$B_{ij} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \prod_{k=2}^{N-1} B_k, \tag{2}$$

Here,

$$B_k = \begin{bmatrix} \cos\beta_k & \frac{-i(\sin\beta_k)}{q_k} \\ -iq_k \sin\beta_k & \cos\beta_k \end{bmatrix} \tag{3}$$

where,

$$q_k = \left(\frac{\mu_k}{E_k}\right)^{\frac{1}{2}} \cos\theta_k = \frac{\left(E_k - n_1^2 \sin^2\theta_1\right)^{\frac{1}{2}}}{E_k} \tag{4}$$

Here,

$$\beta_k = \frac{2\pi}{\lambda} n_k \cos\theta_k (z_k - z_{k-1}) = \frac{2\pi d_k}{\lambda}\left(E_k - n_1^2 \sin^2\theta_1\right)^{\frac{1}{2}} \tag{5}$$

where, n_1, λ, E_k, θ_1 and μ_k are prism RI, wavelength of incident light, dielectric constant, incident angle and permeability of the k^{th} layer, respectively. The net amplitude of reflectivity coefficient, R_P, for p-polarized light is given by:

$$R_P = \left| \frac{(B_{11} + B_{12}q_N)q_1 - (B_{21} + B_{22}q_N)}{(\mathbf{B}_{11} + B_{12}q_N)q_1 + (B_{21} + B_{22}q_N)} \right|^2 \tag{6}$$

2.3 *Performance Parameters*

The performance of a SPR biosensor is examined by some performance parameters. In SPR sensor the sensitivity depends upon resonance angle, $\Delta\theta_{res}$, and refractive index of analyte, Δn_s. Sensitivity (S) is defined as:

$$S = \frac{\Delta\theta_{res}}{\Delta n_s} \tag{7}$$

The detection accuracy (DA) is defined as:

$$DA = \frac{1}{FWHM} \tag{8}$$

Here, FWHM is full width half maximum and is determined by reflectance curve and its unit is in degree.

Figure of merit (FOM) is defined as the product of sensitivity and detection accuracy or ratio of sensitivity to FWHM, i.e.

$$FOM = S \times DA \tag{9}$$

The limit of detection abbreviated as LOD is evaluated for infinitesimal change in sensing analyte which is used to measure the concentration quantitatively of biomolecules in the last layer of biosensor i.e., sensing medium [2], it is specified as:

$$LOD = \frac{\Delta n_s}{\Delta \theta_{res}} \times 0.001^0 \quad (10)$$

3 Results and Discussions

Reflectance curve of structure 1 and structure 2 before the absorption i.e., when refractive index of analyte is 1.330 and after the absorption i.e., when refractive index of analyte is 1.335 are illustrated in Fig. 2(a) & (b) respectively.

Calculated performance parameters of both the structures are arranged in Table 3. From the Table 3, we can analyze that change in RI of sensing medium from 1.330 to 1.335 has shifted the dip in reflectance curve towards higher SPR angle value in both the structure. From Table 3, It has been found that the sensitivity of SPR biosensor without TiO_2 (structure 1) and SPR biosensor with TiO_2 (structure 2) is $208^0 RIU^{-1}$ and $218^0 RIU^{-1}$ respectively. The FOM of structure 2 is higher than structure 1.

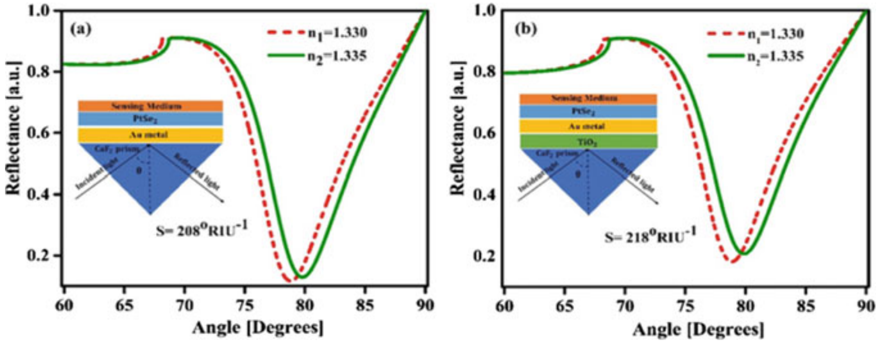


Fig. 2 Reflectance curve before and after the variation in RI of analyte **a** without TiO_2 **b** with TiO_2

Table 3 Performance parameters of each structure

SPR sensor structure	$\Delta\theta_{res}$	Sensitivity ($^O RIU^{-1}$)	FWHM (^{O})	FOM (RIU^{-1})	LOD ($\times 10^{-6}$)	DA (deg^{-1})
Structure 1 (without TiO_2)	1.04^0	208	4.1487	50.128	4.807	0.2410
Structure 2 (with TiO_2)	1.09^0	218	4.08754	53.322	4.587	0.2466

Thus, the SPR biosensor with TiO_2 has the highest sensitivity with high value of figure of merit. Therefore, proposed structure of biosensor shows sensitivity enhancement of 4.80% with respect to proposed structure of biosensor without TiO_2. Figure 3 shows an optimization of different thicknesses of Au metal where at different thicknesses of Au metal value of SPR angle, R_{min}, and sensitivity is plotted. Similarly, Fig. 4 shows the normalized transverse magnetic field intensity plot normal to the prism interface for proposed SPR structure.

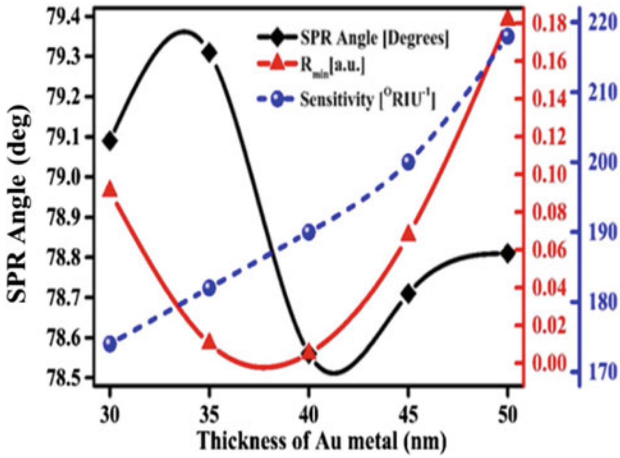


Fig. 3 Study of Au layer thickness to obtain highest sensitivity & maximum SPR angle

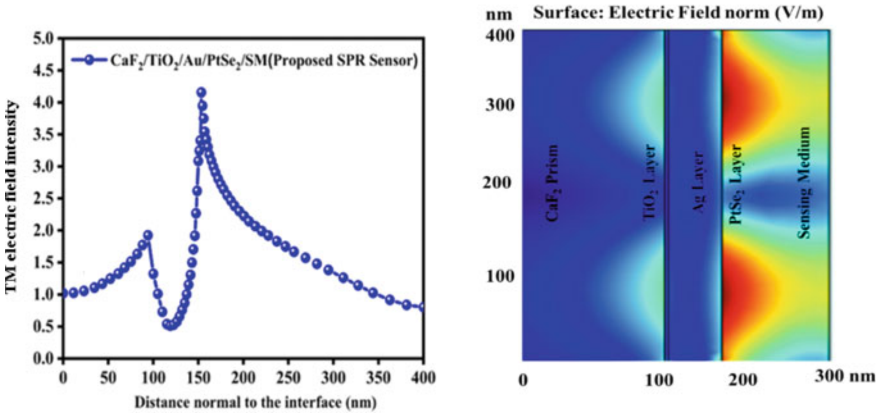


Fig. 4 Plot of TM field intensity variation with distance normal to prism interface for suggested SPR structure

4 Conclusion

A surface plasmon resonance-based biosensor is proposed in this work. The SPR sensor presented in this work contains CaF_2 prism, gold metal, platinum diselenide layer with an adhesive layer of titanium dioxide in between prism and metal. Sensitivity of biosensor without and with TiO_2 was calculated to be 208^0RIU^{-1} and 218^0RIU^{-1} respectively. As compared to the biosensor without the adhesive layer TiO_2 biosensor with the adhesive layer in between glass prism and metal shows 4.80% of sensitivity enhancement.

References

1. Raikwar, S., Srivastava, D.K., Saini, J.P., Prajapati, Y.K.: 2D-antimonene-based surface plasmon resonance sensor for improvement of sensitivity. Appl. Phys. A. **127**(2), 1–8 (2021)
2. Homola, J., Yee, S.S., Gauglitz, G.: Surface Plasmon resonance sensors. Sens. Actuators B Chem. **54**(1–2), 3–15 (1999)
3. Singh, S., Sharma, A.K., Lohia, P., Dwivedi, D.K.: Theoretical analysis of sensitivity enhancement of surface plasmon resonance biosensor with zinc oxide and blue phosphorus/MoS2 heterostructure. Optik **244**, 167618 (2021)
4. Xu, J., et al.: Efficiency enhancements in Ag nanoparticles-SiO2–TiO2 sandwiched structure via Plasmonic effect enhanced light capturing. Nanoscale Res. Lett. **8**(1), 1–5 (2013)
5. Maurya, J.B., Prajapati, Y.K., Singh, V., Saini, J.P.: Sensitivity enhancement of surface Plasmon resonance sensor based on graphene–MoS2 hybrid structure with TiO2–SiO2 composite layer. Appl. Phys. A **121**(2), 525–533 (2015)
6. Singh, S., Prajapati, Y.K.: TiO2/gold-graphene hybrid solid core SPR based PCF RI sensor for sensitivity enhancement. Optik **224**, 165525 (2020)
7. Mahfuz, M.A., Hossain, M., Haque, E., Hai, N.H., Namihira, Y., Ahmed, F.: A bimetallic-coated, low propagation loss, photonic crystal fiber based Plasmonic refractive index sensor. Sensors **19**(17), 3794 (2019)
8. Pal, S., Prajapati, Y.K., Saini, J.P.: Influence of graphene's chemical potential on SPR biosensor using ZnO for DNA hybridization. Opt. Rev. **27**(1), 57–64 (2020)
9. Singh, S., et al.: 2D Nanomaterial based surface Plasmon resonance sensors for biosensing applications. Micromachines **11**(8), 779 (2020)
10. Pal, S., Prajapati, Y.K., Saini, J.P., Singh, V.: Sensitivity enhancement of metamaterial based surface Plasmon resonance biosensor for near infrared. Opt. Appl. **46**(1), 131–143 (2016)
11. Rahman, M.M., Rana, M.M., Rahman, M.S., Anower, M.S., Mollah, M.A., Paul, A.K.: Sensitivity enhancement of SPR biosensors employing heterostructure of PtSe2 and 2D materials. Opt. Mater. **107**, 110123 (2020)
12. Singh, M.K., Pal, S., Prajapati, Y.K., Saini, J.P.: Sensitivity improvement of surface Plasmon resonance sensor on using BlueP/MoS2 heterostructure and antimonene. IEEE Sens. Lett. **4**(7), 1–4 (2020)

Aluminum Nitride Grating Based Plasmonic Sensor Utilizing Enhanced Absorption

Ankit Kumar Pandey

Abstract Aluminum nitride gratings based plasmonic sensor is proposed for refractive index sensing application in near infrared spectral region. The absorbance analysis is performed utilizing rigorous coupled wave analysis (RCWA). The performance parameters are calculated based on the absorbance variation with wavelength. Further, the effect of grating parameters on absorbance properties of the proposed structure is also demonstrated. A sensitivity of 480 nm/RIU is calculated with significant average detection accuracy of 0.028 nm^{-1}. Owning to a large absorbance (~**98**%), the proposed structure can also be used as a perfect absorber for various optoelectronics applications.

Keywords Aluminum nitride · Absorbance · Plasmons · RCWA · Sensitivity

1 Introduction

The emergence of various optical methods such as microscopy, spectroscopy and plasmonics for biosensing applications leads towards the advancement in optical sensing. Now a day, the integration of miniaturized sensors is possible due to combination of microfluidic technologies with microscopic sensing platforms. The phenomenon of oscillations of conduction electrons on coupling with electromagnetic wave (EMW) is commonly referred as surface plasmon resonance (SPR). Cost effective integration, concurrent and label free detection have made SPR based sensors a preferable choice [1]. The most commonly used coupling method in SPR based sensors is prism coupling (Kretschmann configuration) [1, 2]. High sensitivity and ease of fabrication are the main advantages of prism based SPR sensors [3]. However, limited spatial resolution and integration capabilities are their major drawbacks [4]. The grating-based coupling method is another approach that can be utilized for SPR sensors [5]. The last decade has witnessed an extensive advancement in the plasmonic sensors based on metallic nanostructures [6–8]. Large area

A. K. Pandey (✉)
Department of ECE, Bennett University, Greater Noida, India
e-mail: ankitec1016@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_24

patterning for mass production can be achieved by utilizing templating, nanoimprinting and soft lithography techniques [9, 10]. Grating coupling methods usually employ periodic metallic corrugations. In this context, various grating based SPR sensors have been reported so far [6, 7, 11]. Recently, Arora et al. reported figure of merit enhancement (FOM) utilizing dispersion engineered plasmonic nanostructure [12]. Further, a TiO_2 grating based self-referenced SPR sensor has been reported recently [13]. Cao et al. reported metallic grating enhanced plasmonic sensor with a high sensitivity of 1200 nm/RIU. However, maintaining a fixed incident angle (other than normal incidence) is not an easy task. The normal illumination has an advantage of noise reduction and multiplexing of grating nanostructures with miniaturized devices [14, 15]. Moreover, the geometrical parameters (periodicity, shape, height) play an important role in controlling the resonance characteristics of the surface plasmons (SPs).

In the present work, sensor structure having aluminum nitride (AlN) grating on silicon oxide (SiO_2) substrate with gold (Au) as interlayer is proposed. AlN possess several advantages such as erosion and corrosion resistance, chemically stable (at high temperature also) and uniform microstructure. The analysis is based on absorbance spectra. The operating wavelength range is optical communication band enabling the proposed sensor advantageous from integration view point with optical fiber system. In addition, the availability of source and detector operating within the intended wavelength range will be an advantage.

2 Design Consideration and Theory

Figure 1 represents the schematic of proposed multilayered structure consisting of AlN grating on Au deposited over SiO_2. Wavelength interrogation method is used considering a normal incident p-polarized light. SPs excitation relation (bulk material) for grating structure is given as [16]:

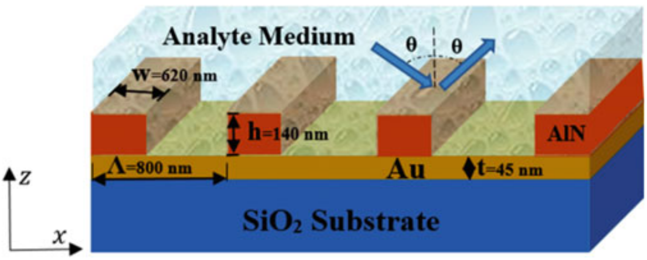


Fig. 1 Structure of proposed SPR sensor. The periodicity, width and height of AlN gratings are represented as ‘Λ’, ‘w’ and ‘h’, respectively. ‘t’ is the thickness of Au layer

$$\frac{2\pi}{\lambda} \times n_s \sin\theta_{inc} + m\frac{2\pi}{\Lambda} = \pm Re\left(\frac{\omega}{c}\sqrt{\frac{\epsilon_m \epsilon_s}{\epsilon_m + \epsilon_s}}\right) \tag{1}$$

In the above expression, the incident light wavelength is λ, the angle of incidence is θ_{inc}, $\omega$and$m$ are the angular frequency and diffraction order, respectively. c is the velocity of light in vacuum. ϵ_m and ϵ_s are the dielectric constant of the metal layer and analyte medium, respectively. The absorbance (A) computation is based on rigorous coupled wave analysis (RCWA) using MATLAB with normal consideration of incident light (i.e., $\theta_{inc} = 0°$) [17]. The resonance wavelength (λ_{SPR}) is the point at which the maximum absorbance is attained. A change in position of λ_{SPR} ($\Delta\lambda_{SPR}$) is experienced on small change in analyte's RI (Δn_s). Experimental dielectric constant values of AlN, Au and SiO_2 are taken from ref [18–20].

Generally, the performance of grating based SPR sensor is theoretically calculated using detection accuracy (D.A.) and sensitivity (S) [13]. These parameters can be defined as follows:

$$D.A. = {}^{1}/_{(FWHM)};\ S = \left({}^{\Delta\lambda_{SPR}}/_{\Delta n_s}\right)\left(in\frac{nm}{RIU}\right) \tag{2}$$

Here, FWHM is the full width at half maximum, i.e., width of SPR curve. D. A. is measured in 1/nm.

3 Results and Discussions

3.1 AlN Grating on Au with SiO_2 Substrate for SPR Sensing

For preliminary calculation of absorbance, the values of 'w', 'Λ', 'h' and 't' are set to be 500 nm, 800 nm, 140 nm and 45 nm, respectively. Initially, pure water (H_2O) is considered as analyte with wavelength-dependent RI values adapted from [21]. Figure 2(a) shows the variation of absorbance curve (A *vs.* λ) analyte medium i.e., H_2O filled grooves. Figure 2(b) shows the magnetic field strength at resonance condition. During simulations, $\theta_{inc} = 0°$ and Floquet periodicity are considered. The maximum absolute value of magnetic field is 757 A/m (towards analyte side) at corresponding $\lambda_{SPR} = 1344.70$ nm which reaffirms the maximum absorbance (A_{max}) at resonance ($A_{max} = 0.9848$). Generally, two peaks are generated in case of such structures with dielectric gratings [22]. However, the second peak (observed for incident angle > 0°) which is considered as substrate mode as discussed in Fig. 6.

The sensing applicability of proposed sensor for analyte RI range is shown in Fig. 3. It is assumed that the analyte has negligible dependency on wavelength. The corresponding λ_{SPR} values for different n_s values of 1.33, 1.34, 1.35 1.36 and 1.37, are 1349.70, 1354.50, 1359.30, 1364.10, and, 1368.90 nm, respectively. Thus, a significant shift is observed in absorbance curve (*i.e.*, $\Delta\lambda_{SPR} = 19.20$ nm) for a small Δn_s (*i.e.*, 0.04 RIU). This shifting indicates towards achieving an average sensitivity

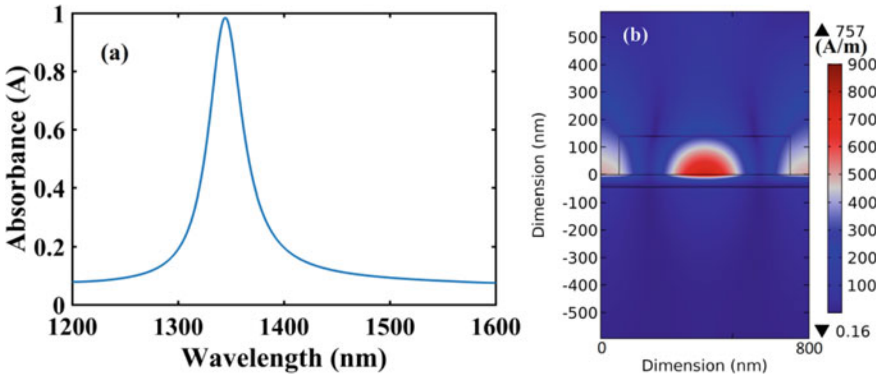


Fig. 2 **a** Simulated absorbance curve for proposed sensor at normal incidence (p-polarized). Here, the wavelength dependent RI of H_2O is taken into consideration. **b** The magnetic field strength at resonance ($\lambda = 1344.70$ nm). The FWHM is 39.0 nm

of 480 nm/RIU with an average FWHM value of 35.32 nm. The obtained value of FWHM value is smaller than previously reported work (FWHM = 55 nm) [12] based on grating scheme which direct towards a high D.A value insuring an accurate sensing behavior of the proposed sensor.

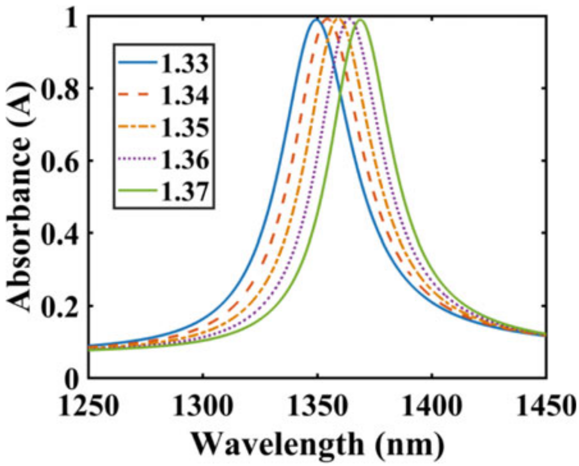


Fig. 3 Simulated absorbance curves for analyte RI range (1.33–1.37) at $\theta = 0°$. Assuming analyte has negligible wavelength dependence RI variation within the considered spectral region

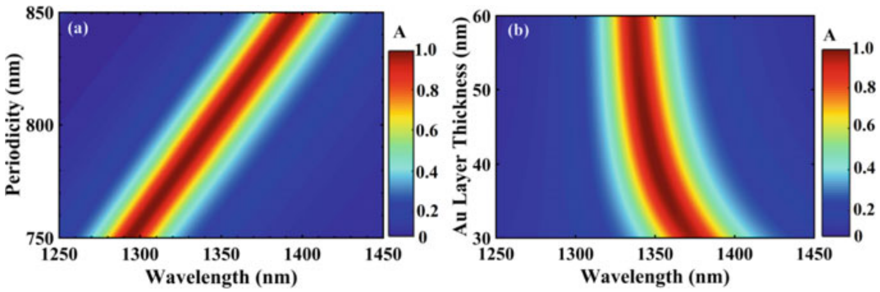


Fig. 4 Simulated 2D absorption spectra variation with wavelength for **a** different periodicity (t = 45 nm) of AlN grating and **b** different t values (Λ = 800 nm). In this case, h = 140 nm and w = 500 nm

3.2 *Effect of AlN Periodicity (Λ) and Gold Layer Thickness (t) Over SiO_2 Substrate*

Figure 4 shows the two-dimensional (2D) variation of absorbance with respect to wavelength for different Λ (750–850 nm) and t (30–50 nm) values keeping other parameters fixed. The analyte is H_2O. It is evident from Fig. 4(a) that, increase in Λ causes a shift in position of λ_{SPR} (identified by maximum A) towards longer wavelengths. Thus, a linear relationship is obtained which can be related with Eq. (1). Further, the FWHM value increases from 36.76 nm (for $\Lambda = 750$ nm) to 38.41 nm (for $\Lambda = 850$ nm). Thus, Λ value in the vicinity of 800 nm may be preferred considering λ_{SPR} position as well as FWHM. Further, A *vs* λ curves are simulated for different thickness of thin Au layer over SiO_2 in order to study its effect on resonance peaks as shown in Fig. 4(b). The absorbance peak shift from $\lambda = 1336.79$ nm (for t = 60 nm) to $\lambda = 1373.51$ nm (for t = 30 nm). The t value in the range 35–50 nm will be the suitable choice in view of high absorbance and operating wavelength.

3.3 *Effect of AlN Grating Width (W) and Height (h)*

In this section the theoretical investigation is performed on the effect of width and height of AlN grating on resonance peaks as shown in Fig. 5 for H_2O analyte. Increasing w from 580 to 660 nm has a small role to play in shifting of resonance peaks corresponding to maximum absorbance as shown in Fig. 5(a). In addition, the sensitivity remains nearly fixed (480 nm/RIU; for analyte RI ~1.33) in the w range (580–660 nm). However, FWHM value decreases from 40.52 nm (for w = 580 nm) to 35.01 nm (for w = 660 nm). Thus a w range of 600–640 nm is preferable in view of maximum absorbance and minimum FWHM.

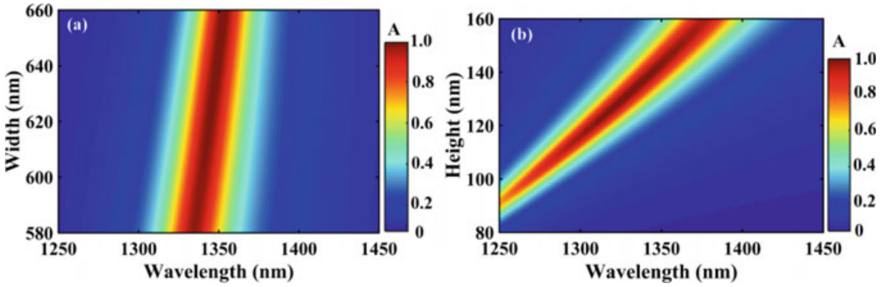


Fig. 5 Simulated absorption spectra with wavelength for **a** different 'w' of AlN grating (h = 140 nm) and **b** different 'h' of AlN grating (w = 500 nm). The λ dependent RI of H_2O is taken as analyte. Λ and t are fixed at 800 and 45 nm respectively. The color bar is showing the absorbance

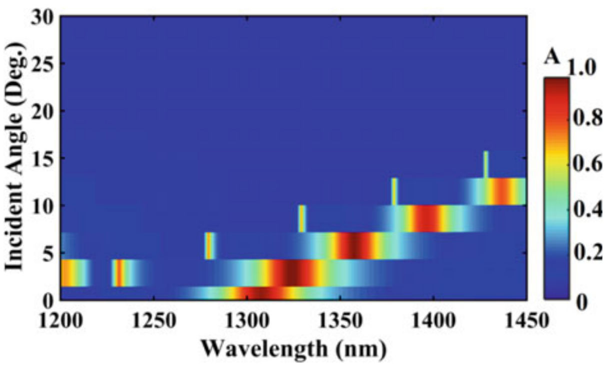


Fig. 6 Simulated absorption spectra with wavelength at different incident angles considering H_2O as analyte. Here, h = 140 nm, w = 500 nm and Λ = 800 nm

Further, it is apparent from Fig. 5(b) that, increasing height of AlN grating (from 80 to 160 nm) keeping other dimensional parameters fixed causes a significant red-shift in the resonance wavelength. Another observation is that the absorbance has maximum value for h range of 110–160 nm. However, on calculation it is obtained that there is a large increase in FWHM value from 21.60 nm (h = 80 nm) to 46.82 nm (h = 160 nm). This increment in FWHM will affect the D.A. Thus, there is a trade of between maximum absorbance and FWHM which can be balanced by selecting the h value as 140 ± 10 nm.

3.4 *Effect Incident Angle (Θ)*

The role of incidence angle is demonstrated in this section. Although, normal incidence light is the preferred choice in most of the cases, effect of incidence angle on

Table 1 Sensitivity comparison

References	Modalities	Maximum S (nm/RIU)
Cinel et al. 2012 [23]	Ag nano-disk arrays	354
Guner et al. 2017 [24]	Au/Ag bi-grating	356
López-muñoz et al. 2017 [25]	Nanostructured chip with Au layer	425
He et al. 2019 [26]	Transmit-array metasurface	192
Proposed scheme	AlN gratings on Au/SiO_2	480

absorbance is depicted in Fig. 6. It is clear that, λ_{SPR} which corresponds to maximum absorbance shifts towards higher wavelength. In addition, the maximum normalized absorbance value decreases from 0.9848 (at 0°) to 0.777 (at 11.46°). Further increase in θ will cause significant reduction in absorbance. It is worthwhile to mention that two absorbance peaks are obtained (as discussed in previous section) which provide significantly less absorbance.

Table 1 presents a comparative analysis of proposed sensor's sensitivity performance with some of the nanostructure based research works. It is evident that the proposed structure is able to provide highest sensitivity.

4 Conclusion

AlN grating on Au/SiO_2 for refractive index sensing in NIR spectral region has been proposed. Apart from a significant sensitivity of 480 nm/RIU, a reasonable average FWHM value is also obtained which is an important factor indicating the preciseness of the sensor. Effects of various grating variable on absorbance are also studied. It is evident that the absorbance behavior is highly affected by the periodicity and the height of AlN grating. The proposed structure with grooves filled with dielectric material (RI = 1.33–1.37) can also be utilized as an absorber in NIR spectral region for various optoelectronics applications.

References

1. Homola, J., Yee, S.S., Gauglitz, G.: Surface plasmon resonance sensors: review. Sens. Actuators B Chem. **54**, 3–15 (1999)
2. Pandey, A.K., Sharma, A.K., Basu, R.: Fluoride glass based surface plasmon resonance sensor in infrared region: performance evaluation. J. Phys. D: Appl. Phys. **50**, 185103 (2017)
3. Sharma, A.K., Pandey, A.K.: Blue Phosphorene/MoS2 Heterostructure based SPR Sensor with Enhanced Sensitivity. IEEE Photonics Technol. Lett. **30**, 595–598 (2018)
4. Tetz, K.A., Pang, L., Fainman, Y.: High-resolution surface plasmon resonance sensor based on linewidth-optimized nanohole array transmittance. Opt. Lett. **31**, 1528–1530 (2006)
5. Sharma, A.K., Pandey, A.K.: Design and analysis of plasmonic sensor in communication band with gold grating on nitride substrate. Superlattices Microstruct. **130**, 369–376 (2019)

6. Valsecchi, C., Brolo, A.G.: Periodic metallic nanostructures as plasmonic chemical sensors. Langmuir **29**, 5638–5649 (2013)
7. Strobbia, P., Languirand, E., Cullum, B.M.: Recent advances in plasmonic nanostructures for sensing: a review. Opt. Eng. **54**, 100902 (2015)
8. Pandey, A.K., Sharma, A.K.: Advancements in grating nanostructure based plasmonic sensors in last two decades: a review. IEEE Sens. J. **21**, 12633–12644 (2021)
9. Odom, T.W., Love, J.C., Wolfe, D.B., Paul, K.E., Whitesides, G.M.: Improved pattern transfer in soft lithography using composite stamps. Langmuir **18**, 5314–5320 (2002)
10. Chen, J., Shi, J., Decanini, D., Cambril, E., Chen, Y., Haghiri-Gosnet, A.M.: Gold nanohole arrays for biochemical sensing fabricated by soft UV nanoimprint lithography. Microelectron. Eng. **86**, 632–635 (2009)
11. Chung, T., Lee, S.Y., Song, E.Y., Chun, H., Lee, B.: Plasmonic nanostructures for nano-scale bio-sensing. Sensors **11**, 10907–10929 (2011)
12. Arora, P., Talker, E., Mazurski, N., Levy, U.: Dispersion engineering with plasmonic nano structures for enhanced surface plasmon resonance sensing. Sci. Rep. **8**, 9060 (2018)
13. Sharma, A.K., Pandey, A.K.: Self-referenced plasmonic sensor with TiO2 grating on thin Au layer: simulated performance analysis in optical communication band. J. Opt. Soc. Am. B. **36**, F25 (2019)
14. Lesuffleur, A., Im, H., Lindquist, N.C., Lim, K.S., Oh, S.-H.: Laser-illuminated nanohole arrays for multiplex plasmonic microarray sensing. Opt. Express. **16**, 219–224 (2008)
15. Menezes, J.W., Ferreira, J., Santos, M.J.L., Cescato, L., Brolo, A.G.: Large-area fabrication of periodic arrays of nanoholes in metal films and their application in biosensing and plasmonic-enhanced photovoltaics. Adv. Funct. Mater. **20**, 3918–3924 (2010)
16. Roh, S., Chung, T., Lee, B.: Overview of the characteristics of micro- and nano-structured surface plasmon resonance sensors. Sensors. **11**, 1565–1588 (2011)
17. Zhao, B.: Thermal radiative properties of micro/nanostructured plasmonic metamaterials including two-dimensional materials (2016)
18. Loughin, S., French, R.H.: Aluminum nitride (AlN). In: Palik, E.D. (ed.) Handbook of Optical Constants of Solids, pp. 373–401. Academic Press, Burlington (1997)
19. Johnson, P.B., Christy, R.W.: Optical constants of the noble metals. Phys. Rev. B **6**, 4370 (1972)
20. Malitson, I.H.: Interspecimen comparison of the refractive index of fused silica. J. Opt. Soc. Am. **55**, 1205–1209 (1965)
21. Kedenburg, S., Vieweg, M., Gissibl, T., Giessen, H.: Linear refractive index and absorption measurements of nonlinear optical liquids in the visible and near-infrared spectral region. Opt. Mater. Express. **2**, 1588–1611 (2012)
22. Pandey, A.K., Sharma, A.K., Marques, C.: On the application of stacked periodic tungsten grating nanostructure in wide-range plasmonic sensing and other photonic devices. Plasmonics **16**, 9–17 (2020).
23. Cinel, N.A., Bütün, S., Özbay, E.: Electron beam lithography designed silver nano-disks used as label free nano-biosensors based on localized surface plasmon resonance. Opt. Express. **20**, 2587 (2012)
24. Guner, H., Ozgur, E., Kokturk, G., Celik, M., Esen, E., Topal, A.E., Ayas, S., Uludag, Y., Elbuken, C., Dana, A.: A smartphone based surface plasmon resonance imaging (SPRi) platform for on-site biodetection. Sens. Actuators B Chem. **239**, 571–577 (2017)
25. López-muñoz, G.A., Estevez, M., Peláez-gutierrez, E.C., Homs-corbera, A., García-hernandez, M.C., Imbaud, J.I., Lechuga, M.: Biosensors and bioelectronics a label-free nanostructured plasmonic biosensor based on Blu-ray discs with integrated micro fluidics for sensitive biodetection. Biosens. Bioelectron. **96**, 260–267 (2017)
26. He, K., Liu, Y., Fu, Y.: Transmit-array, metasurface-based tunable polarizer and high-performance biosensor in the visible regime. Nanomaterials **9**, 1–9 (2019)

Compact Electromagnetic Band Gap Based Filtenna for Ultra Wide Band Application

N. Vidya and M. H. Masood

Abstract This work represents compact Electromagnetic Band Gap based Filtenna for Ultra-Wide Band Application. Coplanar Waveguide (CPW) is used for higher bandwidth. Additional resonance is minimized with the help of Defected Ground Structure (DGS). Two U slots were etched on the ground. Electromagnetic Band Gap Structure (EBG) is adopted to simplify the antenna design as well as for gain improvement. Operating frequency of the antenna is 6.5 GHz. The return loss, radiation pattern and antenna gain are studied. Designed antenna provides return loss about −18 dB and antenna gain is achieved up to 3.4 dBi. The bandwidth is increased to 18% compared to conventional antennas. The measured results agree well with the simulated results. The prototype has return loss about −8.5 dB. The presented method demonstrates the capability of achieving both good filtering and radiating performance in simple structures by etching slots on the patch of a conventional patch antenna. This filtenna is suitable for transceivers operating in ultra-wide band.

Keywords Conventional patch · DGS · EBG · Filtenna · U shaped slots

1 Introduction

Demand for miniaturized compact devices has been increased. Combining devices gives added advantage too. In RF and microwave devices, antenna is the device which resonates/radiates in the particular frequency whereas filter is the device which pass or stop the particular frequency band. combining separate antenna with filter degrade antenna performance. And it will lead to additional losses because of the transmission line connected between antenna and filter. Uniting these two in single PCB will improve performance and cost-effective [1]. Changing last element of the filter as radiating patch of antenna is known as filter synthesis approach. Space occupancy is main drawback in this method [2].

N. Vidya (✉) · M. H. Masood
Department of Electronics Engineering, Madras Institute of Technology, Chennai, Tamilnadu, India
e-mail: nk.vidya1@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_25

In recent days many novel filtering structures are developed using microstrip patch antenna. Because of compactness and low cost, microstrip antennas are most preferable, even though it has lower band width and gain [4]. Here we concentrate on coplanar waveguide antenna with microstrip feeding, for bandwidth enhancement. CPW has low Quality factor (Q-factor) compared to microstrip antenna.

The bandwidth and Q factor are indirectly proportional. Bandwidth is high in lower Q- Factor [5]. For a compact structure normal ground is converted to Defected Ground Structure. Calculatedly creating slots or shapes in the ground plane of an antenna is referred as DGS [3]. Surface waves are unwanted radiation of antenna in ground plane instead of free space which reduce antenna performance, it also reduces the gain and bandwidth of antenna. In order to decrease the surface wave effects and to get higher bandwidth EBG structures are integrated in the patch.

Periodic structures embedded in the antenna substrate is known as Electromagnetic Band Gap structure, which has the ability to change the antenna's electromagnetic properties. In some frequency bands, they suppress surface waves and improve performance. Frequency band ranges from 3.1 to 10.6 GHz are known as ultra-wide band. Its data rate (bandwidth) is greater than 500 Mhz. In ultra-wide band EBG gives notching effect and improve gain of the antenna [6]. EBG structure in microstrip can improve bandwidth and it will reduce the size of the antenna [7]. In recent years novel compact filtenna without additional circuit have more attraction [8–15].

Filtenna is designed to operate in 6.5 GHz. Two U slots were etched on the coplanar ground for notching effect. And 3×3 square EBG patches were embedded on the antenna to suppress surface waves and to improve antenna performance.

2 Filtenna Design

2.1 Filtenna Configurations

Notching effect is produced in the sidebands to get filtering effect in antenna. 3×3 square EBGs are designed on the top of the FR4 substrate. A FR4 substrate with relative permittivity 4.3, loss tangent 0.001 and thickness is about 1.6 mm is used. CPW ground is implanted in front side of the substrate and there will be no separate ground on back side. This made design easy and cost of fabrication also reduced.

Aiming to increase the bandwidth of the antenna, the size of centre square patches are increased slightly. Figure 1 shows the regular microstrip patch antenna with ground on back side and EBG based CPW antenna without separate ground. W1 denotes the width and length of the single square and S is the gap between the square patches shown in Eq. 1 and 2 which are taken in the terms of wavelength [7].

$$\mathrm{W1} = 0.10\,\lambda \tag{1}$$

$$S = 0.02\,\lambda \tag{2}$$

Table 1 shows the dimensions of the proposed filtenna. All the values are represented in millimeter (mm).

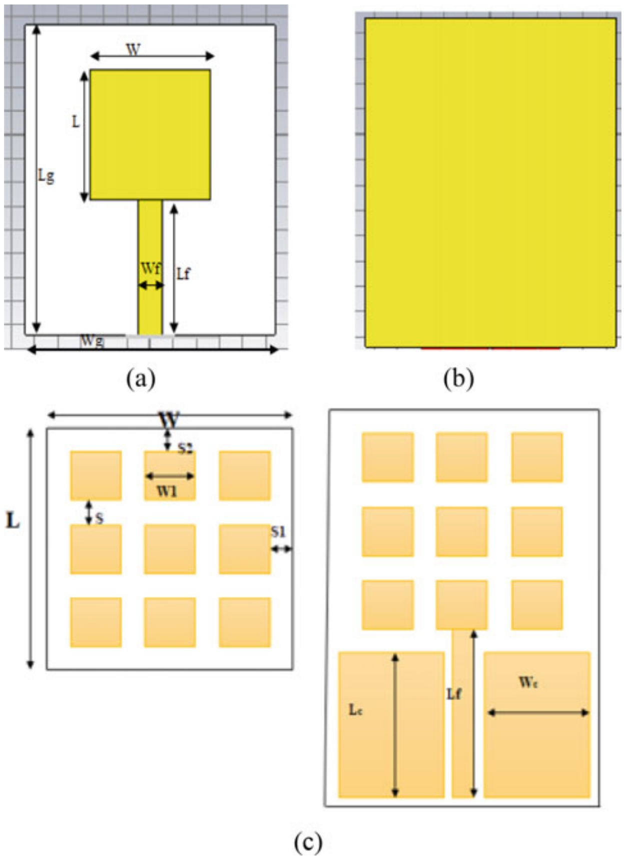


Fig. 1 Geometry of antennas **a** Front View of microstrip patch **b** bottom view of antenna **c** EBG antenna with CPW feed

Table 1 Dimensions of proposed antenna

L	W	Lf	Wf	Wg
10.39	14.04	9.75	1.6	23.26
Lg	S	S1	S2	W1
19.99	2	4	8	7.4
L	Wc	W(cpw)	Lf(cpw)	Lc
34	15	30	12.5	10

Figure 1 Antenna (a) and (c) are designed to operate in 6.5 GHz and simulated using CST Microwave Studio. Figure 2 shows the return loss of the above shown antennas.

When compared to CPW antenna, microstrip patch antenna has multiple resonance and negligibly low band width. The impedance of CPW filtenna is real in 4.2 GHz but the which is designed for 6.5 GHz. Figure 2 shows minor resonance in 6 GHz, by adding slots the additional resonance can be removed and resonant frequency also bring back to 6.5 GHz. Gain pattern of the above shown antenna is displayed in the Fig. 3.

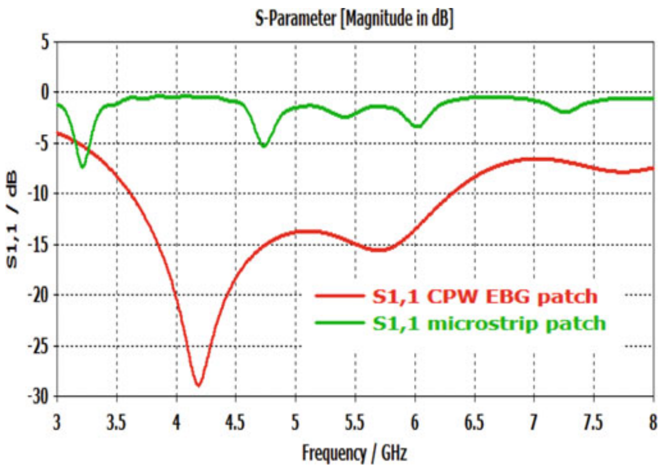


Fig. 2 Simulated results of proposed Antenna with various feed

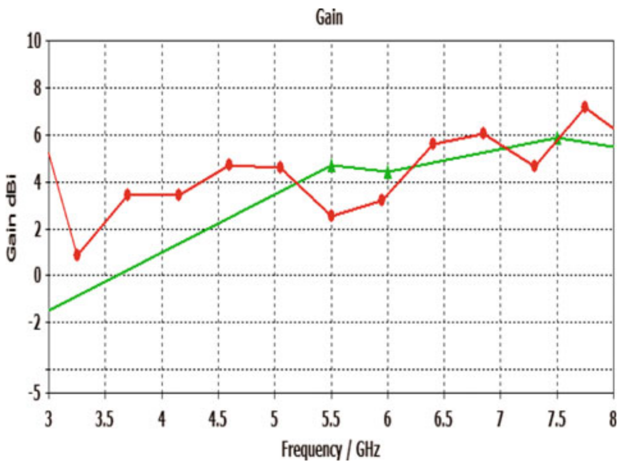


Fig. 3 Gain of proposed Antenna with various feed

The gain is uniform throughout the band. Gain of microstrip and CPW are almost similar, gain at 6.5 GHz is about 4.8 dBi. Bandwidth is negligibly small in microstrip and occupies 4 to 6 GHz in CPW.

2.2 EBG Based Filtenna for UWB Application

Single U-shaped slot is etched in the CPW ground that remove the ripples in the lower frequency band only. Later another U-shaped slot is etched on the ground that removes the ripples throughout the side bands. Figure 4 shows the EBG based filtenna with single and double slots.

For bandwidth improvisation, dimension of center square patch is incresed compared to other squares. Figure 5 shows the return loss of antenna with various slots. The sideband ripples are reduced in the double U slot but return loss is little low compared to single U slot. Single U slot antenna occupies almost 470 MHz at 10 dB impedance bandwidth. Whereas double U slot antenna occupies 550 MHz.

In the lossy FR4 substrate almost 3.4 dBi gain is achieved at resonant frequency 6.5 GHz shown in Fig. 6. Radiation pattern of the antenna is shown in the Fig. 7. Both the filtenna possesses similar radiation characteristics but their main lobe direction varies 10°.

3 Measured Results

Compact Electromagnetic Band Gap based Filtenna for Ultra-Wideband application is fabricated and tested in the antenna design and testing lab of Madras Institute of Technology, Chennai. Entire fabrication takes 60–70 min.

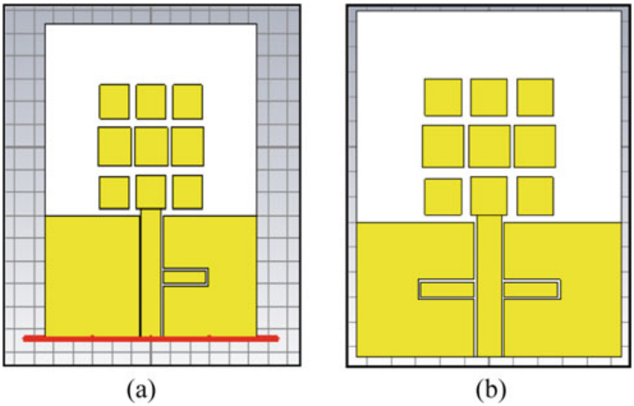


Fig. 4 Designed antennas **a** Single Uslotted DGS **b** Double U sloted DGS

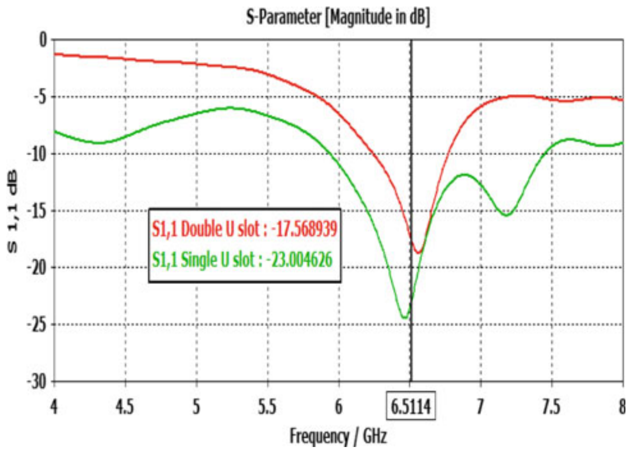


Fig. 5 Return loss of proposed Antenna with single and Double U slot

3.1 Fabrication Procedure

- For fabrication single faced (copper coating is on one side) FR4 substrate is cut with respect to simulated dimension 43 × 54 × 1.6 mm.
- Copper coating on FR4 substrate is secured using insulation tape.
- By soaking antenna in the ferric chloride solution, unwanted copper coatings are removed, it takes duration up to 15 min.
- The insulation tapes are then removed to get the antenna's desired shape and size.

Fabricated antenna is tested using ZVH8 cable and antenna analyzer. This Analyzer can test the devices operating from 100 KHz to 8 GHz. Figure 8 shows the fabricated antenna. And Fig. 9 shows the measurement setup of the antenna. This device can measure the return loss of the antenna.

The fabricated filtenna has additional resonance in sideband because of the fabrication errors, connector mismatch and transmission line (cable) losses. The fabricated antenna resonates in 6.5 GHz as well as in 7.51 GHz. The return loss at 6.5 GHz is −8.52 dB in 7.4 GHz is −18.2 GHz as shown in Fig. 10. Bandwidth of the fabricated antenna is about 19% (3 dB bandwidth at 6.5 GHz) about 12% in 7.5 GHz. Figure 10 shows the return loss of the protype.

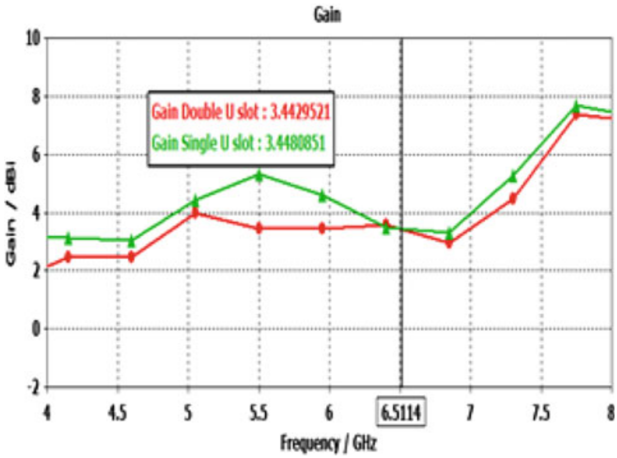


Fig. 6 Gain of the antenna with single and double U slot

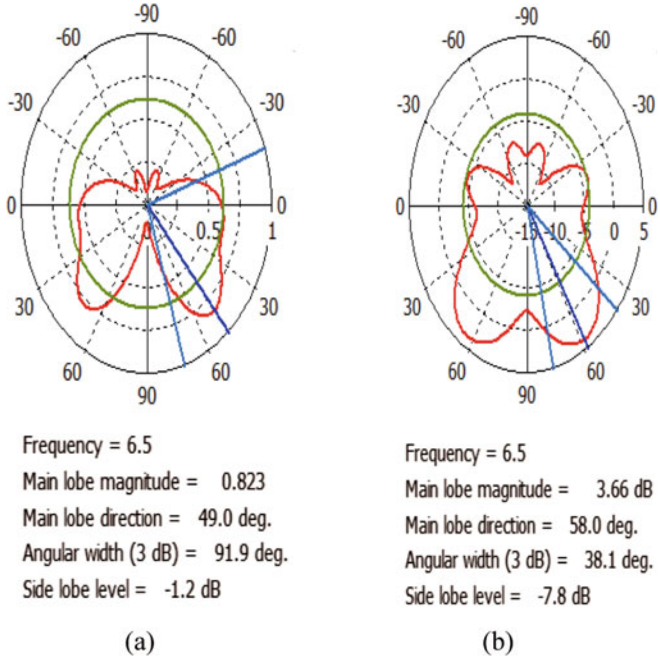


Fig. 7 Radiation pattern of antenna **a** Single U slot and **b** Double U slot

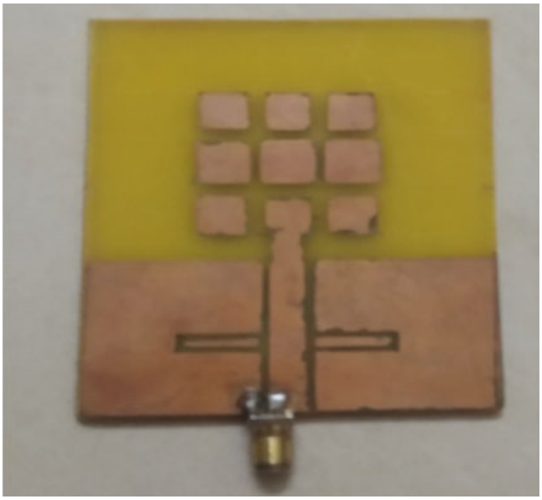


Fig. 8 Prototype of the antenna

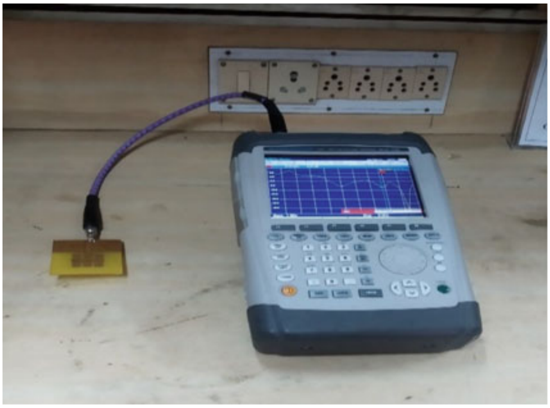


Fig. 9 Measurement setup

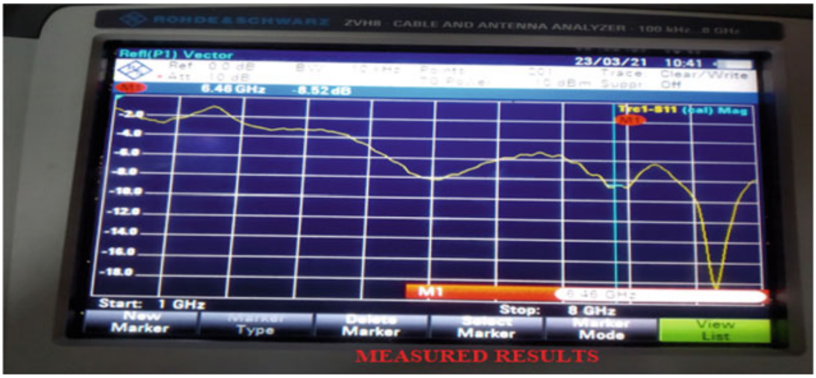


Fig. 10 Measured return loss of the prototype

4 Results and Discussion

When compared to previous works, this filtenna is in compact size, cost effective and easy to fabricate. And the structure is symmetry too. Fabricated prototype also gives good response over the frequency (Table 2).

4.1 Improvement in Structure

The filtenna structure is further improved after fabrication and their characteristics are studied. Primarily the center square patches are replaced with square ring patches. And then entire square patches are converted into square ring patches. Figure 11 shows the improved structures.

For improving the bandwidth of the antenna these changes are done. Figure 12 shows the return loss of the antennas. Return loss is improved in both the cases. Frequency of operation is shifted in the right band. Band width of the single layer square ring patch is increased to 11% compared to the previously designed filtenna.

Converting the entire patches to ring lead to better return loss compared to the previously designed antenna but bandwidth is reduced to 3.75%. Gain pattern also linear over the band.

Figure 13 shows the gain pattern of the antennas gain for center square ring patch is same as square patches but in the second case the gain is decreased. Center square ring patch filtenna has high bandwidth and return loss compared to fabricated antenna. Fabricating this filtenna is easy compared to previous one. Gain also improved.

Table 2 Property comparisons of antenna

	Operational frequency (GHz)	Structure	Size ($\lambda_0 \times \lambda_0$)	Additional network	Fractional bandwidth (%)	Gain (dBi)	Polarization	Return loss (dB)
[1]	2.6	Symmetry	0.8 × 0.8	Antenna itself	23	8.9	Dual	−20
[8]	5	Asymmetry	1.3 × 1.3	Feeding Network	28.4	8.2	Single	−20
[9]	2.45	Asymmetry	1.14 × 1.14	Antenna itself	20	9.74	Single	−19
[10]	3.4	Asymmetry	1.13 × 0.42	Antenna itself	23	6.7	Single	−17
[11]	2.6	Partially Symmetry	1.03 × 1.03	Feeding Network	10.1	9	Dual	−20
This work	6.5	Symmetry	0.5 × 0.43	Antenna itself	Simulated 10.3 Measured 19 (3 dB)	3.44	Single	−17.56 −8.52

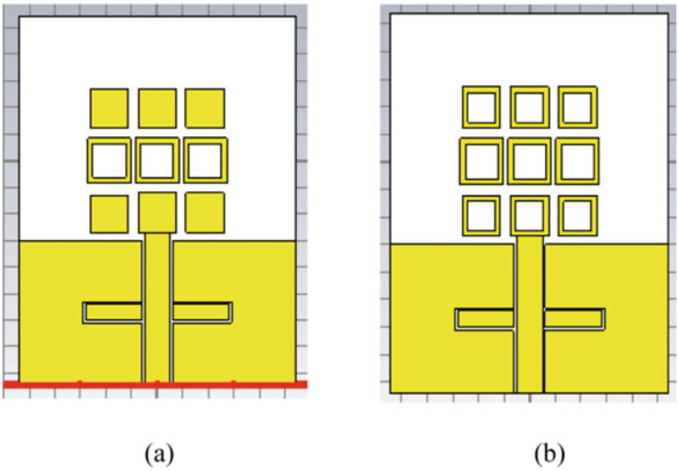


Fig. 11 Improved structure of antenna **a** Single layer square ring and **b** Square ring patch

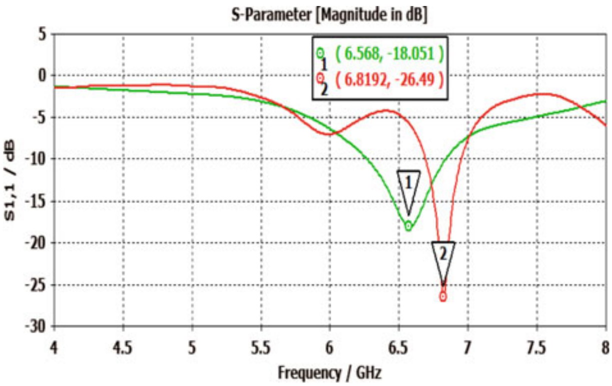


Fig. 12 Simulated results of improved Antenna

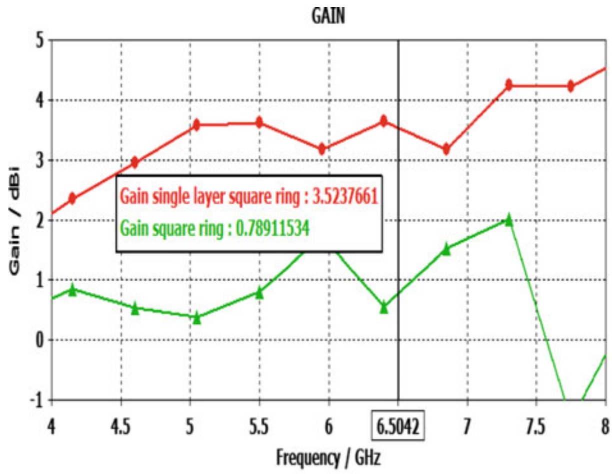


Fig. 13 Gain pattern of antennas

5 Conclusion

This work represents Compact Electromagnetic Band Gap based Filtenna for Ultra-Wideband application. In this work filtering method is combined with antenna. Filtering is achieved by etching some slots on the ground of antenna. In the final electromagnetic Band Gap based CPW fed Ultra-wideband application antenna has return loss about −18 dB. Gain is 3.4 dBi. The wider center square leads to improvement in bandwidth. Band width of the antenna is increased to 550 MHz. The designed antenna is fabricated and tested which has much lower response and it have some additional resonance in nearby frequency because of fabrication losses. The fabricated antenna has −8.5 dB return loss at 6.5 GHZ. The design must be improved so that it can give better response. Size of the antenna can be reduced further more for compactness. And gain of the antenna should be improved. It can be used in wide band wireless devices and application.

References

1. Yang, W., Xun, M., Che, W., Feng, W., Zhang, Y., Xue, Q.: Novel compact high-gain differential-fed dual-polarized filtering patch antenna. IEEE Trans. Ant. Propag. **67**(12), 7261–7271 (2019)
2. Mansour, G., Lancaster, M.J., Hall, P.S., Gardner, P., Nugoolcharoenlap, E.: Design of filtering microstrip antenna using filter synthesis approach. Progr. Electromagn. Res. **145**, 59–67 (2014)
3. Esa, M., Jamaluddin, U., Awang, M.S.: Antenna with DGS for improved performance. In: 2010 IEEE Asia-Pacific Conference on Applied Electromagnetics (APACE), Port Dickson, pp. 1–4 (2010)

4. Huynh, T., Lee, K.F.: Single-layer single-patch wideband microstripantenna. Electron. Lett. **31**(16), 1310–1312 (1995)
5. Balanis, C.A.: Antenna Theory: Analysis and Design. Wiley, Hoboken (2012)
6. Wang, J., Wong, H., Ji, Z., Wu, Y.: Broadband CPW-fed aperture coupled metasurface antenna. IEEE Ant. Wirel. Propag. Lett. **18**(3), 517–520 (2019)
7. Ripin, N., Awang, R.A., Sulaiman, A.A., Baba, N.H., Subahir, S.: Rectangular microstrip patch antenna with EBG structure. In: 2012 IEEE Student Conference on Research and Development (SCOReD), Pulau Pinang, Malaysia, pp. 266–271 (2012)
8. Pan, Y.M., Hu, P.F., Zhang, X.Y., Zheng, S.Y.: A low-profile highgainand wideband filtering antenna with metasurface. IEEE Trans. Ant. Propag. **64**(5), 2010–2016 (2016)
9. Zhang, X.Y., Duan, W., Pan, Y.-M.: High-gain filtering patch antenna without extra circuit. IEEE Trans. Ant. Propag. **63**(12), 5883–5888 (2015)
10. Li, J.-F., Chen, Z.N., Wu, D.-L., Zhang, G., Wu, Y.-J.: Dual-beam filtering patch antennas for wireless communication application. IEEE Trans. Ant. Propag. **66**(7), 3730–3734 (2018)
11. Duan, W., Zhang, X.Y., Pan, Y.-M., Xu, J.-X., Xue, Q.: Dual-polarized filtering antenna with high selectivity and low cross polarization. IEEE Trans. Ant. Propag. **64**(10), 4188–4196 (2016)
12. Lin, C.-K., Chung, S.-J.: A compact simple structured filtering antenna utilizing filter synthesis technique. In: Proceedings of Asia–Pacific Microwave Conference, Yokohama, Japan, pp. 1573–1576 (2010)
13. Zhang, J.-D., Zhu, L., Wu, Q.-S., Liu, N.-W., Wu, W.: A compact microstrip-fed patch antenna with enhanced bandwidth and harmonic suppression. IEEE Trans. Ant. Propag. **64**(12), 5030–5037 (2016)
14. Chen, X., Zhao, F., Yan, L., Zhang, W.: A compact filtering antenna with flat gain response within the passband. IEEE Ant. WirelessPropag. Lett. **12**, 857–860 (2013)
15. Zhang, X.Y., Duan, W., Pan, Y.-M.: High-gain filtering patch antenna without extra circuit. IEEE Trans. Ant. Propag. **63**(12), 5883–5888 (2015)

A Clustering Approach for Sensory-Motor Cortex Signal Classification Using Electroencephalogram Signal for Brain-Computer Interface

Vinay Kulkarni, Yashwant Joshi, and Ramchandra Manthalkar

Abstract The Electroencephalogram (EEG) based Brain-Computer Interface (BCI) provides a robust technique for a physically impaired person and normal humans. The EEG is a complex, nonlinear, non-stationary signal. Therefore, the unlabeled EEG data handling for the classification of hand movements is a focused area for researchers. For addressing this problem, the paper proposes a approach called clustering of unlabeled EEG data. The EEG data were collected from the sensory-motor cortex with proposed protocol while performing elbow movement using the Enobio-20 machine and ARMEO Spring device. The collected EEG data is cleaned using band-pass filter. The Discrete Wavelet Transform (DWT) is used to decompose the filtered EEG signal into multiple frequency bands. The coefficients of each frequency band are clustered by using the K-means algorithm for each sub-band. The feature vector contained first-order moments of wavelet coefficients. The extracted feature vectors were fed to the k-nearest Neighbor (k-NN) with Fine kernel for classification of the elbow movements. The classification accuracy ($94.33 \pm 1.9\%$) of the proposed method is better than some of the recent approaches. The proposed methodology can distinguish and categorize the dynamical motions of the hand, which would be useful for BCI applications.

Keywords Discrete Wavelet Transform (DWT) · Electroencephalogram (EEG) · k-NN · Rehabilitation · K-means Clustering · Brain-Computer Interface (BCI)

V. Kulkarni (✉) · Y. Joshi · R. Manthalkar
Shri Guru Gobind Singhji Institute of Engineering and Technology, Nanded, India
e-mail: kulkarnivinayv@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_26

1 Introduction

The Brain-Computer Interface is an efficient technique to interconnect the human brain with advanced external objects [1]. The BCI keeps track of neural activity and converts rapid movement into commands. The neuro-signals are collected for BCI applications with many technologies, e.g., Electroencephalogram (EEG), functional near-infrared (fNIR), Magnetoencephalogram (MEG), etc. EEG is preferable primarily because of its non-invasive nature, low cost, and easiness in recording brain actions. The framework of BCI is advancing day by day in robotics, Rehabilitation, military, education, and entertainment. These applications are based on many modalities, among which motor activity-based EEG signals are mainly utilized [2, 3]. Though it is easy to capture the motor activity EEG signals, the crucial point is handling the complexity of the signal and capturing the exact period of action. The motor activity based BCI system involves mainly four aspects, signal acquisition, pre-processing of EEG data, features building and finally, the classification of motor activity signals. The recorded EEG signal is complex and nonlinear because of its small amplitude (μV) in the presence of multiple noise sources. Indeed, the feature extraction process helps the system differentiate the motor activity-related EEG signal from the rest of the signal. In motor activity based BCI, most researchers are working on the publicly available datasets [4, 5] in which motor activity is already marked. But significantly less study is conducted on the continuous EEG data without already marked actions. This paper will try to tackle this crucial part of handling continuous EEG signals for BCI applications. EEG-based BCI study has revealed that the attributes of nonstationary signal like EEG, e.g., within a particular frequency zone, zonal spectral area, mainly hidden phasic incidents are expressive and depend on a subject's motor action, cognitive state. These events are limited and having different signs that appear as informative, oscillations [6], brief time segments, phasic events [7], e.g., alpha and beta frequency bands [8]. Because of the nonstationary nature of multiple events present in the signal, EEG needs detailed analytic analysis for features calculations. Nowadays, Motor movement based BCI now works on specific marking identification methods that first recognize motor movement using pattern detection and classification into motor movement tasks.

The novelty of the paper is listed in the following points.

1. unique protocol is designed with ENOBIO-20 and ARMEO Spring rehabilitation device. The combination of these two devices improves the quality of EEG data.
2. The unlabelled EEG data need to be framed to capture the exact motor activity from raw EEG data. Hence time framing technique for EEG data is proposed in the paper.
3. The modified DWT algorithm is implemented on pre-processed EEG data. The filtered EEG is decomposed in selected frequency bands. With modified DWT methodology, K means clustering approach is designed. The clustering approach helps the classifier network to improve performance.

4. The ARMEO Spring rehabilitation device supports a multi-variable load system. The KNN based classifier network is applied on feature sets extracted from filtered EEG data.

The rest of the paper arranged as follow. The resources and procedures used during this paper are clearly described in Sect. 2. It describes how the raw data is collected as per the pre-decided protocol and how it is pre-processed. Section 3 gives the methodology implemented during feature extraction from filtered EEG data. Section 4 provides a detailed explanation of results obtained and discusses the various BCI issues like the effect of elbow movements on the motor cortex area with specific parameters. The last Sect. 5, concludes the work is done and the future scope of the work.

2 Material and Methods

2.1 Participants and Experimental Protocol

The novel protocol is designed to collect the raw EEG data while the subject performing hand movement tasks. Three active subjects (one female) aged between 26–28 years participated in the accompanying analysis for the experimentation. All three were right-handed with no history of any neurological issues. Everyone was instructed about the experiment and signed the consent form. For the experimentation, every subject plays the highflyer game from ArmeoSpring environment. The 10–20 electrode system is implemented with Enobio 20 EEG machine as shown in Fig. 1. The loads A, B, C and D are attached to moving hands in increasing order of weight. The loads are attached while playing the coin collection game. The loads A and B are used while subject was playing the game with right hand.

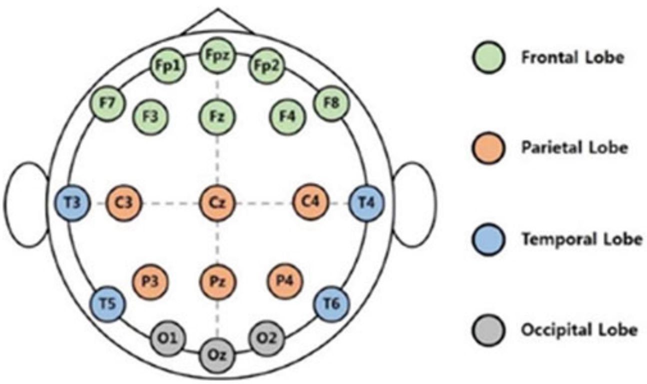


Fig. 1 10–20 electrode placement system for Enobio-20

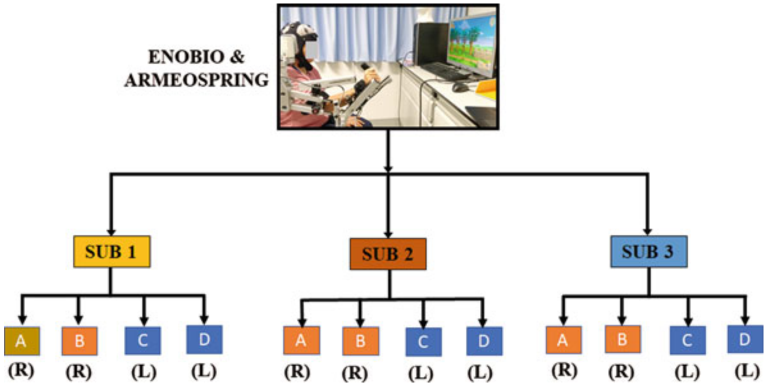


Fig. 2 Experiment setup for EEG data collection

Fig. 3 Protocol to collect hand movement EEG data

Simultaneously, loads C and D are attached while subject was playing the game with left hand. The scenario is displayed in Fig. 2. The experimental paradigm is shown in Fig. 3. The game duration is of 120 s during which subject fully engaged in playing the game. Hence, we have the raw EEG data of left- and right-hand movement of 120 s without actual time markers with sampling rate of 500 samples/sec.

2.2 De-noising of Recorded Raw EEG Data

The following methods are adopted to clean the recorded EEG signal and differentiate the left the right-hand movement. At the same time, collecting EEG data from subjects, typically unwanted signals 50 Hz supply line, eye squinting, even movement of heart while breathing, etc. For example, as shown in Fig. 4, the supply power line disturbs captured signals which can be the primary origin of the noise. Even eye squinting may be caught in recorded data just by observation as it generally has a significant amplitude than standard recording [9]. To remove the noise present in raw EEG data a FIR bandpass filter of [0 32] Hz is implemented in EEGLAB toolbox. The Independent Component Analysis is used from EEGLAB. Mainly, the independent sources are obtained by decomposing the input signal. Noise or artifacts present in EEG data may be produced by other than EEG sources.

3 Methodology

3.1 EEG Data Time Framing

The EEG signal for a complete run contains various types of rhythms like visual effect, cognitive load etc. These tasks are reflected at the respective brain parts in a particular time frame. Accordingly, during EEG signal analysis, the multichannel EEG signal is divided into multiple frames by applying a window of fixed samples. The feature vector is formed from a segment by describing the 1024 time-samples signal. The same methodology is used for each electrode of a 20 channel EEG signal. As shown in above Fig. 5, the length of each frame is 1024 and the overlap of 512 samples. The sampling frequency is 500 samples per second. A better activity is reflected in 1024 samples of EEG time frame. Following feature extraction method is applied over each frame, and features are extracted to form feature vectors.

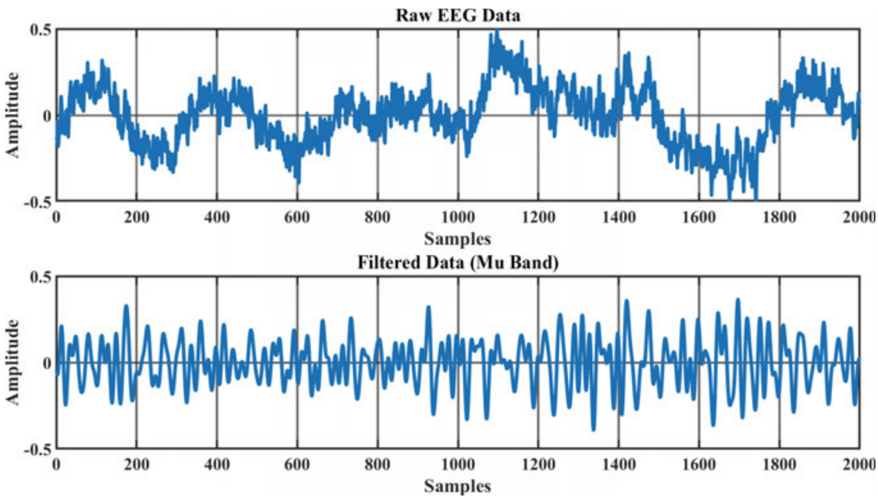


Fig. 4 Preprocessing of EEG signal

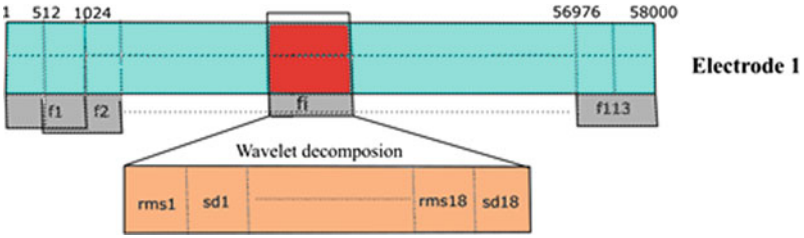


Fig. 5 Segmentation and wavelet decomposition

3.2 Time-Frequency Localization for Feature Extraction

In the case of stationary signals, the Fourier transform gives superior performance. Despite, the EEG time series having a nonstationary or very random nature. Therefore, the Fourier transform method is incompatible with bio signals such as EEG. The wavelet transforms split the signal within the low pass and high pass filter coefficients. This is achieved by performing two operations on the mother wavelet, dilation, and shift. Following is an equation of Continuous wavelet transform (CWT),

$$w_f\ (a, \tau) = \frac{1}{\sqrt{a}} \int_R f(t)\psi * (\frac{t-\tau}{a}) \tag{1}$$

where ψ_t is mother wavelet, a is a scaling parameter, τ is a shifting parameter.

In DWT, τ and a parameters are defined in scales of a logarithm, i.e., k2j and 2j, respectively. The modified equation for DWT can be determined using (2), In the guideline, the CWT delivered an infinite number of coefficients. Hence it gives a repetitive representation of a signal. The Discrete Wavelet Transform (DWT) provides a professional wavelet representation [10]. This representation is constructed with a simple recursive filter function. The initial signal reproduction can be gained by a converse filter.

$$\psi_{j,k} = 2^{\frac{-j}{2}} \psi\left(2^{-j}t - k\right) \tag{2}$$

J = number of decomposition levels, k= constant integer, 2^{-j} is utilized for normalizing $\psi_{j,k}$ i.e., a wavelet function. The Eq. (3) and (4) uses to calculate Approximation coefficients (A) and Detailed coefficients.

$$D = \sum_{k=\infty}^{-\infty} x\,[k]\, h\,[2n-k] \tag{3}$$

$$A = \sum_{k=\infty}^{-\infty} x\,[k]\, g\,[2n-k] \tag{4}$$

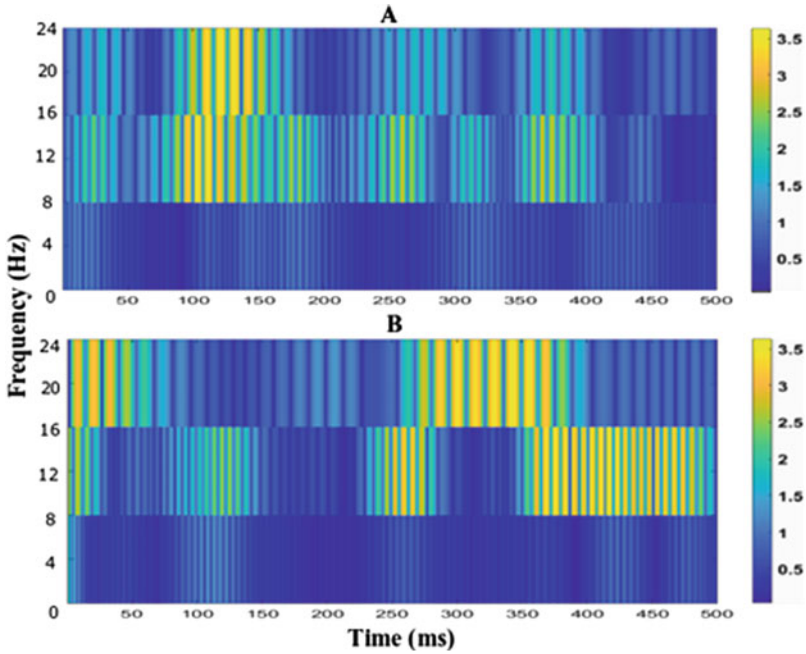


Fig. 6 Time-Frequency decomposition with DWT **A** Left Hand **B** Right Hand

In above equations, h[k] are LPF coefficients, g[k] are HPF coefficients. The method of multi-resolution deterioration of an input signal x[n] is well described in over figure. The level of decomposition (L) is chosen to obtain more coefficients from interested frequency bands. Hence during this work, L = 6 and wavelet function is 'bior 6.8' is chosen. Constructing biorthogonal wavelets permit extra flexibility like having more degrees of freedom than other orthogonal wavelets.

One extra degree of flexibility is the plausibility to develop functions of symmetric wavelets. The significant motor cortex activity is reflected in 8–30 Hz. Therefore, the signal is decomposed up to level six. A few fundamental parameters are calculated to form a final feature vector. It reduces the dimensionality of constructed vectors such as RMS and standard deviation (SD) values derived from wavelet coefficients of each decomposed packet. As shown in Table 1, the wavelet coefficients from nine packets are combined and presented to calculate the RMS and SD values.

The time-frequency graph for right and left-hand movement is shown in Fig. 6. The DWT algorithm is applied to filtered EEG data. The right-hand movement affects the left side of the brain, and the left-hand movement affects the right side of the brain. The BCI activity is reflected in alpha and beta frequency bands. From Fig. 6 (A), the right-hand movement is reflected on the left side of the brain in 8 24 Hz frequency bands. Simultaneously, Fig. 6 (B) gives a reflection of the left-hand movement on right side of the brain. The yellow portion in the figure shows the high motor activity and the blue portion indicates the lower motor activity. In the lower frequency band of 0 4 Hz, zero motor activity is reflected as BCI is more active in alpha (8–12 Hz) and beta (12–24 Hz) bands.

3.3 K-means Clustering

Clustering calculations are utilized broadly not as they were to gather comparative or disparate data but too valuable for data compression and reduction. The foremost utilized clustering methods are fuzzy C-means [11], k-means [12], mountain [13] and subtractive clustering [14]. Here, k-means clustering (KMC) is selected as the weighting process because this method has been widely handled in the literature. In the feature weighting method based on KMC. Initially, the KMC algorithm is used to find the cluster of each feature. Then the distance between the cluster and the average of the feature is calculated. The feature vectors are weighted based on the calculated distance.

The purpose of the feature weighting approach is to match features according to their occurrence in the data set and convert them from a nonlinearly separable data set to a linear separable data set. The feature weighting technique operates on the theory of reducing the variance of the features creating a data set. Through this weighting technique, similar data with the same characteristics can be collected, which improves the discriminating ability of the classifier.

The cluster centers and their elements find out by minimizing the objective function based on the squared error [15]. The algorithm focuses on spotting the cluster center as far as feasible and correlate each data point with the immediate cluster center. In KMC, generally, Euclidean distance is commonly utilized. An objective function (J) is defined:

$$J = \sum_{j=1}^{k} \left(\sum_{k} \|x_k - C_i\|^2 \right) \tag{5}$$

Table 1 Sub-bands using wavelet packet decomposition

Sub-Band	Decomposed signal	Frequency range (Hz)
1	WL61	0–4
2	WH61	4–8
3	**WL511**	**8–12**
4	**WH512**	**12–16**
5	**WL411**	**16–20**
6	**WH412**	**20–24**
7	WL4121	24–28
8	WH4122	28–32

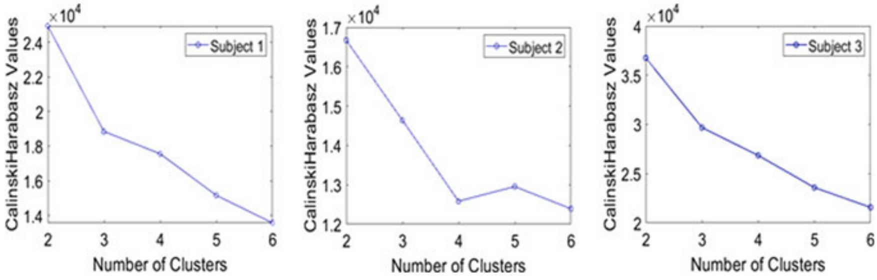


Fig. 7 Calinski-Harabasz criterion for finalize value of K

k = number of clusters, c_i = centers of clusters, and x_k, k_{th} = data-points in i^{th} cluster. A data point belongs to a cluster that is nearest to that data point to its center. Therefore, the binary affiliation vector U represents the clusters. The elements of vector U is formulated below:

$$u_{ij} = \begin{cases} 1 & \left\|x_j - C_i\right\|^2 \leq \left\|x_j - Ct\right\|^2, \forall t \neq i; \\ 0 & \text{otherwise;} \end{cases} \tag{6}$$

u_{ij} implies whether j_{th} data point associates with the i_{th} cluster or not. Each C_i cluster center that minimizes the J objective function is derived as:

$$C_i = \frac{\Sigma_{j=1}^{N} u_{ij} x_i}{\Sigma_{j=1} u_{ij}} \tag{7}$$

N = number of data points.

From [10], the KMC algorithm works as: initially, pick K data points as cluster centres. Secondly, using Eq. (6) calculate membership vector U. In the next step, formulate J from Eq. (5) and with Eq. (7) refresh the values of cluster centers. Repeat the above steps until the value of the cluster center becomes stable. For the calculation of the value of K, Calinski-Harabasz Criterion is applied to the EEG data. Figure 7 indicates that the greatest Calinski-Harabasz value is found at two clusters, implying that two (K=2) is the ideal number of clusters. For the recorded EEG data, it is observed that the Calinski-Harabasz value decreases, as the value of K increases.

The KMC algorithm is calculating the centers to cluster the total EEG data points within the segment. Every segment is consisting of RMS and SD values calculated from wavelet coefficients. These data points are clustered using K = 2. One cluster is for when the subject performs the elbow movement and the other one for a subject is observing the computer screen without doing any elbow movement, i.e., the stationary state. It is observed that one cluster has more data points as compared to others. Above Fig. 7 shows the plot of these data points belongs to two clusters.

4 Results and Discussions

The experimentation for this paper is conducted in Matlab 2020b. The classifier KNN is implemented for classifying left and right-hand movement. The feature vector is divided into training and testing datasets. The 70% of features are provided to classifier as training data and remaining 30% of feature sets given as testing dataset. This will avoid the biasing of feature vectors that leads to give higher accuracy of the proposed framework.

4.1 Silhouette Analysis of Clustering Method

The Silhouette coefficient is a form of explaining and verifying the viscosity within the data cluster [15]. This technology contributes a brief graphical depiction of the classification of each data point. The squared Euclidean distance is adopted to calculate the Silhouette coefficient. For two states, movement and stationary state, Fig. 8 displays the silhouette coefficient value graph. The plot of the silhouette illustrates that the data is divided into two unequal-sized clusters. There are large silhouette values ($\geqslant 0.8$) for all the data points in both the clusters, signifying that the clusters have been isolated correctly.

Fig. 8 Silhouette value plot for two clusters

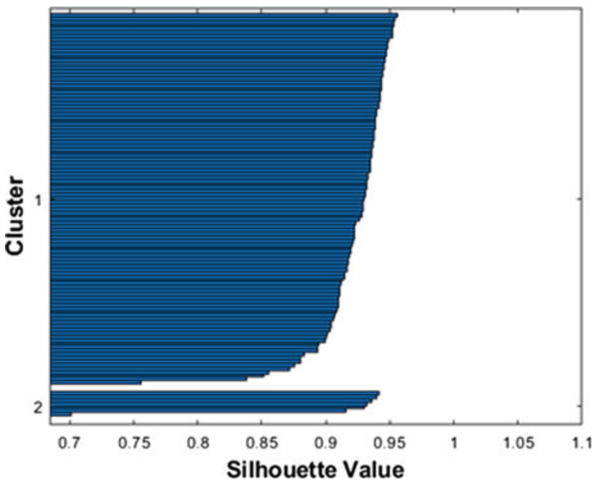


Table 2 Classification performance of k-NN classifier across different subjects and multiple trials

Subject	Accuracy(%)	Precision	Recall	F1-Score
A1	93.82	0.92	0.93	0.93
A2	94.77	0.94	0.98	0.95
B1	92.70	0.90	0.88	0.90
B2	93.10	0.92	0.89	0.92
C1	97.13	0.95	0.98	0.96
C2	94.43	0.93	0.95	0.94

4.2 Classification Results

The k-nearest neighbor (k-NN) classifier is used with Fine kernel for classification of the left-hand elbow and right-hand elbow movement. k-NN is a basic machine learning method [16]. In k-NN, an object is categorized according to a plurality vote of its next surrounding item. Then the categories that frequently appear among its nearest k neighbors are nominated. Due to many benefits, such as effective generalization and simple execution, it has been widely used in machine learning [17]. From the confusion matrix parameters performance of the classifier is checked by calculating the following parameters:

$$accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{8}$$

$$precision = \frac{TP}{TP + FP} \tag{9}$$

$$recall = \frac{TP}{TP + FN} \tag{10}$$

$$F_1 = \frac{2 \times (precision \times recall)}{precision + recall} \tag{11}$$

where TP= True Positive, FN= False Negative, FP= False Positive, TN=True Negative The precision is known as the percentage of relevant outcomes, while recall is described as per cent outcomes that are correctly categorized by the current model. The F1-score, which takes into account both precision and recall to essentially calculate the model's accuracy.

Table 2 gives the classification performance evaluation with parameters like accuracy, precision, recall. The proposed methodology is applied to the pre-processed EEG data. The performance of the classifier is tested in two trials A1 and A2. The outcome of A1 is the average results of three experiments, i.e. easy, medium, and hard. The same scenario is followed for subject B and subject C. Subject C in the

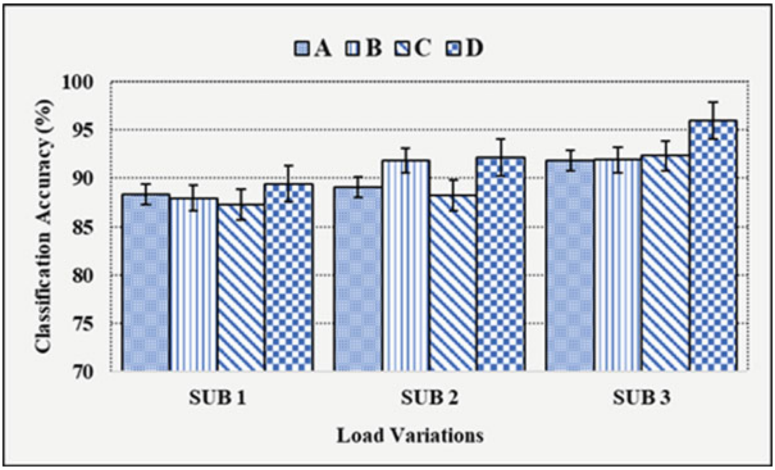


Fig. 9 Classification accuracy of L/R movement across different loads

first experiment gives the best results in terms of accuracy, precision, recall and F1-score. But for subject B, results are a little bit lower as compared to other subjects. In the case of classification, the accuracy is not a sufficient parameter to support the results. Hence here, other parameters like precision, recall and F1-score have been discussed. For subject B2 accuracy is less but the precision is higher. The subject A2 is also achieved less accuracy but a higher recall parameter.

4.3 Performance of Subjects Across Different Loads

There are three types of experiments in a run easy, medium, and hard. The task of classifying these movements is a little bit hard as speed and kinematics are involved in each run. The detailed information of the obtained multiclass classification results with the k-NN classifier is explained in following Fig. 9.

In the easy level, the elbow movement is slow, and elbow movement is moderate in medium and fast in case of a hard level. Hence elbow movements with different kinematics can be classified using a k-NN classifier. The percentage accuracy of every subject across various load experiments have been shown in below Fig. 9.

During experiment B, subject 3 gives the highest accuracy of 95.7%. Overall, during the last experiment E, subjects perform well and improve in concentrating on completing the task. The three kinematic movements of the elbow are somewhat identical. But in the frequency bands, alpha, and beta of sensory-motor cortex, the k-NN classifier gives improved outcomes. The kinematic movements of the elbow have been studied with the help of different experimental conditions. This study will provide a new area of research to the researchers in the field of BCI.

Table 3 Comparison of the proposed method with state-of-the-art methods

Authors	Method	Average CA (%)
Rodríguez-Bermúdez et al. [17]	AAR+HOS+LDA	74.30
Gaur, Pramod, et al. [18]	MEMD+SS-MEMDBF	79.93
Selim, Sahar, et al. [19]	CSP+AM/BA+SVM	86.60
Kevric, Jasmin et al. [20]	MSPCA+WPD+HOS+KNN	92.80
Proposed method	DWT+k-means clustering+ k-NN	**94.33**

4.4 Comparison of the Proposed Method with Previous Work

The comparison is shown with state of the art in Table 3. In [17], the authors Rodríguez-Bermúdez et al. worked on implementing an automatic adaptive classification system. The Adaptive autoregressive (AAR) coefficients are calculated as features and classified with Linear Discriminant Analysis (LDA). The work achieves the classification accuracy of 74.3% only. The work in [18] achieved the classification accuracy of 79.93% with SS-Multivariate Empirical Mode Decomposition (SS-MEMDBF). The features are extracted using Multivariate Empirical Mode Decomposition (MEMD) applied on filtered EEG data. The authors Selim, Sahar, et al. [19] implemented Common Spatial Pattern (CSP) algorithm for feature extraction from EEG data. The hybrid Attractor Metagene (AM) and Bat optimization Algorithm was utilized for feature optimization from all sets of features. The authors Kevric, Jasmin et al. [20] utilized Multiscale Principal Components Analysis (MPCA) for filtering of raw EEG data. The clean EEG data is divided into small frequency bands of a particular range using Wavelet Packet Decomposition (WPD). The Higher-Order Statistics (HOS) are extracted from decomposed frequency bands. The KNN classifier achieves a classification accuracy (CA) of 92.8%.

5 Conclusions

This work focuses on the classification of unlabeled EEG data. The raw EEG data is collected with a designed protocol for classifying elbow movements. For cleaning the raw EEG data, the band-pass filtering method along with Independent Component Analysis (ICA) is adapted. The work aims to classify the elbow movement; a wavelet packet decomposition approach is used on the segmented full-length EEG signal. The clustering algorithm gives the best separation of unlabeled EEG data having overlapping classes of movement and stationary actions. The same clustering method also gives improved results for classifying left-hand elbow and right-hand elbow movement as compared with state of the art. k-NN classifier with Fine kernel gives best results in the classification of binary as well as multiclass classification tasks. The kinematic movements classification tasks provide a new area of research to the researchers. In future work for improving the classification accuracy, a hybrid methodology can be implemented like EEG+fNIRS, etc.

References

1. Wolpaw, J.R., McFarland, D.J., Neat, G.W., Forneris, C.A.: An EEG-based brain-computer interface for cursor control. Electroencephalogr. Clin. Neurophysiol. **78**(3), 252–259 (1991)
2. Nijholt, A., et al.: Brain-computer interfacing for intelligent systems. IEEE Intell. Syst. **23**(3), 72–79 (2008)
3. Yuan, H., He, B.: Brain-computer interfaces using sensorimotor rhythms: current state and future perspectives. IEEE Trans. Biomed. Eng. **61**(5), 1425–1435 (2014)
4. Blankertz, B., et al.: The BCI competition iii: validating alternative approaches to actual BCI problems. IEEE Trans. Neural Syst. Rehabil. Eng. **14**(2), 153–159 (2006)
5. Tangermann, M., et al.: Review of the BCI competition iv. Front. Neurosci. **6**, 55 (2012)
6. Wang, X.-J.: Neurophysiological and computational principles of cortical rhythms in cognition. Physiol. Rev. **90**(3), 1195–1268 (2010)
7. Meurant, G.: Mass Action in the Nervous System. Elsevier, Amsterdam (2012)
8. Adrian, E.D., Matthews, B.H.C.: The Berger rhythm: potential changes from the occipital lobes in man. Brain **57**(4), 355–385 (1934)
9. Kołodziej, M., Majkowski, A., Rak, R.J.: A new method of EEG classification for BCI with feature extraction based on higher order statistics of wavelet components and selection with genetic algorithms. In: Dobnikar, A., Lotrič, U., Šter, B. (eds.) ICANNGA 2011. LNCS, vol. 6593, pp. 280–289. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-20282-7_29
10. Higashi, H., Tanaka, T.: Simultaneous design of fir filter banks and spatial patterns for EEG signal classification. IEEE Trans. Biomed. Eng. **60**(4), 1100–1110 (2012)
11. Liang, S., Choi, K.-S., Qin, J., Pang, W.-M., Wang, Q., Heng, P.-A.: Improving the discrimination of hand motor imagery via virtual reality based visual guidance. Comput. Methods Programs Biomed. **132**, 63–74 (2016)
12. Buch, E., et al.: Think to move: a neuromagnetic brain-computer interface (BCI) system for chronic stroke. Stroke **39**(3), 910–917 (2008)
13. Pfurtscheller, G.: Induced oscillations in the alpha band: functional meaning. Epilepsia **44**, 2–8 (2003)
14. Bablani, A., Edla, D.R., Kuppili, V., Ramesh, D.: A multi stage EEG data classification using k-means and feed forward neural network. Clin. Epidemiol. Glob. Health **8**(3), 718–724 (2020)
15. Rousseeuw, P.J.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. **20**, 53–65 (1987)
16. Asadur Rahman, M., Mahmudul Haque, M., Anjum, A., Nurunnabi Mollah, M., Ahmad, M.: Classification of motor imagery events from prefrontal hemodynamics for BCI application. In: Uddin, M.S., Bansal, J.C. (eds.) Proceedings of International Joint Conference on Computational Intelligence. AIS, pp. 11–23. Springer, Singapore (2020). https://doi.org/10.1007/978-981-13-7564-4_2
17. Rodríguez-Bermúdez, G., García-Laencina, P.J.: Automatic and adaptive classification of electroencephalographic signals for brain computer interfaces. J. Med. Syst. **36**(1), 51–63 (2012)

18. Gaur, P., Pachori, R.B., Wang, H., Prasad, G.: A multi-class EEG-based BCI classification using multivariate empirical mode decomposition based filtering and Riemannian geometry. Expert Syst. Appl. **95**, 201–211 (2018)
19. Selim, S., Tantawi, M.M., Shedeed, H.A., Badr, A.: A csp\am-ba-svm approach for motor imagery BCI system. IEEE Access **6**, 49192–49208 (2018)
20. Kevric, J., Subasi, A.: Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system. Biomed. Signal Process. Control **31**, 398–406 (2017)

Comparative Analysis of Different Encoding Techniques and Types of Sensors for Laser Based Visible Light Wireless Communication

Rushil Kumar, K. Kanishk, Yash Mudgal, and Poornima Mittal

Abstract Wireless communication has been around for several decades, much research has been done in Radio and microwave communication. Exploration of visible light digital communication is a necessity. Laser pulse based digital communication technology is rising rapidly in various industries due to its vast applications especially in space communication or FSO (Free Space optical) communication. The properties of a digital data communication system are greatly affected by the encoding/decoding techniques the system uses during data transmission. In this study, we analyze various encoding techniques and their effects on the quality and speed of data transmission in visible laser communication systems on low-cost, commercially available electronics. Various trade-offs of using analog and digital sensors at various bitrates and different MPU will be discussed to give a more comprehensive view of the factors which affect the line-of-sight laser communication between two points. As a final comparison, we will be demonstrating real time LASER data transfer using different encoding and decoding protocols switching between various sensors on the analog and digital ranges, this data will be projected in a tabular form as we proceed in the paper to draw a conclusion regarding the effects of variation produced by it and study these variations.

Keywords Laser communication · Wireless communication · Free space data transfer · Laser diodes · Sensors · Encoding · Transmission rates

R. Kumar (✉) · K. Kanishk · Y. Mudgal · P. Mittal
Delhi Technological University, New Delhi 110042, India
e-mail: Rushil.kumar18@gmail.com

K. Kanishk
e-mail: k.kanishk001@gmail.com

Y. Mudgal
e-mail: yashmudgal49@gmail.com

P. Mittal
e-mail: poornimamittal@dtu.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_27

1 Introduction

LASERS have been around since 16th May 1960 with the first working and patented laser in Hughes Research Laboratories, Malibu, California [1]. Before this event, optical communication existed Alexander Graham Bell's photophone proves as a prime example yet the domain lacked some fundamental qualities which inhibited its widespread use, such as transmission over great distances and extreme attenuation in any given medium. Problems were faced at the early stages of laser communication in an optical waveguide as well, particularly on attenuation characteristics, for Laser communication in an optical waveguide to be functional and be used as a butter substitute for classical electrical communication the attenuation factor had to be drastically decreased from 1000 dB/km. This task was accomplished by Charles Kao & John Hawkings in 1966 dropping the losses to 20 dB/km [2]. Currently, the losses stand at 0.28 dB/Km at 1550 nm inserted in a hollow-core optical waveguide [3]. Apart from the usage of optical waveguide for transmitting data via Lasers, the utilization of FSO (Free Space Optical) communication picked up pace circa 1990 with the relay mirror experiment (RME) [4], and with the uplink established to a deep-space vehicle demonstrated through Galileo optical experiment (GOPEX) which transmitted pulsed laser signal from two optical ground stations (OGS) mounted at California and New Mexico, in 1992 [5]. The meander around this idea ultimately led to the first ground-to-space.

2-way comms link used in the ground/orbiter Laser-COMM demonstration (GOLD) implementing unique argon (Ar) ion laser [6].

The current scenario of space comms advancements pertains to the domain of LASER communication leaving behind the classical RF communication techniques. One major factor of free-space Laser communication systems over the classical RF system is the transmission wavelength, which in atmospheric (clear condition over 20 km visibility) lies in the range of 785 and 1550 nm [7, 8]. This is significantly smaller than the wavelength of RF comms. Resulting in various advantages of free-space Laser communication systems over the classical RF system:

1. Greater bandwidth allocation: It is common knowledge that as the frequency of the carrier waveform increases, we can increase the number of signals being superimposed on the carrier wave. Hence, Laser comms have a greater Bandwidth allocation allowance than RF comms.
2. High directivity: The advantage of the optical carrier over the RF carrier can be seen from the antenna gain as the gain is closely related to directivity as the deviation angle (theta) is proportional to λ/DR. The optical wavelength is nanoscale. Thus, high directivity and improved gain are obtained in Laser communication systems [9].

After analyzing the existing literature [10, 11, 19–21], we can say, various other advantages of Free space Laser communication, this gives us more than enough evidence to develop and analyze the domain of Laser communication. Laser pulse-based digital communication technology is rising rapidly in various industries due to its vast applications especially in space communication or FSO (Free Space optical) communication. The properties of a digital data communication system are greatly affected by the encoding/decoding techniques the system uses during data transmission. Therefore, in this paper, we propose an analysis of all the mentioned encoding techniques (Manchester, RZ & NRZ) and their effect on the efficiency of data communication. This analysis will consider various factors and trade-offs by using different types of sensors available commercially to better demonstrate this analysis we will be transmitting text/image files to conduct an extensive analysis of the data transfer.

This paper is arranged in a total of 4 sections with the current introduction as Sect. 1. Further, schematics and characteristics of different current mirrors are discussed in Sect. 2, whereas, the analysis results are illustrated in Sect. 3. Finally, important outcomes are summarized in Sect. 4.

2 Methodology

2.1 Laser Sources Used in FSL Communication

Since their invention, lasers have transformed through the decades from gas to solid-state. The type of laser sources used in industrial applications is typically very high-power lasers with almost no beam divergence. This is because the beam has to travel a very far distance so the power loss due to scattering should be minimum and to maintain the high intensity or the concentration of power on the receiver, the beam should not spread much. The typical working range of terrestrial FSL systems is in the range of infrared i.e. 800–3000 nm. Satellite to satellite-based FSL has been utilizing the broad spectrum of 500–2000 nm [10].

To emulate the space transmission, we are utilizing a 5 mW, 650 nm laser diode. The diode of choice is very economical and best for demonstration purposes. When coupled with a breakout board, direct communication from the microcontroller can be done easily to engage and disengage the laser diode on command.

From the specifications (Table 1), it can be concluded that the laser which is being used for demonstration can perform well enough for smaller distances ranging from 1 to 10 m.

2.2 Loss During Transmission

In terrestrial and terrestrial-space communication, atmospheric losses are the main cause of beam attenuation. Scattering is the most prominent loss factor. The traveling beam loses its power successively when it travels through the medium (air). Various factors including the humidity, pollution level, dust concentration, and turbulent convection currents contribute towards the scattering. Some factors like Rayleigh

Table 1 Laser diode specification

Specification	Value
Operating voltage	2.8–5.2 V
Max current draw	25 mA
Optical power output	5 mW
Beam divergence (Parallel)	8° − 12°
Beam divergence (Perpendicular)	23° to 32°
Beam tolerance	− 3° to 3°
Operating temperature	− 10 °C to 40 °C

scattering which depend on the size and concentration of particles in the air can become wavelength-dependent as well. The resultant transmission coefficient can be written as the product of transmission coefficients resulting from the above-mentioned causes. As described by Zakharov and Shakanov, these coefficients differ due to the composition of the atmosphere at various heights [11]. The resultant transmission coefficient for ‘m’ discrete atmospheric layers can be written as:

$$K_{res} = \prod_{i=1}^{m} K_m \tag{1}$$

The relation between the transmission coefficient and the scattering coefficient per unit layer thickness is given by:

$$K_{res} = e^{kl} \tag{2}$$

where ‘l’ is the atmospheric layer thickness in meters.

This atmospheric loss is non-existent in space-to-space communication. For our use case and demonstration, we are considering this attenuation to be negligible and we are assuming that the transmitted power and the received power are equal and no loss whatsoever is introduced to the laser beam when being transmitted.

2.3 Receivers

Analog Sensors. Light sensors which depending upon the intensity generally work on the principle of changing their electrical characteristics. Light Dependent Resistor (LDR) is a common sensor that changes the resistance across its terminals according to the intensity of the light incident on the sensor. The analog voltage across its terminals is measured and converted to digital signals via the onboard analog-to-digital converter of the microcontroller board. The rise and fall times of such sensors play a huge role in accepting them as suitable candidates for data transmission purposes.

Table 2 Digital sensor specifications

Specification	PT334-6C [13]	TSL250R [14]	Unit
Operating wavelengths	400–1100	400–700	nm
Wavelength of peak responsivity	775	940	nm
Operating voltage	5	5	V
Max output voltage	3.3	5	V
Avg. Rise/Fall Time	15	7	s

The rise and fall time of LDR is huge which is the main disadvantage of such sensors [12]. This is mainly due to the huge variance of resistance in low light conditions, for e.g., in NSL-19M51 photoresistor, the resistance varies from 20–100 KΩ for 10 lx and takes a total of 10 s to go to the huge resistance of 20 MΩ in totally dark conditions. An NSL-19M51 analog sensor is used as it is economical and readily available. The typical rise time of LDR ranges from 2.8 to 18 ms and the typical fall time ranges from 48 to 120 ms [15].

Digital Sensors. Digital sensors are effective in data transmission as they output discrete signals in either a HIGH or LOW signal as they use a voltage comparator circuit. They are very sensitive and have significantly less rise and fall times than analog sensors like LDR. Two Photodiode sensors are utilized in this analysis, PT334-6C and TSL250R. PT334-6C is a typical photodiode and is utilized in combination with a breakout board which allows setting the threshold. The breakout has separate pins for the power and one pin for the sensor output, either HIGH (5 V) or LOW (0 V), depending on the sensor state. Whereas TSL250R is a 3-pin device with an inbuilt amplifier and readily gives an output pin and two power pins (See Table 2).

2.4 Types of Encoding Techniques Used

There are many types of encoding techniques available to encode our data to be transmitted and received as signal successfully, every encoding technique has its advantages and its drawbacks associated with it according to its architecture. The list of the following techniques is going to be included for our analysis:

Manchester Encoding. Manchester encoding (aka Phase Encoding) as per the IEEE 802.3 convention, we define Manchester encoding convention of 0's as transition logic High Level to logic Low Level and convention of 1's as transition logic Low Level to logic High Level. Manchester encoding uses a simple algorithm that can be simply interpreted as the XOR between transmitted data and the clock. The advantage of Manchester is that the signal is self-synchronizing, this, in turn, minimizes the error rate and increments the fidelity of the data. Drawbacks of this encoding technique cause more bits to be transmitted, and it needs more bandwidth [16].

Table 3 Encoded data bits generated from various encoding techniques depending upon the data and clock bits

Data bit	Clock bit	Manchester encoded data	RZ unipolar encoded data	NRZ unipolar encoded data
0	0	0	0	0
1	0	1	0	1
0	1	1	0	0
1	1	0	1	1

RZ Unipolar Encoding. Unipolar level encoding refers to the following Logic values; it uses only 2 values for data bit representation; it generally takes either the positive logic value (or Logic HIGH) as 1's and the reference value of 0 V as 0's in the data bit. Whereas in Bipolar, the 0's and 1's are negative and positive logic values. The Unipolar Return-to-Zero (RZ) encoding technique defines 1's as a transition from logic High Level to logic Low Level and 0's as 0 V during the entire clock cycle which can be described as the logical AND between Data bit and Clock bit [17].

NRZ Unipolar Encoding. Similarly, the unipolar Not-Return-to-Zero (NRZ) uses the standard encoding algorithm that is the same as on–off keying modulation of data, which is the most basic form of encoding data. The unipolar Not-Return-to-Zero (NRZ) defines the 1's in its data bits as logical HIGH (or set positive voltage) throughout the entire clock cycle and 0's as the reference voltage of 0 V throughout the entire clock cycle. One cannot recover the clock from unipolar NRZ encoded data as there is no reference of pulse length for that reason the receiver and transmitter need to be tuned to transmit and receive data successfully [17] (see Table 3).

2.5 Experimental Setup

Components of the Setup. Various components are being used in this setup, which are listed below:

- MPU (ATMEGA328P Arduino UNO board)
- Light Sensitive Sensors (LDR/Photodiode/TSL250R)
- Transmitter and receiver station (Laptop on serial communication with the board to process incoming text and image)
- Laser source (5 mW; 650 nm breakout board laser module focused at 1 m length by the lens attached in front of the module)
- Voltage divider configuration to the Light Sensitive Sensors (to read the analog voltage from the receiver board trim potentiometer set at 10 K ohm.)

Circuitry. (See Fig. 2) There are 2 different MPU in the setup connected to separate laptops by their respective jumper ports. We had used a TTL converter to act as a USB to UART bridge for serial communication between the laptop and MPU, the speed of the data transfer is determined within the sender and receiver Arduino, in order to avoid the delay in laptop-Arduino communication and affect the actual data speed. One laptop serves as the sender station where the data file (text/image) to be sent is picked from and the other serves as the receiver station where the data file (text/image) received is translated and stored. To detect the transmitted light, we have used voltage divider configuration from a light-dependent resistor, which is otherwise switched with a voltage comparator in the photodiode (PT334 -6C) and the Laser Non-modulator Tub Sensor module with TSL250R. There is roughly a 2-m gap in between the Laser and receivers.

2.6 Data Flow Description

The data is being taken from a text file saved on the source end station (laptop), which is firstly opened with python on the terminal and python reads the text file character by character. This character is then communicated to the MPU via a COM port through the serial communication protocol. The MPU receives the character as a byte of data, it then further gives this as the signal input to the Laser Diode mounted on the MPU pin 11.

The laser diode operates and transmits signals encoded as light waves. The light-wave gets transmitted via air as a medium. The light-emitting out of the other end falls on the photosensitive receiver (LDR/ photodiode). The reading from the sensor is measured by the receiving end MPU which then synthesizes it to a 1 or 0(as we have specified in the code a threshold above which we term the sensor values as a logical HIGH).

This data is fed into another COM port of the receiving end station (another laptop); we use python at the receiving end station to read the COM port values via serial read protocol. Python synthesizes the data, decodes it from bits and bytes to utf-8(Unicode) characters. Hence, receiving the data we had sent from the sender station. We can achieve our goal for data transmission and reception analysis via the time interrupt protocols inside the MPU and python.

3 Results and Discussion

The results obtained while conducting this experiment to analyze and verify the data and its various parameters are stated as follows (see Table 4).

Firstly, while analyzing the data speeds we made the BER (Bit Error Rate) constant by implementing Manchester and Unipolar RZ in a way that at receiving end, instead

Table 4 Data transfer speeds for selected sensors on different encoding techniques

Encoding technique used	Sensor	Baud rate (Bps)	Data speed (In Kbps)
NRZ	LDR	1015.9104	0.9921
	PT334-6C	27,443.2	26.80
	TSL250R	142,848	139.50
Unipolar RZ	LDR	534.9376	0.5224
	PT334-6C	14,592	14.25
	TSL250R	90,900.48	88.77
Manchester	LDR	500.0192	0.4883
	PT334-6C	14,489.6	14.15
	TSL250R	86,947.84	84.91

of measuring the rising and falling data pulse, we measured the two different levels sent in the same clock pulse of these techniques, which causes the BER of all 3 techniques to be almost similar in terms of reading the data reliably, the only drawback is, to transfer the same amount of data we need twice as much sampling of the incoming bitstream.

From Fig. 3 it can be seen that NRZ is faster than both Unipolar RZ and Manchester, as we mentioned, that we measured the data in the two different levels sent in the same clock pulse of these techniques causing twice as much sampling of the incoming bitstream than in NRZ for a singular byte.

Unipolar RZ and Manchester have comparable Data speeds due to the data pulse structure similarities in the encoding techniques and similarity in our implementation of the encoding as well. However, Unipolar RZ is slightly faster than Manchester. Probably due to the fact, that 0's in Unipolar RZ and Manchester are depicted by different Logic values, as in Manchester, encoded data is depicted by XOR of data with the clock pulse which in the case of a 0(data bit) is the clock cycle itself as shown in Fig. 1. Whereas, in Unipolar RZ, encoded data is depicted by AND of data with the clock pulse which in the case of a 0(data bit) is 0 throughout the clock cycle as shown in Fig. 1, this 0 Logic value throughout the clock cycle gives the sensor some relaxation. Due to the reduced number of rising and fall cycles in RZ, the system becomes more tolerant to errors and hence the clock width can be contracted to some extent resulting in an increased bit rate.

We also visually compared sent and received bits by recursively sending a string of bits and mapping it on the received signal to the set signal bits. As shown in Fig. 4, we can observe the contrast between faithfully receiving a signal Fig. 4(a) and when the data is not successfully interpreted Fig. 4(b). To distinguish between successful and unsuccessful data transmission.

From the analysis of the experiments, we can conclude that Free-Space Laser Communication proves to be an effective way of transmitting data even on a low-cost system using commercial off-the-shelf components. It can be seen that the sensor rises and fall time greatly affects the data transmission rates. Sensors like LDRs cannot

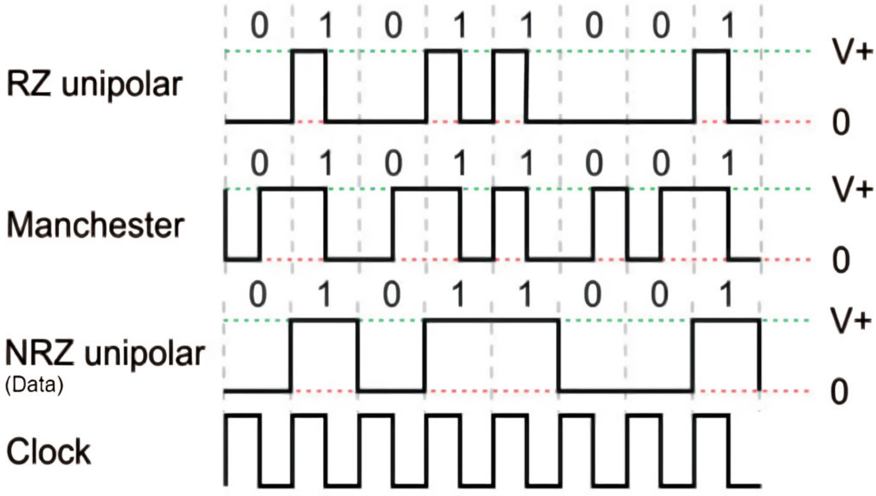


Fig. 1 Effect of encoding techniques on a stream of data along with the clock [18]

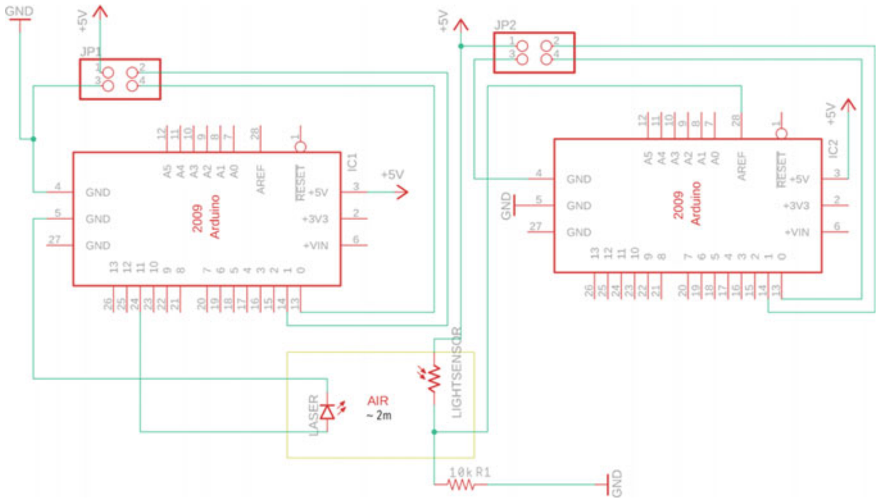


Fig. 2 Circuit schematics of the setup made on autodesk EAGLE

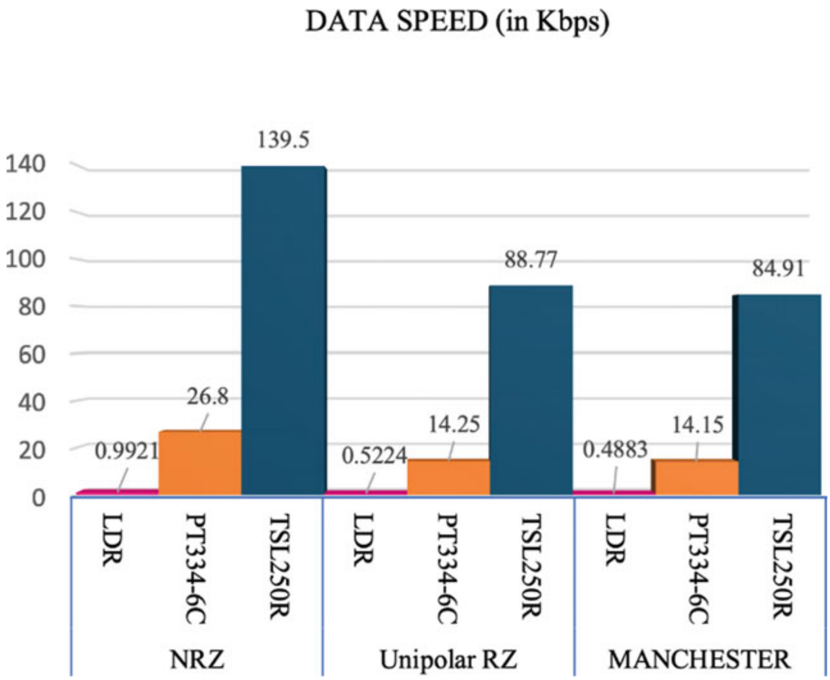


Fig. 3 Speed comparison for encoding techniques on selected sensors

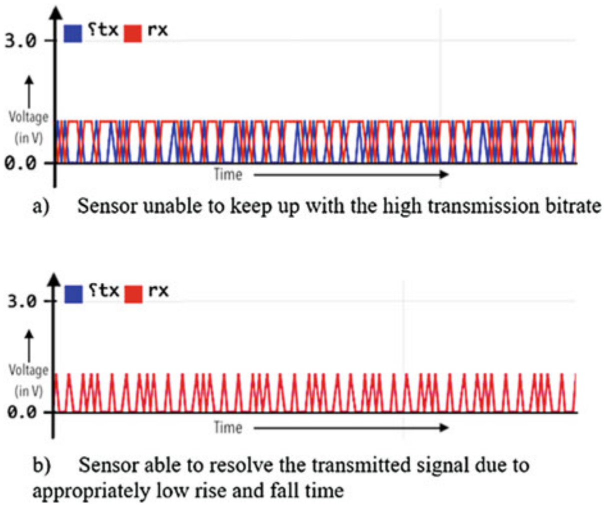


Fig. 4 **a** Tx and Rx waveform mismatched representing error in the data received. **b** the Tx and Rx waveform overlaps giving rise to an orange color representing the same bit transmitted and received

keep up with high bitrates, whereas photodiodes perform very well at receiving data at similar rates. In CdS based LDR sensors, the junction capacitance is greater than the pin photodiodes with an intrinsic layer placed between the junctions that reduce the junction capacitance. The lower junction capacitance in the photodiodes increases the response time drastically.

4 Conclusion

Through the experiments, we can deduce that NRZ encoding technique proves to be the fastest considering our implementations of the same. Manchester encoding encapsulates the clock in the transmitted data itself, proving to be the most reliable. The data transfer speed of RZ encoding is slightly greater than Manchester due to the reduced rise and fall cycles and hence RZ becomes more tolerant to errors.

The study we have done here on the COTS components can be used to provide us insights into a High budget (Highly sensitive and responsive system), In recent times FSL communication is on the cusp of a breakthrough with many space agencies including SpaceX, and NASA's Artemis equipped with FSL Communication systems [19, 20], the findings of this paper might help in making remote and compact systems.

References

1. Hecht, J.: Beam: the race to make the laser. Technical report: Optics & Photonics News (2005)
2. Kao, K.C., Hockham, G.A.: Dielectric-fibre surface waveguides for optical frequencies. In: Proceedings of the Institution of Electrical Engineers, vol. 113, no. 7, p. 1151, July 1966
3. Garg, A.K., Janyani, V., Batagelj, B., Abidin, N.Z., Bakar, M.A.: Hybrid FSO/fiber optic link based reliable & energy efficient WDM optical network architecture. Opt. Fiber Technol. **61**, 102422 (2021)
4. Lightsey, P.: Scintillation in ground-to-space and retroreflected laser beams. Opt. Eng. **33**(8), 2535–2543 (1994)
5. Wilson, K.E., Lesh, J.R.: An overview of galileo optical experiment (GOPEX), Technical report: TDA progress Report 42-114, Communication Systems Research Section, NASA (1993)
6. Jeganathan, M., Wilson, K., Lesh, J.: Preliminary analysis of fluctuations in the received uplink-beacon-power data obtained from the GOLD experiments. Telecommun. Data Acquis. Prog. Rep. **124**, 20–32 (1995)
7. Mecherle, G.S., Holcomb, T.: Videobeam TM portable laser communicator, optical wireless communications. Proc. SPIE **3532**, 22–28 (1998)
8. Medved, D.B.: Full throughput wireless ATM, wireless technologies and systems: millimeter wave and optical. Proc. SPIE **3232**, 160–166 (1997)
9. Majumdar, A., Ricklin, J.: Free-Space Laser Communications: Principles and Advances, vol. 2. Springer, New York (2008). https://doi.org/10.1007/978-0-387-28677-8
10. Kaushal, H., Kaddoum, G.: Optical communication in space: challenges and mitigation techniques. IEEE Commun. Surv. Tutor. **19**(1), 57–96 (2017)
11. Zaponov, A.E., Sakharov, M.V.: Estimation of propagation losses for infrared laser beam in turbulent atmosphere. J. Phys. Conf. Ser. **774**, 012123 (2016)
12. LUNA Optoelectronics: Ceramic Package Photocells NSL-19M51 REV 01-04-16 (2016)

13. EVERLIGHT Electronics Co.: Phototransistor PT334-6C, DPT-0000185 Rev. 4 (2016)
14. AMS OSRAM: TSL250R, TSL251R,TSL252R Light-to-Voltage Optical Sensors, AMS Datasheet [v 1-00] (2016)
15. Light Dependent Resistor (LDR)| LEDnique. http://lednique.com/opto-isolators-2/light-dependent-resistor-ldr. Accessed 25 July 2021
16. Forster, R.: Manchester encoding: opposing definitions resolved. Eng. Sci. Educ. J. **9**(6), 278–280 (2000)
17. Chand, N., Loriz, J.J., Hunton, A.J., Eteson, B.M.: Performance comparison of NRZ and RZ modulations, vol. 5892, p. 58920U. ISOP (2005)
18. Wong, W.G.: https://www.electronicdesign.com/technologies/communications/article/21802271/whats-the-difference-between-nrz-nrzi-and-manchester-encoding. Accessed 02 Sep 2021
19. Goorjian, P.M.: A new laser beam pointing method using laser arrays NASA/TM-2018-58815, November 2018
20. Jewett, R.: Latest starlink satellites equipped with laser communications, musk confirms, Via Satellite, 25 January 2021. https://cutt.ly/eWcY4Uc. Accessed 02 Sep 2021
21. Chen, A., et al.: Underwater optical wireless communications link for short-range data transmission: a proof-of-concept study (Doctoral dissertation) (2020)
22. Kumar, B., Raj, K., Mittal, P.: FPGA implementation and mask level CMOS layout design of redundant binary signed digit comparator. Int. J. Comput. Sci. Netw. Secur. **9**(9), 107–115 (2009). ISSN: 1738-7906
23. Rathod, A.P.S., Lakhera, P., Baliga, A.K., Mittal, P., Kumar, B.: Performance comparison of pass transistor and CMOS logic configuration based De-multiplexers. In: International Conference on Computing Communication and Automation (ICCCA–2015), pp. 1433–1437, Galgotias University, Noida, India, 15–16 May 2015

Lead Free Perovskite Solar Cell Using TiO_2 as an Electron Transport Materials and Cu_2O as a Hole Transport Materials

Vaibhava Srivastava, R. K. Chauhan, and Pooja Lohia

Abstract With an efficiency reaching 26.1% in a minuscule span of eleven years (2008–2019), Perovskite solar cells (PSC) have become a premier alternate for the traditional solar cells. Here simulation is done through SCAPS-1D software wherein the electrical parameters of PSC with TiO_2 as an Electron Transport Material (ETMs) and Cu_2O as Hole Transport Material (HTMs) and Au as back contact is simulated. Having studied the thickness of absorber layer, doping concentration of absorber layer and working temperature of the architecture by investigating at varied conditions observed that the perovskite solar cell showed the efficiency as 28.67%, FF = 83.20% $J_{SC} = 35.33$ mA/cm^2 and $V_{OC} = 0.9754$ V. The thermal stability study indicates that these devices configuration are constant at 300 K.

Keywords Perovskite solar cell · ZnO:Al · Simulation · CH3NH3SnI3,Cu2O · TiO_2 etc.

1 Introduction

As Absorber material, Perovskite solar cells (PSC) uses organometal halide perovskites which are the novel third generation solar cells. Perovskite solar cells (PSC) materials have numerous advantages including low exciton binding energy, large diffusion lengths, tunable bandgap and high absorption coefficients, that helps to cause in the power conversion efficiency of PSC from 3.9% in 2009 to 24.3% in 2019 [3, 4]. To convert a significant percentage of the incident solar power to electrical power in a direct way is due to photovoltaic effect some factors such as fill Factor (FF), high values of PCE, open circuit voltage (Voc) and short circuit current density (J_{SC}) by a commercially viable solar cell. In particular perovskite solar cell structure, a perovskite absorber layer is inserted between HTM and an ETM [7]. $CH_3NH_3SnI_3$ absorbed broad range of solar spectrum and can also be fabricated

V. Srivastava · R. K. Chauhan · P. Lohia (✉)
Department of Electronics and Communication Engineering, Madan Mohan Malaviya University of Technology, Gorakhpur 273010, India
e-mail: lohia.pooja6@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_28

at low temperature which reduces the cost of fabrication [8]. The suitable bandgap alignment is the possible reason for the selection of TiO_2 material as ETL and moreover, it is less expensive, non-toxic and it is highly chemically stable. It is necessary to select suitable HTM and ETM so as to enhance the stability and reproducibility of PSC [6].

Lead PSC is the most used hybrid material for perovskite. But the toxicity of lead causes health and ecological challenge. To overcome with the problem of human health related complications due to Lead toxicity problems, lead-based perovskite material is replaced with tin-based perovskite material [12, 13]. The use of Sn instead of Pb in LHPSC has been demonstrated in an attainable way to fabricate stable solar cells with reasonable efficiencies [1, 2]. Due to smaller band gap of $MASnI_3$, it covers a broad range of the visible wavelength [14].

The current work briefs a simulation of proposed device structure numerically for the optimization of the electrical parameters has been performed by the help of SCAPS-1D simulator.

In this simulation works, Glass/ZnO:Al/TiO_2/$CH_3NH_3SnI_3$/Cu_2O/Au planar PSC device with n-i-p architecture has been examined and the impact of thickness along with the concentration level of doping of absorbing layer and temperature have been studied. ZnO:Al is used as window layer. This layer transports the light towards the active layer. For this simulation Gold (Au) is a metal back contact. The main aim of the present work is to get the optimum output parameters for a highly efficient PSC with reduced cost.

2 Device Structure and Material Parameters

The Department of Electronics and Information Systems of Gent University had developed SCAPS in Belgium. It works on various profile of defined architecture such as carrier transport and recombination, steady state diagram & bulk defects [5, 6]. It has been seen that it adapts for various modelling of photonic structures, micro and polycrystalline thin film devices. The basic setting used to determine light or dark illumination. The parameter used to describe the performance of PSC parameters are J_{SC}, V_{OC}, FF & PCE. Schematic representation of device is shown in Fig. 1 (Table 1).

3 Results and Discussion

3.1 Impact of Working Point Temperature

For this particular study, the temperature considerations for the optimal device performance are increased from 280 to 480 K to observe temperature dependency of the device performance. The normal working temperature of the model is 300 K, but for the condition of outside, it is higher than that of normal working temperature

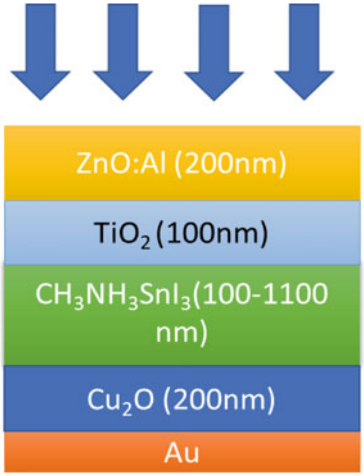


Fig. 1 Schematic representation of device

Table 1 Device parameters of different layers [5–7, 9]

Parameters	Cu_2O	$CH_3NH_3SnI_3$	ZnO:Al	TiO_2
Thickness (nm)	200	varied	200	100
Bandgap (eV)	2.17	1.3	3.3	3.260
Electron affinity (eV)	3.2	4.17	4.6	4.2
Relative Permittivity	7.11	6.5	9.0	10.0
Mobility of electron (cm^2 V s^{-1})	200	1.6	100	100
Mobility of hole (cm^2 V s^{-1})	80	1.6	25	25
CB effective density of states (cm^{-3})	2.02E + 17	1.00E + 20	2.2E + 18	2.2E + 18
VB effective density of states (cm^{-3})	1.10E + 19	1.00E + 20	1.80E + 19	1.8E + 19
Donor density, N_D (cm^3)	0	0	1.00E + 19	1.00E + 19
Acceptor density, N_A (cm^3)	1.00E + 18	3.25E + 15	0	0
Thermal velocity of electron (cm/S)	1.00E + 7	1.00E + 7	1.00E + 7	1.00E + 7
Thermal velocity of hole (cm/S)	1.00E + 7	1.00E + 7	1.00E + 7	1.00E + 7
Defect density, N_t (cm^3)	4.00E + 14	1.00E + 16	–	1.00E + 15

parameter [10, 11]. The plotted graphs below represent the obtained results. For the increase in temperature from 280 to 480 K, there can be seen an efficiency drop to 18.23% and correspondingly the FF is 71.50%, J_{SC} is 35.34 mA/cm^2 and V_{OC} = 0.7216 V. It can be clearly observed that the device performance and efficiency are inversely proportional to the increase of temperature. For optimal performance the working temperature of the architecture is taken as 300 K. At this particular temperature the maximum efficiency which is achievable of the architecture is 28.67%, FF is 83.20%, J_{SC} =35.338 mA/cm^2 and V_{OC} = 0.9754 V (Fig. 2).

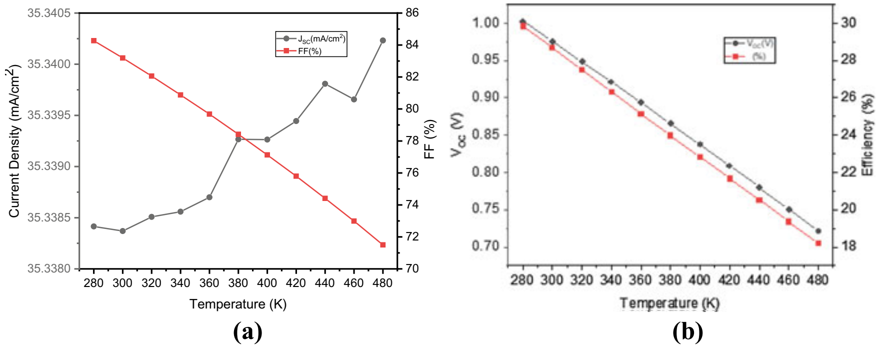


Fig. 2 Impact of working point temperature **a** PCE, V_{OC} **b** J_{SC}, FF

3.2 *Impact of the Thickness of Absorber Layer*

The thickness of the absorber layer plays an important role in determining the device efficiency. For the simulation results, we consider electrical parameters of the model (such as J_{SC}, FF, V_{OC}, and PCE) of thickness of absorber layer (scaling from 100 to 1100 nm). If the thickness is increased, the efficiency increases from 26.41 to 28.67%. Surprisingly at the absorber thickness of 100 nm, the highest reasonable parameters for photovoltaic are the PCE is 28.67%, FF is 83.2%, J_{SC} is 35.33 mA/cm^2 and V_{OC} is 0.9754 V.

From 700 nm to higher thickness the V_{OC} and J_{SC} both go into saturation owing to the recombination of charges interior of the absorber layer before approaching to the contacts. For larger thickness of perovskite solar cell V_{OC} becomes a function of I_{SC} and Io which ultimately depends upon the perovskite layer thickness. The dependence of solar cell parameters on thickness are shown in Fig. 3(a) and (b).

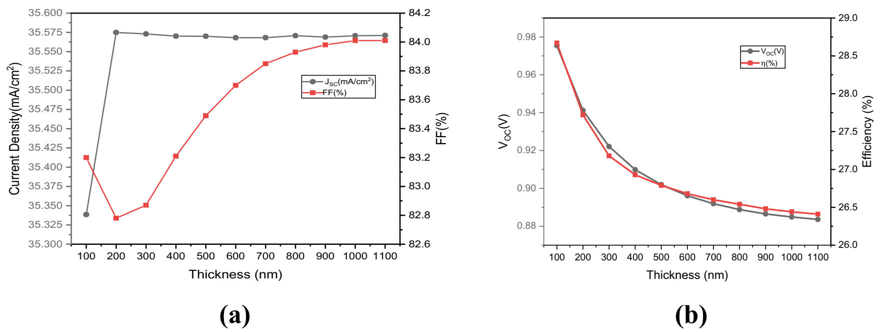
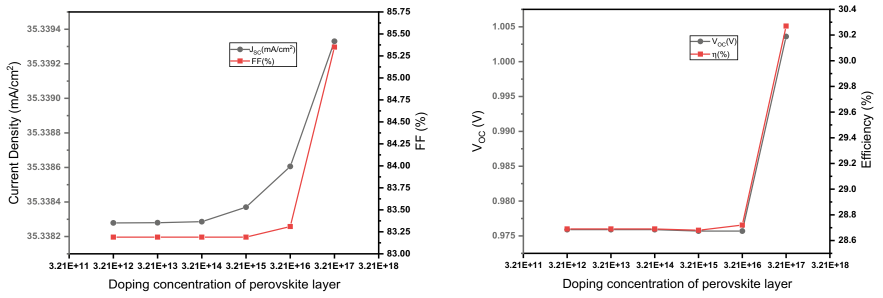


Fig. 3 Influence of the different thickness of absorber sheet **a** J_{SC}, FF. **b** PCE, V_{OC}



(a) **(b)**

Fig. 4 Impact of doping concentration of absorber layer on **a** J_{SC}, FF. **b** PCE, V_{OC}

Table 2 Comparison table between recent reported result and proposed result

ETL/Absorber/HTL	V_{OC} (V)	J_{SC} (mA/cm^2)	FF (%)	η (%)	Ref.
$TiO_2/CH_3NH_3SnI_3/CuI$	1.26	25.67	75.24	24.82	[6]
$WS_2/CH_3NH_3SnI_3/Cu_2O$	0.80	34.46	83.29	23.03	[15]
$TiO_2/CH_3NH_3SnI_3/Cu_2O$	0.97	35.33	83.20	28.67	Proposed work

3.3 Influence on Acceptor Concentration of Absorber Layer

To increase the properties of semiconductors devices like photovoltaic, doping process plays very important role. In solar cell architecture doping of photovoltaic material will decide the layers' electrical parameters which in turn would influence the overall results of the model.

The Acceptor concentration was increased from 1E+12 to 1E+17 cm^{-3} in order to examine the consequences of level of doping concentration of Perovskite thin layer on the overall device performance. It is clear from the figure that the efficiency saturates is 28.69% up to doping concentration 1E+15 cm^{-3} and then increases (from 28.68 to 30.27%) up to 1E+17 cm^{-3}. As we increase the doping concentration of absorber layer V_{OC} from 0.9759 to 1.0036 V, the FF and J_{SC} also same as to PCE curve. Impact of acceptor concentration on electrical parameters of photovoltaic cell are shown in Fig. 4(a) and (b) (Table 2).

4 Conclusion

In this simulation work, PSC cell simulator SCAPS 1D has been used to explore the device architecture of Glass/ZnO:Al/$TiO_2/CH_3NH_3Snl_3/Cu_2O$/Au PSC . Having varied the thickness of absorber layer form 100 to 1100 nm, the best performance of the device has been obtained at 100 nm thickness. In order to witness the changes in the PV parameters of the Cell, temperature has been altered from 280 to 480 K.

Non-toxicity, easy manufacture and low cost are some of the astounding characteristics shown by the structure of the tin-based PSC. Inferring on the result we can say that Cu_2O is the best alternate organic hole transport material for high PCE devices. During this work optimum PCE of 28.67%, J_{SC} = 35.33% mA/cm^2, FF = 83.20% and V_{OC} = 0.9754 V are attained at 100 nm and 1E+15 cm^{-3} acceptor concentration of absorber material. Our results are in good agreement with recently reported work of other researchers.

References

1. Green, M.A., Baillie, A.H., Snaith, H.J.: The emergence of perovskite solar cells. Nat. Photonics **8**, 506–514 (2014). https://doi.org/10.1038/nphoton.2014.134
2. Khattak, Y.K., Baig, F., Toura. H., Beg. S., Soucase, B.M.: CZTSe Kesterite as an alternative hole transport layer for $MASnI_3$ perovskite. J. Electron. Mater. **48** (2019). https://doi.org/10.1007/s11664-019-07374-5
3. Minemoto, T., Murata, M.: Impact of back contact of perovskite solar cells without hole transport material analyzed by device simulation. Current Appl. Phys. **S1567–1739**(14), 00243 (2014). https://doi.org/10.1016/j.cap.2014.08.002
4. Rai, N., Rai, S., Singh, P.K., Lohia, P., Dwivedi, D.K.: Analysis of various ETL materials for an efficient perovskite solar cell by numerical simulation. J. Mater. Sci. Mater. Electron. **31**, 16269–16280 (2020). https://doi.org/10.1007/s10854-020-04175-z
5. Jayan, D., Sebastian, V.: Comprehensive device modelling and performance analysis of $MASnI_3$ based perovskite solar cells with diverse ETM, HTM and back metal contacts. Solar Energy 40–48 (2021). https://doi.org/10.1016/j.solener.2021.01.058
6. Mandadapu, U., Vedanayakam, S.V., Thyagarajan, K., Babu, B.J.: Optimization of high efficiency tin halide perovskite solar cells using SCAP-D. Int. J. Simul. Process. Model. **13** (2018). https://doi.org/10.1504/IJSPM.2018.093097
7. Patel, A.K., Rao, P.K., Mishra, R., Soni, S.K.: Numerical study of a high-performance thin film CIGS solar cell with a-Si and $MoTe_2$ hole transport layer. Optik Int. J. Light Electron. Opt. 167498 (2021). https://doi.org/10.1016/j.ijleo.2021.167498
8. Tan, K., Lin, P., Wang, G., Liu, Y., Zongchang, X.U., Lin, Y.: Controllable design of solid -state perovskite solar cells by SCAPS device simulation. Solid-State Electron. **S0038–1101**(16), 30142–3 (2016)
9. Rai, S., Pandey, B.K., Dwivedi, D.K.: Modeling of highly efficient and low cost CH_3NH_3Pb (I1-x Clx)$_3$ based perovskite solar cell by numerical simulation. Opt. Mater. **100**, 109631 (2020). https://doi.org/10.1016/j.optmat.2019.109631
10. Salah, M.M., Hasan, K.M., Abouelatta, M., Shaker, A.: A comparative study of different ETMs in perovskite solar cell with inorganic copper Iodide as HTM. Optik **S0030–4026**(18), 31551–31561 (2019). https://doi.org/10.1016/j.ijleo.2018.10.052
11. Husainat, A., Ali, W., Cofie, P., Attia, J.O.: Simulation and analysis of methylammonium lead iodide ($CH_3NH_3PbI_3$) perovskite solar cell with Au contact using SCAPS 1D Simulator. Am. J. Opt. Photonics **7**(2), 33–40. (2019). https://doi.org/10.11648/j.ajop.20190702.12
12. Azri, F., Sengouga, N., Meftah, A., Meftah, A.: Electron and hole transport layers optimization by numerical simulation of a perovskite solar cell. Solar Energy **181**, 372–378 (2019). https://doi.org/10.1016/j.solener.2019.02.017
13. Mandadapu, U., Vedanayakam, S.V., Thyagarajan, K., Reddy, M.K., Babu, B.J.: Design and simulation of high efficiency tin halide perovskite solar cell. Int. J. Renew. Energy Res. **7**(4), 1604–1612 (2017). https://doi.org/10.20508/ijrer.v7i4.6182.g7270

14. Dwivedi, D.K.: Modeling of CZTSSe solar cell for window layer optimization. Optik **222**, 16540 (2020). https://doi.org/10.1016/j.ijleo.2020.165407
15. Kumar, A., Singh, S.: Numerical modeling of planar lead free perovskite solar cell using tungsten disulfide (WS_2) as an electron transport layer and Cu_2O as a hole transport layer. Mod. Phys. Lett. B **34**(24), 2050258 (2020). https://doi.org/10.1142/S0217984920502589

Routing and Congestion in Vehicular Ad-Hoc Networks (VANET's): Characteristics, Challenges and Solutions

Lokesh M. Giripunje, Abhay Vidyarthi, and Shishir Kumar Shandilya

Abstract Vehicular Ad hoc Networks (VANETs) are the special types of Mobile Ad hoc Networks (MANETs) designed to assist drivers in emergency situations and improve the overall driving experience. The VANET is an infrastructure-less environment with communication-enabled highly mobile vehicular nodes and has capabilities of sending and receiving various messages using Dedicated Short Range Communication (DSRC). VANET uses IEEE 802.11p standard for Wireless Access in Vehicular Environment (WAVE) for sharing traffic and safety information. VANET is playing a crucial role in Intelligent Transport Systems (ITS) by avoiding accidents and reducing pollution to achieve accident-free and pollution-less motorways. Highly mobile vehicular nodes and rapidly changing network topology throws many challenges in the dissemination of critical messages in VANET. Improper routing in emergency and Safety applications will result in congestion and degradation of network performance and QOS. This survey paper presents an overview of VANET with the help of VANET communications, characteristics, and applications of VANET. This article discusses the problem of congestion and reviews the researcher's contribution to solving congestion problems as it plays a major role in the Quality of Service (QOS) of VANET.

Keywords Quality of Service (QOS) · Vehicular Ad hoc Network (VANET) · Intelligent Transport Systems (ITS) · Routing · Congestion control

1 Introduction

Considering the increase in count of automobiles on roads in recent years, governments are imposing strict traffic rules and regulations as drivers safety is of utmost importance while driving. With technological advancements, there have been contributions to make driving experience safer and pollution free. Even after these efforts, every year there is huge number of road accidents and almost 1.25 million deaths

L. M. Giripunje (✉) · A. Vidyarthi · S. K. Shandilya
VIT Bhopal University, Bhopal, India
e-mail: lokeshgiripunje@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_29

in accidents on roads [1]. Vehicle Ad Hoc Network (VANET) since its introduction is contributing in actual realization Intelligent Transportation System (ITS). Hence, academia, governments, telecommunication and car manufacturing industries are showing interest in addressing problems of VANET and ITS to fulfill the target of safety and congestion free motorways [2].

The VANET is infrastructure less environment of high mobility Vehicles with transmission and reception capabilities for sending and receiving various messages using Dedicated Short Range Communication (DSRC). VANET uses IEEE 802.11p standard for Wireless Access in Vehicular Environment (WAVE) for sharing traffic and safety information. Federal Communications Commission (FCC) in United States assigned frequency from 5.850 to 5.925 GHz (75 MHz) band for DSRC [3]. Till date various approaches were proposed for the both safety and non-safety communications for various vehicular network applications. The DSRC channel uses short and medium wave communication [4] and band allocation is different in the Europe and the Japan is shown in Fig. 2 [5] (Fig. 1).

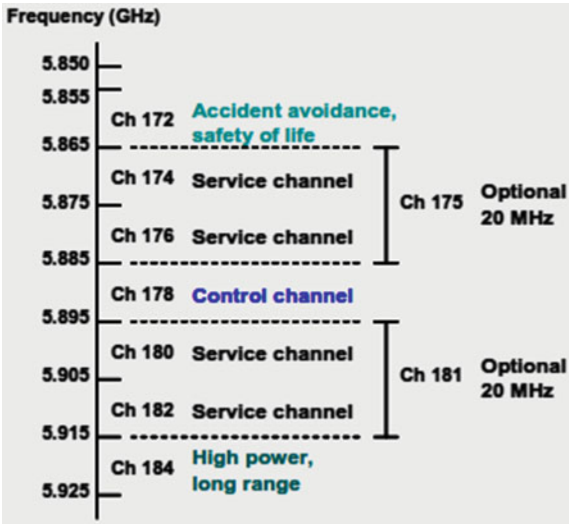


Fig. 1 DSRC channel allocation in US [3]

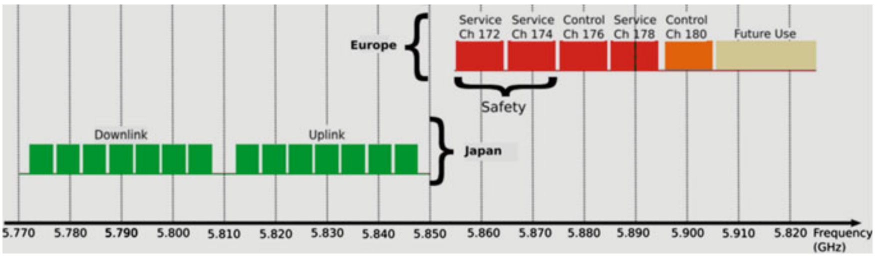


Fig. 2 DSRC channel allocation in Europe and Japan [5]

DSRC channels are used for all types of VANET communications. VANET Communication involves real time as well as non-real time data transfer for verity of its applications. Real time data transfer messages are Cooperative Awareness Messages (CAMs) and Basic Safety Message (BSM). These messages include Periodic beacon and event-driven emergency messages. While non-real data transfer includes e-maps and meteorological information [6]. In vehicular environment, periodic beacon messages are broadcasted at higher rate which leads to congestion of Control Channel (CCH). Maintaining QOS and performance parameters is a challenging task in mobile wireless network scenario like VANET with constrained environment [7].

In vehicular network with constrained resources, transferring of data within network with the help of vehicular nodes is a challenging task. The routing protocol is possible solution as it can identify the intermediate nodes for reliable transmission of data via multi-hopping. Performance of routing protocol is measured quantitatively with the help of QOS parameters [8]. Use of optimization methodologies in designing of routing protocol helps to satisfy QOS requirement and improve network performance parameters [9]. This paper presents comprehensive overview of VANET with its features, challenges, requirements, QOS metrics and review of state of art approaches for designing a congestion aware routing protocol.

The Structure of remaining article as follows. VANET communications types are given in Sect. 2, Sect. 3 presents features of VANET, while Sect. 4 illustrates VANET applications. In Sect. 5 challenging issues and possibilities in VANET are demonstrated. In Subsequent sections we discussed researcher's contribution to address congestion control in VANETs. Section 6 talks about Congestion and Sect. 7 discuss parameters affecting congestion. Section 8 demonstrates congestion control strategies for VANETs, followed by QOS parameters in Sect. 9. State of art research in the VANET congestion control is presented in Sect. 10. Section 11 presents conclusion of review work with remark on the future work.

2 Vanet Communications

The recent improvement in networking technologies helps to improve safety on roads and hence Vehicular networks are deployed through a combined effort of various stake holders like industries, governmental organizations and academia [10].

VANET communications are broadly categorized into V2V and V2I. V2I is again divided into vehicle-to roadside (V2R) communication and communication using cellular networks (V2X). Figure 2 illustrates example of VANET with various V2V, V2R/R2V and V2X/X2V communications. To carry out various communications the VANET should be enabled with modern communication capabilities to gather and route traffic information with internet connectivity [11].

2.1 Vehicle-To-Vehicle (V2V) Communications

This is pure VANET ad-hoc architecture in which vehicles (mobile nodes) with Onboard Units (OBUs) communicates with surrounding nodes without infrastructure support. In this, only vehicular nodes control entire communication.

2.2 Vehicle-To-Infrastructure (V2R)/Infrastructure to Vehicle (R2V) Communications

In this architecture, vehicular nodes can exchange information with the roadside unit RSU/Infrastructure in single-hop or multi-hop fashion.

2.3 Vehicle-To-Cellular Network (V2X)/Cellular Network to Vehicle (X2V) Communications

In V2X, information is exchanged between a cellular network and OBU. A vehicle exchanging messages with a gas station as shown in Fig. 3 is example of V2X Communication [11].

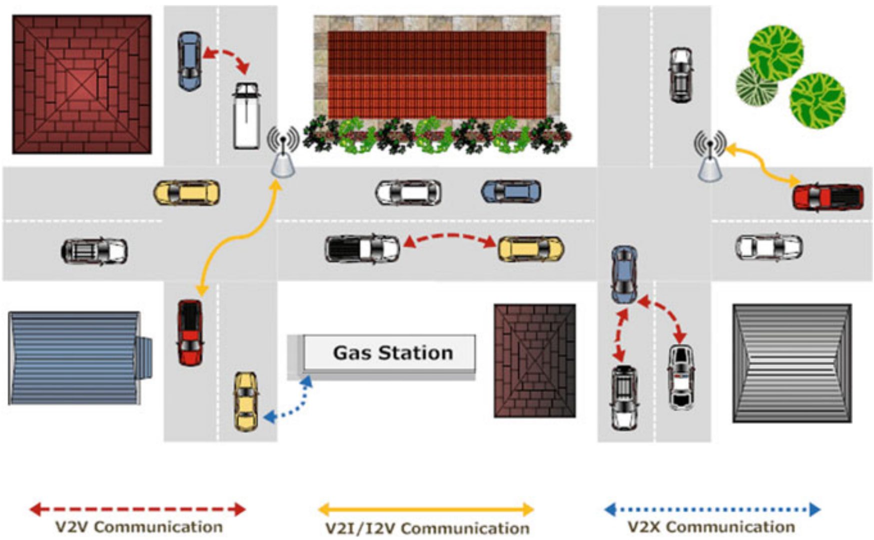


Fig. 3 Example of VANET with various VANET communications [11]

3 Features of Vanet

Special features of VANET which makes VANET different from MANET are briefed as follows:

3.1 Highly Mobile

Speed of Vehicular nodes in VANET leads to highly mobile nodes which cause sudden and quick network topology change and change in channel condition. This characteristic of VANET makes it very difficult to maintain any structures because these structures change rapidly as the topology changed. Due to high mobility, it is very difficult to protect node privacy and track of a node's position in motion [12].

3.2 Rapid Changes in Network Topology

As Vehicular nodes in VANET are moving at high and random speed, the node position changes frequently with movement. Hence, network topology in VANETs changes continuously [12].

3.3 Geographically Unbounded Network

The geographical size of VANET can be of the size of one city, several cities, one state, and one country or for the world [12].

3.4 Rapid and Frequent Exchange of Information

In VANET, the exchange of information is very rapid and frequent as nodes are moving at high speed and it needs data with no latency [12].

3.5 Secured Wireless Communication

Data exchanged in VANET is of critical importance. In VANET, the nodes in motion keep on exchanging information from the other vehicular nodes and roadside units in the wireless environment. Hence, for all VANET Communications security must be examined [12].

3.6 Time Critical

The exchange of information in VANET should be with least possible delay so that the node can take decision in time and carry out action in an appropriate manner [12].

3.7 Sufficient Energy

In VANET, nodes (vehicle) are having Sufficient Energy for all VANET operations [12].

4 Vanet Applications

Based on types of messages exchanged there are two broad categories for VANET applications. These two application classes are Safety applications and non-safety applications.

4.1 Safety Applications

VANET Safety applications focus mainly on intimating drivers about potential hazard by sending emergency messages. VANET provides all safety-related services by the using appropriate routing schemes to communicate. These applications need the minimum delay. These applications include [5].

Collision Avoidance. Research shows that if driver gets an intimated about potential hazard by sending warning (emergency) messages 60% of accidents can be avoided [13, 14].

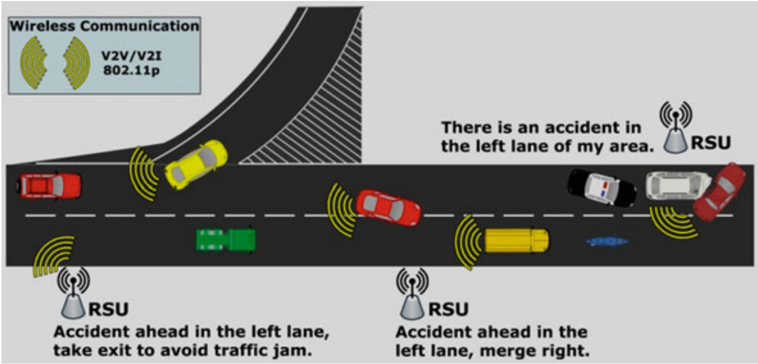


Fig. 4 Example of traffic management application [5]

Cooperative Driving. The drivers have a huge role to play in effective implementation of this application by sharing received warning signals to all neighbor drivers [15, 16]. For example sharing information related to curve speed and lane change.

Traffic Optimization. A vehicular node plays the role of knowledge collector and passes the traffic related information in VANETs in order to manage the traffic. The vehicular traffic can be controlled by broadcasting signals like traffic jam due to congestion on roads, emergency situation or accidents etc. to all vehicles within the network so that vehicles can opt for different path to avoid jam which results in saving time [14] (Fig. 4).

4.2 Non-safety Applications

These applications increases driving experience along with safety-oriented applications e.g. Infotainment applications, sharing of data while driving etc. Apart from safety applications, following are Non-safety-Oriented applications services [14].

Comfort Applications. These applications includes services like sharing music, movies etc. between the nodes of VANET. Unicast routing protocols are used for implementing these applications [5] (Fig. 5).

Continuous Internet Connectivity. These applications permit the clients to share information with anywhere on the web. Users in VANET always want to connect with the Internet.

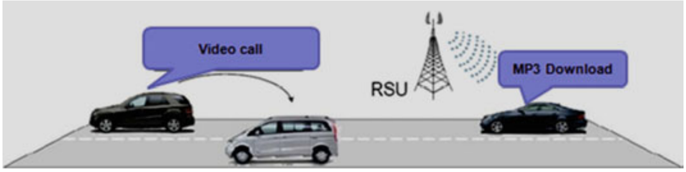


Fig. 5 Example of comfort application [5]

Payment Services. Payment service includes collection of toll, location based services etc. GPS plays an important role in functioning of these VANET services [16].

5 Challenging Issues in Vanet and Possibilities

To maximize the safety and effectiveness of VANET, it is necessary to address problems associated with the design, deployment and implementation of VANET. The research challenges confronting VANET technology are broadly classified as Technical challenges, Social and economic challenges [12].

5.1 Technical Challenges

The challenges should be addressed before the real time implementation of VANET. Some technical challenges are given below:

Management of Network. High mobility of vehicles in the network leads to Sudden and quick changes in channel condition and topology of network. Hence it is difficult to set up and maintain structures rapidly. This leads to improper network management, rapidly changing network topology and unbounded network size [17].

High speed Wireless Communication Technologies. To fulfill the need of vehicular environment these technologies provide air-interface protocols and support. New age communication technologies are also availed for VANET by many researchers. IEEE 802.11p is cost effective, and more relevant for highly mobile and fast changing VANET communications.

Issues in Frequency Spectrum Allocation. 75 MHz bandwidth at 5.850–5.925 GHz is allocated for VANET communications in United States. But, complete and continuous band of 75 MHz is not available in the communication spectrum of Europe as in US. In Europe, for Car to Car communication/Safety messages 2 channels of 10 MHz each are allocated at 5.875–5.925 GHz as a derivative of the approach used in US. This allotment leads to synchronization among all other countries. The remaining spectrum for can be used for non-safety-oriented applications [2–7].

Issues in VANET Routing. MANET routing protocols do not satisfy the requirements in VANET. Apart from special characteristics of VANET, most of the applications of VANET need multi-hop communications between vehicular nodes. Hence, routing has been deeply studied by researchers for addressing issues in VANET.

Over last few years various researchers have been working on design optimal VANET routing solution with minimum latency and minimum dropped packed during routing process without increase in the network overhead [14, 17].

Environmental Impact. Environmental impact must be considered during deployment of VANET as electromagnetic waves used by VANET for communication are affected by the environment [12].

MAC Design. For Communication, VANET uses the shared medium and Medium Access Control design plays important role in sharing of resources. Therefore, VANET should have proper medium assessment strategies.

IEEE has developed IEEE 802.11p/WAVE standard for VANET as existing 802.11 did not meet the VANET challenges and requirements such as low latency, high packet delivery, and multichannel communications. Modifications in 802.11eand 802.11a with an in-built QOS is implemented in IEEE 802.11p/WAVE standard for VANET communications [13].

Security. VANET is majorly used for providing safety on roads to avoid accidents. Therefore, security of these communications must be satisfied considering the importance of life critical road safety applications [17].

Congestion and Collision Control. The density of traffic in rural area and during night in urban area is very less this causes network partition and packet losses. While is city scenarios during rush hours, the traffic density is too high which leads to congestion. So, adaptive and scalable congestion control is demand of growing VANET [2].

5.2 *Social and Economic Challenges*

Although the IEEE 802.11p standard is accepted around the globe by automobile industries [13]. A consumer rejects monitoring for the traffic signal violation and hence it's impossible to convince manufacturers to build a system which causes problems for customer. On contrary, customer likes system which warning message of police trap. Hence, there are very few reasons for motivating the manufacturer to deploy VANET [17, 18].

Hence, social and economic challenges should be considered along with technical challenges to deploy the VANET.

6 Congestion in Vanet

The network congestion in VANET leads degradation of network performance and QOS [19, 20]. To maintain the network performance in real time dense traffic is difficult. The congestion control mechanism mainly focuses on efficient use of network with least latency and packet dropping having no network overhead to provide maximum throughput [21, 22].

In vehicular environment, periodic beacon messages are broadcasted at higher rate which leads to congestion of Control Channel (CCH). These messages contain Speed, direction, identity of each vehicle in network which is can delay safe delivery of emergency messages and hence CCH availability is very important for safety applications. A adaptive congestion aware routing protocol is needed safe delivery of data to the intended destination [2, 22].

7 Parameters Affecting Congestion

Primary parameters causing congestion are beacon frequency, transmit rate and transmit power, while utility function, fairness, prioritization, and node density are derived parameters as they depend on primary parameters [2] (Fig. 6).

7.1 Beacon Frequency

The number of periodic beacon messages broadcasted per unit time by vehicle in VANET is called as beacon frequency. The frequency of beacon messages is directly proportional to vehicle density. This reduces channel load and congestion [23, 24].

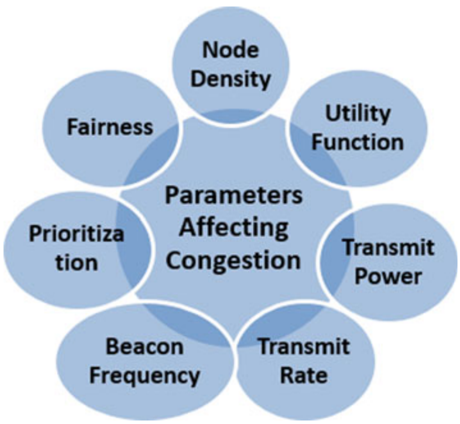


Fig. 6 Parameters affecting congestion

7.2 Transmit Rate

The rate of transmission of packets from a vehicular node in VANET over the channel is called as Transmit rate. Reductions in channel load and channel busy time (CBT) with transmit rate control increases probability of event driven messages and reduce congestion [23, 24].

7.3 Transmit Power

The power content of a vehicular node in VANET is called as Transit Power. By controlling transmission power, nodes beacon message transmission can be controlled. This improves transmission of event driven messages because of reduction in channel busy time (CBT) and thus congestion reduction [23, 24].

7.4 Utility Function

Utility function of sent data packet for an application is measure of determining its utility at current time for that application. The utility function is calculated for each packet arriving at a node in an application. Based on the value of utility function, the rate of particular data packets is controlled [24].

7.5 Fairness

The ability of providing all available resources equally to each vehicular node in VANET is called as Fairness. Fairness gives equal opportunity to access channel, maximizes the network performance by increasing transmit power and coverage to provide reliable and safe environment [23, 24].

7.6 Prioritization

The way of assigning higher priority to emergency safety messages than periodic beacon message is called as prioritization. This increases probability of event driven messages and reduce congestion [24].

7.7 Node Density

The number of vehicular nodes in an area is called as node density. Traffic flow affects node density in VANET. The increase density of vehicles affects network performance as packet transmission flow increases which causes channel load and ultimately channel congestion. The node density plays an important in simulation and analysis purpose [24].

Following subsections examine the role of these parameters in performance of different congestion control algorithms in VANET.

8 Congestion Control Strategies for Vanets

8.1 Proactive, Reactive, and Hybrid Strategies

The congestion control strategies in VANETs can be classified into three classes viz; reactive congestion control, proactive congestion control, and hybrid congestion control strategies. This classification is based on how the strategies are decided to prevent or control the congestion [18, 25–27] (Fig. 7).

In reactive congestion control strategy, first channel gets congested then actions to reduce channel load planned. In proactive strategy, well defined system calculates channel load based on various network parameters and makes sure that is below threshold level and hence it avoids congestion of channel. Proactive and reactive approach advantages are combined in hybrid congestion control [25, 26].

In VANET, the information exchange in vehicular nodes and road side units is mainly used for safety applications. Congested channel conditions leads to degradation in performance. Out of three classes, Proactive approaches are very suitable for VANET as channel congestion can be avoided.

A proactive approach has two drawbacks. One it needs either dynamic or statistical estimation of propagation condition for different environments and second, to

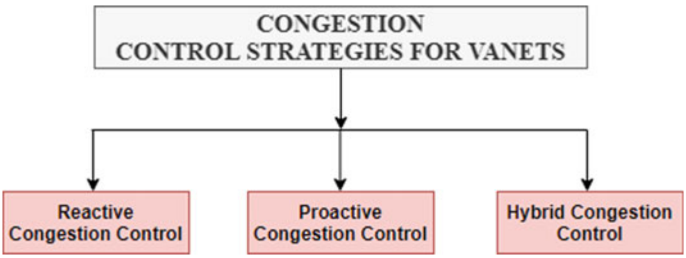


Fig. 7 Classification of congestion control strategies

calculate traffic in a certain period of time. The disadvantage of reactive approach is that it takes action to control only after detection of channel congestion [25, 26].

8.2 Congestion Detection Methods

The two methodologies for VANET congestion detection are, one event driven and second is a measurement-based method. The congestion is detected when an event-driven safety messages are broadcasted in the network. With the detection of congestion, this method stops all communication except the event-driven safety messages [10, 18]. Thus, it improves the performance of safety applications. Cooperative lane change warning is an example of event driven method for congestion detection.

The measurement-based methods, if the channel parameters exceeds predefined thresholds then the congestion is detected [18, 20]. In [28], channel usage based on packet transmission process detects congestion if it exceeds predefined level. The congestion can be controlled by decreasing the transmission rate in case of exceeding service channel queue threshold [23].

9 Performance and QoS Evaluation Parameters

VANET Communication involves real time and non-real time data transfer for verity of its applications. Real time data transfer messages are BSM/CAM. These messages include Periodic beacon and event-driven emergency messages. While non-real data transfer includes e-maps and meteorological information. VANET applications classified under this have multiple strict requirements in terms of QOS and performance. Maintaining QOS and performance parameters is a challenging task in VANET scenario with constrained environment [29].

Till date, different researchers have suggested different solutions to the problem of congestion in VANET. Each approach is having different environment and targets in achieving congestion control in VANET. We can justify the comparison of two approaches only if we compare them on same parameters having same scenarios in same simulation environment. So, we should first decide on performance evaluation parameters to have meaningful comparison. The different performance metrics are discussed in this section, these parameters will help in comparison of various congestion control algorithms.

9.1 Channel Busy Time (CBT)

Load of channel is measured in terms of CBT (Channel Busy Time) and defined as ratio of time the channel is sensed as busy and the total observation time. CBT is taken as input parameter in process of congestion control Vehicles channel load depends on two factors, vehicle density in the coverage area and rate of message generation. It is also referred as Channel CBR (Channel Busy Ratio) or Busy Percentage (CBP) or Channel Busy Function (CBF) [5, 6, 30].

$$\text{Channel Busy Time (CBT)} = \text{channel busy time} \,/\, total \ \text{observation} \ time$$

9.2 Packet Delivery Ratio (PDR)

In network, collisions, fading or control channel availability and other factors are responsible for Packet loss. A total received packet to the total packets issued from the source is termed as PDR. The PDR value is based on many parameters like packet and cluster size, coverage area, traffic scenario and vehicular node mobility. PDR defines the efficiency of the protocol to deliver packets to its destination and assures reliable communication. The inverse of PDR is Packet Loss Ratio (PLR) [5, 6, 30].

$$Packet \ \text{Delivered} = \sum transmitted \ Packets - \sum \text{Lost} \ Packets$$
$$Packet \ \text{Delivery} \ Ratio = \sum Packets \ \text{received}/\sum transmitted \ Packets$$

9.3 Packet Loss Ratio (PLR)

The PLR is the ratio of dropped packets to the total packets generated by the source. PLR measures surrounding vehicular awareness by estimating ratio of packets reaching a vehicle which is just an overview while inter-packet delay (IPD) is in depth analysis. PDR and PLR are interrelated to each other [5, 6, 30].

$$Packetlost = \sum transmitted \ Packets - \sum received \ Packets$$
$$Packet \ loss \ Ratio = \sum Packets \ losts/\sum transmitted \ Packets$$
$$\%\text{PLR} = (1 - \text{PDR}) * 100$$

9.4 Inter-Packet Delay (IPD)

The Inter Packet Delay (IPD) or Inter Reception Time (IRT) is defined as the average delay between two successive packets receptions for a particular sender-receiver pair in VANET. IPD helps to design Transmission power control algorithm. If IPD is more then old information is received which can create problems in emergency situations and hence must be discarded. IPD measures surrounding vehicular awareness by providing more through view. As awareness varies since same PLR may have different IPD. IPD helps to determine the reliability of VANET with increasing car density for timely information exchange [6].

9.5 End to End Delay (E2E Delay)

The end-to-end delay is defined as the average amount of time taken by each vehicle packets to travel from source node to destination node. End to End delay/Latency includes channel access delay, next node forwarding delay, transmission time and propagation delay. In VANET, routing overhead or links failure leads to increase in E2E Delay.E2E delay is calculated as sum of mean of delays for all vehicles in the VANET. The average E2E delay of VANET can also be calculated as,

$$E2E\,\text{Delay} = (TR - TS) * number\ of\ packets$$

where, T_R is time at which packet is received at destination node, T_S is time at which packet is sent from source node [30].

9.6 Bandwidth Utilization

Bandwidth specifies how much data can be transferred through the link between sender and receiver in the network. Bandwidth utilization is defined as amount of frequency band used from the total bandwidth available. Bandwidth sharing plays very important role in network performance. Major cause of congestion in VANET is high utilization of Bandwidth [31].

9.7 Network Throughput

Network throughput in bits per second (bit/s or bps) is actual measure of how much data is sent successfully while bandwidth is how much data can be sent over channel theoretically [5]. Higher throughput signifies better network performance [30].

9.8 Overhead

Beacon messages in VANET are major cause of the network overhead. Overhead is ratio of non-useful data transmitted to the total data transmitted [5].

Above discussed performance parameters are main entities in deciding performance of Network, along with these parameters few other parameters also plays vital role in QOS estimation which are discussed below.

9.9 Scalability

Scalability is referred as adaptability to efficiently work in several situations and scenarios. The performance should not degrade with changing situations and scenarios. The scalable networks are in high demand but it's very difficult achieve. For better network performance, the network should be highly scalable [5].

9.10 Reachability

Reachability is a ratio of number of vehicles receiving broadcasted data to the total number of vehicles in VANET. More the reachability, the better is the broadcasting method. Ideally broadcasting scheme should have 100% reachability [5].

9.11 Redundancy

Redundancy calculated as total number of re-transmissions required in order to send a single message from sender to ultimate receiver in broadcasting scheme. This is needed to recover data in case of failed data transmission. To overcome the problem and retrieve data, the sender node has to retransmit the same message. The network should have less redundancy for better network performance [5].

9.12 Propagation Speed

Propagation speed is measured in terms of time taken to reach from source to intended destination. In wireless environment, the transmissions are mostly multihop. So, in VANET its related required hops for data transmission from sender to furthest receiver [5].

Additional matrices such as Packet Collision Rate (PCR), Information Dissemination Rate (IDR), error rate, noise, and end-to-end connection cost can be considered depending on VANET service [32].

10 State of Art Approaches for Congestion Control

Rate-based, power-based, CSMA/CA-based, scheduling and prioritization-based, and hybrid strategies are 5 commonly used approaches for congestion control in VANET [2, 20].

10.1 Rate-Based Strategies

Following are the few approaches employing rate based strategies in which rate of generation and transmission of packets is varied in accordance with channel load for congestion control.

Wischhof [34] suggested decentralized Utility-based Packet Forwarding and Congestion Control (UBPFCC) while On-Demand Rate Control (ODRC) proposed by Huang et al. [35] adjusts transmission rate (Event driven and Beacon Messages) to control congestion.

Seo et al. [36], R. K. Schmidt et al. [37] and Sommer et al. [38], controls transmission rate for beacon messages which reduces channel load to control congestion. He et al. [39] transmissions rate of event-driven messages transmission range is kept more than beacon messages. If congestion is detected, MAC blocking method blocks all beacon messages and control channel is made available for emergency/event driven messages only.

In the work proposed by Ye et al. [40], rate of packets flow is decided using broadcasting efficiency and reliability. T. Tielert et al. [13] suggests that transmission range is decided by channel load using Periodically Updated Load Sensitive Adaptive Rate (PULSAR). A dual queue scheduler suggested by Shwetha A. et al. [41] sends packet data via less congested paths and hence controls traffic.

Model based on two modules is suggested by Darus, M. Y et al. [10] ensure high reliability and timely delivery. First module is measurement based while second module is event driven detection based. This deals with both event driven messages and periodic messages.

10.2 Power-Based Strategies

Following are the few approaches employing power based strategies in which transmission power is adjusted to have Communication fairness to all nodes dynamically [25, 27].

Distributed-Fair Power Adjustment for Vehicular environment (D-FPAV) overcomes disadvantages of FPAV. In D-FPAV, event driven messages are given more priority than beacon messages as traffic density decides range for beacon transmission [42] while it was same in FPAV [33]. Part of total bandwidth is kept event-driven communications in FPAV and D- FPAV [42].

Mathematical model to estimate channel occupancy is given by Fallah et al. [43]. This model uses vehicle density, transmission range, contention window size as input to decide channel occupancy.

The original idea of congestion control via dynamic transmit power control is to reduce power for low priority messages and to increase power for high priority messages. These approaches will provide increased communication range till 1 km with increased transmit power level with in infrastructure equipped with Road Site Units (RSUs) [22].

10.3 CSMA/CA-Based Strategies

Following are the few approaches employing CSMA/CA based strategies in which window size and frame spacing is adjusted to avoid congestion [25, 27].

H. C. Jang et al. [44] suggests reduction in channel overhead by increasing contention window size. C. W. Hsu et al. [45] suggests adjustment based on density of vehicles. The adjustment in contention window size and Arbitration inter-frame spacing (AIFS) shows reduction in collision. Stanica et al. [46] suggests Safety Range CSMA (SR-CSMA), in this higher reception rate as transmission range for safety messages is kept less than other messages.

Back-off window size and AIFS adjustment helps to overcome drawbacks of 802.11p's Enhanced Distributed Channel Access (EDCA). Vehicular network performance is improved by avoiding retransmission of messages through MAC congestion control strategy proposed by Barradi et al. [47].

10.4 Scheduling and Prioritization-Based Strategies

Following are the few approaches employing prioritizing and scheduling strategies in which proper planning for prioritizing and scheduling prevents channel saturation which ultimately controls the congestion [2, 22].

Suthaputchakun in [48] suggests priority is decided based on the average delay and urgency. Bouassida and M. Shawky in [49] suggests size of the messages along with static and dynamic parameters decides priority of messages.

Smallest Data-size First (SDF),Longest Wait Time (LWT),First-in-First out (FIFO), First Deadline First (FDF), Longest Total Stretch First (LTSF) and Deadline and Size (D*S) algorithm are various approaches which decides priority [50].Flooding because of beacon messages is addressed by Context Aware Beacon Scheduling (CABS) suggested by Bai et al. [51].

Switching between multi-channel VANETs is enabled in the work suggested by Felice et al. [52] through proposed WAVE-enhanced Safety message Delivery strategy. This helps in reducing delay. Collision free safe broadcasting is achieved in VeMAC [20] suggested by Hasan et al. This approach allocates different time slots to all nearby nodes.

Researches in [18] were proposed limiting the forwarded packets to avoid congestion but this leads to loss of some useful information. The researchers in [17] suggested efficient rebroadcasting for vehicles in the same lane which reduces the transmission rate with minimum overhead. On roads, accidents also involve other lanes.

10.5 Hybrid Strategies

The combination four congestion control strategies mentioned above is used in hybrid strategies to control two or more parameters and reduce congestion [2, 27].

In the research work [23], K. Ravikumar et al., suggests two unit (hybrid) methods. In Tabu search based control unit, first, congestion is detected with first event driven message in network. Second, if threshold level of 70% is crossed by channel load. The Proactive Application-based system suggested by Sepulcre et al. [25] uses a combination of power and rate based approach. The exchanging CAM's helps to reduce channel overload and contributes in reduction of congestion.

Channel fairness is provided by Vehicle Oriented Congestion Control Algorithm suggested in [53]. In AVOCA, parameters controlling congestion gets reset when vehicle enters the zone and gets freeze when vehicle leaves coverage area with stoppage of packet transmission. Reception rate of the emergency message is increased as power and transmission rate is shared as suggested by Djahel et al. in [54] but leads to increase in overhead in channel and delay.

Uni-Objective Tabu (UOTabu) approach explained in [55] is based on metaheuristic techniques. Transmission rate and range depending on channel usage is decided by Tabu search. This improves overall vehicular network performance. Time interval in control channel and transmission rate is decided by utility of packets as discussed in Adaptive Message Rate Control (AMRC) [56].

Table 1 Comparison of VANETs congestion control strategies

Congestion Control Strategy	Methodology	Advantages	Disadvantages
Rate-based strategies	Adjustment in transmission rate and generation rate based on the channel load	QOS for Safety applications improved	Causes congestion of control channel
Power-based strategies	Adjustment in transmission power	Communication fairness for all nodes with minimum collision	Causes channel collision and saturation
CSMA/CA-based Strategies	Adjustment in window size and frame spacing	Retransmission is avoided and reception ratio is increased	No consideration of real time issues of VANET
Scheduling and prioritization-based Strategies	Scheduling and priority allotment for messages and channels	Channel utilization is more with no wastage of bandwidth. Data transmitted with minimum delay	Increases overhead
Hybrid Strategies	Combination of above strategies	Suitable for complex and real time VANET	Increases overhead

In [57], author's fine-tuned D-FPAV by using hybrid approach based on transmission power and rate control mechanism. Maximum Beaconing load (MBL) assigned dynamically on the basis of traffic and non-traffic situations.

Bhagyashri P. Deshpande et al. in [58] proposed a congestion control technique which has three modules. In the first module, from all available nodes, only few nodes are activated which are responsible to transmit data. It saves energy and also reduces collision. Second module uses Relaxation Theory and Max–Min fairness (RT-MMF) algorithm for controlling congestion. The distance between sender and receiver is calculated using minimum distance algorithm. The packet transmission rate at each node controlled by RT is called Engineering Level (EL). EL allocates fair bandwidth (number of slots) to each packet without any additional loss or delay to reduce the probability of congestion. If threshold bandwidth exceeds, communication between nodes will stop and EL will be reset. For the security purpose messages are encrypted using SHA-256 algorithm.

The hybrid strategies presented in [22–25] and [53–58]. Some recent fuzzy-based methods which use hybrid strategies [59–65] were designed to predict congestion and make reliable transmission of security messages in VANET.

Below is compression table for congestion control strategies discussed above (Table 1).

11 Conclusions and Future Works

Routing and congestion control in VANET is one of the topics of interest topic for researchers. Routing is one of the solution to achieve congestion control and improve network performance and QOS. The special features of VANET such as highly mobile vehicular nodes which causes rapid topology changes and link disconnections causes congestion and makes routing a challenging task. This paper reviews all possible aspects of VANET, Communications in VANET, Features of VANET, Application areas of VANET, challenging issues and possibilities in VANET before discussing the problem of congestion and its possible solutions. Various QOS and performance features are also discussed as selection of routing protocol requires satisfying the expectations with respect to QOS evaluation. In this paper we present comparative study based on some selected literature in the field of routing and congestion control.

This survey work helped us in identifying major challenges faced in VANET research and clears that researcher uses quantitative approach in VANET QOS Context. This review work suggest that researchers have contributed in overcoming VANET challenges but still we need to improvement considering ITS QOS fulfillment. Although, VANET is extensively studied till date we have not achieved what is expected from VANET in terms of ITS. Still there is scope of improvement in VANET overall performance. As discussed in this review article, our concern is to contribute to existing research on VANET QOS improvement.

Our future work will focus on new congestion aware routing solutions to enhance QOS and the overall system performance. We are proposing adaptive solution to predict congestion control and establish more reliable and safe data transmission routes in vehicular environment for QOS improvement.

References

1. World Health Organization: Global Status Report on Road Safety 2015, World Health Organization, Geneva, Switzerland (2015)
2. Giripunje, L.M., Masand, D., Shandilya, S.K.: Congestion Control in vehicular ad-hoc networks (VANET's): a review. In: Abraham, A., Shandilya, S., Garcia-Hernandez, L., Varela, M. (eds.) Hybrid Intelligent Systems. HIS 2019, Advances in Intelligent Systems and Computing, vol. 1179. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-49336-3_26
3. Jiang, D., Delgrossi, L.: IEEE 802.11p: towards an international standard for wireless access in vehicular environments. In: VTC Spring 2008–IEEE Vehicular Technology Conference, 2008, pp. 2036–2040 (2008). https://doi.org/10.1109/VETECS.2008.458
4. IEEE Draft Standard for Wireless Access in Vehicular Environments (WAVE)–Multi-Channel Operation. In: IEEE Unapproved Draft Std P1609.4/D6.0, vol., no., 2010, March 2010
5. Mchergui, A., Moulahi, T., Alaya, B., Nasri, S.: A survey and comparative study of QOS aware broadcasting techniques in VANET. Telecommun. Syst. https://doi.org/10.1007/s11235-017-0280-9
6. Liu, X., Jaekel, A.: Congestion control in V2V safety communication: problem, analysis approaches. Electronics **8**(5), 540 (2019). https://doi.org/10.3390/electronics8050540
7. Toh, C.: Ad Hoc Mobile Wireless Networks: Protocols and Systems, Pearson Education: London, UK (2001)

8. Darwish, T., Abu Bakar, K.: Traffic aware routing in vehicular ad hoc networks: characteristics and challenges. Telecommun. Syst. **61**, 489–513 (2016). https://doi.org/10.1007/s11235-015-0008-7
9. Kaur, P., Singh, A.: Nature-inspired optimization techniques in VANETs and FANETs: a survey. In: Bhattacharyya, S., Chaki, N., Konar, D., Chakraborty, U., Singh, C. (eds.) Advanced Computational and Communication Paradigms. Advances in Intelligent Systems and Computing, vol. 706. Springer, Singapore (2018). https://doi.org/10.1007/978-981-10-8237-5_63
10. Darus, M.Y., Bakar, K.A.: Congestion control algorithm in VANETs. World Appl. Sci. J. **21**(7), 1057–1061 (2013). https://doi.org/10.5829/idosi.wasj.2013.21.7.242
11. Abdelhamid, S., Benkoczi, R., Hassanein, H.S.: Vehicular clouds: ubiquitous computing on wheels. In: Adamatzky, A. (eds.) Emergent Computation. Emergence, Complexity and Computation, vol. 24. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-46376-6_20
12. Raw, R.S., Kumar, M., Singh, N.: Security challenges, issues and their solutions for VANET. Int. J. Netw. Secur. Appl. (IJNSA) **5**(5), 95–105 (2013)
13. Jeremiah, C., Nneka, A.J.: Issues and possibilities in vehicular ad-hoc networks (VANETs). In: 2015 International Conference on Computing, Control, Networking, Electronics and Embedded Systems Engineering (ICCNEEE) (2015)
14. Singh, S., Agrawal, S.: VANET routing protocols: issues and challenges. In: Proceedings of 2014 RAECS UIET Panjab University Chandigarh, 06–08 March 2014
15. Karagiannis, G., Altintas, O., Ekici, E., Heijenk, G., Jarupan, B., Lin, K., et al.: Vehicular networking: a survey and tutorial on requirements, architectures, challenges, standards and solutions. IEEE Commun. Surv. Tutor. **13**(4), 584–616 (2011)
16. Cristofer, E., Chen, L., Vinel, A., Lin, S.Y.: Future applications of VANETs. In: Campolo, C., Molinaro, A., Scopigno, R. (eds.) Vehicular ad hoc Networks, pp. 525–544. Springer International Publishing, Cham (2015). https://doi.org/10.1007/978-3-319-15497-8-18
17. Zhang, W., Festag, A., et al.: Congestion control for safety messages in VANETs: concepts and framework. In: Proceeding 8th Conference on ITS Telecommunications (ITST), pp. 199–203, Thailand (2008)
18. Zang, J., Stibor, L., et al.: Congestion control in wireless networks for vehicular safety applications. In: Proceeding the 8th European Wireless Conference, Paris, France (2007)
19. Sattari, M.R.J., Noor, R.M., Keshavarz, H.: A taxonomy for congestion control algorithms in vehicular ad hoc networks. In: COMNETSAT 2012, pp. 44–49. IEEE (2012). 978-1-4673-0889-2/12/
20. Goswami, B., Asadollahi, S.: Novel approach to improvise congestion control over vehicular ad hoc networks (VANET). In: 2016 International Conference on Computing for Sustainable Global Development (INDIACom). IEEE (2016). 978-9-3805-4421-2/16/
21. Rostami, A., et al.: Stability challenges and enhancements for vehicular channel congestion control approaches. IEEE Trans. Intell. Transp. Syst. **17**(10), 2935–2947 (2016)
22. Taherkhani, N.: Congestion control in vehicular ad hoc networks (doctoral dissertation), University of Montreal, Canada (2015)
23. Ravikumar, K., Vishvaroobi, T.: Congestion control in vehicular ad hoc networks (VANET) using meta-heuristic techniques. Int. J. Comput. Sci. Trends Technol. (IJCST) **5**(4), 66–72 (2017). ISSN: 2347-8578
24. Vyas, I.B., Dandekar, D.R.: Review on congestion control algorithm for VANET. In: International Conference on Quality Up-gradation in Engineering, Science and Technology (ICQUEST-2014) (2014). International Journal of Computer Applications (0975 – 8887)
25. Sepulcre, M., Mittag, J., Santi, P., Hartenstein, H., Gozalvez, J.: Congestion and awareness control in cooperative vehicular systems. In: Proceedings of the IEEE, vol. 99, pp. 1260–1279 (2011)
26. Jabbarpour, M.R., Noor, R.M., Khokhar, R.H., Ke, C.H.: Cross-layer congestion control model for urban vehicular environments. J. Netw. Comput. Appl. **44**, 1–16 (2014)
27. Shen, X., Cheng, X., Zhang, R., Jiao, B., Yang, Y.: Distributed congestion control approaches for the IEEE 802.11 p vehicular networks. IEEE Intell. Transp. Syst. Mag. **5**, 50–61 (2013)

28. Baldessari, R., Scanferla, D., Le, L., Zhang, W., Festag, A.: Joining forces for vanets: acombined transmit power and rate control algorithm. In: 6th International Workshop on Intelligent Transportation (WIT) (2010)
29. Hossain, E., et al.: Vehicular telematics over heterogeneous wireless networks: a survey. Comput. Commun. **33**(7), 775–793 (2010)
30. Alwan, M.H., Ramli, K.N.: Performance evaluation for high speed vehicle in VANET. Int. J. Appl. Eng. Res. **13**(10), 7937–7941 (2018). ISSN 0973-4562
31. Kamini, R.K.: VANET parameters and applications: a review. Global J. Comp. Sci. Technol. **10**(07), 72–77 (2010)
32. Kassem, A.: QOS provisioning in VANET infrastructure, pp. 1–11. Lebanon: Faculty of Engineering, IUL University (2015)
33. Torrent-Moreno, M., Mittag, J., Santi, P., Hartenstein, H.: Vehicle-to-vehicle communication: fair transmit power control for safety-critical information. IEEE Trans. Veh. Technol. **58**, 3684–3703 (2009)
34. Wischhof, L., Rohling, H.: Congestion control in vehicular ad hoc networks. In: Proceeding of IEEE International Conference on Vehicular Electronics and Safety, Germany, pp. 58–63 (2005)
35. Huang, C.-L., Fallah, Y.P., Sengupta, R., Krishnan, H.: Information dissemination control for cooperative active safety applications in vehicular ad-hoc networks. In: IEEE Global Telecommunications Conference, 2009, GLOBECOM 2009, pp. 1–6 (2009)
36. Seo, H., Yun, S., Kim, H.: Solving the coupon collector's problem for the safety beaconing in the IEEE 802.11 p WAVE. In: IEEE 72nd Vehicular Technology Conference Fall (VTC 2010-Fall), pp.1–6 (2010)
37. Schmidt, R.K., Leinmüller, T., Schoch, E., Kargl, F., Schäfer, G.: Exploration of adaptive beaconing for efficient inter vehicle safety communication. IEEE Netw. **24**(1), 14–19 (2010)
38. Sommer, C., Tonguz, O.K., Dressler, F.: Traffic information systems: efficient message dissemination via adaptive beaconing. IEEE Commun. Mag. **49**(5), 173–179 (2011)
39. He, J., et al.: Adaptive congestion control for DSRC vehicle networks. IEEE Commun. Lett. **14**(2), 127–129 (2010)
40. Ye, F., Yim, R., Roy, S., Zhang, J.: Efficiency and reliability of one-hop broadcasting in vehicular ad hoc networks. IEEE J. Sel. Areas Commun. **29**(1), 151–160 (2011)
41. Shwetha, A., Sankar, P.: Queue management scheme to control congestion in a vehicular based sensor network In: Proceedings of the Second International Conference on Inventive Systems and Control (ICISC 2018), pp. 917–921. IEEE (2018)
42. Torrent-Moreno, M., Santi, P., Hartenstein, H.: Distributed fair transmit power adjustment for vehicular ad hoc networks. In: 2006 3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks, pp. 479–488 (2006). https://doi.org/10.1109/SAHCN.2006.288504
43. Fallah, Y., Huang, C., Sengupta, R., Krishnan, H.: Congestion control based on channel occupancy in vehicular broadcast networks. In: IEEE Vehicular Technology Conference fall (VTC), Canada, pp.1–5, September 2010
44. Jang, H.-C., Feng, W.-C.: Network status detection-based dynamic adaptation of contention window in IEEE 802.11 p. In: IEEE 71st Vehicular Technology Conference (VTC 2010-Spring), pp. 1–5 (2010)
45. Hsu, C.-W., Hsu, C.-H., Tseng, H.-R.: MAC channel congestion control mechanism in IEEE 802.11 p/WAVE vehicle networks. In: IEEE Vehicular Technology Conference (VTC Fall), pp. 1–5 (2011)
46. Stanica, R., Chaput, E., Beylot, A.L.: Congestion control in CSMA-based vehicular networks: Do not forget the carrier sensing. In: 9th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON), pp. 650–658 (2012)
47. Barradi, M., Hafid, A.S., Gallardo, J.R.: Establishing strict priorities in IEEE 802.11 p WAVE vehicular networks. In: IEEE Global Telecommunications Conference, pp. 1–6 (2010)
48. Suthaputchakun, C.: Priority-based inter-vehicle communication for highway safetymessaging using IEEE 802.11 e. Int. J. Veh. Technol. **2009**, 1–13 (2009)

49. Bouassida, M.S., Shawky, M.: A cooperative congestion control approach with in VANETs: formal verification and performance evaluation. EURASIP J. Wirel. Commun. Netw. **2010**(11), 399–403 (2010)
50. Kumar, V., Chand, N.: Data scheduling in VANETs: a review. Int. J. Comput. Sci. Commun. **1**, 399–403 (2010)
51. Bai, S., Oh, J., Jung, J.-I.: Context awareness beacon scheduling scheme for congestion control in vehicle to vehicle safety communication. Ad Hoc Netw. **11**, 2049–2058 (2013)
52. Felice, M.D., Ghandour, A.J., Artail, H., Bononi, L.: Enhancing the performance of safety applications in IEEE 802.11 p/WAVE vehicular networks. In: IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), pp. 1–9 (2012)
53. Huang, Y., Fallon, E., Qiao, Y., Rahilly, M., Lee, B.: AVOCA–A vehicle oriented congestion control algorithm. In: ISSC, Trinity College Dublin (2011)
54. Djahel, S., Ghamri-Doudane, Y.: A robust congestion control scheme for fast andreliable dissemination of safety messages in VANETS. In: IEEE Wireless Communications and Networking Conference (WCNC), pp. 2264–2269 (2012)
55. Taherkhani, N., Pierre, S.: Congestion control in vehicular ad hoc networks using Meta-heuristic techniques. In: Proceedings of the Second ACM International Symposium on Design and Analysis of Intelligent Vehicular Networks and Applications, pp. 47–54 (2012)
56. Guan, W., He, J., Ma, C., Tang, Z., Li, Y.: Adaptive message rate control of infrastructured DSRC vehicle networks for coexisting road safety and non-safety applications. Int. J. Distrib. Sens. Netw. **2012**, 95–108 (2012)
57. Sattari, M.R.J., Noor, R.M., Ghahremani, S.: Dynamic congestion control algorithm for vehicular ad-hoc networks. Int. J. Softw. Eng. Appl. **7**(3), 95–108
58. Bhagyashri, P., Deshpande, B.P., Alvi, A.S.: Congestion control technique with safety transmission of messages in vehicular ad hoc network. Int. Res. J. Eng. Technol. **04**(02) (2017). e-ISSN: 2395-0056
59. Amiri, E., Hooshmand, R.: Improved AODV based on TOPSIS and fuzzy algorithms in vehicular ad-hoc networks. Wirel. Pers. Commun. (2019). https://doi.org/10.1007/s11277-019-06894-x
60. Mchergui, A., Moulahi, T., Nasri, S.: QoS evaluation model based on intelligent fuzzy system for vehicular ad hoc networks. Computing **102**, 2501–2520 (2020). https://doi.org/10.1007/s00607-020-00820-x
61. Ahmed, M., Iqbal, S., Awan, K., Sattar, K., Khan, Z.A., Sherazi, H.: A congestion aware route suggestion protocol for traffic management in internet of vehicles. Arab. J. Sci. Eng. (2019). https://doi.org/10.1007/s13369-019-04099-9
62. Shelke, M., Malhotra, A., Mahalle, P.N.: Fuzzy priority based intelligent traffic congestion control and emergency vehicle management using congestion-aware routing algorithm. J. Ambient Intell. Humaniz. Comput. (2019). https://doi.org/10.1007/s12652-019-01523-8
63. Santamaria, A.F., Fazio, P., Tropea, M., Raimondo, P., De Rango, F.: A new distributed predictive congestion aware re-routing algorithm for ${\rm CO}_2$ emissions reduction. IEEE Trans. Veh. Technol. (1) (2019). https://doi.org/10.1109/tvt.2019.2905753
64. Rahnamaei Yahiabadi, S., Barekatain, B., Raahemifar, K.: TIHOO: an enhanced hybrid routing protocol in vehicular ad-hoc networks. EURASIP J. Wirel. Commun. Netw. **2019**(1) (2019). https://doi.org/10.1186/s13638-019-1503-4
65. Miri, S.T., Tabatabaei, S.: Improved routing vehicular ad-hoc networks (VANETs) based on mobility and bandwidth available criteria using fuzzy logic. Wirel. Pers. Commun. (2020). https://doi.org/10.1007/s11277-020-07278-2

Design and Analysis of Fast Frequency Acquisition Phase Frequency Detector

Nigidita Pradhan and Sanjay Kumar Jana

Abstract In this paper, we present the analysis of the conventional phase detector (PD) and phase frequency detector (PFD). Then, we have proposed the modified PFD using D-Flip Flop (DFF) based on true single-phase clock (TSPC) topology. The conventional PFD generates the UP/DN signal with respect to the phase difference between the two inputs. However, the proposed technique of immediate reset path improves the operating frequency in the order of 1.5 MHz–2.4 GHz. Also, it shows the minimum power consumption as compared with the traditional PFD architecture @ 416. 83 μW. Along with this, it shows the improvement in terms of phase noise i.e., –102.3 dBc/Hz at 1 MHz offset frequency. The design is simulated in a standard 0.18 μm CMOS technology node with a 1.8 V supply voltage. Furthermore, the achieved frequency band is applicable for high-speed and low-power PLL applications such as Zigbee, Wi-Fi, and Bluetooth.

Keywords Phase Detector (PD) · Phase Frequency Detector (PFD) · Phase noise · High speed · Phase Locked Loop (PLL)

1 Introduction

Phase Locked Loop (PLL) is one of the important blocks in modern communication [1, 2]. This circuit becomes highly convenient wherever there is a need for the local oscillator to synchronize with an independent incoming signal, such as serial data links and RF wireless communications [1, 6–9]. The traditional PLL showed in Fig. 1(a) consists of PD, Charge pump (CP), Loop Filter (LP), Voltage Controlled Oscillator (VCO), and Frequency Divider (FD) in the reset path. The PD block is the key element of the PLL [1–3]. This PD detects the phase difference between the

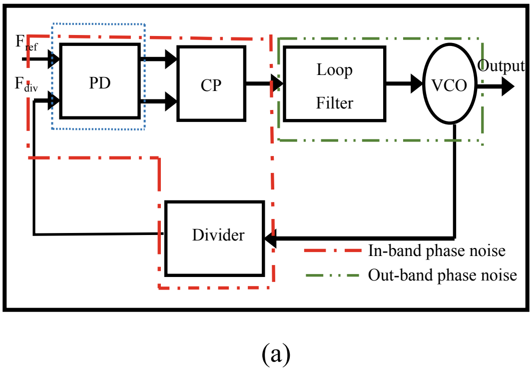
N. Pradhan (✉) · S. K. Jana
Department of Electronics and Communication Engineering NIT SIKKIM, Ravangla, India
e-mail: nigidita@nitsikkim.ac.in

S. K. Jana
e-mail: skjnit@nitsikkim.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_30

two inputs i.e., the reference signal (F_{ref}) and the divider output frequency (F_{div}), and generates the error signal corresponding to the duration of the mismatch between the two inputs. However, the conventional PD suffers from low speed and small phase comparison range i.e., $-\pi$ to $+\pi$ [3–5]. Therefore, there was the need to design a high-speed PD that can correctly function in the high frequency range [6, 7]. With the advancement in technology, there was the requirement to detect the difference between both the phase and frequency as well. In this regard, a phase frequency detector was designed which is the type of PD having the ability to detect both the phase and frequency. Figure 1(b) shows the conventional PFD with two resettable DFF and AND gate to perform reset operation. This, PFD block detects the difference between the phase and frequency of the two inputs (F_{ref} and F_{div}) and passes the information to their corresponding outputs (UP/DN). Several researchers have proposed different PD architectures to improve its performance [2, 3]. However, with the evolution in technology, there was a certain issue in PD which was overcome by PFD. Following the rise in demand for the PFD, various architecture has proposed in the state of art work such as nc-pfd [1], precharge PFD [3], latch based PFD [6], pass transistor PFD [6], and dynamic logic tri-state PFD [8], but all this design consist of the reset path which elongates the reset time. Thus, we present the analysis of conventional EX-OR based PD and PFD designed using the TSPC based DFF in an open-loop configuration. From the analysis, the issue in the conventional PFD and PD was understood. In this regard, the modified PFD architecture with an immediate reset path is proposed to improve the overall performance of the PLL.

The remaining sections of the paper are organized as follows: Sect. 2 presents the EX-OR based PD and its analysis, Sect. 3 details the architecture of true single-phase (TSPC) based DFF and its working operation, Sect. 4 presents the architecture of conventional PFD using TSPC DFF, Sect. 5 briefly presents the proposed modified PFD architecture and its analysis, and finally, Sect. 6 draws the conclusion.



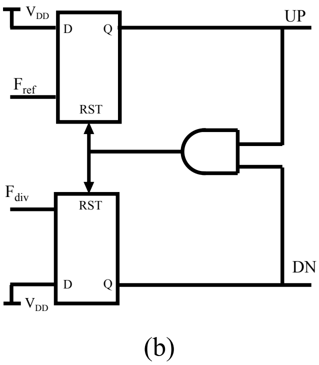


Fig. 1 **a** Conventional phase locked loop [5] **b** D-Flip Flop PFD

2 EX-OR Based Phase Detector and Its Analysis

The phase detector is the core element of a PLL [2, 8]. One of the simplest forms of PD is EX-OR based which is also known as comparator [2]. Basic EX-OR based PD has been implemented in the cadence virtuoso environment using SCL 180 nm technology node with a supply voltage of 1.8 V. Figure 2(a) shows the circuit of PD consisting of three stages where the variation of the two inputs (D and CLK) presents the output. The working is such that when both the inputs are at logic '1' the output of 1st stage is '0', which makes the output of 2nd stage also '0'.

Moreover, in the 3rd stage, the transistor M6 and M12 is OFF and transistor M8 and M10 is ON, therefore the output of the 3rd stage is logic low i.e., '0'. However, the transient simulation is performed for various cases of phase error to verify the functionality of the PD. Figure 2(b) shows the transient response when the phase difference between the two inputs is '0'. Here, we can observe that the output is '0' but it consists of the reference spur which is the unwanted disturbances in the PLL. This spur causes the problem in the PLL applications, for example in the radio receivers they can mix the interference into the channel bandwidth which in turn degrades the signal to noise ratio (SNR) [9]. Further, Figs. 3 and 4 show the transient response with the phase difference of π and π/2 respectively. At the output we can see, when both the inputs (D and CLK) are at the same frequency then, if the phase of two inputs changes, the output varies accordingly. However, the PD fails when there is a frequency difference between two inputs, as the EX-OR PD provides the output which does not give any information about the change in frequency. Thus, there will be false locking during the PLL acquisition time.

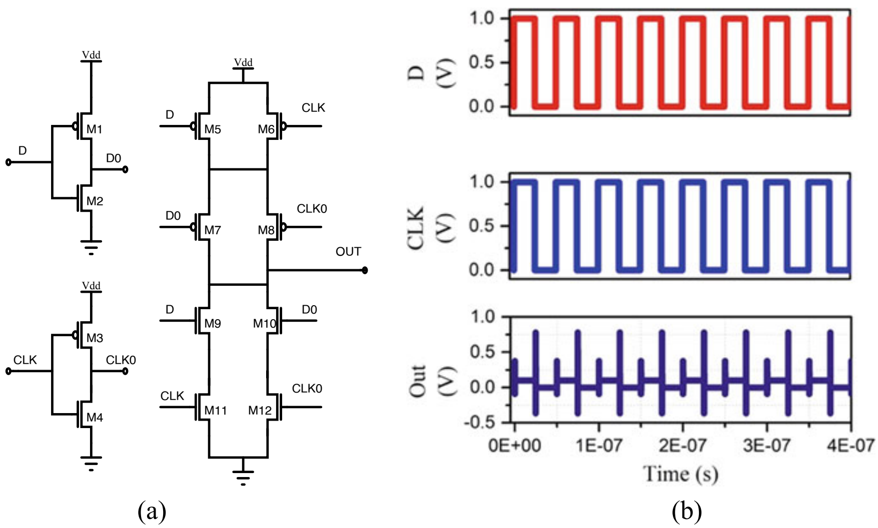


Fig. 2 **a** EX-OR based phase detector [2] **b** Timing diagram showing the working of PD with phase difference = 0

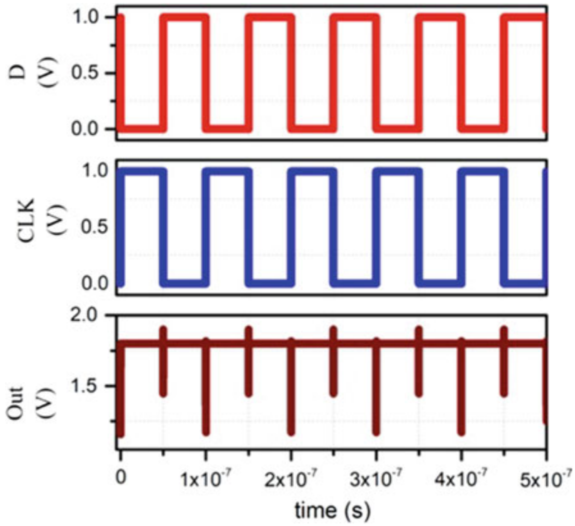


Fig. 3 Transient response with phase difference = π

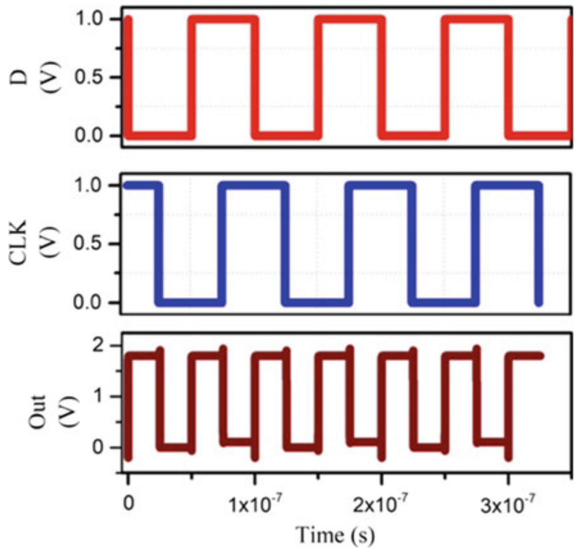
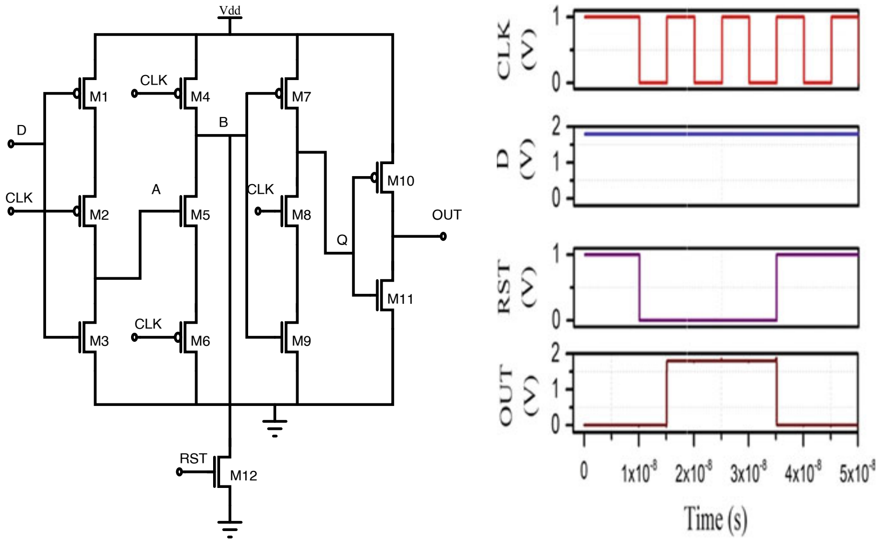


Fig. 4 Transient response with phase difference = π/2

3 True Single-Phase Clock (TSPC) Based DFF

Figure 5(a) shows the architecture of TSPC based DFF with two inputs (D and CLK) and RST designed at 180 nm CMOS technology node with 1.8 V power supply. Then, the performance of DFF is analyzed [13]. The advantage of this topology is that it turns on the power only during switching, which in turn reduces the power consumption. Figure 5(b) shows the transient response of the TSPC DFF. In the



(a) (b)

Fig. 5 TSPC based DFF **a** Architecture [13] **b** Transient response

circuit, the input D is always at logic ‘high’. At, time t = 0, when CLK is at logic ‘high’ the output goes to logic ‘low’ because RST is also at logic high. Then at, time t = t + 1, when, CLK and D are at logic ‘high’ and RST is at logic ‘low’ then the output is switched to logic ‘high’.

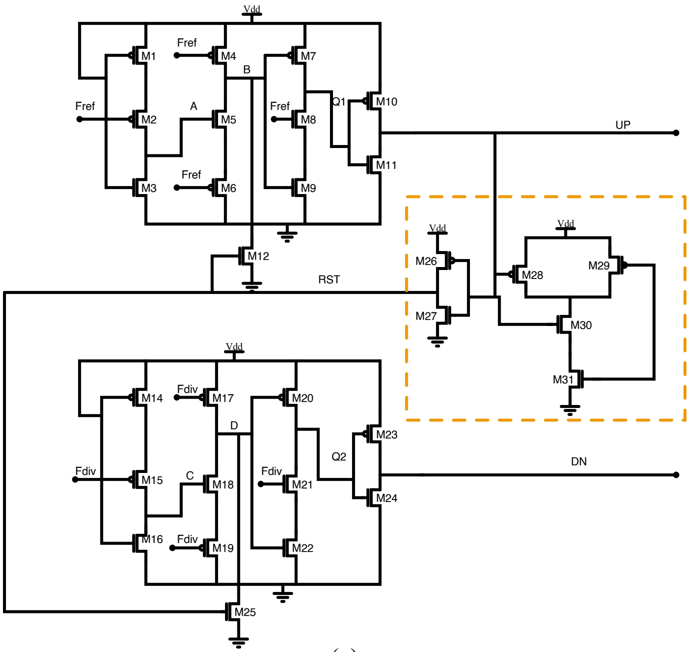
4 Conventional Phase Frequency Detector Architecture

Figure 6(a) shows the schematic of the conventional PFD which has been designed using the TSPC based DFF (see Fig. 5(a)), where AND gate is used for the reset operation. The input of the two Flip-Flops (FF) is F_{ref} and F_{div}. Input ‘D’ in both the FF is connected to logic high (Vdd). To understand the working of the conventional PFD, let us assume that initially both UP and DN are at their initial state i.e., ‘0’. When F_{ref} is high then M1 and M2 turn OFF. In this case, the output of the 1st stage is ‘0’ and that of the 2nd stage is high impedance meaning the output is not connected to either of the power rails. As a result, the charge at the 2nd stage does not change making the output of the 3rd stage as ‘0’. Thus, the final output will be logic ‘1’. However, if F_{ref} and F_{div} both are logic high then the corresponding output UP and DN will be at logic ‘1’. This generates the unwanted state at the output. At this time, AND gate switches ON and resets the output to its initial state. The NMOS transistor in the DFF is connected with the output of the AND gate. Whenever the output of AND is logic high transistor M12 and M25 turn ON and pull down the high state of

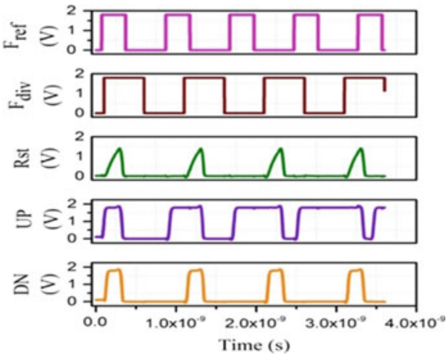
2[nd] stage to logic low. As a result, the unwanted situation at the output is removed. Thus, this special detector is allowed to detect both phase and frequency differences [4–6]. Moreover, the phase frequency detector has an advantage over phase detector as it provides the larger phase comparison range (–2π to +2π) and also has the capability to distinguish the lead or lag of the inputs signal [6]. However regardless of the type, the PFD/ PD generates the ripple in the control voltage of the VCO, as a result, it increases the phase noise of the PLL [5, 11, 16]. It has been understood that PFD contributes to the phase noise more significantly as compared to other PLL blocks [5]. Hence, Fig. 6(b) shows the transient response of the conventional PFD which shows when F_{ref} is leading F_{div} then corresponding UP is high until F_{div} turns DN high. When both the output is high then the AND gate becomes active and pulls down both the output to their initial condition. Thus because of the reset path, the PFD does not operate at high frequency as $f_{ckref} \leq \frac{1}{2.t_{res}}$ [6].

5 Proposed Phase Frequency Detector Architecture

From the above analysis, it has been found that 't_{res}' is the factor that degrades the PFD performance. Hence, to improve the maximum frequency of operation the simple architecture of PFD is proposed without the traditional reset path as shown in Fig. 7. The operation of the DFF is the same as that of the conventional, but here the AND reset is replaced by the immediate reset approach. This technique uses four transistors M12, M13, M24, M25 in the 2[nd] stage (B and D) to serve the purpose of immediate reset. The working of the proposed PFD is such that during normal operation when both F_{ref} and F_{div} is at logic 'high' the node A and C of 1[st] stage are logic 'low' then the 2[nd] stage (node B and D) gives the output as logic 'high'. Next, the 3[rd] stage output will be logic 'low' which switches the final output (UP and DN) to logic 'high'. Thus, during such conditions, the pulldown network M12, M13, M24, M25 turns ON making the output of the 2[nd] stage as logic 'low'. Hence node Q1 and Q2 become logic 'high' which reset the output to its initial condition. Therefore, due to the elimination of the traditional reset path and addition of the immediate reset as the pull-down network in the proposed architecture, has reduced the reset time as of which the maximum operating frequency is increased [6]. In addition, using the higher reference frequency (F_{ref}) helps to improve the loop bandwidth, which in turn improves the locking time of the PLL [12, 14]. Figure 8 shows the transient response showing the working of the PFD when F_{ref} is leading F_{div}. From the figure, it has been observed that the proposed PFD works efficiently at different conditions. Next, the set parameters for the proposed PFDs are shown in Table 1. Further, detail circuit operation is given in Table 2. Thus, the power consumption of the proposed design is minimized i.e., 416.83 μW @ 2 GHz. This makes the design an appropriate choice for low-power and high-speed PFD. Figure 9(a) shows the transient response when F_{div} is leading F_{ref}. Here, we can see that both UP and DN signals carry the phase error information in terms of the pulse width of the signal. However, phase noise is classified as in-band or out of band (see Fig. 1(a)) [5]. And PFD contributes



(a)



(b)

Fig. 6 Conventional phase frequency detector **a** Architecture **b** Timing diagram showing the working of PFD

a significant amount of in-band phase noise. Therefore, Fig. 9(b) shows the phase noise performance of the three PD/PFDs (EX-OR based PD, conventional PFD, and proposed PFD) at 2 GHz and it can be seen that the proposed PFD gives better phase noise performance by 3.4 dB from conventional PFD and 22.5 dB from EX-OR based PD. The performance summary of the proposed designs and other state of art works have been summarized in Table 3. From the table, we can see that the proposed PFD offers good operating frequency due to minimum reset time and power consumption

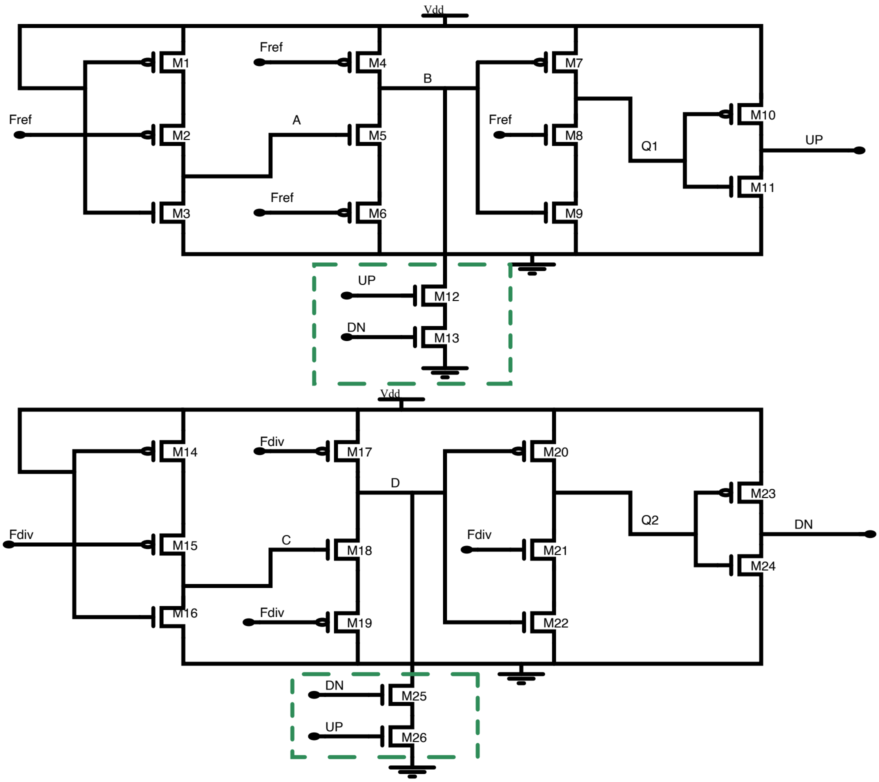


Fig. 7 Proposed PFD architecture

is highly minimized compared to ref. [6, 15]. Along with this, the operating frequency and phase noise are improved.

6 Conclusion

We have presented three designs, which have been analyzed at 180 nm CMOS technology node at 1.8 V power supply. The reset time reduces the maximum frequency of operation and increases the power consumption. Therefore, in the proposed architecture the traditional reset path is removed. In addition, the implemented immediate reset path has benefitted the design to improve the overall PLL performance. However, in comparison between the three designs, the proposed PFD shows better performance in terms of frequency, phase noise, and power consumption. Hence, the proposed PFD works efficiently in the range of 1.5 MHz–2.4 GHz offering a good range of bandwidth for Bluetooth, Wi-Fi, and Zigbee application.

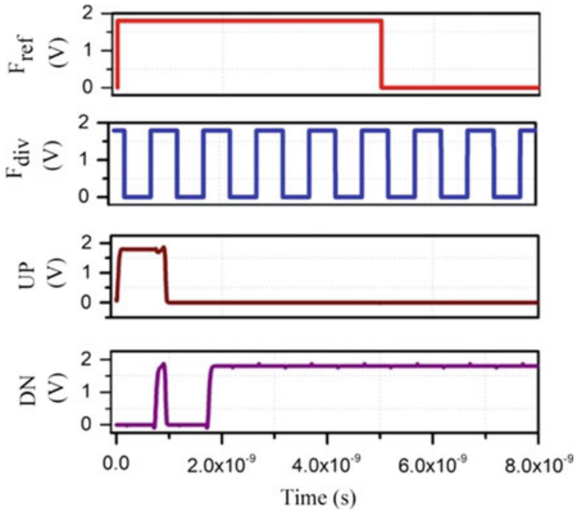


Fig. 8 Transient response when F_{ref} leading F_{div}

Table 1 Set parameters

SL. No.	Parameters	Value
1	Maximum operating Frequency (F_{max}) (GHz)	1.5– 2.4 MHz
2	T_{res} (ps)	113.4
3	Power dissipation (μW)	416.83
4	Technology node (μm)	180
5	Supply Voltage (V)	1.8

Table 2 UP and DN w.r.t F_{ref} and F_{div}

State (F_{ref})	State (F_{div})	F_{ref}	F_{div}	Stage-I		Stage-II		Stage-III		Stage-IV	
				A	C	B	D	Q1	Q2	UP	DN
Lag	Lag	L	L	0	0	1	1	1	1	0	0
Lag	Lead	L	H	0	0	1	1	1	0	0	0
Lead	Lag	H	L	0	0	1	1	0	1	1	0
Lead	Lead	H	H	0	0	1	1	0	0	1	1
						0	0	1	1	0	0

'H' = High, 'L' = Low

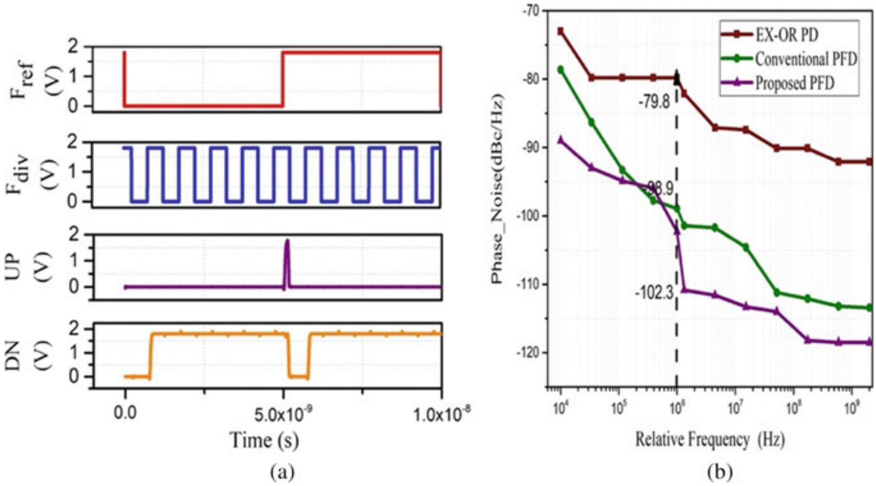


Fig. 9 **a** Transient response when F_{div} is leading F_{ref} **b** Phase noise @ 2 GHz

Table 3 Performance and comparison of PFD

Parameters	[6]*		[10]O	[15]Δ	This Work		
	PASS Transistor PFD	Latch Based PFD			EX-OR Phase Detector	Conventional PFD	Proposed PFD
Maximum Frequency (GHz)	1.28	1.53	0.8	1	0.8	1.4	1.5–2.4 MHz
Process Technology (nm)	250	250	500	90	180	180	180
Supply Voltage (V)	–	–	3	1.8	1.8	1.8	1.8
Phase Noise (dBc/Hz)	–	–	–	−109.5 @ 1 MHz	−78.9 @ 1 MHz	−98.9 @1 MHz	−102.2 @1 MHz
Power Consumption (uW)	620	1400	–	1730	381.09 @ 0.8	464.95 @1.4	416.83 @ 2 GHz

*Measured *Post Layout °Simulation

Acknowledgements The authors would like to thank, Ministry of Electronics and Information Technology (MeitY) Govt. of India for providing financial support under the SMDP-C2SD Project.

References

1. Tang, Y., Geiger, R.L.: Phase detector for PLL-based high-speed data recovery. Electron. Lett. **38**(23), 1 (2002)
2. Harikrushna, D., Tiwari, M., Singh, J.K., Khare, A.: Design Implementation and characterization of XOR phase detector for DPLL in 45nm CMOS technology. ACIJ **6**, 45 (2011)
3. Kumar, N., Kumar, M.: Design of low power and high-speed phase detector. IEEE Conf. (2016)
4. Johansson, H.O.: A simple precharged CMOS phase frequency detector. IEEE J. Solid-State Circuits **3**, 295 (1998)
5. Brennan, P.V., Thompson, I.: Phase/frequency detector phase noise contribution in PLL frequency synthesizer. Electron. Lett. **37**, 15 (2001)
6. Mansuri, M., Liu, D., Yang, C.K.K.: Fast frequency acquisition phase-frequency detectors for GSamples/s phase-locked loops. IEEE J. Solid-State Circuits **37**(10), 1331 (2002)
7. Gholami, M.: Phase detector with minimal blind zone and reset time for GSamples/s DLLs. Circuits Syst. Signal Process. (2016). https://doi.org/10.1007/s00034-016-0485-2
8. Johnson, T., Fard, A., Aberg, D.: An improved low voltage phase-frequency detector with extended frequency capability. In: 47th IEEE International Midwest Symposium on Circuits and Systems Conference (2004)
9. Charles, C.T., Allstot, D.J.: A calibrated phase/frequency detector for reference spur reduction in charge-pump PLLs. IEEE Trans. Circuits Syst. II **53**(9), 822–826 (2006)
10. Hu, W., Chunglen, L., Wang, X.: Fast frequency acquisition phase-frequency detector with zero blind zone in PLL. Electron. Lett. **43**(19), 1018 (2007)
11. Majeed, A.K.K., Kailath, B.J.: Low power, high frequency, free dead zone PFD for a PLL design. IEEE Conf. (2013)
12. Homayoun, A., Razavi, B.: Analysis of phase noise in phase/frequency detectors. IEEE Trans. Circuits Syst. I **60**, 529 (2013)
13. Shaikh, J., Rahaman, H.: High Speed and Low power preset-able modified TSPC D Flip-flop design and performance comparison with TSPC D Flip-flop. IEEE Conf.
14. Razavi, B.: Design of Analog COMS Integrated Circuits. Mc Graw Hill, New York (2001)
15. Nanda, U., Acharya, D.P., Patra, S.K.: Design of an efficient phase frequency detector to reduce blind zone in a PLL. Microsyst. Technol. **23**, 533–539 (2017)
16. Pradhan, N., Jana, S.K.: Design of phase frequency detector with improved output characteristics operating in the range of 1.25 MHz–3.8 GHz. Analog Integr. Circuits Signal Process. **107**(1), 101–108 (2021)

Design of Low Power Two-Stage OTA Based Second-Order G_m-C Filter for Fast Locking PLL Based Frequency Synthesizer

Priti Gupta, Sanjay Kumar Jana, and Mahesh Chandra Govil

Abstract With the development of high-speed low power integrated circuits, fast locking PLL based frequency synthesizer has huge demand. This work deals with the design of the low power and high-performance two-stage operational transconductance amplifier (OTA). A Gate-driven quasi floating gate (GD-QFG) technique has been incorporated to minimize the power consumption of the OTA. The proposed OTA has been simulated with the help of a 180 nm CMOS process at 1 V supply. The dc gain, unity-gain bandwidth (UGB), and power consumption have achieved 82.26 dB and 38.37 MHz and 189.55 µW, respectively. Further, the proposed low power OTA has been used for the designing of a high frequency second-order G_m-C filter. The cut-off frequency and third-order intermodulation (IIP3) of the G_m-C filter is 9.15 MHz, and –3.17 dm respectively. The power consumption is 340 µW at the 1 V voltage supply. The proposed G_m-C filter can be utilized for fast-locking PLL based frequency synthesizer.

Keywords CMOS · OTA · GD-QFG MOS · Secord-order G_m-C filter

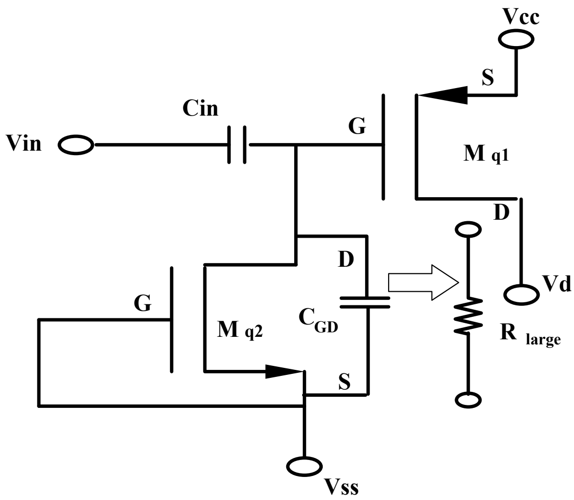
1 Introduction

Advancement in the VLSI industry has increased the requirement of low power and better performance analog circuits [1–9]. Many power-efficient circuit design configurations such as dynamic threshold (DT) MOS [4], BD (bulk-driven) MOS [5–7], floating gate (FD) MOS [8], quasi-floating bulk (QFB) MOS [8], and quasi-floating gate (QFG) [6] are used for less power and high-speed applications. A high-speed frequency synthesizer requires the fast locking PLL for clock generation and clock distribution circuits. There are many fast-speed PLLs are available [9–11]. However, a further reduction in the locking time is possible for high-speed PLL applications. G_m-C filter can be used for the high-frequency PLL [11]. Operational transconductance

P. Gupta (✉) · S. K. Jana · M. C. Govil
Electronics and Communication Engineering Department, Computer Science and Engineering Department, National Institute of Technology Sikkim, Ravangla, India
e-mail: pritiguptakit23@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_31

Fig. 1 Gate-driven quasi floating gate MOSFET



amplifier (OTA) is one of the widely used basic elements for analog and mixed-signal design such as G_m-C filters (OTA-C), analog-to-digital (A/D) converters, oscillators, and variable gain amplifiers, etc. [1–3]. Low power and high-performance OTAs are needed for designing the high-frequency G_m-C filters. Many structures of the OTA are available that are used for different applications [10–15]. It is required to design the OTA with improved performances for low power G_m-C filter.

1.1 Gate-Driven Quasi Floating Gate (GD-QFG)

GD-QFG MOS is one of the low power design techniques where the gate is made in quasi-floating state with the help of capacitance (Cin) at the gate terminal as depicted in Fig. 1. The Transistor Mq2 develops a higher resistance with operating in the cut-off region.

2 Proposed GD-QFG Two-Stage OTA Design and Analysis

The proposed GD-QFG based two-stage OTA is depicted in Fig. 2, where M1 and M2 are common source differential stages. Identical transistors $M1_Q$ and $M2_Q$ are GD-QFG MOS's where C1 and C2 capacitances have made gate in the quasi floating state that allows for the low voltage operation. Transistors M_P and M_N provide high resistance by working in the cut-off region. With the help of large resistance, less voltage is required at the gate terminal. Transistors M3, M4, and M6, M7 form the current mirrors. The second stage is formed with the help of M5 and M6 transistors.

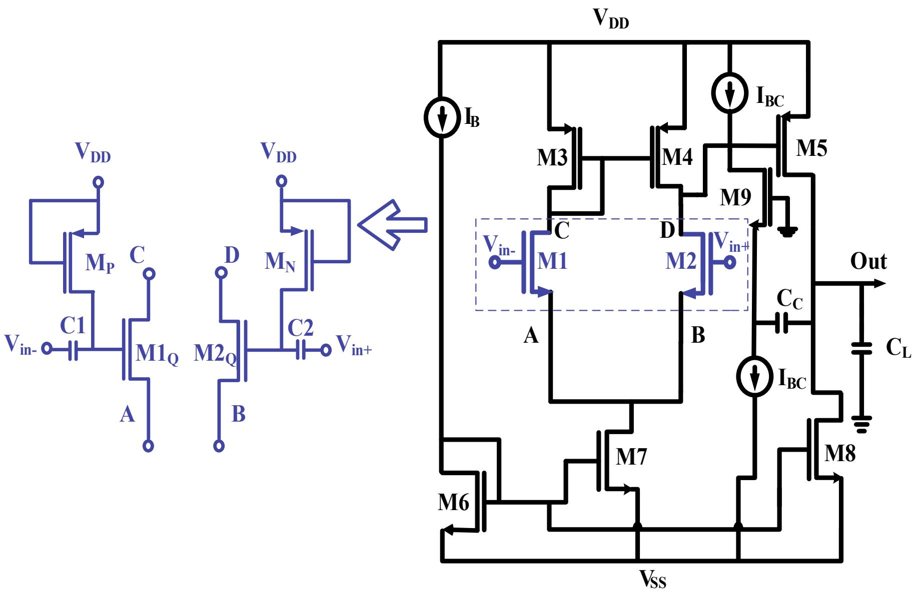


Fig. 2 Proposed GD-QFG OTA

The current buffer compensation has formed with the help of transistor M9, current source I_{BC} and coupling capacitance CC. CL represents the load capacitance.

The proposed OTA is simulated with the help of 180 nm CMOS technology using cadence virtuoso IC616. The magnitude and phase analysis of GD-QFG OTA is depicted in Fig. 3a and b that achieved the dc gain and unity gain bandwidth (UGB) of 82.26 dB and 38.37 MHz, respectively. The power requirement for the circuit is approximately 189.55 at 1 V supply voltage that shows 34.245% reduction in the power consumption than conventional OTA structure as depicted in Table 1. The power consumption is also less than the reported results in [13, 15].

3 Design of Second-Order G_m-C Filter and Its Simulation Results

The structure of the G_m-C-based second-order filter is depicted in Fig. 4. Here, two GD-QFG OTAs with the same transconductance have been considered with transconductance G_{m1} and G_{m2}. The two capacitances are named as C1 and C2 with different values. The GD-QFG OTA allows the filter to consume less power. The transfer function is given per Eq. 1.

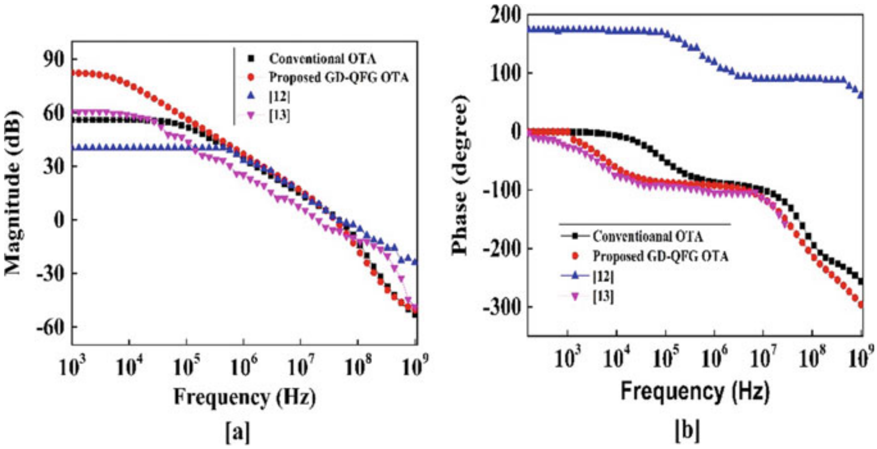
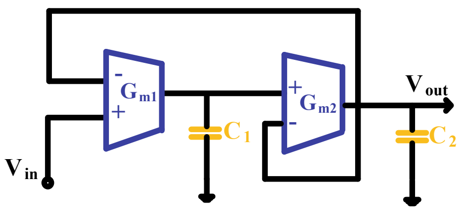


Fig. 3 Frequency response of GD-QFG based OTA **a** magnitude **b** phase

Table 1 Performance comparison of GD-QFG two-stage OTA with GD-OTA and existing work

Parameter	Conventional GD-OTA	GD-QFG OTA	[13]	[14]	[15]
Technology (μm)	180	180	0.5	0.18	0.18
Supply voltage	1.8	1	±1.25	0.7	1.8
DC gain (dB)	57	82.26	63.4	57.5	72
UGB (MHz)	63	38.37	—	—	—
GBW (MHz)	34.43	36.27	4.9	3	86.5
PM (degree)	47.07	15.13	83	60	50
CMRR (dB)	41	95.36	80	19	—
Power Consumption (μW)	287.3	189.55	437.5	25.4	11,900
Load capacitance (pF)	4	4	25	20	200

Fig. 4 Proposed second order G_m-C filter using GD-QFG MOS OTA



$$\frac{V_{out}}{V_{in}} = \frac{\frac{g_{m1}g_{m2}}{C_1C_2}}{s^2 + \frac{g_{m2}}{C_2}s + \frac{g_{m1}g_{m2}}{C_1C_2}} \tag{1}$$

The filter cut-off frequency is depicted in Eq. 2

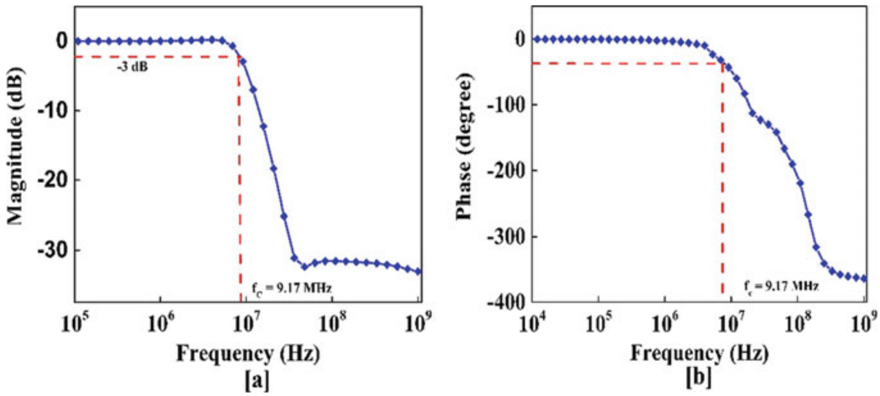


Fig. 5 G_m-C filter frequency response **a** magnitude **b** phase

$$w_n = \sqrt{\frac{G_{m1}G_{m2}}{C_1C_2}} \tag{2}$$

The quality factor Q is represented as in Eq. 3

$$Q = \sqrt{\frac{G_{m1}C_2}{G_{m2}C_1}} \tag{3}$$

Considering the transconductance $G_{m1} = G_{m2} = G_m = 420\ \mu℧$, and $Q = 0.7$, the circuit has been simulated with the help of the 180 nm CMOS technology. The frequency response is depicted in Fig. 5 that has achieved a cut-off frequency of 9.17 MHz.

The total output noise at the output of the filter is given as 126.55 pV/sqrt (Hz) as is depicted in Fig. 6a. The third-order intermodulation (IIP3) has measured by applying the two signal of frequencies 9 and 9.5 MHz that has archived the IIP3 of –3.17 dB as per Fig. 6b. The performance of the GD-QFG OTA type G_m-C filter has been compared with the reported results in Table 2. The power consumption of 340 μW has been achieved that shows the proposed filter has consumed less power than the reported results in [16–18].

4 Conclusion

This work presents the high gain and low power of the two-stage OTA. The low voltage operation has been achieved with the help of the GD-QFG OTA. The proposed OTA has achieved a 34.245% reduction in power consumption than the conventional OTA at a voltage supply of 1.8 V. Further, a second-order low power filter has

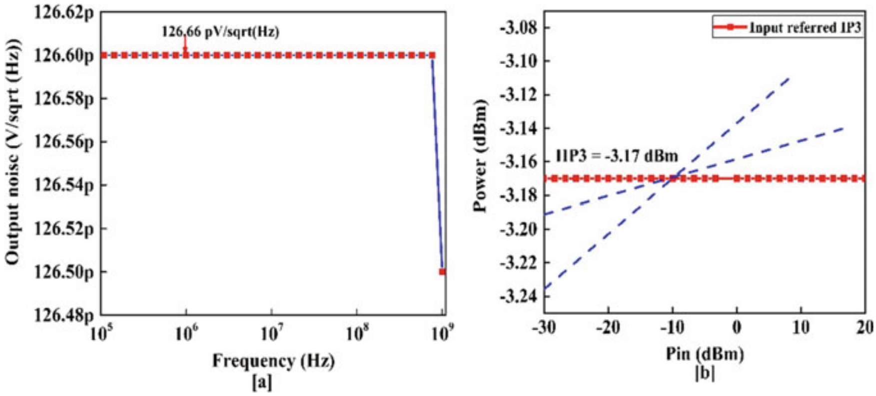


Fig. 6 **a** Output noise **b** Third-order intermodulation (IIP3)

Table 2 Performance comparison of second-order G_m-C filter

Parameter	GD-QFG OTA type G_m-C filter	[16]	[17]	[18]
Technology (nm)	180	180	250	90
Supply voltage (V)	1	1.8	±1.25	1
Cut-off frequency (MHz)	9.17	0.10–0.300	1.59	1000
Filter type	G_m-C	G_m-C	G_m-C	G_m-C
Power consumption (μW)	340	3200	390	2500
IIP3 (dB)	−3.17	−52	—	−42

been designed with the help of the proposed GD-QFG two-stage OTA. The cut-off frequency of the filter has achieved 9.17 MHz with a reduced power consumption of 340 μW. The proposed filter has been compared with the reported work and found better performance in terms of power consumption. The proposed high-frequency G_m-C filter can be used as an active loop filter for the fast-locking PLL.

Acknowledgements Authors would like to thank, Ministry of Electronics and Information Technology (MeitY), Govt. of India for providing financial support under SMDP-C2SD Project.

References

1. Kumar, T.B., Kar, S.K., Boolchandani, D.: A wide linear range CMOS OTA and its application in continuous-time filters. Analog Integr. Circuits Signal Process. **103**, 283–290 (2020)
2. Gao, C., Liu, H., Xie, N.: A low power double-sampling extended counting ADC with Class-AB OTA for sensor arrays. IEEE Trans. Circuits Syst. **62**, 29–38 (2015)

3. Elamien, M.B., Mahmoud, S.A.: On the design of highly linear CMOS digitally programmable operational transconductance amplifiers for low and high-frequency applications. Analog Integr. Circuits Signal Process. **97**, 225–241 (2018). 103
4. Gupta, P., Jana, S.K.: Design of dynamic threshold OTA-based transconductance-capacitance loop filter for PLL applications. Netw. Syst. **147**, 476–483 (2021)
5. Gupta, P., Tripathi, S. L.: Low power design of bulk driven operational transconductance amplifier. In: Devices for Integrated Circuit (DevIC), pp. 241–246 (2017)
6. Khateb, F.: Bulk-driven floating-gate and bulk-driven quasi-floating-gate techniques for low-voltage low-power analog circuits design. AEU Int. J. Electron. Commun. **68**(1), 64–72 (2014)
7. Zuo, L., Islam, S.K.: Low-voltage bulk-driven operational amplifier with improved transconductance. IEEE Trans. Circuits Syst. I **60**(8), 2084–2091 (2013)
8. Raj, N., Singh, A.K., Gupta, A.K.: High Performance current mirrors using Quasi floating bulk. Microelectron. J. **52**, 11–22 (2016)
9. Abolhasani, A., Mousazadeh, M., Khoei, A.: Fast-locking PLL based on a novel PFD-CP structure and reconfigurable loop filter. IET Circuits Devices Syst. **14**(8), 1235–1242 (2020)
10. Chiu, W.H., Huang, Y.H., Lin, T.H.: A dynamic phase error compensation technique for fast-locking phase-locked loops. IEEE Trans. Circuits Syst. **45**(6), 1137–2114 (2010)
11. Cao, Z., Li, Y., Yan, S.: A 0.4 ps-RMS-jitter 1–3 GHz ring-oscillator PLL using phase noise preamplification. IEEE J. Solid-State Circuits **43**(9), 2079–2089 (2008)
12. Aghaee, T., Biabanifard, S., Golmakani, A.: Gain boosting of recycling folded cascode OTA using positive feedback and introducing new input path. Analog Integr. Circuits Signal Process. **90**, 237–2461 (2017)
13. Surkanti, P.R., Furth, P.M.: Converting a three-stage pseudoclass-AB amplifier to a true-class-AB amplifier. IEEE Trans. Circuits Syst. II **59**(4), 229–233 (2012)
14. Cabrera-Bernal, E., Pennisi, S., Grasso, A.D., Torralba, A., Carvajal, R.G.: 0.7-V three-stage class-AB CMOS operational transconductance amplifier. IEEE Trans. Circuits Syst. I **63**(11), 1807–1815 (2016)
15. Sutula, S., Dei, M., Teres, L., Serra-Graells, F.: Variable-mirror amplifier: a new family of process/. independent class-AB single-stage OTAs for low-power SC circuits. IEEE Trans. Circuits Syst. I **63**(8), 1101–1110 (2016)
16. Lee, C.N.: Multiple-mode OTA-C universal biquad filters. Circuits Syst. Signal Process. **29**, 263–274 (2010)
17. Kar, S.K., Sen, S.: Linearity improvement of source degenerated transconductance amplifiers. Analog Integr. Circuits Signal Process. **74**, 339–407 (2013)
18. Abdolmaleki, M., Dousti, M., Tavakoli, M.: Design and simulation of tunable low-pass G_m-C filter with 1 GHz cut-off frequency based on CMOS inventers for high speed telecommunication applications. Analog Integr. Circuits Signal Process. **100**, 279–286 (2019)

A Temperature Dependent Modified TEAM Model

Manmath Suryavanshi, P. Karuppanan, Abhay Kumar Gautam, and Sreeteja Reddy Kotha

Abstract Memristor is a novel parasitic device which is used for several applications such as memories, logic and neuromorphic systems. One of the most attractive properties of the memristive device is the nonvolatility combined with their scalability and compatibility with CMOS. Bipolar resistive switches based on metal oxides such as TiO_2 have been identified as memristive devices. These are based on the pinched hysteresis loop observed in the current–voltage (IV) characteristics. Several models of memristor have been reported in the literature. Most notable are the Linear Ion Drift (LID), Simmon Tunnel Barrier (STB) and TEAM Models. TEAM Model has been identified to be most suitable for large scale simulations. However, most models completely neglect the effect of temperature variation. In this paper we present possible modification and propose a modified TEAM (MT) model which can account for temperature changes, offering a higher accuracy, control and simulation options while preserving the simplicity of the TEAM Model. The Modified Model is capable of accommodating temperature variation effects and shows an excellent fit to the STB fitted TEAM model, reporting temperature dependent fitting resistance deviation of 5.45% compared with the fabricated Memristor in STB Model. Moreover, the modification opens potential for development of large scale length-dependent libraries for simulation of memristors.

Keywords Memristor · Modelling · TEAM · Temperature effects · Modified TEAM · Simulation

1 Introduction

Many memristor models have been proposed and compared in literature [1]. Memristors were first proposed by Leon Chua but it remained a concept until Hewlett Packard

M. Suryavanshi (✉) · P. Karuppanan · A. K. Gautam · S. R. Kotha
ECED, Motilal Nehru Institute of Technology Allahabad, Prayagraj, India
e-mail: suryavanshimanmath@gmail.com

P. Karuppanan
e-mail: pkaru@mnnit.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_32

Labs were able to fabricate the first memristor in 2008 [2]. A mathematical model was proposed which used a state variable to determine the current exhibited resistance of the memristor [3]. This became the basic structure to define the memristor mathematically and more models have been developed based on the idea of using a state variable to define the state of memristor [4, 5]. These models have also been compared, in which the TEAM Model, developed by Kvatinsky et al. [5], proves to be a suitable choice for medium to large scale simulations. It is therefore imperative to study the drawbacks of this model and identify the opportunities to enhance this model. One such limitation is absence of temperature effects. Memristor operations are heavily affected by temperature, however the effects of temperature are not discussed in popular and accurate models of the memristor. The semiconductor metal oxide shows several effects related to drift, diffusion and thermophoresis including how the diffusion of ions and vacancies is impacted with change in temperature [6, 7]. 1-D models of memristor have been developed which take into account the temperature effects along with other important aspects of the semiconductor that may influence its behavior. This includes the fuse anti-fuse mechanisms [8, 9], the electrochemical metallization memory [10–13] and the valance change memory [14–19]. These models often require considerable simulation time to deliver results. Resistance switching is perhaps one of the most important properties of memristors and raises interest due to its potential applications [20]. These applications have inspired development of new logic families and potential logical circuits [21, 22]. However, simulations and development of bigger and more complex memristor circuits warrants the need for simple yet effective models that can be used for large scale simulations. To the best of our knowledge, only handful of such models exist and they are limited in capability or are simply too complex to be practical for large scale computations. Several models do not consider effect of temperature – an important parameter for any semiconductor designing process. Several parameters of a memristor, such as the ON resistance, OFF resistance, Conductivity, Charge mobility and diffusion coefficient rely heavily on the operating temperature [23] but these are not accounted for in the present low computational time models. In this work, we identify the drawback of one of such models and suggest a low resource-intensive modified model.

2 Memristor Models

Many memristor models have been proposed by researchers. Most notable models for this work are discussed in this section.

A. **Linear Ion Drift (LID) Model**

The Linear Ion Drift Model (LID) [2] uses highly approximated model dynamics to simulate a memristor. Linear Ion motion is assumed. This model uses the linear relation:-

$$\frac{dw}{dt} = \mu_v.\frac{R_{on}}{D}.i(t) \tag{1}$$

B. **Simmon Tunnel barrier (STB) Model**

A much more accurate model is the Simmons Tunnel Barrier Model (STB) [3], which offers high accuracy for the price of complexity and simulation speed. The STB Model uses the following equations:-

$$\frac{dx(t)}{dt} = c_{off}\sinh\left(\frac{i}{i_{off}}\right)\exp\left[-exp\left(\frac{x - a_{off}}{w_c} - \frac{|i|}{b}\right) - \frac{x}{w_c}\right], \quad i > 0 \tag{2a}$$

$$\frac{dx(t)}{dt} = c_{on}\sinh\left(\frac{i}{i_{on}}\right)\exp\left[-exp\left(\frac{x - a_{on}}{w_c} - \frac{|i|}{b}\right) - \frac{x}{w_c}\right], \quad i < 0 \tag{2b}$$

where c_{off}, c_{on}, i_{off}, i_{on}, a_{off}, a_{on}, w_c, and b are appropriate parameters.

C. **The ThrEshold Adaptive Model (TEAM)**

The TEAM model is one of the most versatile models which offers various advantages, such as flexibility and ease of use. This model handles the simulation and the corresponding basic characteristics of memristor very well. TEAM model is governed by following equations: -

$$\frac{dx(t)}{dt} = k_{off}.\left(\frac{i(t)}{i_{off}} - 1\right)^{a_{off}}.f_{off}(x), \quad 0 < i_{off} < i \tag{3a}$$

$$\frac{dx(t)}{dt} = 0 \qquad i_{on} < i < i_{off} \tag{3b}$$

$$\frac{dx(t)}{dt} = k_{on}.\left(\frac{i(t)}{i_{on}} - 1\right)^{a_{on}}.f_{on}(x), \quad i < i_{on} < 0 \tag{3c}$$

The voltage current relationship used by the TEAM model is given as:

$$v(t) = R_{ON}\left(e^{\left(\frac{\lambda}{x_{off}-x_{on}}\right)(x-x_{on})}\right).i(t) \tag{3d}$$

where,

$$\frac{R_{OFF}}{R_{ON}} = e^{\lambda} \tag{3e}$$

where k_{on}, k_{off}, a_{on}, a_{off}, are constants and i_{off} and i_{on} are current thresholds. The functions $f_{off}(x)$ and $f_{on}(x)$ act as the window functions [4]. The doped TiO_2 region of the model offers low resistance, indicated by R_s in Fig. 1. The undoped region, whose width is represented by x (used as state variable), offers high resistance. This

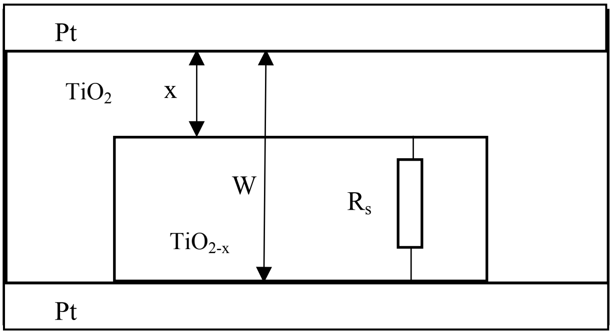


Fig. 1 The TiO_2 memristor model used in STB and TEAM model

width is able to change based on the applied voltage and current flowing through the memristor.

A Comparison of these models is presented in Table 1 and state variables are compared in Fig. 2. The linear ion drift model is the simplest model. However, it lacks in accuracy and assumes linear ion drift inside the memristor resulting in linear change in state variable (Fig. 2). The Simon Tunnel Barrier Model (STB) is by far the most accurate model. However, it uses complex equations which put stress on simulations and is not ideal choice for large scale simulations. The TEAM Model is an approximation of STB model which is shown by the slightly approximated curve of TEAM state variable in Fig. 2. It is preferred for simulations due to its simplicity. The Eqs. (3a–3e) are readily modelled in simulators and makes this model an ideal choice for large scale simulations.

As is evident from the model Eqs. (3a–e), the TEAM model does not take into account the effects of temperature. Therefore, a need for a simple model arises which can also accommodate the effects of temperature change. This is a major limitation which is addressed in the next sections.

Table 1 Comparison of memristor models

Model name	Linear ion	Simmons tunnel barrier [3]	TEAM [4]
Flexibility	NO	NO	YES
Asymmetry	NO	YES	YES
Accuracy	Very low	Very high	Sufficiently high
Window function	Required	Not Required	Required
Threshold exists	NO	YES	YES

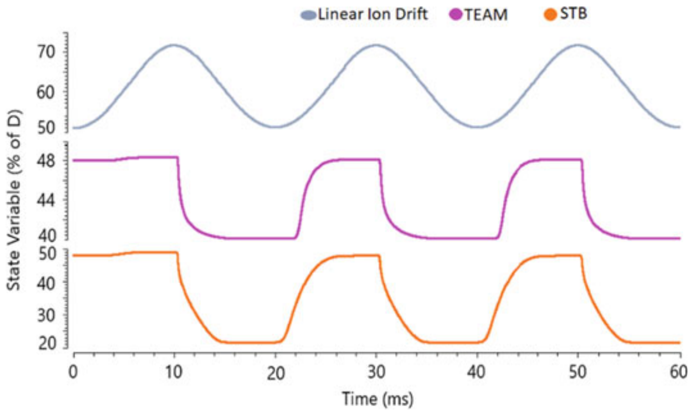


Fig. 2 Comparison of Linear Ion Drift (LID), Simmon Tunnel Barrier (STB) and TEAM model variation of state variable with applied sinusoidal voltage

3 Modified TEAM Model

In semiconductors, estimation of charge carrier concentration and their distribution is a vital step towards ascertaining their nature and properties. The non-linear Valance Charge Migration model [24] (VCM) can be utilized for ascertaining the ON resistance of the memristor depending on temperature.

$$
\begin{aligned}
Jv(x) = {} & 2q_o f a exp\left(-\frac{U_a}{K_B T}\right)\sinh\left(\frac{q_o E_v a}{2K_B T}\right)N_v(x) \\
& - q_o f a^2 \exp\left(-\frac{U_a}{K_B T}\right)\cosh\left(\frac{q_o E_v a}{2K_B T}\right)\left(\frac{\partial N_v}{\partial x}\right)
\end{aligned} \tag{4}
$$

$$
\frac{\Delta N_v}{\Delta t} = -\left(\frac{1}{q_o}\right)\nabla . J_v \tag{5}
$$

where, q_o is the charge of oxygen vacancies, f is escape attempt frequency, U_a is activation energy, a is hopping parameter and N_v is the initial TiO_2 doping profile in the memristor. The memristor is assumed to be 2-D rectangular block. Equations (4) and (5) are used on an initial doped TiO_2 concentration profile. These equations use the high electric field which arises inside the memristor. In addition, they are sensitive to doping concentration. However, to sense the resistance of the device, these equations are coupled with Poisson's Eq. (6) and Drift current (7) equation to calculate the current flowing through the device.

$$
\nabla^2 P(x) = -\frac{q}{\varepsilon_r \varepsilon_o}(N_v(x) - n(x)) \tag{6}
$$

$$
J_n(x) = -q.n(x).\mu_n.\nabla P(x) \tag{7}
$$

Further modifications of this procedure apply the use of Poisson-Boltzmann Eq. (8) to simulate electrostatic potential when the device is at equilibrium [25].

$$\nabla^2 P(x) = -\frac{q}{\varepsilon_r \varepsilon_o}\left(N_v(x) - nexp\left(\frac{qP(x)}{K_B T}\right)\right) \tag{8}$$

[25] describes steps to simulate the vacancy migration. The most time-consuming process of the algorithm is the process of simulating the migration of oxygen vacancies.

In this work, we bypass this step of the algorithm to obtain R_{on} by assuming an initial vacancy profile in the ON state (low resistance). Subsequent calculations of the zero-bias voltage and the current using a reference voltage contribute negligible time. The steps of this reduced algorithm are as follows:

1. Equation (8) is used at boundary conditions to calculate zero biasing potential and free electron distribution inside the memristor using the vacancy profile provided.
2. Equations (6) and (7) are used to obtain the current. A small voltage (testing voltage) is used here to test the resistance of the device.

The initial vacancy profile provided in step 1 is such that it leads to low resistance across the memristor. Furthermore, the R_{OFF}/R_{ON} resistance ratio (Ψ (T)) determined by the algorithm is presented in [25] and tabulated in Table 2 and is used to determine temperature dependent resistance ratio function Ψ(T). The best fitting curve from the limited number of points available in Table 2 is generated. The curve fitting procedure is bounded by the constraint that the function should be always decreasing.

Ψ(T) is given by Eq. (9) and plotted in Fig. 3.

$$\Psi_{10}(T) = 6263.4306 * 10^{30.345} * T^{-12.11} \tag{9}$$

where T is temperature in degree Kelvin and Ψ_{10} denotes 10 nm length. The steps of the mentioned algorithm can be used to obtain R_{ON}. Once R_{ON} is obtained, Eq. (9) is used to obtain corresponding R_{OFF} resistance for the operating temperature and length of the memristor. The obtained values of R_{ON} and R_{OFF} can be used to simulated the memristor via the TEAM Model.

Table 2 OFF/ON resistance ratios for different temperatures

Device temperature (K)	R_{OFF}/R_{ON} Ratio
298	15e3
338	4e3
358	1e3
398	500

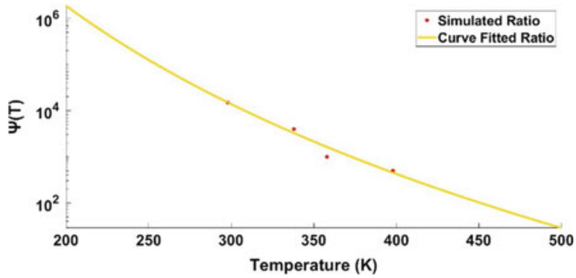


Fig. 3 Curve fitted ROFF/RON ratio function for 10 nm device. The red dots represent the simulated values given in Table 2 [25]

4 Simulation Results and Discussion

The simulation of the model is performed using MATLAB. The memristor is assumed to have length 10 nm. A 2.5 V amplitude sinusoidal voltage source with 50 Hz frequency is used to excite the memristor for 3 cycles. Simulation is performed for seven temperatures, starting from 200 to 500 K with 50 K intervals. Current and state variable change are recorded. The initial vacancy concentration in Fig. 4 is selected based on 10 nm vacancy profile in [25]. The memristor is considered to be in ON state in this vacancy distribution profile. The next sections describe the effect of temperature on different parameters of the memristor and simulated results.

A. **Effect on the ON resistance of Memristor**

The R_{ON} resistance values obtained for different temperatures is shown in Fig. 5(a). It can be observed that change in R_{ON} is not significant as the temperature reaches to 500 K. In fact, R_{ON} drops by only ~175 Ω. This result agrees with [26]. It is shown that significant change in temperature does not cause the

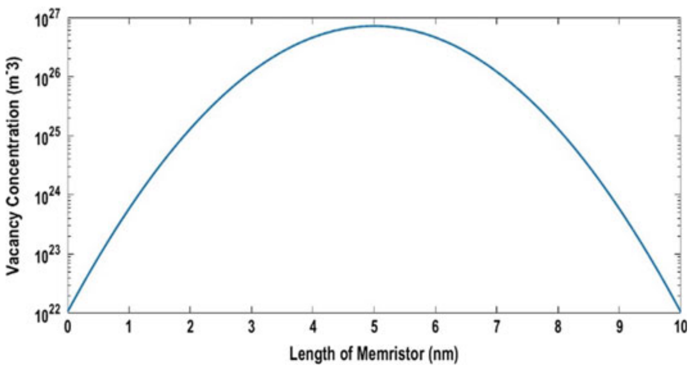


Fig. 4 Vacancy concentration profile for 10 nm memristor

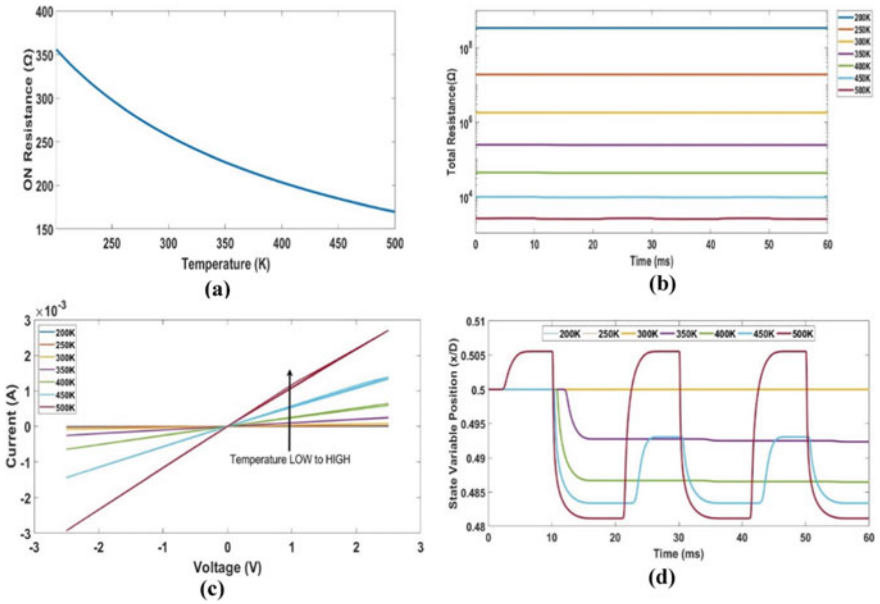


Fig. 5 Variation of different values with temperature. **a** ON resistance, **b** Total Resistance, **c** current–voltage characteristic, **d** State variable change

device ON resistance to change much. However, the OFF-resistance changes considerably. This is reinforced by Eq. (9).

B. **Effect on total resistance and current of Memristor**

The net R_{OFF} or R_{ON} resistance depends on the position of the state variable in the TEAM Model. This means that the net resistance of memristor keeps changing as the voltage applied across it changes. The obtained total resistances at different temperatures are presented in Fig. 5(b). It is clear that the net resistance decreases as the temperature is increased. This can be explained due to the fact that the vacancy hopping tendencies in the doped TiO_2 increase with increase in temperature. The oxygen vacancies act as dopants and assist in local electrical and thermal conductivity inside the memristor [26]. This is further reinforced by the jitter in the resistance plot towards the higher temperature (or lower resistance) ends. This indicates that the resistance switching is more pronounced at higher temperatures. Furthermore, the effect of resistance switching with temperature can be observed in I-V curves in Fig. 5(c). It can be seen that the current is negligible at lower temperatures due to the high R_{OFF} resistance at lower temperatures. As the temperature is increased, more current can flow through the device at same applied voltage due to significant lowering of the R_{OFF} resistance. For temperatures >400 K, slight bending in the current plot for OFF cycle (V > 0) can be observed. This effect is more pronounced at 500 K. The curve changes slope due to sudden change in resistance, resulting in change of the curve.

C. **Effect on switching speed of memristor**

As temperature increases, the switching speed of memristor also increases. This is due to the lowered R_{OFF} resistance of the memristor, allowing more current to pass through and thus, more shift in the vacancies, changing the overall resistance of the memristor quite fast when compared to lower temperatures. In Modified TEAM Model, this change of switching speed can be seen in Fig. 5(d). At lower temperatures, negligible change in state variable is observed. As the temperature is raised, some switching behavior is observed at 300–400 K. On further raising the temperature (>400 K) we observe large changes in state variable position. Also of interest is the increase in speed of change, which is easily observed between 450 and 500 K.

D. **Fitting Modified TEAM to Simmon fitted TEAM Model**

The modified TEAM Model demonstrates expected behavior when operating temperature of the memristor is changed. However, the effects have been demonstrated without concern of practicality and the parameters have not been adjust to reflect any known model. Thus, it is necessary to fit this model for practical applications. Kvatinsky et al. have achieved excellent fitting for TEAM Model [4]. Since the modified team is an extension of this work, it is natural to consider the Simmon fitted TEAM (SFT) Model as reference for the modified Temperature dependent Modified TEAM (MT) model. Although the STB Model has been developed by fabrication and analysis of 50 nm TiO_2 memristor device [3], the STB model does not account for any dependence on device length within the model itself. Similarly, the SFT Model does not have any effect with changing length of the memristor. In the model presented in this work, however, the memristor length plays a crucial role along with temperature and doping profile in determining the ON resistance of the device. Device length for this model is 10 nm. Fitting current for Modified TEAM and SFT Model is shown in Fig. 6(a).

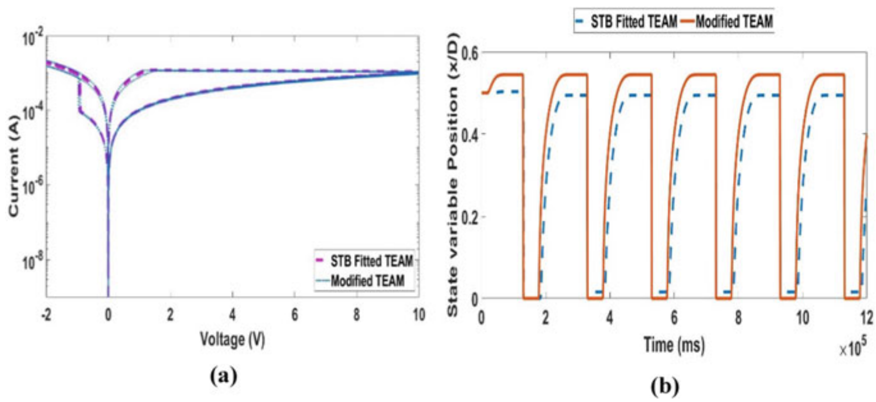


Fig. 6 Fitting results. **a** Current–Voltage characteristics, **b** State Variable Variation

Table 3 Comparison of TEAM and modified TEAM memristive models

Model	TEAM	Modified TEAM
Temp dependency	NO	YES
Vacancy concentration dependency	NO	YES
Length dependency	NO	YES
ON resistance	Fitted for best results, may not be accurate	Determined by physical parameters. More accurate
R_{ON} value (Ω)	1e3	226.7
Rs of STB Model (Ω)	215	215
R_{ON} deviation (%)	365.11	5.45

The parameters for this fit are kept same as that of the STB fitted TEAM Model. However, the modified model does not take inputs from user for R_{ON} and R_{OFF}. Thus, the fitting process consists of determining the right temperature for the Modified TEAM at which it shows the same current–voltage characteristics as the SFT Model. The fitting procedure concludes that the MT Model is best fitted at operating temperature of 350 K. The ON resistance of the 10 nm device is 226.7 Ω which is very close to the value obtained in STB Model (215 ± 6 Ω) [3]. The high value of fitting Temperature can be explained due to the fact that there is considerable difference in device length in STB model and the length used in this work. Device length can significantly impact the ON resistance as well as change the relation given by Eq. (9), requiring different set of data and curve fitting to obtain resistance ratio, due to which 50 nm length device has not been used in this work. Furthermore, ON resistance can be modified by changing the doping concentration profile of the memristor which is the result of fabrication process. The vacancy distribution profile used here is not according to the STB model for 50 nm device, hence the increased temperature. However, the fitting process of MT Model shows that a 10 nm memristor device with given vacancy distribution profile can be expected to operate same as the 50 nm device used to develop the STB Model at temperature 350 K as the device will exhibit R_{ON} very close to R_s determined in STB Model. Since TEAM model heavily relies on the ratio of OFF and ON resistance, this fitting procedure reinforces the accuracy of curve fitted Eq. (9). The differences between the SFT Model and Modified TEAM are summarized in Table 3 and state variable comparison is shown in Fig. 6(b).

5 Conclusion

In this work, the sensitivity list of the TEAM Model is expanded to include temperature, vacancy concentration as well as device length. The proposed TEAM Model performs well in different temperature ranges and displays all expected effects with variation in temperature. Effect on various parameters, such as effect on total resistance, current, switching speed were analyzed with variation in temperature. Additionally, the modified Model is fitted to STB Fitted TEAM Model. The fitting ON resistance calculated was found to be more accurate than the one used in TEAM Model. Moreover, this model eliminates the need for user provided ON and OFF resistance values and determines them from much more reliable methods based on physical parameters—vacancy concentration profile, length and temperature. This model is also length dependent and can be useful for development of standardized length based memristor libraries for large scale simulations much like MOSFET Libraries.

References

1. Suryavanshi, M., Karuppanan, P., Gautam, A.K., Kotha, S.R., Mishra, A.: "Implementation and parametric analysis of memristor models–comparative study. In: 2021 4th Biennial International Conference on Nascent Technologies in Engineering (ICNTE), pp. 1–6 (2021). https://doi.org/10.1109/ICNTE51185.2021.9487680
2. Chua, L.O.: Memristor—the missing circuit element. IEEE Trans. Circuit Theory **18**(5), 507–519 (1971)
3. Strukov, D.B., Snider, G.S., Stewart, D.R., Williams, R.S.: The missing memristor found. Nature **453**, 80–83 (2008)
4. Pickett, M.D., et al.: Switching dynamics in titanium Dioxide memristive devices. J. Appl. Phys. **106**, 074508 (2009)
5. Kvatinsky, S., Friedman, E.G., Kolodny, A., Weiser, U.C.: TEAM: threshold adaptive memristor model. IEEE Trans. Circuits Syst. I Regul. Pap. **60**(1), 211–221 (2013). https://doi.org/10.1109/TCSI.2012.2215714
6. Strukov, D., Borghetti, J., Williams, S.: Coupled ionic and electronic transport model of thin-film semiconductor memristive behavior. Small **5**, 1058–1063 (2009). https://doi.org/10.1002/smll.200801323
7. Sawa, A.: Resistive switching in transition metal oxides. Mater. Today **11**(6), 28–36 (2008). https://doi.org/10.1016/s1369-7021(08)70119-6
8. Strukov, D.B., Alibart, F., Stanley Williams, R.: Thermophoresis/diffusion as a plausible mechanism for unipolar resistive switching in metal–oxide–metal memristors. Appl. Phys. A **107**, 509–518 (2012). https://doi.org/10.1007/s00339-012-6902-x
9. Russo, U., Ielmini, D., Cagli, C., Lacaita, A.L.: Self-accelerated thermal dissolution model for reset programming in unipolar resistiveswitching memory (RRAM) devices. IEEE Trans. Electron. Devices **56**(2), 193–200 (2009)
10. Valov, I., Waser, R., Jameson, J.R., Kozicki, M.N.: Electrochemical metallization memories—fundamentals, applications, prospects. Nanotechnology **22**(25), 254003 (2011)
11. Liu, D., Cheng, H., Zhu, X., Wang, G., Wang, N.: Analog memristors based on thickening/thinning of Ag nanofilaments in amorphous manganite thin films. ACS Appl. Mater. Interfaces **5**(21), 11258–11264 (2013)
12. Liu, D., et al.: Programmable metallization cells based on amorphous La0.79Sr0.21MnO3 thin films for memory applications. J. Alloys Compd. **580**, 354–357 (2013)
13. Liu, D., et al.: Nonvolatile bipolar resistive switching in amorphous SR-doped LaMnO3 thin films deposited by radio frequency magnetron sputtering. Appl. Phys. Lett. **102**(13), 134105 (2013)
14. Choi, B.J., et al.: Resistive switching mechanism of TiO2 thin films grown by atomic-layer deposition. J. Appl. Phys. **98**(3), 033715 (2005)

15. Seo, S., et al.: Reproducible resistance switching in polycrystalline NiO films. Appl. Phys. Lett. **85**(23), 5655–5657 (2004)
16. Szot, K., Speier, W., Bihlmayer, G., Waser, R.: Switching the electrical resistance of individual dislocations in single-crystalline SrTiO3. Nat. Mater. **5**, 312–320 (2006)
17. Liu, S.Q., Wu, N.J., Ignatiev, A.: Electric-pulse-induced reversible resistance change effect in magnetoresistive films. Appl. Phys. Lett. **76**(19), 2749–2751 (2000)
18. Quintero, M., Levy, P., Leyva, A.G., Rozenberg, M.J.: Mechanism of electric-pulse-induced resistance switching in manganites. Phys. Rev. Lett. **98**, 116601 (2007)
19. Choi, B.J., et al.: Nitride memristors. Appl. Phys. A **109**(1), 1–4 (2012)
20. Hu, M., Li, H., Chen, Y., Wu, Q., Rose, G.S., Linderman, R.W.: Memristor crossbar-based neuromorphic computing system: a case study. IEEE Trans. Neural Netw. Learn. Syst. **25**(10), 1864–1878 (2014). https://doi.org/10.1109/TNNLS.2013.2296777
21. Kvatinsky, S., Satat, G., Wald, N., Friedman, E.G., Kolodny, A., Weiser, U.C.: Memristor-based material implication (IMPLY) logic: design principles and methodologies. IEEE Trans. Very Large Scale Integr. (VLSI) Syst. **22**(10), 2054–2066 (2014). https://doi.org/10.1109/TVLSI.2013.2282132
22. Strukov, D.S., et al.: Hybrid CMOS/memristor circuits. In: IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1967–1970 (2010). https://doi.org/10.1109/ISCAS.2010.5537020
23. Singh, J., Raj, B.: Temperature dependent analytical modeling and simulations of nanoscale memristor. Eng. Sci. Technol. Int. J. **21**(5), 862–868 (2018). https://doi.org/10.1016/j.jestch.2018.07.016
24. Hashem, N., Das, S.: Switching-time analysis of binary-oxide memristors via a nonlinear model. Appl. Phys. Lett. **100**(26), 262106 (2012)
25. Abunahla, H., Mohammad, B., Homouz, D., Okelly, C.J.: Modeling valance change memristor device: oxide thickness, material type, and temperature effects. IEEE Trans. Circuits Syst. I Regul. Pap. **63**(12), 2139–2148 (2016). https://doi.org/10.1109/tcsi.2016.2622225
26. Kim, S., et al.: Physical electro-thermal model of resistive switching in bi-layered resistance-change memory. Sci. Rep. **3**, Art. no. 1680 (2013)

Design and Performance Analysis of Full Adder Circuit Using Different Logic Styles

Mitul Kumar Choudhary, Ashish Raturi, and Poornima Mittal

Abstract A significant development has been seen in the field of Integrated circuits (IC) technology. With the fast-growing world, we want our ICs to be fast, should occupy less area by less transistor count, and should have low power consumption. One circuit that is seen in most of ICs is the Arithmetic and Logic Unit (ALU) circuit. Adder and Multipliers are considered to be an integral part of ALU. On the LTSpice simulator, this study presents an effective approach for designing complete adder circuits, which are the foundation of unit ALUs. First we designed 24 Transistor adder commonly called as Mirror adder and then calculated its propagation delay and power on LTSpice simulator and then build the adder using 14 Transistor logic, 6 Transistor Logic and then build the Memristor based logic adder having zero transistors in it and have done the comparative study and analysis on their propagation delay (including rise and fall time) and power consumption by the circuit.

Keywords ALU · Mirror adder · 14 transistor · 6 transistor · Memristor based adder

1 Introduction

In today's world, almost entire system are designed on advanced technology called as system of chip (SOC) in which all the components and circuits are built on a solitary chip which increases the intricacy of the system. VLSI plays a very critical role in the development of those areas. Since majority of the chip circuit include adders as a heart of ALU circuit, so minimizing the delay, power and area of the full adders are most important part of our concern [1–4].

M. K. Choudhary (✉) · A. Raturi · P. Mittal
Department of Electronics and Communication Engineering, Delhi Technological University, Delhi 110042, India
e-mail: mitulkumarchoudhary_2k18ec101@dtu.ac.in

P. Mittal
e-mail: poornimamittal@dtu.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_33

Adders are mostly used in arithmetic operations. It is often a speed-limiting element as well. Therefore careful optimization of full adder is very essential. The full adder is a digital circuit that processes addition of three 1-bit inputs A, B, and C_{in} (carry input) and outputs one-bit sum (S) and carry output (Cout). Two EX-OR gates, two AND gates, and one OR gate make up the logical circuit. The Boolean logic for the Sum of an adder is given as follows

$$Sum = A \wedge B \wedge C_{in} \tag{1}$$

The Boolean logic for the cout of an adder is given as follows (Fig. 1 and Table 1)

$$Cout = (A.B) + (Cin.(AB)). \tag{2}$$

The truth table for 1-bit full adder circuit is shown in the table above. In terms of CMOS logic and without CMOS logic, there are several approaches to construct

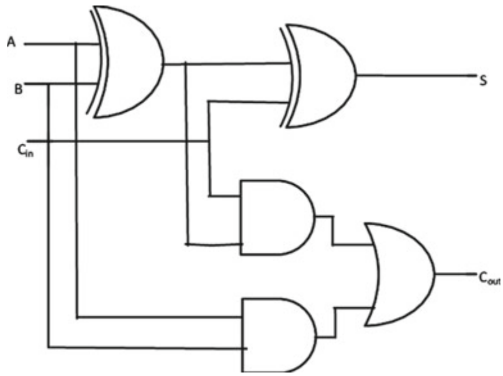


Fig. 1 Circuit diagram of full adder

Table 1 Truth table for full adder

A	B	C_{in}	Sum	C_out
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

a full adder circuit. With the fast advancement of technology, the rising demand for speed and power has compelled us to build optimized and low-energy full-adder circuits, which are an essential component of our ALU. The data route of a circuit, which consists of transistors, consumes the majority of the power in any circuit. As a result, by lowering the number of transistor, we may minimize power consumption while simultaneously shortening the data route, making the circuit faster [5–8].

2 Mirror Adder

Only 24 transistors are required for this full adder. The NMOS and PMOS chains are mirror image of each other, and produces the authenticated results because of self-dual nature of both the sum and carry output. As a result, carry generation circuitry can only have a maximum of two series transistors. PDN and PUN networks are not complementary rather they are the mirror reflection of each other [9]. The main advantage of mirror adder is that we just have to optimize the carry stage of our transistors for speed. All the transistors in the sum stage can be of minimum size. The major drawback of mirror adder is that it requires 28 transistors to build the full adder circuit which consumes lots of power and area by the transistor count and also has a very high propagation delay (Figs. 2 and 3).

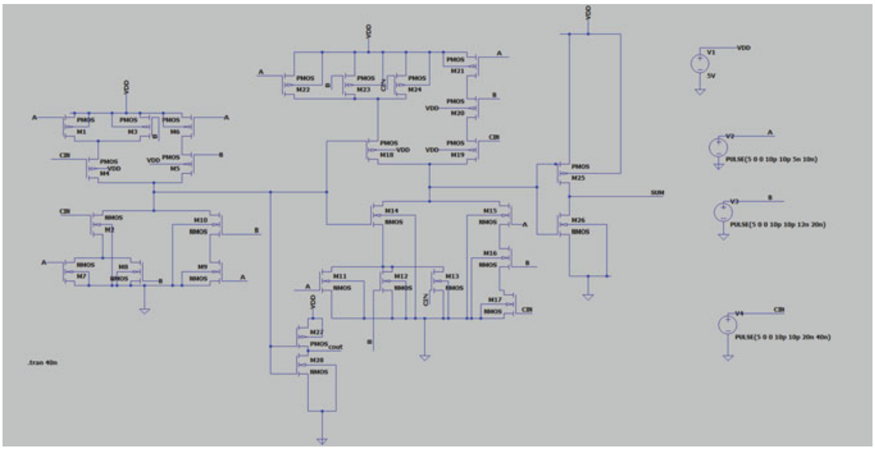


Fig. 2 Simulation diagram of mirror adder

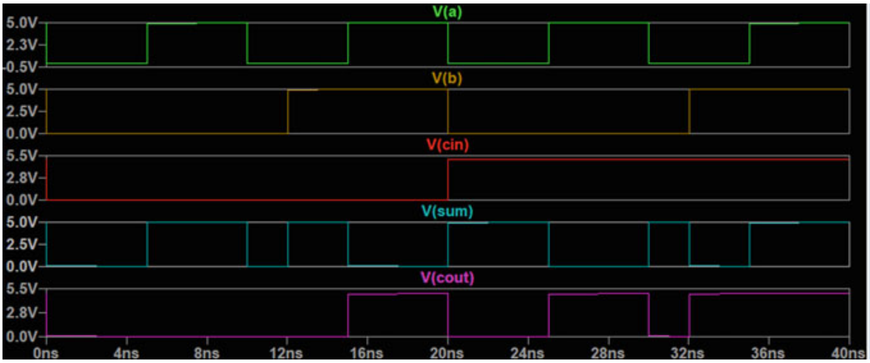


Fig. 3 Simulated output of mirror adder

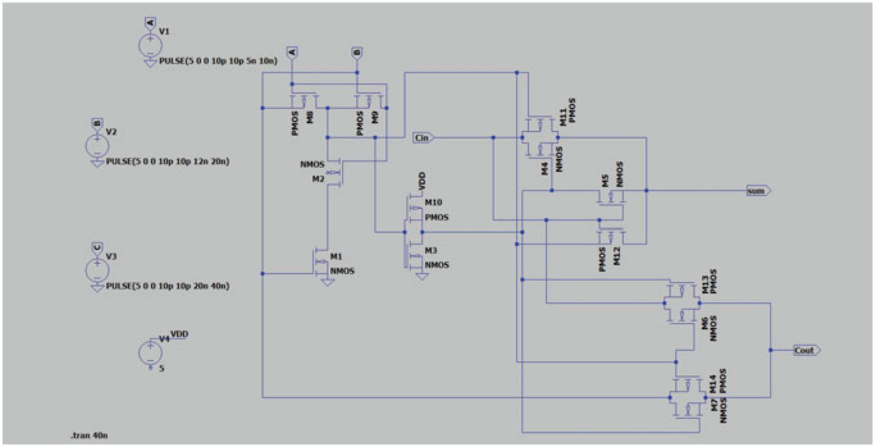


Fig. 4 Simulated circuit diagram of 14T Adder

3 14 Transistor Adder

The 14 Transistor adder is designed using 4 Transistor logic based XOR gate which not only reduces the total transistor count but also consumes less power and also possess the high speed. The 14 Transistor adder design has only one inverter in it but still it cannot overcome the occurrence of glitches at the output node. Another disadvantage is introduction of static power dissipation at the output node of inverter. When A = 0 and B = 0 the xor gate does not produces the full rail to rail voltage swing and because of this both N and P transistors will be weakly on and thus will lead to direct path from supply voltage to ground causing the static power dissipation. The consequences of this drawback is that a single unit of this full adder will burn lot of power but still it remains good in respect to other logic based full adder circuit

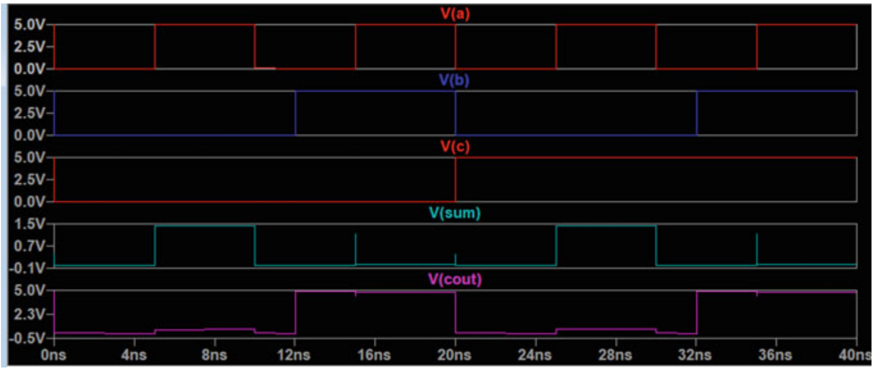


Fig. 5 Simulated output of 14T Adder

having higher number of transistors. So, using 4 transistor will help in reducing area by count and since less transistors are used to make the 14 Transistor full adder, node capacitance will also be less and dynamic power consumption and propagation delay will be less as compared to the Mirror adder circuit [10, 11] (Figs. 4 and 5).

4 Six Transistor Adder

The gate diffusion input method is used in six transistor adder. G (common gate input of tiny N and P type Mosfet), P (input to source/drain of P channel Mosfet), and N (input to the source/drain of N channel Mosfet) are the three input terminals of a GDI cell. Both PMOS and NMOS have their bulk terminals linked to N or P, respectively to have the zero threshold voltage.

Using this GDI cell, first, we build the XOR gate circuit, and then with help of that, we build six transistor adder. The area consumed in this technique is less as compared to other methods because it is composed of lesser number of transistors. And as we know since transistor count is less, so total node capacitance will be less

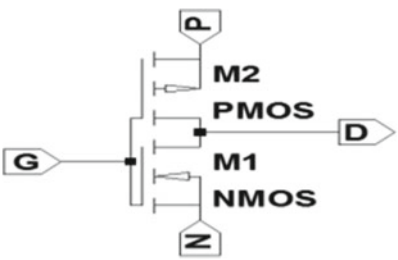


Fig. 6 GDI cell

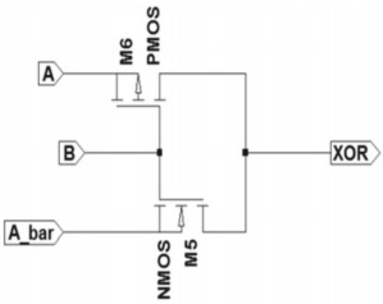


Fig. 7 XOR gate using GDI technique

and hence charging and discharging time will also be lesser. And as we all know that the dynamic charging/discharging power depends on load capacitance and in this case since total load capacitance will be lesser so dynamic power will also get reduced. Here in our six-transistor logic based adder circuit, the XOR gate is built using the GDI technique for producing the sum output, and the AND gate and OR gate in a full adder are replaced by the multiplexer to generate Cout [12, 13] (Fig. 7).

4.1 XOR Gate Using GDI Technique

The B input is linked to both the PMOS and NMOS gates, whereas the terminal A_bar is tied up to the NMOS source and the terminal A is tied up to the PMOS source. When B is zero, PMOS is on and NMOS is off, and the output is A; however, when B = 1, PMOS is off and NMOS is on, and A_bar is linked to the output. As a result, the GDI cell functions as an XOR gate circuit [14].

4.2 Six Transistor Logic Adder Using GDI Technique

In Fig. 8, N1, N2, N3, P1, P2, and P3 are the six transistors shown. N1 and P1 operate as a first XOR gate, generating A XOR B, and this A XOR B terminal serves as an input to the following GDI cell, where N2 and P2 acts as a second EX-OR gate, with input C_{in} linked to P2's source and Cin_bar to N2's source. The sum output is produced at the common drain terminal of transistor N2 and P2 tied together. This means that when $A\hat{}B = 1$ then NMOS N2 will be on and input Cin_bar will be connected to the output and when $A\hat{}B = 0$ then PMOS P2 will be on and input Cin will be passed to the output so based on this logic the 4 Transistors of the circuit will act as 3 input XOR gate producing the result $A\hat{}B\hat{}C_{in}$. In the six-transistor logic adder, the two AND gates and OR gate of the static logic-based complete adder are replaced by the multiplexer architecture [8, 15–17]. If we look at the truth table of

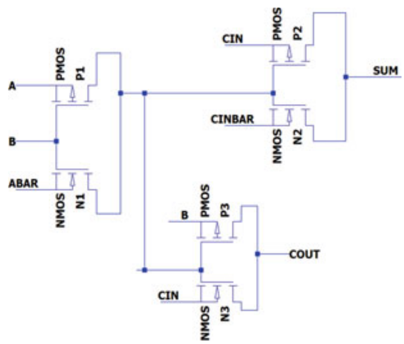


Fig. 8 Full adder circuit using GDI technique

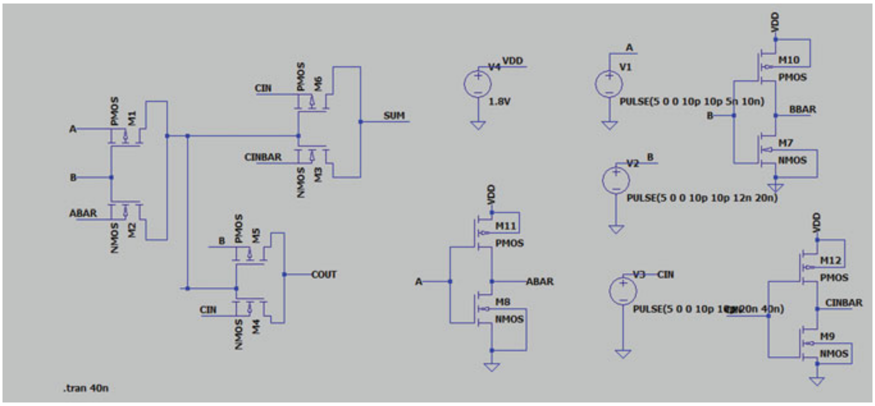


Fig. 9 Simulated circuit diagram of 6T adder

the full adder circuit carefully, we observe that When AˆB = 1 the output of the full adder Cout = Cin and when AˆB = 0 then the Cout of the circuit is equal to input B. So based on this analysis, the output of the first EX-OR gate is applied to the common gate terminals of N3 and P3, input B is linked to P3's source, input Cin is connected to N3's source, and this logic behaves as an multiplexer and generates Cout at the common drain terminal of transistor N3 and P3 (Figs. 9 and 10).

5 Memristor Ratioed Logic Adder

A memristor is a non-linear two-terminal passive electrical component linking the electrical charge and magnetic flux linkage. In 1971, for the first time, it came into existence as a fourth device after resistor, capacitor, and inductor. Daintly titanium dioxide (TiOi) is layered in middle of platinum (Pt) on both the top and lower of the

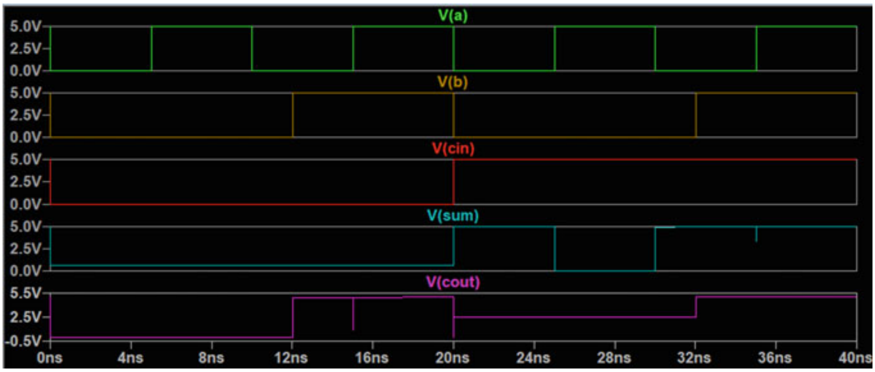


Fig. 10 Simulated output of 6T adder

electrode in the design of the memristor. Two layers of TiO2 thin film are present. The first layer is an electrically insulating layer of undoped titanium oxide (TiO2). The second layer is an electrically conductive doped titanium oxide (TiO2-x). R_{on} denotes the decreased resistance while R_{off} denotes the increased resistance [16]. On LTSpice Simulator, R_{on} value of the memristor is 100 Ω and R_{off} value is 10 Kohm. Since the main cause of delay in any logic circuit is the CMOS transistors, as they contain the capacitors so charging and discharging them causes the delay in the circuit. The memristor-based logic design contains no CMOS transistors so total propagation delay is fewest in the memristor which is its biggest advantage. Like a resistor, it creates and maintains a safe flow of electric current across a device, but it can also remember the last charge that was flow through it. This means that logic bit in memristor are stored as resistance states in memristors.

5.1 *Designing the AND and OR Gate Using Memristor*

When the positive polarities of two memristors are linked together and fed into the negative polarity, an AND gate is created. The higher resistance is caused by the positive input that feeds into the negative. Similarly, an OR gate is created by connecting the negative polarities of two memristors and feeding them into the plus polarity. The reduced resistance is due to the positive input that feeds into the plus polarity (Fig. 11).

For AND gate, when any of the logic is 1 i.e. either V1 = 1 and V2 = 0 or V1 = 0 and V2 = 1 in these cases current flow from higher potential voltage to lower potential voltage i.e. from supply voltage to Ground. When current flows through the from memristor U1, the resistance of memristor U1 increases to value R_{off} and resistance of memristor U2 decreases to value R_{on} and current moves through the GND. So the resistances R_{on} and R_{off} will form the voltage divider circuit at the output terminal. So the equation for the Output Out is given as follows

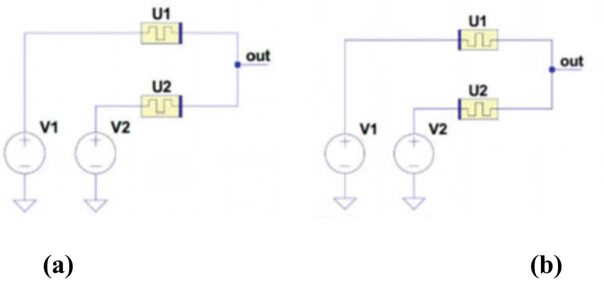


Fig. 11 AND gate using memristor logic **b** OR gate using memristor logic

$$Out = \frac{Ron}{Ron + Roff} \times Vdd \approx GND \tag{3}$$

When both the inputs applied are identical i.e. (V1 = 0, V2 = 0 or V1 = 1, V2 = 1) its output of the circuit becomes equal to that of input. Similarly operation of an OR gate can be understood with help of AND gate logic in memristor.

5.2 Design of Full ADDER Circuit Using Memristor Logic

Here in this Memristor-based Ratioed logic circuit, we have used Sum of Maxterms to produce the sum and carry output. For this purpose, we will use 25 memristors as shown in the above Fig. 6. As we all know that sum is given by (Fig. 12)

$$Sum = (A\_bar + B\_bar + Cin).(A\_bar + B + Cin\_bar).(A + B\_bar + Cin\_bar).(A + B + Cin) \tag{4}$$

Carry Output is given by

$$Cout = (A + B).(B + Cin).(Cin + A) \tag{5}$$

When the input is A = 0, B = 0, and Cin = 0, MEM1, MEM2, MEM4, MEM6, MEM8, MEM9, MEM13, MEM14, and MEM15 become Ron. And Resistor value of other memristors becomes high i.e. Roff. And so the output Sum = 0 and carry out C_out = 0. When the input is A = 0, B = 1, and Cin = 0, MEM1, MEM4, MEM6, MEM9, MEM8, MEM9, MEM10, MEM11, MEM13, MEM14, MEM15, MEM16, MEM18, MEM21, MEM23, MEM25 become Ron. And Resistor value of other memristors becomes high i.e. Roff. And so the output Sum = 1 and carry out C_out = 0. When the input is A = 0, B = 0, and C_{in} = 1, MEM1, MEM2, MEM3, MEM4, MEM8, MEM12, MEM13, MEM14, MEM16, MEM21, MEM22, MEM24 and MEM25 become Ron. And Resistor value of other memristors becomes high i.e. Roff. And so the output Sum = 0 and carry out C_out = 0.

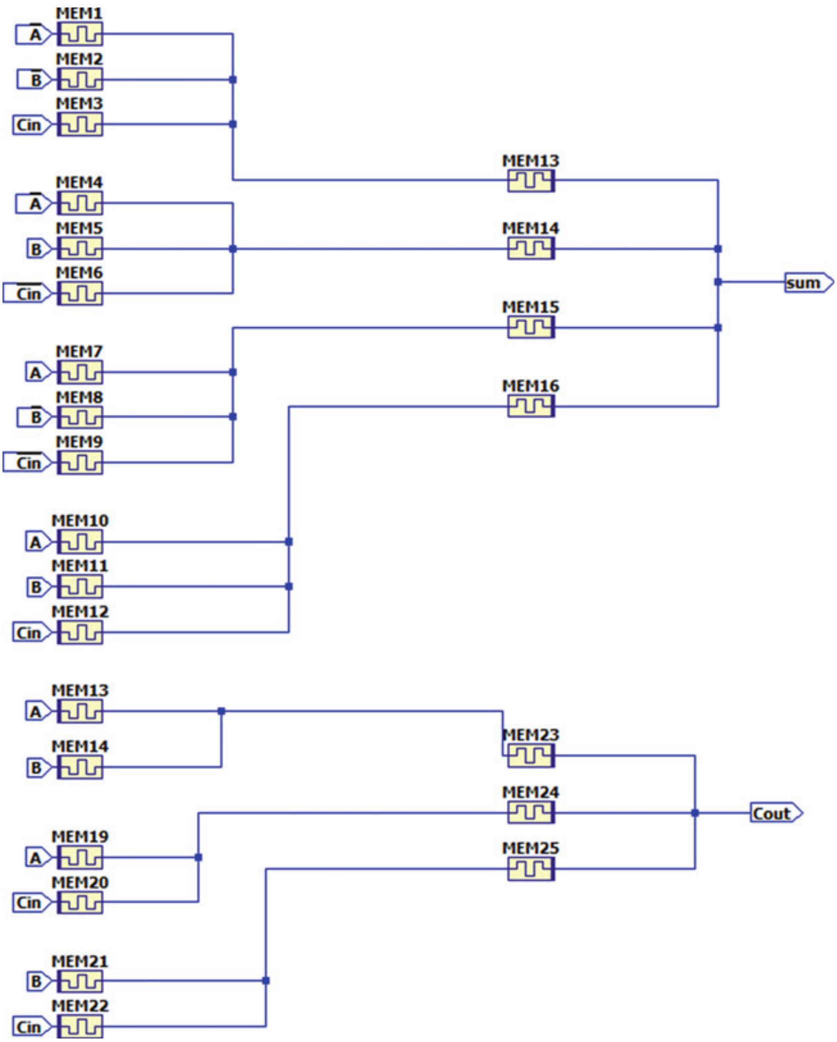


Fig. 12 Proposed memristor ratioed logic adder

When the input is A = 0, B = 1, and Cin = 1, MEM1, MEM3, MEM4, MEM5, MEM11, MEM12, MEM13, MEM14, MEM16, MEM18, MEM20, MEM21, MEM22, MEM23, MEM24, MEM25 become Ron. And Resistor value of other memristors becomes high i.e. Roff. And so the output Sum = 0 and carry out C_out = 1.

When the input is A = 1, B = 0, and Cin = 0, MEM2, MEM6, MEM7, MEM8, MEM9, MEM10, MEM10, MEM13, MEM14, MEM15, MEM16, MEM17, MEM19, MEM23, MEM24 become Ron. And Resistor value of other memristors becomes high i.e. Roff. And so the output Sum = 1 and carry out C_out = 0.

When the input is A = 1, B = 0, and C_{in} = 1, MEM2, MEM3, MEM7, MEM8, MEM10, MEM12, MEM13, MEM15, MEM16, MEM17, MEM19, MEM20, MEM22, MEM23, MEM24, MEM25 become Ron. And Resistor value of other memristors becomes high i.e. Roff. And so the output Sum = 0 and carry out C_out = 1. When the input is A = 1, B = 1, and C_{in} = 0, MEM5, MEM6, MEM7, MEM9, MEM10, MEM11, MEM14, MEM15, MEM16, MEM17, MEM18, MEM19, MEM21, MEM23, MEM24, MEM25 become Ron. And Resistor value of other memristors becomes high i.e. Roff. And so the output Sum = 0 and carry out C_out = 1. When the input is A = 1, B = 1, and Cin = 1, MEM3, MEM5, MEM7, MEM10, MEM11, MEM12, MEM13, MEM14, MEM15, MEM16, MEM17, MEM18, MEM19, MEM20, MEM21, MEM22, MEM23, MEM24, MEM25 become Ron. And Resistor value of other memristors becomes high i.e. Roff. And so the output Sum = 1 and carry out C_out = 1. Voltage Output range is between 2.5 to 5 V. If we feed the input 1, 1, 0 to A, B, Cin and apply the nodal analysis in the below figure we get (sum-5)/1/3 + (sum)/(1/3) = 0 then sum = 2.5 V (Figs. 13, 14 and 15).

The table is given below shows improvement in propagation delay of the circuit as we advance our design from the 28-transistor logic circuit to Memristor based logic

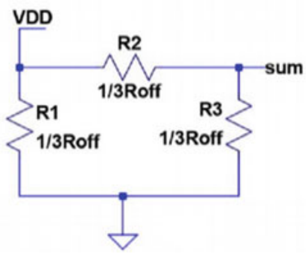


Fig. 13 Resistor implementation of memristor based adder

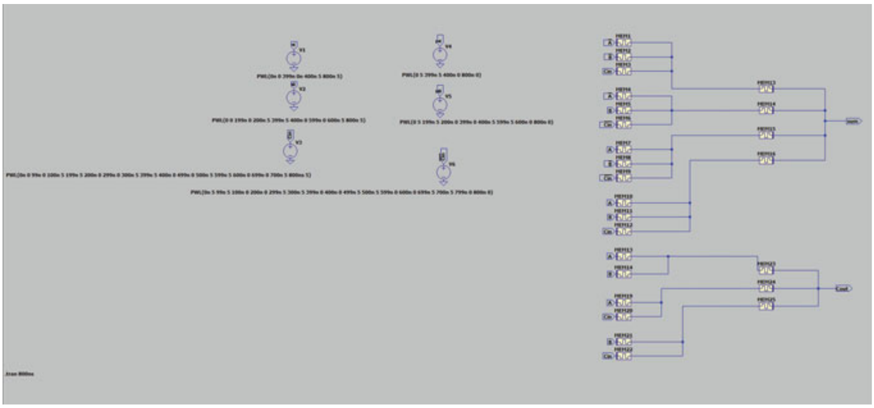


Fig. 14 Simulated circuit diagram of memristor based logic circuit

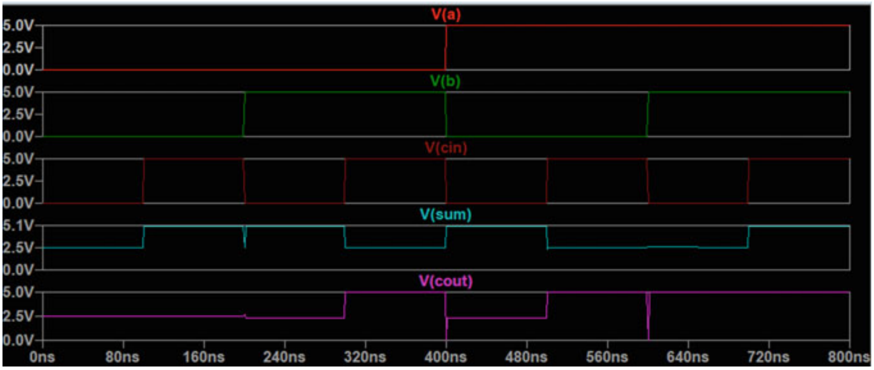


Fig. 15 Simulated output of memristor based logic circuit

Table 2 Comparative study table for propagation delay and power consumption of various full adder logic design

Adder style	Transistor count	Time (Ps)	Power (nW)
24T adder	24	310	540.5
14T adder	14	106.66	272.04
6T adder	6	93.55	109.175
(MRL) adder	0	29.51	88.64

circuit. It gives an improvement in the area by having less transistor count in each of our modified circuit designs as compared to the conventional logic based 1-bit full.adder circuit. Time.delay.and.power consumption are calculated with help of the LTspice Simulator (Table 2).

6 Conclusion

In this paper, we have designed a 24 Transistor adder, 14 Transistor adder,6 Transistor adder and memristor-based adder circuit and have done the comparative study on their propagation delay, power consumption using LTspice Simulator. As we build a 1-bit full adder logic circuit with lesser number of transistors, the propagation delay as well as power dissipation in the full adder logic design keeps on reducing. Memristor Ratioed Logic Adder is the fastest among all the adders designed. Its propagation delay reduces upto 90% as compared to Mirror adder circuit having 24 Transistors in it thus resulting in a high speed logic circuit and power consumption gets reduced upto 83% in memristor based full adder circuit as compared to Mirror Adder. And hence the major advantage of this full adder is it can improve the efficiency of the processor to the greater extent.

References

1. Bui, H.T., Wang, Y.: New 4-transistors XOR, and XNOR designs. In: Proceedings of the 2nd IEEE Asia Pacific Conference, ASIC, pp. 25–28 (2000)
2. Kaarthik, K., Vivek, C.: Hybrid Han Carlson adder architecture for reducing power and delay. Middle-East J. Sci. Res. **24**(Special Issue on Innovations in Information, Embedded and Communication Systems), 308–313 (2016)
3. Zimmermann, R., Fichtner, W.: Low-power logic styles: CMOS versus pass-transistor logic. Solid-State Circ. IEEE J. **32**, 1079–1090 (1997)
4. Wairya, S., Nagaria, R.K., Tiwari, S.: Comparative performance analysis of XOR-XNOR function based high-speed CMOS full adder circuits for low voltage VLSI design. Int. J. VLSI Des. Commun. Syst. (VLSICS) **3**(2) (2012)
5. Qu, L., Cui, X., Xu, X., Cui, X., Ma, Y.: The multi-input MRL logic gate and its application. In: IEEE International Conference on Electron Devices and Solid-State Circuits (EDSSC), pp. 1–2, June 2019
6. Rawat, G., Rathod, K., Goyal, S., Kala, S., Mittal, P.: Design and analysis of ALU: vedic mathematics approach. In: International Conference on Computing Communication and Automation (ICCCA 2015), Galgotias University, Noida, India, 15–16 May 2015, pp. 1372–1376 (2015)
7. Shams, A.M., Bayoumi, M.A.: Performance analysis of low-power 1-bit CMOS full adder cells. IEEE Trans. Very Large Scale Integr. (VLSI) Syst. **10**(1), 20–29 (2002)
8. Lin, J.-F., Hwang, Y.T., Sheu, M.H., Ho, C.C.: A novel high-speed and energy efficient10-transistor full adder design. IEEE Trans. Circ. Syst. **54**(5):1050–1059 (2007)
9. Yano, K., Sasaki, Y., Rikino, K., Seki, K.: Top-down pass-transistor logic design. IEEE J. Solid-State Circuits **31**, 792–803 (1996)
10. Chandrakasan, A.P., Brodersen, R.W.: Minimizing power consumption in digital CMOS circuits. Proc. IEEE **83**(4), 498–523 (1995)
11. Kumar, B., Raj, K., Mittal, P.: FPGA implementation and mask level CMOS layout design of redundant binary signed digit comparator. Int. J. Comput. Sci. Netw. Secur. **9**(9), 107–115 (2009). ISSN 1738-7906
12. LTspice IV user guide. Linear Technology Corporation, 1998–2012 (2012)
13. Singh, T.: Hybrid memristor-CMOS (MeMOS) based logic gates and adder circuits CoRR, arXiv:1506.06735 [cs.ET], pp. 1–11, June 2015 [Pre-Print]
14. Yamtim, S., Tooprakai, S.: Low power hybrid memristor-CMOS full adder circuit using NAND gate. In: 2017 International Symposium on Multimedia and Communication Technology, pp. 49–52, 23–25 August 2017
15. Shams, A.M., Darwish, T.K., Bayoumi, M.A.: Performance analysis of low-power 1-bit CMOS full adder cells. IEEE Trans. Very Large Scale Integr. (VLSI) Syst. **10**(1):20–29 (2002)
16. Palanivel Rajan, S.: Experimental explorations on EOG signal processing for real time applications in lab VIEW. IEEE Digital Library Xplore (2012)
17. Rathod, A.P.S., Lakhera, P., Baliga, A.K., Mittal, P., Kumar, B.: Performance comparison of pass transistor and CMOS logic configuration based de-multiplexers. In: International Conference on Computing Communication and Automation (ICCCA 2015). Galgotias University, Noida, India, 15–16 May 2015, pp. 1433–1437

Reconfigurable Microstrip Patch Antenna for Multiband Wireless Applications

Piyush Kr. Mishra, J. A. Ansari, Amrees Pandey, and Iqra Masroor

Abstract This paper presented a new holistic design of a reconfigurable microstrip patch antenna for multiband applications. This design consists of an Fr4 substrate and a new patch design loaded with slots and line feed. Slots and other design structures are compact and optimized for the application. Frequency reconfigurability achieved through PIN diode. These modifications are providing multiple bands of frequencies with wide bandwidth and stable gain. This antenna covers a large range of microwave frequency bands such as L-band (1.4 GHz), S-band (3.4 GHz), C-band (5.9, 6.7, and 6.8 GHz), and X-band (8.7 GHz) and makes this design suitable for a wide range of wireless communication applications. The antenna is design and simulated using HFSS-17 simulation software.

Keywords Reconfigurable antenna · PIN diode · Switching

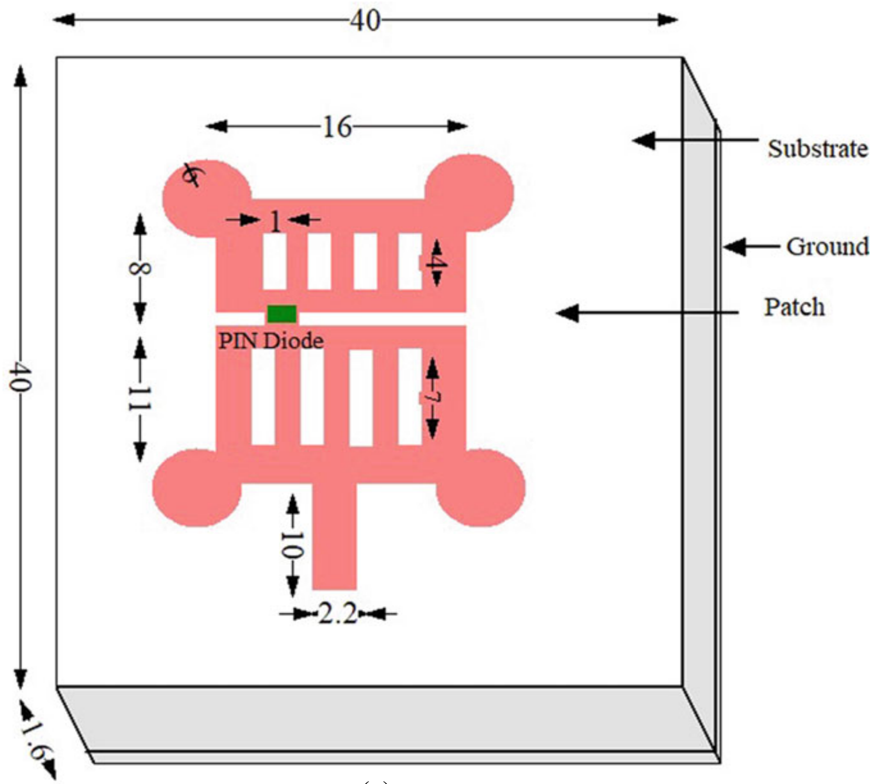
1 Introduction

The present fast-growing era of wireless communication is very dynamic and faces new challenges every day. Wireless applications such as WiMax, Wi-fi, WLAN, GPS are works on passband frequency and the antenna is a very crucial part of these communication methods [1, 2]. New techniques and applications of wireless technology require a smart and multifunctional antenna that can efficiently utilize the allocated resources. [3] Rapid developments require a single antenna with multiple bands of operation and resource utilization simultaneously [4, 5]. These requirements are not fulfilled by fixed characteristics and performance type antenna such as the antenna for which frequency, radiation pattern, polarization remain unchanged [4]. To overcome these limitations reconfigurable antennas are presented as a solution [6]. Reconfigurable antennas are designed by changing their physical structure by connecting and disconnecting the structure that leads to a change in current distribution in the radiator [7]. For this purpose different types of on and off, switches are

P. Kr. Mishra (✉) · J. A. Ansari · A. Pandey · I. Masroor
JK Institute of Applied Physics and Technology, University of Allahabad, Prayagraj, India
e-mail: ec.piyush21@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_34

used such as Varactor diode, RF MEMS switch, PIN diode, capacitance switches, and lumped elements. PIN diode or varactor diode is faster and a suitable replacement to RF-MEMS [8]. The PIN diode has a switching speed of 1–100 nsec. PIN diode provides the more dynamic ability to reconfigurable antennas. In this presented antenna a reconfigurable antenna is designed on FR4 substrate using PIN diode between the two sections of patch design to get frequency reconfigurable structure. This design consists of some basic concepts of multiband antennas to get multiple bands of operation in both the case of switching. It resonates at multiple frequencies with wide bandwidth, stable gain, and radiation pattern.



(a)

Fig. 1 **a** Schematic diagram of the proposed reconfigurable antenna **b**, **c**, **d** Step wise evolution of proposed Antenna 1, Antenna 2, Antenna 3 (proposed) **e** Antenna with diode off state **f** Antenna with diode on state

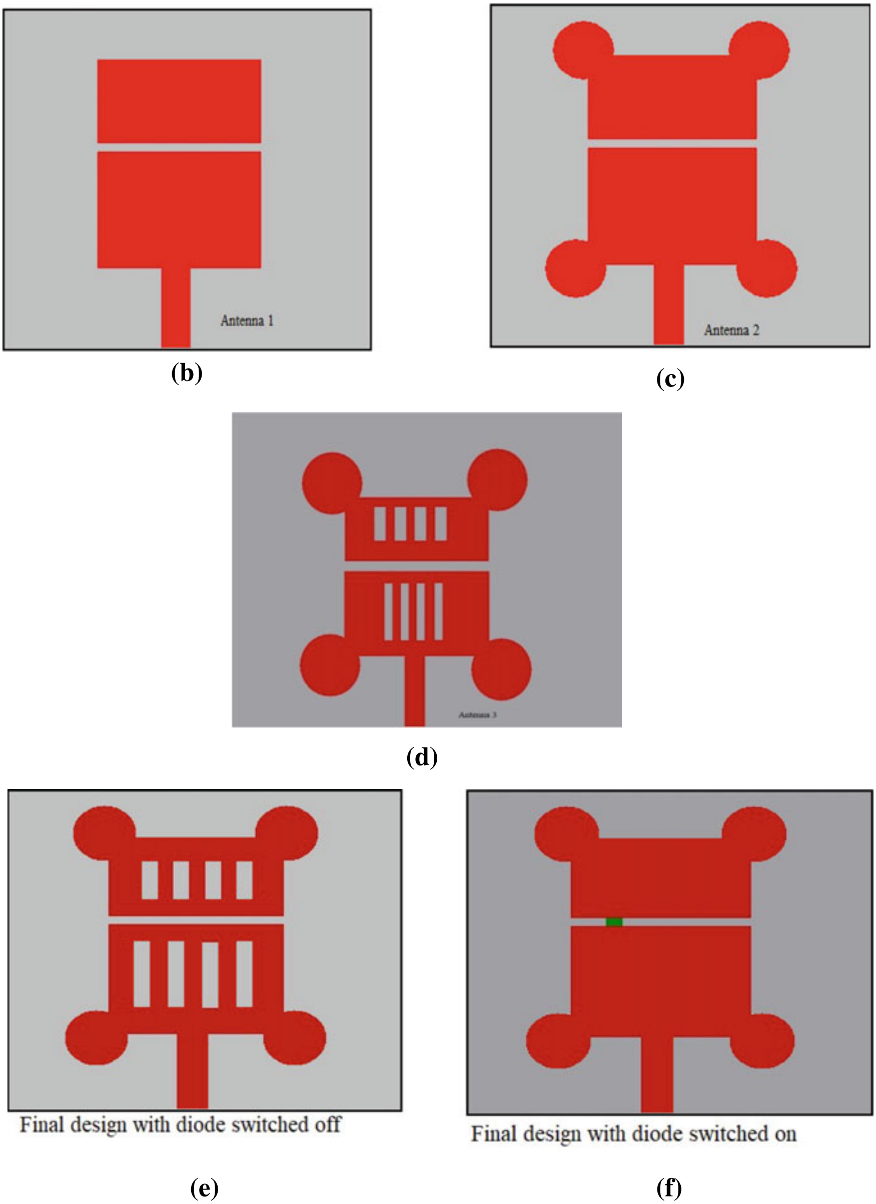


Fig. 1 (continued)

2 Evolution of Antenna Design and Analysis

It consists of a radiating patch that has two different sections. The first one is a rectangular patch of dimension 11×16 mm^2 with two sub circle of radius 6 mm as a radiating subsection. There are four slots of dimension 7×1 mm^2 cut out from the rectangular patch. This section of the patch is a line feed with a strip of dimension 10×2.2 mm^2. The second section of the radiating patch consists of a rectangular patch of dimension 8×16 mm^2 which is separated from the first one with a gap of 1 mm strip. It also consists of two subsection circles of radius 6 mm. Four slots of dimension 4×1 mm^2 were cut out of the slot.

To get the reconfigurable behavior a PIN diode is applied in the gap between two subsections of the patch. This pin diode is has switched on and off to reconfigure the presented design. The two-state of the PIN diode is shown in Fig. 1(e) and (f) provide different structures for radiation.

The overall patch design works in two different modes.

2.1 Case-1 When the PIN Diode Switched off

In this case, the only first section of the patch is mainly contributed to overall radiation. The radiating patch and the slots give multiband operation while subsection circles provide wideband to the resonant frequency. Here the second section also works as parasitic elements to provide good gain at wideband resonant frequencies. This arrangement polarizes two frequencies within the band at 6.8 GHz with a gain of 3.35 and 8.7 GHz with a gain of 1.5. The first resonant frequency has a bandwidth of 1 GHz with positive gain and the next resonant frequency has a bandwidth of 0.1 GHz and also has a positive gain.

2.2 Case 2 When the PIN Diode is Switched on

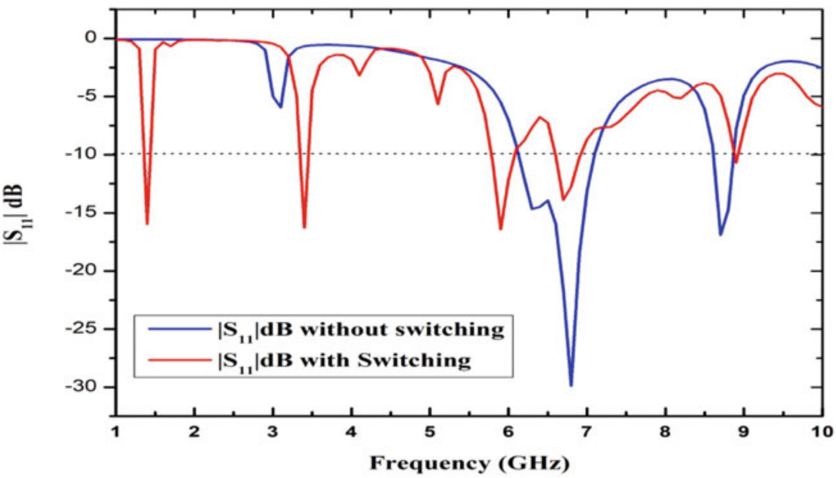
In this case, the second section of the patch is connected to the first one and increases the overall patch dimension. With the help of this switching redistribution of the current takes place within the patch and this reconfigured patch structure resonant at another band of frequencies. The second section has four slots and these slots contribute again to getting multiple resonant frequencies within the band. The gap area between the two sections also works as a parasitic element and enhances the characteristics of the antenna. This reconfigured design polarizes four different frequencies within the band at 1.4 GHz with a gain of 2, 3.4 GHz with a gain of 2.4, 5.9 GHz with a gain of 2.5, and bandwidth of 0.1 and 6.7 GHz with a gain of 3.9 and bandwidth of 0.4 GHz.

The antenna is fabricated on Fr4 substrate which has a dielectric constant of 4.4. All the antennas design and simulated using HFSS-17 software.

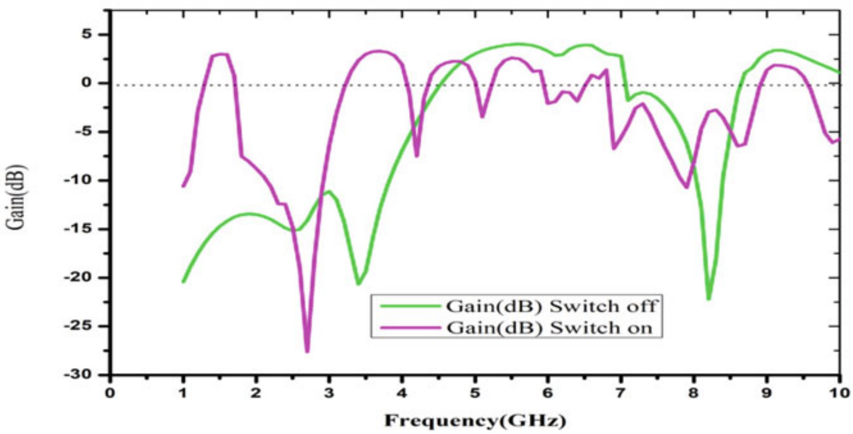
3 Results and Discussion

The presented reconfigurable antenna design operates on two different bands of frequencies and both the band of operation exhibits multiband characteristics. The return loss (S_{11}) Characteristics are depicted in Fig. 2(a) and showed that in the first case C-band and X-band of operation come into the picture. The 6.8 GHz frequency with a good gain is used in radar and satellite applications [9]. At the frequency of 6.8 GHz, a wide bandwidth of 1 GHz also achieves which makes it suitable for a large section of C-band operation. The radiation field has two major lobes at this frequency. The next polarize frequency in this case in 8.7 GHz as shown in Fig. 2(a). The radiation pattern has a major lobe and some minor lobes at this frequency. This frequency utilizes in X-band operations. This frequency is useful in cognitive radio applications [10]. In the second case when the diode is switched on, overall characteristics are reconfigured, and return loss characteristics exhibits that there are four frequencies polarizes. The first polarization frequency 1.4 GHz and radiation pattern shows a major lobe at this frequency. This frequency is useful for L-band operations and is suitable for GPS and IRNSS applications [11].

The second frequency polarizes, in this case, is 3.4 GHz works for S-band applications. The radiation pattern has a major lobe at this point of operation. The third polarized frequency is 5.9 GHz fall in the application range of C-band and is used in WiMax Technology. The radiation pattern for this frequency shows a major lobe and some minor lobes. The fourth polarized frequency is 6.7 GHz in also used for C-band operations such as satellite and radar applications and the radiation pattern for this frequency shows to lobes of radiation. The Gain and radiation Pattern of the design in both the case of operation shown in Fig. 2(d) which exhibits that the antenna gives stable operation at all the resonant frequencies in both the case.

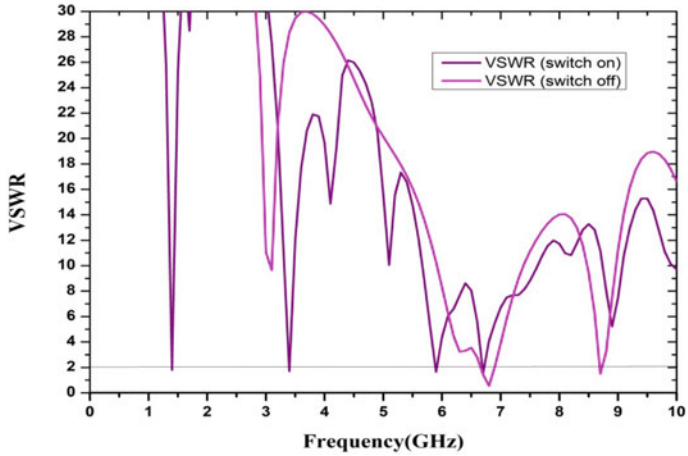


(a)

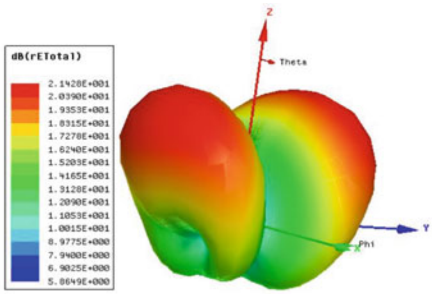


(b)

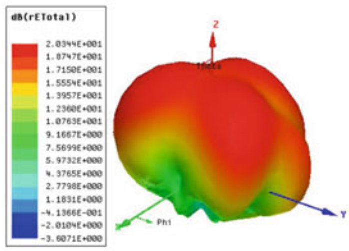
Fig. 2 **a** Return loss characteristics, **b** Gain vs. Frequency plot **c** VSWR plot **d** Radiation characteristics at polarized frequencies **e** Current distribution at different state



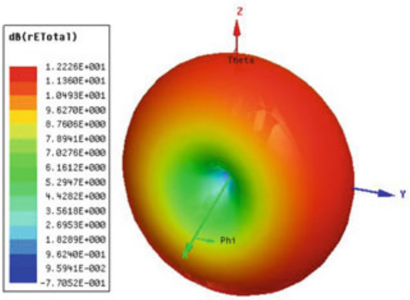
(c)



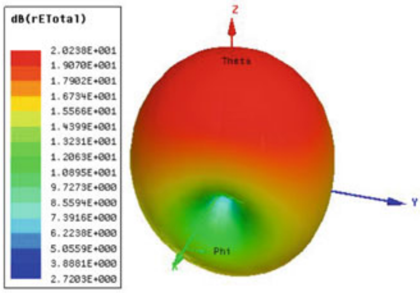
(d_1) at 6.8GHz



(d_2) at 8.7GHz

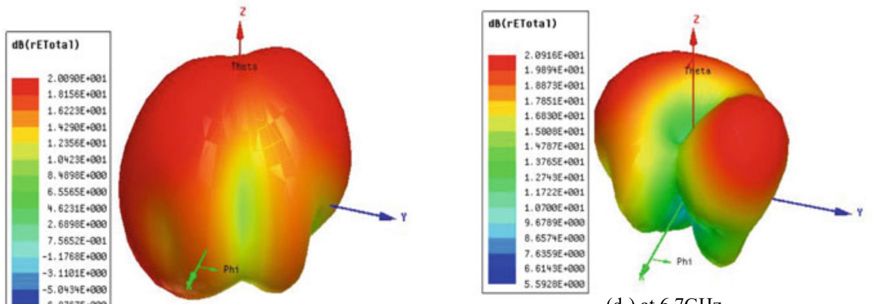


(d_3) at 1.4GHz



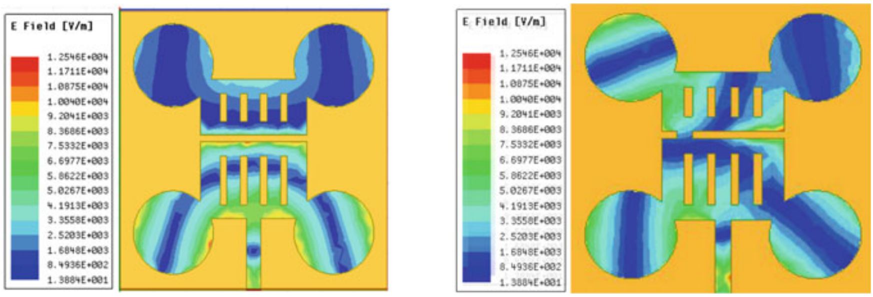
(d_4) at 3.4GHz

Fig. 2 (continued)



(d_5) at 5.9GHz (d_6) at 6.7GHz

(d)



(e_1) switches off state (e_2) switches on state

(e)

Fig. 2 (continued)

4 Conclusions

The presented reconfigurable antenna with two sections coupled with a PIN diode is a simple structure, low cost, lightweight, easy to fabricate antenna with multiple bands of applications. It is suitable for L- band, S, band, C-band, and X-band of operation. This design can also be tested with different type of substrate material to get some more area of applications. The significant contribution of the proposed work is that this antenna design gives multiple operating frequencies with stable gain and radiation pattern; make it suitable for a wide range of applications such as location tracking, health monitoring, defense and radar, sports, IoT, etc.

References

1. Shi, S.J., Ding, W.P.: Radiation pattern reconfigurable microstrip antenna for WiMAX application. Electron. Lett. **51**(9), 662–664 (2015)
2. Lin, S.Y., Huang, K.C.: A compact microstrip antenna for GPS and DCS application. IEEE Trans. Antennas Propag. **53**(3) (2005)
3. Alexiou, A., Haardt, M.: Smart antenna technologies for future wireless systems: trends and challenges. IEEE Commun. Mag. **42**(9), 90–97 (2004). https://doi.org/10.1109/MCOM.2004.1336725
4. Dubal, S., Chaudhari, A.: Mechanisms of reconfigurable antenna: a review. In: 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence), pp. 576–580. IEEE, January 2020
5. Debele, A.S.: Review of multiband antenna for mobile communication. J. Electr. Eng. Electron. Control Comput. Sci. (2020)
6. Sharma, S., Tripathi, C.C.: Reconfigurable antennae: a review 1 (2011)
7. Jenath, M., Nagarajan, V.: Review on frequency reconfigurable antenna for wireless applications. In: 2017 International Conference on Communication and Signal Processing (ICCSP), pp. 2240–2245 (2017). https://doi.org/10.1109/ICCSP.2017.8286815
8. Kholapure, A.S., Karandikar, R.G.: Emerging techniques for printed reconfigurable antenna: a review. In: 2016 Second International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), pp. 57–61. IEEE, September 2016
9. Sathikbasha, M.J., Nagarajan, V.: DGS based multiband frequency reconfigurable antenna for wireless applications. In: 2019 International Conference on Communication and Signal Processing (ICCSP), pp. 0908–0912. IEEE, April 2019
10. Tawk, Y., Christodoulou, C.G.: A new reconfigurable antenna design for cognitive radio. IEEE Antennas Wirel. Propag. Lett. **8**, 1378–1381 (2009). https://doi.org/10.1109/LAWP.2009.2039461
11. Singh, J., Lohar, F.L., Sohi, B.S.: Design of circular polarized patch antenna for NaviC receiver applications. IOP Conf. Ser. Mater. Sci. Eng. **1033**(1), 012039 (2021)

An Ultra Efficient QCA SRAM Cell for Nanotechnology Applications

Divya Tripathi and Subodh Wairya

Abstract Quantum-dot cellular automata is a cutting edge enumeration methodology that suggests less area and high speed compare with CMOS technology. The CMOS circuitry having issues related to short channel and device density so QCA is better and powerful alternative to reduce the area as well as increase the speed of the circuitry. QCA could be a modern computing innovation that's made of quantum cell containing two electrons and dots. In this article the QCA D flip-flop is suggested and further this QCA D flip-flop (DFF) is utilized to design QCA SRAM cell and these QCA structures will be utilized to design interesting and ideal designs of QCA SRAM cell. The comprehensive investigation of the planned design suggests that is optimum D flip-flop shown 20.0% and proposed optimum QCA SRAM cell shown 60.84% improvement in cell count and the proposed optimum D flip-flop shown 91.7% enhancement and the proposed optimum SRAM cell architecture is 69.3% enhancement in the cost efficient as associated to its existing past layout. The suggested circuits are tested with the QCA Designer simulation tool. The simulation results clarify the suggested designs outperform the best past layout in specifications of quantum cell, area, latency and cost.

Keywords CMOS · QCA · DFF · SRAM

1 Introduction

Moore's law suggests, the chip design industry has advanced remarkably up to nano scale. Quantum Dot Cellular Automata (QCA) is a nano architecture paradigm computation method for digital circuitry designing [1, 2]. The chip plan industry has advancement over conventional CMOS and demonstrate to be a better alternate

D. Tripathi (✉)
Electronics and Communication Engineering, Institute of Engineering and Technology, Lucknow, India
e-mail: divyatripathi.cest@gmail.com

S. Wairya
Dr. APJ. Abdul Kalam Technical University, Lucknow, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_35

than past methods. Presently circuit, researchers are driving towards a progressed connected science subordinate on electron polarization, which gives beneficial comes around the decrement and the quantum cost to the intemperate less level [3–5].

Various SRAM cells are used to design complex processors to store information. Memory covers a huge region in the system on the chip industry [6, 7]. Scaling of the transistor estimation is the necessity of the present day advancement. It makes a contrast in reducing the area of the design which comes around in diminishment of the degree of SRAM and advances integration thickness, but it increases leakage current [8]. SRAM cluster that comprises various cells, gets to be an enormous way of leakage current. This issue of information stability in the SRAM cells emerges due to voltage scaling and devices dimension in a new strategy [9]. The RSNM is SNM at the examined mode is deliberate due to more transistor using and a lessening voltage supply. This issue can be settled by utilizing a separate read circuit [10]. This reads circuit increment the sum of transistors inside the cell which impact in an increase within the space of the SRAM cell. To move forward SNM and diminish in the region of SRAM, diverse SRAM cells are laid out [11]. The SNM of the SRAM cell can as well progress by utilizing legitimate measuring the proportion of the transistor. Read–Write circuit is required in memory design that can alter the information on memory cluster at the time of type in mode and hold information in perused mode. SRAM is inactive due to holding its data as long as the control supply is connected without requiring occasional refreshment [12]. These are the methods through which data stability can improve during read and write mode of operation. Data stability of the cell in read mode and hold mode is a very main restriction in advanced technology. So, this paper is based on cell optimization and the realization of a QCA D flip-flop that is further used in designing of the QCA SRAM cell.

1.1 The Essential Contribution

The essential contribution of this article is as follows

a) Designing of an efficient QCA D flip-flop (DFF).
b) Designing of QCASRAM cell topology using a proposed cost-efficient DFF.
c) The suggested layout is summarized based on cell calculation, the area occupied, latency, and the quantum cost that confirms that the suggested layouts have a lesser arena and better quantum cost associated with its past best layout.

1.2 Paper Structure

The article's layout is classified as follows: Sect. 2 deals with QCA nomenclature. The proposed layout of the QCA D flip-flop and QCA SRAM cell architecture is presented in Sect. 3. In Sect. 4 proposed methodology of QCA SRAM cell has been explained. Section 5 presented the simulation results and was compared to the previous best ideas. Section 6 explained the energy dissipation of proposed QCA SRAM cell. Finally, Sect. 7 shows the conclusion.

2 Quantum Dot Cellular Automata (QCA) Terminology

A QCA is a collection of quantum cells in which each cell has received information electrostatically from its neighboring cell [6, 7] (Fig. 1).

2.1 Majority Gate

Majority gate is a combination of 5 cells in such a way that 2 cells are input 1 is polarization cell 1 is normal cell and 1 output cells are used in majority gate. It plays a vital role in QCA computing (Fig. 2).

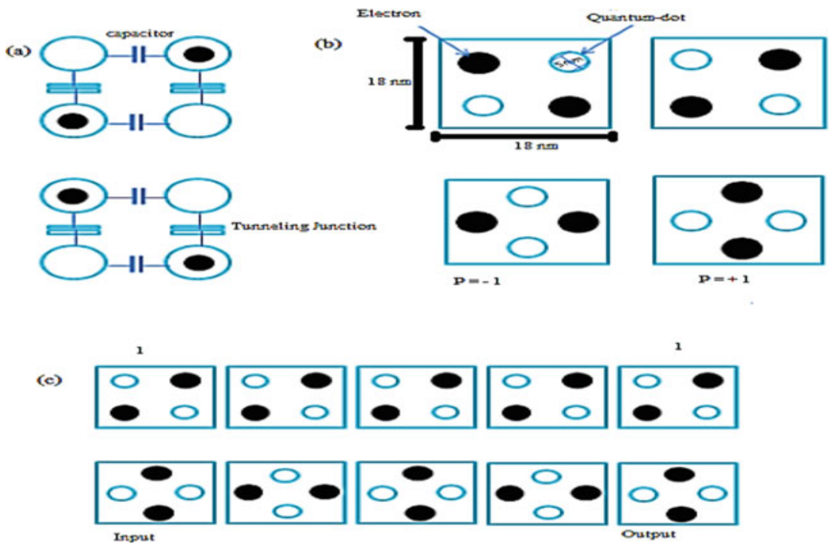


Fig. 1 Cells in QCA **a** Operational diagram **b** Cells polarization **c** QCA wire

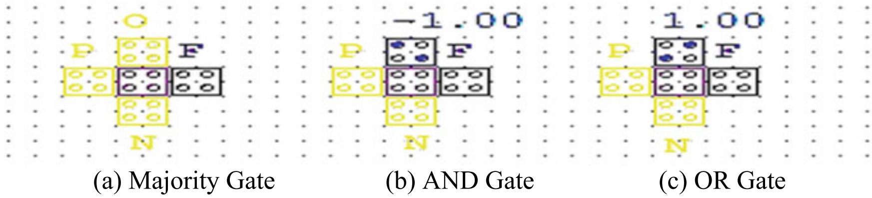


Fig. 2 QCA layout of basic logic gates (Majority Gate)

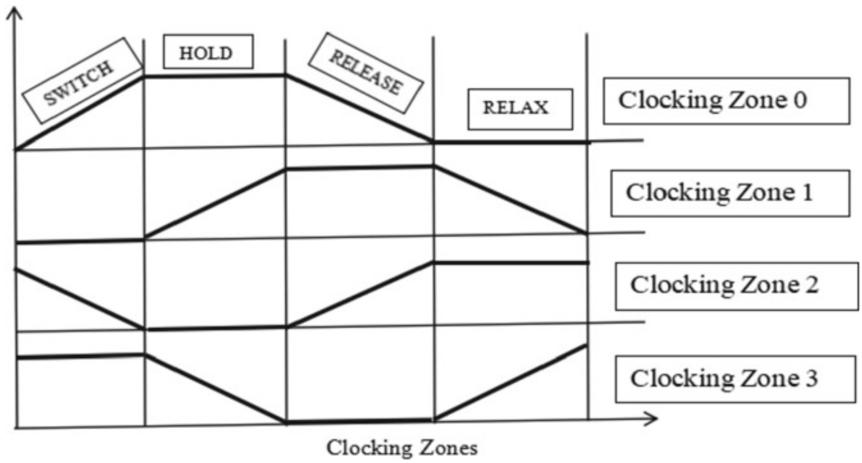


Fig. 3 Clocking zones of QCA

2.2 QCA Clocking

The ability to make and eliminate the meta stable phase is suggested by QCA clock [7–10]. The QCA clock consists four phases: Switch, Control, Release and Relax [8, 9] (Fig. 3).

3 Proposed QCA D Flip Flop Architectures

In sequential logic, flip flops are the most important capacity component. In this segment, an optimum QCA D flip-flop is proposed. Computers, telecommunication and a variety of other digital devices all employ flip flops. The D flip-flop (DFF) is the most basic flip flop, that takes the input signal D at an unequivocal stage of the clock cycle. The output Q unchanged. It is also named as a “data” or “delay” flip-flop. The basic block diagram of memory cell using D flip-flop is presented n in Fig. 4 (Figs. 5 and 6).

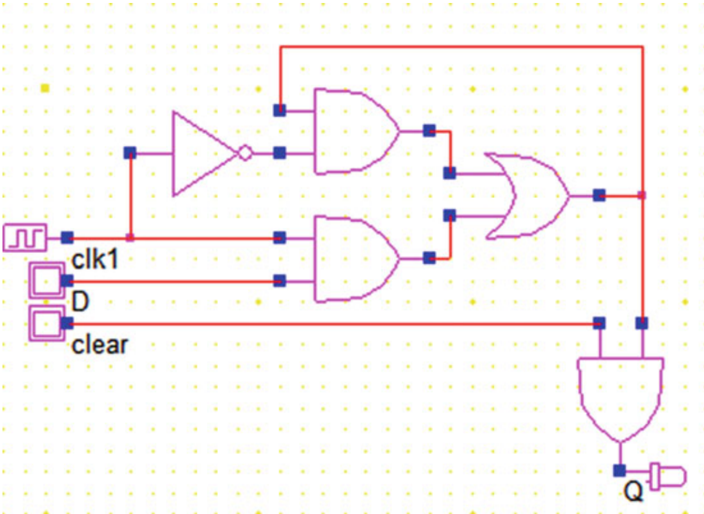


Fig. 4 Conventional logic diagram using D flip-flop

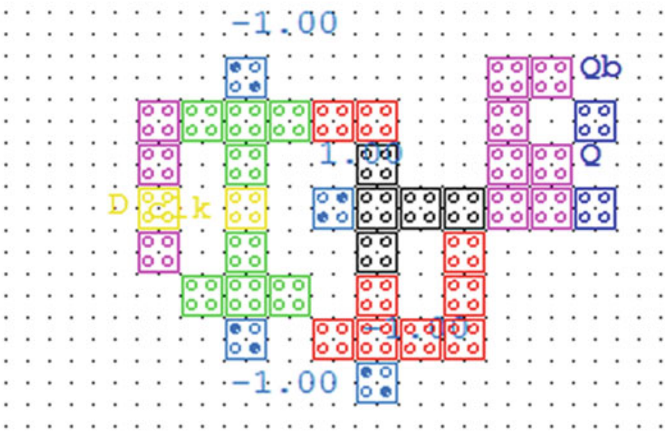


Fig. 5 QCA layout of proposed D flip-flop

4 Proposed QCA SRAM Cell

In VLSI, thousands of devices are fabricated on a single chip which has a smaller surface area [9]. Low power in portable devices becomes a primary factor [10]. The total operation of suggested QCA SRAM design is when R /w is ‘0’ the output of the circuit will be uncharged. The output layout distinguishes the impact of input and set/reset by taking R/W = ‘1’. Additionally, W is ‘1’ and set is ‘1’ gives an

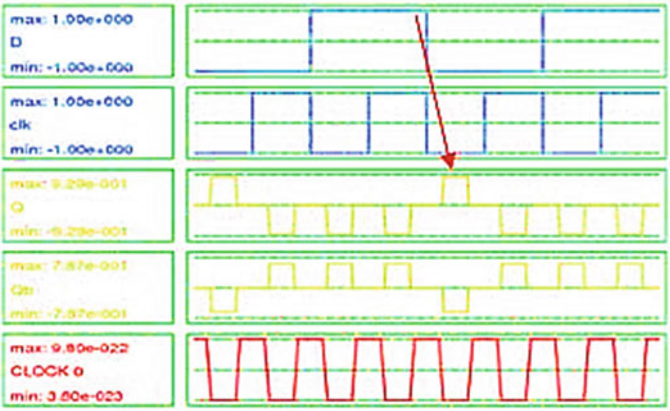


Fig. 6 Simulation waveform of QCA proposed D flip-flop

Table 1 Logic operational table of suggested D flip-flop

clk	D	Q	Qb
0	0	Q	Qb
0	1	Q	Qb
1	0	0	1
1	1	1	0

Table 2 Logic operational table of suggested SRAM cell

Select	R/W	Set/Reset	Out
0	0	x	Out
0	1	x	Input
1	x	1	1
1	x	0	0

input flag to be exchanged to the output, and for the sel is '0' the output will be charged as per set/reset flag as summarized in Table 1 (Table 2 and Figs. 7, 8 and 9).

Simulation results are presented, while the positive edge of clock the input proliferates to the output and enroll a substantial output after 2 clock cycles, something else no alters within the output. This shows rectify operation of suggested structure.

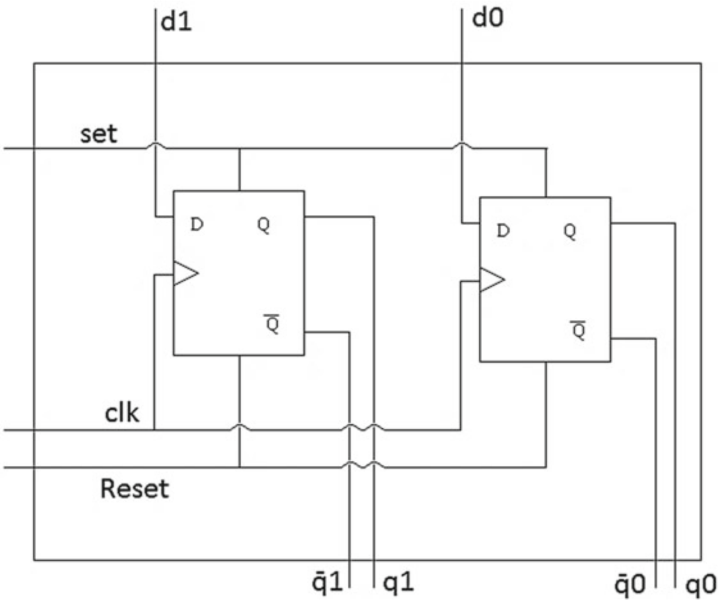


Fig. 7 Conventional block Diagram of SRAM cell using D flip-flop [12]

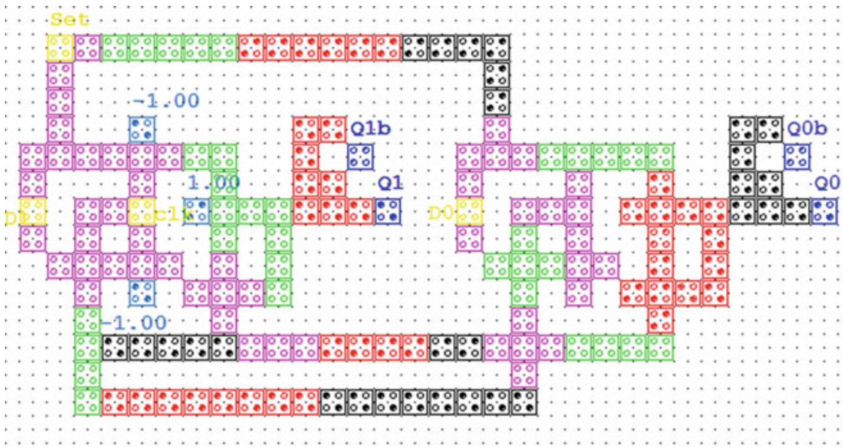


Fig. 8 Proposed QCA SRAM cell layout

5 Simulation Result and Discussion

The functionality and performance of the suggested layouts are verified by a QCA designer tool and shows that the suggested layouts are better as related to the best existing past design. To assess the effectiveness and adaptability of suggested

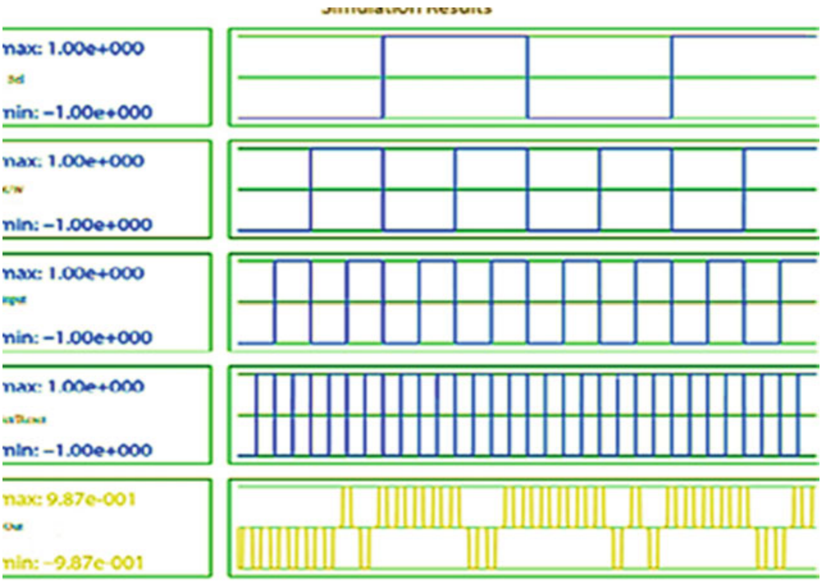


Fig. 9 Simulation Waveform of **proposed** QCA SRAM cell

Table 3 Performance parameter comparison of QCA D flip flop

QCA D flip flop	Cells	Area (μm^2)	Latency (Clocking cycles)	Quantum cost (Area * Latency)
[8]	48	0.05	1.00	0.05
[9]	84	0.09	2.75	0.24
[10]	86	0.09	2.70	0.25
[11]	120	0.14	3.25	0.45
[12]	74	0.10	1.50	0.15
[13]	46	0.05	1.50	0.75
PD	40	0.04	1.00	0.04

QCA layouts the comparison with its best existing designs in parameters of cell calculations, area, latency, and quantum cost is presented in Tables 3 and 4 (Fig. 10).

6 Energy Dissipation of the Proposed QCA SRAM Cell

Energy dissipation of the suggested QCA RAM cell is calculated by QCADesigner E tool that is an extension of QCA Designer [14]. It executes the valuation of the energy dissipation of suggested architectures built on the study. The parameters like

Table 4 Performance parameter comparison of QCA SRAM Cell

QCA memory cell designs	Cells	Area (μm^2)	Latency clock cycle	Quantum cost (Area * Latency)
[26]	233	0.31	1.73	0.53
[27]	173	0.25	1.50	0.37
[28]	158	0.16	2.00	0.32
[29]	100	0.11	3.00	0.33
[30]	63	0.07	2.00	0.14
[31]	109	0.13	1.75	0.05
[32]	75	0.06	1.50	0.09
[33]	71	0.09	1.25	0.11
[34]	177	0.22	0.60	0.13
PD	69	0.09	0.50	0.04

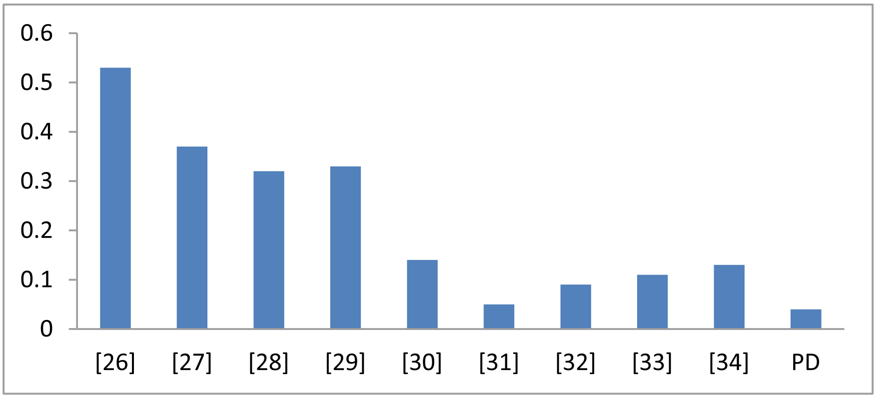


Fig. 10 Quantum cost analysis of numerous QCA SRAM cells

E_bath_total is the energy dissipation in every cycle, Sum_bath is total energy dissipation, Ab is the average bath and Ac average clock energy dissipation while simulation. The suggested QCA SRAM cell takes 6 iterations for the complete simulation process as mentioned in Table 5.

Table 5 Energy dissipation of proposed QCA SRAM Cell (N69)

E_bath_total (Ebtx)	E_clk_total (Ectx)	E_Error_total (EEtx)	Sum_bath (Sb)	Avg_bath (Ab)	Sum_clk (Sc)	Avg_clk (Ac)
1.85e−003	1.04e−003	−1.82e−004				
2.06e−003	1.05e−003	−3.18e−004				
2.39e−003	3.16e−004	−2.34e−004				
2.16e−003	8.29e−004	−1.03e–004				
1.87e−003	1.03e−003	−1.80e−004	2.32e−002 (Er:−2.41e−003)	2.11e−003 (Er:−2.19e−004)	8.03e−003	8.61e−004
2.03e−003	1.05e−003	−3.46e−004				
2.39e−003	3.26e−003	−2.83e−004				
2.16e−003	879e−004	−1.09e–004				
1.87e−003	1.06e−004	−1.80e–004				
2.03e−003	3.15e−003	−3.16e−004				
2.39e−003	2.19e−003	−2.65e−004				

7 Conclusion

QCA has pulled in researchers as modern rising nanotechnology for substitution the recent CMOS innovation. QCA innovation could be designed among them since it has no current flow and employments a one of a kind worldview for advanced representation and transferring information other than other merits accessible within the other technologies. In this article the QCA D flip flop is suggested and further this QCA D flip-flop is utilized to design QCA SRAM cell and these QCA structures will be utilized to design interesting and ideal designs of QCA SRAM cell topology. The comprehensive investigation of the suggested layout that the proposed designs is optimum D flip-flop shown 20.0% and proposed optimum QCA SRAM cell shown 60.84% improvement in cell count and the proposed optimum D flip-flop shown 91.7% enhancement and the proposed optimum SRAM cell architecture is 69.3% enhancement in the cost efficient as associated to its existing past layout. It is observed that the proposed plan has been displayed with Set/Reset capacity, and on the off chance that we need to concern the highlight for RAM cells, the suggested plan is ideal in parameter of the number of cells and possesses range and quantum fetches related to its best past layouts.

References

1. Taugaw, P.D., Lent, C.S.: Logical devices implemented using QCA. J. Appl. Phys. **75**, 1818–1825 (1994)
2. Shafi, M.A., Bahar, A.N., Bhuiyan, M.M.R., Shamim, S.M.: The average output polarization dataset for signifying the temperature influence for QCA designed reversible logic circuits. Elsevier Inc. **409**, 42–48 (2018)
3. Balijepalli, H.: Design, implementation and test of novel quantum-dot cellular automata FPGAs for the beyond CMOS Era. Thesis and Dissertations, p. 260 (2012)
4. Sridharan, K., Pudi, V.: Design of arithmetic circuits in quantum dot cellular automata nanotechnology. Stud. Comput. Intell. 599 (2015)
5. Hashemi, S., Navi, K.: New robust QCA D flip flop and memory structures. Microelectron. J. **43**, 929–940 (2012)
6. Pandey, S., Singh, S., Wairya, S.: Designing an efficient approach for JK and T flip-flop with power dissipation analysis using QCA. Int. J. VLSI Des. Commun. Syst. **7**, 29–48 (2016)
7. Hashemi, S., Farazkish, R., Navi, K.: New quantum dot cellular automata cell arrangements. J. Comput. Theor. Nanosci. **10**, 798–809 (2013)
8. Lim, L.A., Ghazali, A., Yan, S.C.T., Fat, C.C.: Sequential circuit design using Quantum-dot Cellular Automata (QCA). In: IEEE International Conference on Circuits and Systems (ICCAS), pp. 162–167 (2012)
9. Gladshtein, M.: Quantum-dot cellular automata serial decimal processing-in-wire: run-time reconfigurable wiring approach. Microelectron. J. **55**, 152–161 (2016)
10. Rezaei, A., Saharkhiz, H.: Design of low power random number generators for quantum-dot cellular automata. Int. J. Nano Dimens. **7**(4), 308–320 (2016)
11. Angizi, Sh., Moaiyeri, M.H., Farrokhi, Sh., Navi, K., Bagherzadeh, N.: Designing quantum-dot cellular automata counters with energy consumption analysis. Microprocess. Microsyst. **39**(7), 512–520 (2015)

12. Hashemi, S., Navi, K.: New robust QCA D flip flop and memory structures. Microelectron. J. **43**, 929–940 (2012)
13. Abutaleb, M.: Robust and efficient quantum-dot cellular automata synchronous counters. Microelectron. J. **61**, 6–14 (2017)
14. Tripathi, D., Wairya, S.: Energy efficient code converter for nanotechnology application. J. Crit. Rev. **7**, 2916–2925 (2020)
15. Wang, L., Xie, G.: A novel XOR/XNOR structure for modular design of QCA circuits. Trans. Circ. Syst. **II**, 1–5 (2020)
16. Roohi, A., Khademolhosseini, H., Sayedsalehi, S., Navi, K.: A novel architecture for quantum-dot cellular automata multiplexer. Int. J. Comput. Sci. **8**, 55–60 (2011)
17. Singh, R., Sharma, D.K.: QCA based RAM design using a resilient reversible gate with improved performance. J. Circ. Syst. Comput. **29** (2020)
18. Majeed, A., Alkaldy, E., Zainal, S., Navi, K., Nor, D.: Optimal design of RAM cell using novel 2:1 multiplexer in QCA technology. Circuit World. (2019). https://doi.org/10.1108/CW-06-2019-0062
19. Khosroshahy, M.B., Moaiyeri, M.H., Angizi, S., Bagherzadeh, N., Navi, K.: Quantum-dot cellular automata circuits with reduced external fixed inputs. Microprocess. Microsyst. **50**, 154–163 (2017)
20. Ahmadpour, S.S., Mosleh, M.: A novel fault-tolerant multiplexer in QCA technology. J. Superco. **74**, 4696–4716 (2018)
21. Li, X., Zhiwei, S., Hongping, C.: A new design of QCA based nanoscale multiplexer and its usage in communications. Int. J. Commun. Syst. **33**, 11–12 (2020)
22. Tripathi, D.S., Wairya, S.: Cell optimization and realization of MGDI and QCA based combinational logic circuits for nanotechnology applications. In: 2020 17th IEEE India Council International Conference. (INDICON 2020). IEEE Delhi Section Netaji Subhas University, New Delhi, India (2020)
23. Singh, G., Sarin, R.K., Raj, B.: A novel robust exclusive-or function implementation in QCA nanotechnology with energy dissipation analysis. J. Comput. Electron. **15**, 455–465 (2016)
24. Sheikhfaal, S.: A QCA layout design methodology. Part I Emerg. Technol. J., 1–6 (2018)
25. Chabi, A.M., Roohi, A., Khademolhosseini, H., Sheikhfaal, S., Angizi, S., Navi, K., DeMara, R.F.: Towards ultra-efficient QCA reversible circuits. Microprocess. Microsyst. **49**, 127–138 (2017)
26. Tripathi, D., Wairya, S.: A cost efficient QCA RAM cell for nanotechnology applications. In: 3rd International Conference on VLSI, Microwave & Wireless Technologies (ICVMWT 2021), MMMUT, Gorakhpur (2021)
27. Vankamamidi, V., Ottavi, M., Lombardi, F.: A line-based parallel memory for QCA implementation. IEEE Trans. Nanotechnol. **4**(6), 690–698 (2005)
28. Taskin, B., Hong, B.: Dual-phase line-based QCA memory design. In: Proceedings of the IEEE Conference on Nanotechnology, Cincinnati, OH, U.S.A., pp. 302–305 (2006)
29. Walus, K., Vetteth, A., Jullien G., Dimitrov, V.S.: RAM design using quantum-dot cellular automata. In: Proceedings of the Nanotechnology Conference and Trade Show, vol. 2, pp. 160–163 (2003)
30. Dehkordi, M.A., Shamsabadi, A., Ghahfarokhi, B.S., Vafaei, A.: Novel RAM Cell designs based on inherent capabilities of quantum-dot cellular automata. Microelectron. J. **42**, 701–708 (2011)
31. Hashemi, S., Navi, K.: New robust QCA D flip-flop and memory structures. Microelectron. J. **43**, 929–940 (2012)
32. Azimi, S., Moaiyeri, M.H.: Efficient and robust SRAM cell design based on quantum-dot cellular automata. ECS J. Solid State Sci. Technol. **7**(3), 38–45 (2018)
33. Khosroshahy, M.B., Moaiyeri, M.H. Navi, K., Bagherzadeh, N.: An energy and cost efficient majority-based RAM cell in quantum-dot cellular automata. Results Phys. **7**, 3543–3551 (2017)
34. Sasamal, T.N., Singh, A.K., Ghanekar, U.: Design and implementation of QCA D-flip flops and RAM cell using majority gates. J. Circ. Syst. Comput. **28**(5) (2019)

Optimal Guaranteed Cost Control of an Uncertain and Shift-Delayed 2-D Discrete FM First Model via Memory State Feedback

Govind Prasad Pandiya, Abhay Vidyarthi, Manish Tiwari, Amit Dhawan, and R. Maheswar

Abstract In the addressed issue, we aimed in ensuring the stability of an uncertain two-dimensional discrete system as represented by the Fonasni-Marchesini first model. The uncertainty considered in the addressed issue is norm-bounded uncertainty. Using Lyapunov method and memory state feedback technique, a new criterion has been developed in terms of the solution of certain linear matrix inequalities. By recasting the problem as a convex optimization problem, optimal guaranteed cost controllers have been selected to minimize the upper bound on the closed-loop cost function. The proposed technique's effectiveness has also been demonstrated by considering two different examples of Industrial processes.

Keywords Memory state feedback · Shift delay · LMI · Robust stability · Guaranteed cost control

G. P. Pandiya (✉) · A. Vidyarthi · R. Maheswar
SEEE, Vellore Institute of Technology (VIT) University, Bhopal-Indore Highway Kothrikalan, Sehore 466114, Madhya Pradesh, India
e-mail: govind.pandiya2020@vitbhopal.ac.in

A. Vidyarthi
e-mail: abhay.vidyarthi@vitbhopal.ac.in

R. Maheswar
e-mail: maheswar.r@vitbhopal.ac.in

M. Tiwari · A. Dhawan
ECED, Motilal Nehru National Institute of Technology Allahabad, Pryagraj 211004, Uttar Pradesh, India
e-mail: mtiwari@mnnit.ac.in

A. Dhawan
e-mail: dhawan@mnnit.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_36

1 Introduction

In the last few decades, there has been an increasing interest in the study and analysis of two-dimensional discrete systems. 2-D discrete systems applications are in heat exchangers, thermal processes, image data processing, etc. In two-dimensional discrete systems, information is available in two independent directions. Fornasini-Marchesini's first model is one of the popular two-dimensional discrete models [1]. The two-dimensional discrete systems when subjected to uncertainties may result in the deterioration of their performance.

In order to ensure the 2-D discrete model stability, the technique which has gained a lot of importance. In guaranteed cost control approach not only, the stability of the model under consideration has been ensured but also it helps in achieving the minimized cost function value even when the model has been subjected to the worst parametric uncertainties. Based on this approach various researchers have contributed a lot [2, 3, 4, 12–16, 4–11].

In [17] via memory-less state feedback, the problem of robust guaranteed cost control for 2-D discrete systems as represented by the Fornasini-Marchesini second model with shift delays has been considered. In [18], the same problem has been addressed using the guaranteed cost control approach and it has been observed that the result obtained using the memory state feedback is less conservative than those of [17].

In the various reported criteria so far, an effort has been made in ensuring the robust stability of the two-dimensional discrete model as represented by the Fornasini-Marchesini first model using the memory-less state feedback technique. If a state feedback controller with feedback provisions not only on the present but also on the past history of the state is designed particularly for a shift delayed system, we are expected to achieve better results and this helps in achieving the overall improvement in the performance of the system under consideration. To the best of the author's knowledge, the problem of designing an optimal guaranteed cost controller for an uncertain and shift delayed Fornasini-Marchesini first model via memory state feedback has not been addressed by any researcher so far and this motivates us to carry out this work [19]. Our aim is to select the optimal guaranteed cost controller which not only ensures the asymptotic stability but also helps in maintaining the least guaranteed cost upper bound. In our proposed technique we extend the result of [3].

The paper is organized as follows. The problem of robust guaranteed cost control for an uncertain and shift delayed two-dimensional discrete system as described by the Fornasini-Marchesini first model has been formulated and in addition to that we have also recalled some useful results in Sect. 2. In Sect. 3 using the Lyapunov approach as was earlier applied in [18], a convex optimization problem has been formulated. In Sect. 4, an application of the presented technique has been demonstrated in two different examples. The previously reported approach [3]. Finally, In Sect. 5 some concluding remarks are given.

2 Problem Formulation

Consider a class of two-dimensional discrete systems described Fornasini-Marchasini's first Model by the shift delays.

$$\begin{aligned}\mathfrak{x}_{(i+1,j+1)} = (O_1 + O_1)\mathfrak{x}_{(i,j+1)} + (O_{1d} + \triangle O_{1d})\mathfrak{x}_{(i-d,j+1)} \\ +(O_2 + \triangle O_2)\,\mathfrak{x}_{(i+1,j)} + (O_{2k} + \triangle O_{2k})\mathfrak{x}_{(i+1,j-k)} \\ +(O_3 + \triangle O_3)\mathfrak{x}_{(i,j)} + (O_{3dk} + \triangle O_{3dk})\,\mathfrak{x}_{(i-d,j-k)} \\ +(P + \triangle \mathrm{P})\mathfrak{u}_{(i,j)}\end{aligned} \tag{1a}$$

$$O = [O_1\; O_{1d}\; O_2\; O_{2k}\; O_3\; O_{3dk}], \tag{1b}$$

where $\mathfrak{x}_{(i,j)}$ is a $n \times 1$ state vector, O_1,O_{1d},O_2, O_{2k} , O_3,O_{3dk} $\in$ $\mathrm{R}^{n\times n}$,$\mathfrak{u}_{(i,j)}$ *is an* $m \times 1$ input vector, $P \in \mathrm{R}^{n\times m}$. The matrices $\triangle O_{\mathfrak{p}}(\mathfrak{p} = 1,2,3)$ and $\triangle$ consider parameter uncertainties that are assumed to be of the form.

$$[\triangle O\; \triangle P] = \mathfrak{D} Q_{(i,j)}[\mathfrak{m}_1\; \mathfrak{m}_2], \tag{1c}$$

where,

$$\begin{aligned}&\triangle O = [\triangle O_1\; \triangle O_{1d}\; \triangle O_2\; \triangle O_{2k}\; \triangle O_3\; \triangle O_{3dk}],\\ &\mathfrak{m}_1 = [\mathfrak{m}_{11}\; \mathfrak{m}_{11d}\; \mathfrak{m}_{12}\; \mathfrak{m}_{12k}\; \mathfrak{m}_{13}\; \mathfrak{m}_{13dk}]\;\; \mathfrak{m}_2 = [\mathfrak{m}_{21}\; \mathfrak{m}_{22}\; \mathfrak{m}_{23}\;],\end{aligned}$$

In the above, $\mathfrak{D} \in \mathrm{R}^{\mathrm{n}\times\mathrm{p}}$, $\mathfrak{m}_{11}$, $\mathfrak{m}_{11d}$, $\mathfrak{m}_{12}$, $\mathfrak{m}_{12k}$, $\mathfrak{m}_{13}$, $\mathfrak{m}_{13dk} \in \mathrm{R}^{q\times n}$ and $\mathfrak{m}_2 \in \mathrm{R}^{q\times m}$ along with known structural matrices of uncertainty and $Q_{(i,j)} \in \mathrm{R}^{p\times q}$ is the unknown matrix, which represents parameter uncertainty and satisfies

$$Q^T_{(i,j)}Q_{(i,j)} \leq\; \Upsilon \text{ (Or equivalent } \left\|Q_{(i,j)}\right\| \leq 1). \tag{1d}$$

there exist two integers that are positive $\mathfrak{d}_1$ and $\mathfrak{d}_2$

$$\begin{aligned}&\mathfrak{x}_{(i,j)} = 0\;\; \forall\; j \geq \mathfrak{d}_1, -d \leq\; i \leq 0, \mathfrak{x}_{(i,j)} = \varphi_{ij}, 0 \leq j < \mathfrak{d}_1, -d \leq i \leq 0,\\ &\mathfrak{x}_{(i,j)} = 0\;\;\; \forall\;\; i \geq \mathfrak{d}_{2,} - k \leq i \leq 0, \mathfrak{x}_{(i,j)} = \varphi_{ij}, 0 \leq i < \mathfrak{d}_2, -k \leq i \leq 0,\end{aligned} \tag{1e}$$

In the future, the condition is assumed to be arbitrary, but it belongs to a set

$$\begin{aligned}\Omega = \{\mathfrak{x}_{(i,j)} \epsilon\, \mathrm{R}^n : (i,j) = \mathfrak{M}\mathcal{N}_j, \mathcal{N}_j^T\mathcal{N}_j\; \leq \Upsilon\, , -d \leq \mathbf{\Upsilon} \leq 0, 0 \leq j \leq \mathfrak{d}_1\},\\ \cup\, \{\mathfrak{x}_{(i,j)} \epsilon\, \mathrm{R}^n : (i,j) = \mathfrak{M}\mathcal{N}_i, \mathcal{N}_i^T\mathcal{N}_i\; \leq \Upsilon, -k \leq \Upsilon \leq 0, 0 \leq i \leq \mathfrak{d}_2\},\end{aligned} \tag{1f}$$

The system cost function is represented by

$$\omega = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\left[u_{(i,j)}^{T}\Xi u_{(i,j)}\right] + \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\mathfrak{T}_{ij}^{T}W_2\mathfrak{T}_{ij} \tag{2a}$$

$$0 < \Xi_k = \Xi_k^T \in R^{m\times m} \ (k = 1,2,3) \tag{2b}$$

$$\mathfrak{T}_{ij} = \left[x_{(i,j+1)}^T x_{(i-d,j+1)}^T x_{(i+1,j)}^T x_{(i+1,j-K)}^T x_{(i,j)}^T x_{(i-d,j-K)}^T\right]^T \tag{2c}$$

$$W_1 = \begin{bmatrix} \Omega_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \Omega_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \Omega_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & \Omega_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & \Omega_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & \Omega_6 \end{bmatrix}, \tag{2d}$$

where $0 < \Omega_t = \Omega_t^T \in R^{n\times n}$ $(t = 1,2,3,4,5,6)$.

We can design a memory controller

$$\begin{aligned} u_{(i,j)} &= \Psi_1 x_{(i,j+1)} + \Psi_2 x_{(i-d,j+1)} + \Psi_3 x_{(i+1,j)} \\ &\quad + \Psi_4 x_{(i+1,j-k)} + \Psi_5 x_{(i,j)} + \Psi_6 x_{(i-d,j-k)} \end{aligned} \tag{3}$$

the cost function of the closed-loop system is represented by

$$\begin{aligned} x_{(i+1,j+1)} &= [(O_1 + \Delta O_1) + (P + \Delta P)\Psi_1]x_{(i,j+1)} \\ &\quad + [(O_{1d} + \Delta O_{1d}) + (P + \Delta P)\Psi_2]\, x_{(i-d,j+1)} \\ &\quad + [(O_2 + \Delta O_2) + (P + \Delta P)\Psi_3]x_{(i+1,j)} \\ &\quad + [(O_{2k} + \Delta O_{2k}) + (P + \Delta P)\Psi_4]x_{(i+1,j-k)} \\ &\quad + [(O_3 + \Delta O_3) + (P + \Delta P)\Psi_5]x_{(i,j)} \\ &\quad + [(O_{3dk} + \Delta O_{3dk}) + (P + \Delta P)\Psi_6]x_{(i-d,j-k)} \end{aligned} \tag{4}$$

Moreover, cost function

$$\omega = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\mathfrak{T}_{ij}^{T}W_2\mathfrak{T}_{ij}.$$

where

$$W_2 = W_1 + \Psi\Xi\Psi. \tag{5a}$$

$$\text{Ᵽ} = [\text{Ᵽ}_1 \text{Ᵽ}_2 \text{Ᵽ}_3 \text{Ᵽ}_4 \text{Ᵽ}_5 \text{Ᵽ}_6] \tag{5b}$$

Now $\omega \leq \omega^*$, where ω^*, is some pre-set constant,

Lemma 1. [14] if system (1a–1f) is asymptotically stable and only

$$de\,t\{\Upsilon - \text{Ⱬ}_1\text{Ⱬ}_2(O_3 + \mathfrak{D}Q_{(i,j)}\text{ⲙ}_{13}) - \text{Ⱬ}_2(O_2 + \mathfrak{D}Q_{(i,j)}\text{ⲙ}_{12}) - \text{Ⱬ}_1(O_1 + \mathfrak{D}Q_{(i,j)}\text{ⲙ}_{11}) - \text{Ⱬ}_1^{d+1}\text{Ⱬ}_2^{k+1}(O_{3dk} + \mathfrak{D}Q_{(i,j)}\text{ⲙ}_{13dk}) - \text{Ⱬ}_2^{k+1}(O_{2k} + \mathfrak{D}Q_{(i,j)}\text{ⲙ}_{12k}) - \text{Ⱬ}_1^{d+1}(O_{1d} + \mathfrak{D}Q_{(i,j)}\text{ⲙ}_{11d})\} \neq 0. \tag{6}$$

for all $\text{Ⱬ}_1, \text{Ⱬ}_2, Q_{(i,j)} \in \text{Ч}^2$,
where $\text{Ч}^2 = \{(\text{Ⱬ}_1, \text{Ⱬ}_2, Q_{(i,j)}): |\text{Ⱬ}_1| \leq 1, |\text{Ⱬ}_1| \leq 1, \|Q_{(i,j)}\| \leq 1\}$..

Remarks 1. With $\triangle O = \mathfrak{D}Q_{(i,j)}[\text{ⲙ}_{11}\ \text{ⲙ}_{11d}\ \text{ⲙ}_{12}\ \text{ⲙ}_{12}\ \ \text{ⲙ}_{13}\ \text{ⲙ}_{13dk}] = 0$ recognized as an asymptotic condition, one can easily arrive at the following lemma.

Lemma 2 [18] **.** A positive definite value $\text{Ʒ} > 0$ with symmetric matrices. The closed-loop system is asymptotic stable such that

$$\text{Ꞓ} = [\mathrm{G}_{\triangle 1} + \mathrm{H}_{\triangle}\text{Ᵽ}_1\ \ \mathrm{G}_{\triangle 1d} + \mathrm{H}_{\triangle}\text{Ᵽ}_2\ \ \mathrm{G}_{\triangle 2} + \mathrm{H}_{\triangle}\text{Ᵽ}_3\ \ \mathrm{G}_{\triangle 2k} + \mathrm{H}_{\triangle}\text{Ᵽ}_4\ \ \mathrm{G}_{\triangle 3} + \mathrm{H}_{\triangle}\text{Ᵽ}_5 + \mathrm{G}_{\triangle 3dk} + \mathrm{H}_{\triangle}\text{Ᵽ}_6]^{\mathrm{T}} \times \text{Ʒ}[\mathrm{G}_{\triangle 1} + \mathrm{H}_{\triangle}\text{Ᵽ}_1\ \ \mathrm{G}_{\triangle 1d} + \mathrm{H}_{\triangle}\text{Ᵽ}_2\ \ \mathrm{G}_{\triangle 2} + \mathrm{H}_{\triangle}\text{Ᵽ}_3\ \ \mathrm{G}_{\triangle 2k} + \mathrm{H}_{\triangle}\text{Ᵽ}_4\ \ \mathrm{G}_{\triangle 3} + \mathrm{H}_{\triangle}\text{Ᵽ}_5\ + \mathrm{G}_{\triangle 3dk} + \mathrm{H}_{\triangle}\text{Ᵽ}_6]$$

$$-\begin{bmatrix} \text{Ʒ}\text{-}\text{Ʒ}_1\text{-}\text{Ʒ}_2\text{-}\text{Ʒ}_3\text{-}\ \text{Ʒ}_4\text{-}\ \text{Ʒ}_5 & 0 & 0 & 0 & 0 & 0 \\ 0 & \text{Ʒ}_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \text{Ʒ}_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \text{Ʒ}_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & \text{Ʒ}_4 & 0 \\ 0 & 0 & 0 & 0 & 0 & \text{Ʒ}_5 \end{bmatrix} < 0\,. \tag{7}$$

for all $\|Q_{(i,j)}\| \leq 1$
where

$$O_{\Delta 1} = O_1 + \triangle O_1, O_{\Delta 1d} = O_{1d} + \triangle O_{1d}, O_{\Delta 2} = O_2 + \triangle O_2,$$

$$O_{\Delta 2k} = O_{2k} + \triangle O_{2k}, O_{\Delta 3} = O_3 + \triangle O_3, O_{\Delta 3dk} = O_{3dk} + \triangle O_{3d}\ ,$$

$$P_{\Delta} = P + \triangle P$$

The following definition is based on the above lemma.

$$\text{Ꞓ} + \mathrm{W}_2 < 0 \text{ for all } \|Q_{(i,j)}\| \leq 1 \tag{8}$$

Lemma 3 [18] **.** Let $\mathfrak{D} \in \mathrm{R}^{n\times p}$, $\text{ⲙ} \in \mathrm{R}^{q\times n}$, $\boldsymbol{\mathcal{G}} \in \mathrm{R}^{n\times n}$ and $\mathfrak{B}_t = \mathfrak{B}_t^T \in \mathrm{R}^{n\times n}$ be given matrices, and a positive definite matrix exists Ʒsuch that

$$[\boldsymbol{\mathcal{G}} + \mathfrak{D}\,\boldsymbol{Q}\text{ⲙ}]^{\mathrm{T}}\,\text{Ʒ}[\boldsymbol{\mathcal{G}} + \mathfrak{D}\,\boldsymbol{Q}\text{ⲙ}] + \mathfrak{B}_{\mathrm{t}} < 0 \tag{9}$$

for all satisfying $Q^T Q \leq \Upsilon$ a scalar $\varepsilon > 0$ such that

$$\begin{bmatrix} -\mho^{-1} + \varepsilon \mathfrak{D}\mathfrak{D}^T & O \\ O^T & \varepsilon^{-1} \pitchfork^T \pitchfork + \mathfrak{B}_t \end{bmatrix} < 0. \tag{10}$$

3 Main Result

In this section, we are trying to perform quadratic stability and cost function.

Lemma 4 [18] **.** If there exists matrices $\mho \succ 0$ is satisfying (8) holds then the system is quadratically stable and the closed-loop cost function satisfies the bound.

$$\begin{aligned} \omega \leq \{ & (\eth_1 - 1)\, \lambda_{max}\, (\mathcal{M}^T(\mho - \mho_1 - \mho_2 - \mho_3 - \mho_4 - \mho_5)\mathcal{M}) \\ & + (\eth_1 - 1)(d + 1)\, \lambda_{max}\, (\mathcal{M}^T \mho_1 \mathcal{M}) + (\eth_2 - 1)\, \lambda_{max}\, (\mathcal{M}^T \mho_2 \mathcal{M}) \\ & + (\eth_1 - 1)(k + 1)\, \lambda_{max}\, (\mathcal{M}^T \mho_3 \mathcal{M}) + (\eth_1 + \eth_2 - 1)\, \lambda_{max}\, (\mathcal{M}^T \mho_4 \mathcal{M}) \\ & + [(1 + k)(1 + d) - (\eth_1 - 1)(\eth_2 - 1)]\, \lambda_{max}\, (\mathcal{M}^T \mho_5 \mathcal{M}) \} \end{aligned} \tag{11}$$

all parametric uncertainties that are admissible.

Proof 1: Defines proof and Lemma 2.

In quadratic 2-D Lyapunov function represented by

$$V\left(\chi_{(i,j)}\right) = \chi^T_{(i,j)} \mho \chi_{(i,j)}) \tag{12}$$

Let $\triangle V\left(\chi_{(i,j)}\right)$ be defined as

$$\begin{aligned} \triangle V\left(\chi_{(i,j)}\right) = {} & \chi^T_{(i+1,j+1)} \mho \chi_{(i+1,j+1)} \\ & - \chi^T_{(i,j+1)} (\mho - \mho_1 - \mho_2 - \mho_3 - \mho_4 - \mho_5) \chi_{(i,j+1)} \\ & - \chi^T_{(i-d,j+1)} \mho_{\mathbf{1}} \chi_{(i-d,j+1)} - \chi^T_{(i+1,j)} \mho_{\mathbf{2}} \chi_{(i+1,j)} \\ & - \chi^T_{(i+1,j-K)} \mho_{\mathbf{3}} \chi_{(i+1,j-k)} - \chi^T_{(i,j)} \mho_{\mathbf{4}} \chi_{(i,j)} \\ & - \chi^T_{(i-d,j-K)} \mho_{\mathbf{5}} \chi_{(i-d,j-k)} \end{aligned} \tag{13}$$

Adding the Eq. (13) both sides

$$
\begin{aligned}
\omega = -\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \triangle V\left(\mathfrak{x}_{(i,j)}\right) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} & [\mathfrak{x}^T_{(i,j+1)}(\mathfrak{V} - \mathfrak{V}_1 - \mathfrak{V}_2 - \mathfrak{V}_3 - \mathfrak{V}_4 - \mathfrak{V}_5)\mathfrak{x}_{(i,j+1)} \\
& -\mathfrak{x}^T_{(i+1,j+1)}(\mathfrak{V} - \mathfrak{V}_1 - \mathfrak{V}_2 - \mathfrak{V}_3 - \mathfrak{V}_4 - \mathfrak{V}_5)\mathfrak{x}_{(i+1,j+1)} \\
& +\mathfrak{x}^T_{(i-d,j+1)}\mathfrak{V}_1\mathfrak{x}_{(i-d,j+1)} + \mathfrak{x}^T_{(i+1,j)}\mathfrak{V}_2\mathfrak{x}_{(i+1,j)} \\
& -\mathfrak{x}^T_{(i+1,j+1)}\mathfrak{V}_2\mathfrak{x}_{(i+1,j+1)} + \mathfrak{x}^T_{(i+1,j-K)}\mathfrak{V}_3\mathfrak{x}_{(i+1,j+1)} \\
& -\mathfrak{x}^T_{(i+1,j+1)}\mathfrak{V}_3\mathfrak{x}_{(i+1,j+1)} + \mathfrak{x}^T_{(i,j)}\mathfrak{V}_4\mathfrak{x}_{(i,j)} \\
& -\mathfrak{x}^T_{(i+1,j+1)}\mathfrak{V}_4\mathfrak{x}_{(i+1,j+1)} \\
& +\mathfrak{x}^T_{(i-d,j-K)}\mathfrak{V}_5\mathfrak{x}_{(i-d,j-k)} - \mathfrak{x}^T_{(i+1,j+1)}\mathfrak{V}_5\mathfrak{x}_{(i+1,j+1)}.
\end{aligned} \tag{14}
$$

Applying the initial condition (1f)

$$
\begin{aligned}
\omega \le & \sum_{j=1}^{\mathfrak{d}_1-1}[\mathfrak{x}^T_{(0,j)}(\mathfrak{V} - \mathfrak{V}_1 - \mathfrak{V}_2 - \mathfrak{V}_3 - \mathfrak{V}_4 - \mathfrak{V}_5)\mathfrak{x}_{(0,j)}] \\
& + \sum_{j=1}^{\mathfrak{d}_1-1}\sum_{i=-d}^{0}[\mathfrak{x}^T_{(i,j)}\mathfrak{V}_1\mathfrak{x}_{(i,j)}] + \sum_{i=1}^{\mathfrak{d}_2-1}[\mathfrak{x}^T_{(i,0)}\mathfrak{V}_2\mathfrak{x}_{(i,0)}] \\
& + \sum_{i=1}^{\mathfrak{d}_2-1}\sum_{j=-k}^{0}[\mathfrak{x}^T_{(i,j)}\mathfrak{V}_3\mathfrak{x}_{(i,j)} + \sum_{i=0}^{\mathfrak{d}_2-1}\sum_{j=0}^{\mathfrak{d}_1-1}[\mathfrak{x}^T_{(i,j)}\mathfrak{V}_4\mathfrak{x}_{(i,j)}] \\
& - \sum_{i=1}^{\mathfrak{d}_2-1}\sum_{j=1}^{\mathfrak{d}_1-1}[\mathfrak{x}^T_{(i,j)}\mathfrak{V}_4\mathfrak{x}_{(i,j)}] + \sum_{i=-d}^{0}\sum_{j=-k}^{0}[\mathfrak{x}^T_{(i,j)}\mathfrak{V}_5\mathfrak{x}_{(i,j)}] \\
& - \sum_{i=1}^{\mathfrak{d}_2-1}\sum_{j=1}^{\mathfrak{d}_1-1}[\mathfrak{x}^T_{(i,j)}\mathfrak{V}_5\mathfrak{x}_{(i,j)}].
\end{aligned} \tag{15}
$$

Theorem 1 [14] **.** there exists a positive scalar $\varepsilon > 0$, matrices $\text{Ч}_1, \text{Ч}_2, \text{Ч}_3, \text{Ч}_4, \text{Ч}_5, \text{Ч}_6$ and the following LMI is feasible:

$$\begin{bmatrix} \Gamma_1 & \Gamma_2 \\ * & \Gamma_3 \end{bmatrix} < 0, \tag{16}$$

where

$$\Gamma_1 = \begin{bmatrix} (-V+\varepsilon\mathfrak{D}\mathfrak{D}^T) & (O_1V+P\text{Ч}_1) & (O_{1d}V+P\text{Ч}_2) & (O_2V+P\text{Ч}_3) & (O_{2k}V+P\text{Ч}_4) & (O_3V+P\text{Ч}_5) & (O_{3dk}V+P\text{Ч}_6) \\ (O_1V+P\text{Ч}_1)^T & -V+\mathbb{V}_1+\mathbb{V}_2+\mathbb{V}_3+\mathbb{V}_4+\mathbb{V}_5 & 0 & 0 & 0 & 0 & 0 \\ (O_{1d}V+P\text{Ч}_2)^T & 0 & -\mathbb{V}_1 & 0 & 0 & 0 & 0 \\ (O_2V+P\text{Ч}_3)^T & 0 & 0 & -\mathbb{V}_2 & 0 & 0 & 0 \\ (O_{2k}V+P\text{Ч}_4)^T & 0 & 0 & 0 & -\mathbb{V}_3 & 0 & 0 \\ (O_3V+P\text{Ч}_5)^T & 0 & 0 & 0 & 0 & -\mathbb{V}_4 & 0 \\ (O_{3dk}V+P\text{Ч}_6)^T & 0 & 0 & 0 & 0 & 0 & -\mathbb{V}_5 \end{bmatrix}$$

$$\Gamma_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ V\overline{\pitchfork_{11}^T} & \text{Ч}_1^T\Xi_1^{1/2} & V\Omega_1^{1/2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ V\overline{\pitchfork_{11d}^T} & \text{Ч}_2^T\Xi_1^{1/2} & 0 & V\Omega_2^{1/2} & 0 & 0 & 0 & 0 & 0 & 0 \\ V\overline{\pitchfork_{12}^T} & 0 & 0 & 0 & \text{Ч}_3^T\Xi_2^{1/2} & V\Omega_3^{1/2} & 0 & 0 & 0 & 0 \\ V\overline{\pitchfork_{12k}^T} & 0 & 0 & 0 & \text{Ч}_4^T\Xi_2^{1/2} & 0 & V\Omega_4^{1/2} & 0 & 0 & 0 \\ V\overline{\pitchfork_{13}^T} & 0 & 0 & 0 & 0 & 0 & 0 & \text{Ч}_5^T\Xi_3^{1/2} & V\Omega_5^{1/2} & 0 \\ V\overline{\pitchfork_{13dk}^T} & 0 & 0 & 0 & 0 & 0 & 0 & \text{Ч}_6^T\Xi_3^{1/2} & 0 & V\Omega_6^{1/2} \end{bmatrix} \tag{17}$$

$\Gamma_3 = \{-\varepsilon\Upsilon, \Upsilon, \Upsilon, \Upsilon, \Upsilon, \Upsilon, \Upsilon, \Upsilon, \Upsilon, \Upsilon\}$,

$\overline{\pitchfork_{11}} = \pitchfork_{11} + \pitchfork_{21}\text{Ѱ}_1$, $\overline{\pitchfork_{11d}} = \pitchfork_{11d} + \pitchfork_{21}\text{Ѱ}_2$, $\overline{\pitchfork_{12}} = \pitchfork_{12} + \pitchfork_{22}\text{Ѱ}_3$,

$\overline{\pitchfork_{12k}} = \pitchfork_{12k} + \pitchfork_{22}\text{Ѱ}_4$, $\overline{\pitchfork_{13}} = \pitchfork_{13} + \pitchfork_{23}\text{Ѱ}_5$, $\overline{\pitchfork_{13dk}} = \pitchfork_{13dk} + \pitchfork_{23}\text{Ѱ}_6$

$V^{-1} = \mho$, $\mho^{-1} = V$, $V\mho_1V = \mathbb{V}_1$, $V\mho_2V = \mathbb{V}_2$, $V\mho_3V = \mathbb{V}_3$, $V\mho_4V = \mathbb{V}_4$, $V\mho_5V = \mathbb{V}_5$.

In this condition, stabilizing control law matrices are

$$\mathfrak{P}_1 = \text{Ч}_1 V^{-1}, \mathfrak{P}_2 = \text{Ч}_2 V^{-1}, \mathfrak{P}_3 = \text{Ч}_3 V^{-1}, \mathfrak{P}_4 = \text{Ч}_4 V^{-1}, \mathfrak{P}_5 = \text{Ч}_5 V^{-1}, \mathfrak{P}_6 = \text{Ч}_6 V^{-1} \tag{18}$$

Proof 2: Lemma 3 and (1a–1f), (5a and 5b), (8), (10) can be repositioned as

$$\begin{bmatrix}
-\mathfrak{Z}^{-1} + \varepsilon \mathfrak{D}\mathfrak{D}^T & (O_1 + P\text{Ч}_1) & (O_{1d} + P\text{Ч}_2) & (O_2 + P\text{Ч}_3) & (O_{2k} + P\text{Ч}_4) & (O_3 + P\text{Ч}_5) & (O_{3dk} + P\text{Ч}_6) \\
(O_1 + P\text{Ч}_1)^T & \varepsilon^{-1}\overline{m}_{11}^T\overline{m}_{11} + \mathfrak{P}_1^T \Xi_1 \mathfrak{P}_1 + \Omega_1 - \mathfrak{Z}_1 & \varepsilon^{-1}\overline{m}_{11}^T\overline{m}_{11d} + \mathfrak{P}_1^T \Xi_1 \mathfrak{P}_2 & \varepsilon^{-1}\overline{m}_{11}^T\overline{m}_{12} & \varepsilon^{-1}\overline{m}_{11}^T\overline{m}_{12k} & \varepsilon^{-1}\overline{m}_{11}^T\overline{m}_{13} & \varepsilon^{-1}\overline{m}_{11}^T\overline{m}_{13dk} \\
(O_{1d} + P\text{Ч}_2)^T & \varepsilon^{-1}\overline{m}_{11d}^T\overline{m}_{11} + \mathfrak{P}_2^T \Xi_1 \mathfrak{P}_1 & \varepsilon^{-1}\overline{m}_{11d}^T\overline{m}_{11d} + \mathfrak{P}_2^T \Xi_1 \mathfrak{P}_2 + \Omega_2 - \mathfrak{Z}_2 & \varepsilon^{-1}\overline{m}_{11d}^T\overline{m}_{12} & \varepsilon^{-1}\overline{m}_{11d}^T\overline{m}_{12k} & \varepsilon^{-1}\overline{m}_{11d}^T\overline{m}_{13} & \varepsilon^{-1}\overline{m}_{11d}^T\overline{m}_{13dk} \\
(O_2 + P\text{Ч}_3)^T & \varepsilon^{-1}\overline{m}_{12}^T\overline{m}_{11} & \varepsilon^{-1}\overline{m}_{12}^T\overline{m}_{11d} & \varepsilon^{-1}\overline{m}_{12}^T\overline{m}_{12} + \mathfrak{P}_3^T \Xi_2 \mathfrak{P}_3 + \Omega_3 - \mathfrak{Z}_3 & \varepsilon^{-1}\overline{m}_{12}^T\overline{m}_{12k} + \mathfrak{P}_3^T \Xi_2 \mathfrak{P}_4 & \varepsilon^{-1}\overline{m}_{12}^T\overline{m}_{13} & \varepsilon^{-1}\overline{m}_{12}^T\overline{m}_{13dk} \\
(O_{2k} + P\text{Ч}_4)^T & \varepsilon^{-1}\overline{m}_{12k}^T\overline{m}_{11} & \varepsilon^{-1}\overline{m}_{12k}^T\overline{m}_{11d} & \varepsilon^{-1}\overline{m}_{12}^T\overline{m}_{12} + \mathfrak{P}_4^T \Xi_2 \mathfrak{P}_3 & \varepsilon^{-1}\overline{m}_{12k}^T\overline{m}_{12k} + \mathfrak{P}_4^T \Xi_2 \mathfrak{P}_4 + \Omega_4 - \mathfrak{Z}_4 & \varepsilon^{-1}\overline{m}_{12k}^T\overline{m}_{13} & \varepsilon^{-1}\overline{m}_{12k}^T\overline{m}_{13dk} \\
(O_3 + P\text{Ч}_5)^T & \varepsilon^{-1}\overline{m}_{13}^T\overline{m}_{11} & \varepsilon^{-1}\overline{\mathfrak{E}}_{13}^T\overline{\mathfrak{E}}_{11d} & \varepsilon^{-1}\overline{m}_{13}^T\overline{m}_{12} & \varepsilon^{-1}\overline{m}_{13}^T\overline{m}_{12k} & \varepsilon^{-1}\overline{m}_{13}^T\overline{m}_{13} + \mathfrak{P}_5^T \Xi_3 \mathfrak{P}_5 + \Omega_5 - \mathfrak{Z}_5 & \varepsilon^{-1}\overline{m}_{13}^T\overline{m}_{13dk} + \mathfrak{P}_5^T \Xi_3 \mathfrak{P}_6 \\
(O_{3dk} + P\text{Ч}_6)^T & \varepsilon^{-1}\overline{m}_{13dk}^T\overline{m}_{11} & \varepsilon^{-1}\overline{m}_{13dk}^T\overline{m}_{11d} & \varepsilon^{-1}\overline{m}_{13dk}^T\overline{m}_{12} & \varepsilon^{-1}\overline{m}_{13dk}^T\overline{m}_{12k} & \varepsilon^{-1}\overline{m}_{13dk}^T\overline{m}_{13} + \mathfrak{P}_6^T \Xi_3 \mathfrak{P}_5 & \varepsilon^{-1}\overline{m}_{13dk}^T\overline{m}_{13dk} + \mathfrak{P}_6^T \Xi_3 \mathfrak{P}_6 + \Omega_6 - \mathfrak{Z}
\end{bmatrix} \tag{19}$$

$diag\{\Upsilon, \mathfrak{Z}^{-1}, \mathfrak{Z}^{-1}, \mathfrak{Z}^{-1}, \mathfrak{Z}^{-1}, \mathfrak{Z}^{-1}, \mathfrak{Z}^{-1}\}$ and applying Schur,
Moreover, to satisfy (16) it requires that

$$-(\mathbb{V}_1 + \mathbb{V}_2 + \mathbb{V}_3 + \mathbb{V}_4 + \mathbb{V}_5) \geq 0 \tag{20}$$

Equation (20) can be reframed as

$$V^{-1} \geq V^{-1} (\mathbb{V}_1 + \mathbb{V}_2 + \mathbb{V}_3 + \mathbb{V}_4 + \mathbb{V}_5) V^{-1} \tag{21}$$

we obtain from the above equation

$$\lambda_{max} (\mathrm{M}^T V^{-1} \mathrm{M}) \geq \lambda_{max} \left(\mathrm{M}^T V^{-1} (\mathbb{V}_1 + \mathbb{V}_2 + \mathbb{V}_3 + \mathbb{V}_4 + \mathbb{V}_5) V^{-1} \mathrm{M}\right). \tag{22}$$

Now this also means that

$$\lambda_{max} (\mathrm{M}^T \mathfrak{Z} \mathrm{M}) \geq \lambda_{max} (\mathrm{M}^T (\mathfrak{Z}_1 + \mathfrak{Z}_2 + \mathfrak{Z}_3 + \mathfrak{Z}_4 + \mathfrak{Z}_5) \mathrm{M}). \tag{23}$$

This leads to.

$$\begin{aligned}
&\lambda_{max}\,(M^T \Im M) \geq \lambda_{max}\,(M^T \Im_1 M), \lambda_{max}\,(M^T \Im M) \geq \lambda_{max}\,(M^T \Im_2 M),\\
&\lambda_{max}\,(M^T \Im M) \geq \lambda_{max}\,(M^T \Im_3 M), \lambda_{max}\,(M^T \Im M) \geq \lambda_{max}\,(M^T \Im_4 M),\\
&\lambda_{max}\,(M^T \Im M) \geq \lambda_{max}\,(M^T \Im_5 M).
\end{aligned} \tag{24}$$

Simply the bound of the cost function obtained from (11) as,

$$\omega \leq [(\eth_1 - 1) + (\eth_1 - 1)(d+1) + (\eth_2 - 1) + (\eth_2 - 1)(k+1) + (\eth_1 + \eth_2 - 1) + (1+k)\,(1+d) - (\eth_1 - 1)(\eth_2 - 1)]\ \lambda_{max}\,(M^T \Im M).$$

The proof of Theorem 1 has been completed.

Remarks 2
The above results are strictly LMI form, the matrix inequality (16) has linear variables $\varepsilon, \Psi_1, \Psi_2, \Psi_3, \Psi_4, \Psi_5, \Psi_6, V, V_1, V_2, V_3, V_4$ and V_5. Theorem 1 presented a technique for construct the guaranteed cost controllers for memory state feedback by (16) the solution is feasible. the method of selecting suitable controllers minimizing the guaranteed cost in (22).

Remarks 3
It is important to note that, Theorem 1 used to design two-dimensional discrete systems via memoryless state feedback with shift delayed guaranteed cost controller described by the Fornasini-Marchisini First model simply by putting $\wp_2 = \wp_4 = \wp_6 = 0$

Remarks 4
Further, if $(O_3 + \Delta O_3) = 0, (O_{3dk} + \Delta O_{3dk}) = 0, (P + \Delta P) = 0$ together with $\wp_5 = \wp_6 = 0$, coincides the result along with $\boldsymbol{\wp_2 = \wp_4 = \wp_6 = 0}$ Theorem 1 corresponds to the existing result of a two-dimensional discrete system, which is given in [20].

Theorem 2 [17] **.** Then the system robust asymptotically stable with the initial condition (1f) and the cost function (2a–2d) is, if the following optimization problem. minimize λ

$$s.t \begin{cases} (i). \qquad (16), \\ (ii). \begin{bmatrix} -\lambda\Upsilon & M^T \\ M & -V \end{bmatrix} < 0 \end{cases} \tag{25}$$

has feasible solution $\varepsilon > 0, \lambda > 0$ matrices $\Psi_1, \Psi_2, \Psi_3, \Psi_4, \Psi_5, \Psi_6$ and positive definite symmetric matrices $V, V_1, V_2, V_3, V_4, V_5$ formerly the control law $u_{(i,j)} = \wp x_{(i,j)}$. The optimal guaranteed cost control law, which ensures the minimization of guarantee the cost given by (22).

Proof 3: According to Theorem 1 the stabilizing control law (3) designed $\mathfrak{u}_{(i,j)} = \mathfrak{P}\mathfrak{x}_{(i,j)}$ in terms of any feasible solution. To achieve the optimum value of the upper bound of the guaranteed cost, the term λ_{max} $(\mathfrak{M}^T \mathfrak{Z} \mathfrak{M})$ in (24) is changed to λ_{max} $(\mathfrak{M}^T\, V^{-1}\, \mathfrak{M}) < \lambda\, \Upsilon$ which, implies the constraint (ii) in (24). The minimization of $(\varepsilon + \lambda)$ implies the minimization of the guaranteed cost in (24). The proof of Theorem 2 has become complete.

4 Illustrative Examples

In this section, some different examples of industrial processes that illustrate the effectiveness of our proposed method by Theorem 2.

Example I
We will consider a shift delay system (1a–1f) and cost function (2a–2d) and demonstrate the effectiveness of Theorem 2 with a specific industrial example.

$$
\begin{aligned}
&O_1 = \begin{bmatrix} 0 & -0.7500 \\ -1.0000 & 0 \end{bmatrix}, O_{1d} = \begin{bmatrix} 0 & 0 \\ 0 & -0.0533 \end{bmatrix} \\
&O_2 = \begin{bmatrix} 0.2700 & 0 \\ 0 & 0 \end{bmatrix}, O_{2k} = \begin{bmatrix} 0 & -0.0152 \\ 0 & 0 \end{bmatrix} \\
&O_3 = \begin{bmatrix} 0 & -0.001 \\ 0 & 0 \end{bmatrix}, O_{3dk} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\
&P = \begin{bmatrix} 0.5000 \\ 0.1000 \end{bmatrix}, \mathfrak{D} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \\
&\mathfrak{m}_{11} = [-0.050 \quad -0.007], \mathfrak{m}_{11} = [0.02 \quad 0.05], \\
&\mathfrak{m}_{12} = [-0.060 \quad -0.002], \mathfrak{m}_{12k} = [0.01 \quad 0.03], \\
&\mathfrak{m}_{13} = [0.120 \quad 0.0010], \mathfrak{m}_{13dk} = [0 \quad 0], \\
&\mathfrak{m}_2 = [0.0075], \mathfrak{M} = [0.1;0.06] \\
&\Xi_1 = [0.06], \Xi_2 = [0.01], \Xi_3 = [0.02], \\
&\Omega_1 = \text{diag}\{6\ 6\}, \Omega_2 = \text{diag}\{0.60\ 0.60\}, \\
&\Omega_3 = \text{diag}\{0.060\ 0.060\}, \Omega_4 = \text{diag}\{0.060\ 0.060\}, \\
&\Omega_5 = \text{diag}\{0\ 0\}, \Omega_6 = \text{diag}\{0\ 0\}, \\
&\text{d=2, k=1}
\end{aligned}
\tag{26}
$$

Using the MATLAB LMI toolbox [21], if feasible solution is given by

$$
\begin{aligned}
&V=\begin{bmatrix}1.0143 & 0.1751\\ 0.1751 & 2.1377\end{bmatrix}, Ѵ_1=\begin{bmatrix}0.0412 & 0.0300\\ 0.0300 & 0.0288\end{bmatrix}\\
&Ѵ_2=\begin{bmatrix}0.0355 & 0.0076\\ 0.0076 & 0.0185\end{bmatrix}, Ѵ_3=\begin{bmatrix}0.0042 & 0.0030\\ 0.0030 & 0.0256\end{bmatrix}\\
&Ѵ_4=\begin{bmatrix}0.0018 & 0.0005\\ 0.0005 & 0.0004\end{bmatrix}, Ѵ_5=\begin{bmatrix}0.0997 & 0.0465\\ 0.0465 & 0.3260\end{bmatrix}\\
&Ч_1=[0.0183 \quad 3.2308], Ч_2=[-0.0016 \quad -0.0189],\\
&Ч_3=[-0.5570 \quad -0.0962], Ч_4=[0.0054 \quad 0.0661],\\
&Ч_5=[0.0003 \quad 0.0043], Ч_6=[0 \quad 0],\\
&\varepsilon = 15.7493, ⅄=0.0107
\end{aligned}
\tag{27}
$$

According to Theorem 2, for this system, the optimal guaranteed cost controllers are.

$$
\begin{aligned}
&Ꝕ_1=[-0.1664 \quad 1.5250], Ꝕ_2=[0.000 \quad -0.0089],\\
&Ꝕ_3=[-0.5492 \quad 0.000], Ꝕ_4=[-0.000 \quad 0.0309],\\
&Ꝕ_5=[-0.0000 \quad 0.0020], Ꝕ_6=[0 \quad 0],
\end{aligned}
\tag{28}
$$

The corresponding cost function has the least upper bound

$$\omega^{*} = 0.5071 \tag{29}$$

Example II

We will consider a shift delay system (1a–1f) and cost function (2a–2d) and demonstrate the effectiveness of Theorem 2 with a specific industrial example.

$$
\begin{aligned}
&O_1 = \begin{bmatrix} 0.081 & 1 \\ 0 & -0.2200 \end{bmatrix}, O_{1d} = \begin{bmatrix} -0.0100 & 0 \\ 0 & 0.0300 \end{bmatrix} \\
&O_2 = \begin{bmatrix} 0.10000 & 1.0000 \\ 0 & 0.7000 \end{bmatrix}, O_{2k} = \begin{bmatrix} 0.0200 & 0 \\ 0 & -0.0200 \end{bmatrix} \\
&O_3 = \begin{bmatrix} 0 & -0.001 \\ 0 & 0 \end{bmatrix}, O_{3d} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\
&\mathrm{P} = \begin{bmatrix} 0.5000 \\ 0.1000 \end{bmatrix}, \mathfrak{D} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \text{Ɱ} = \begin{bmatrix} 0.0210 \\ 0.1600 \end{bmatrix}, \\
&\pitchfork_{11} = [-0.0500 \quad 0], \pitchfork_{11d} = [0.0100 \quad 0.0050], \\
&\pitchfork_{12} = [0 \quad 0.0050], \pitchfork_{12k} = [0.0030 \quad 0.0030], \\
&\pitchfork_{13} = [0.1000 \quad 0.1000], \pitchfork_{13dk} = [0 \quad 0], \pitchfork_{2} = [-2.0000], \\
&\varXi_1 = [0.0500], \varXi_2 = [0.0100], \varXi_3 = [0.0100], \\
&\Omega_1 = \mathrm{diag}\{0.0020 \quad 0.0020\}, \Omega_2 = \mathrm{diag}\{1.0000 \quad 1.0000\}, \\
&\Omega_3 = \mathrm{diag}\{0.0600\ 0.0600\}, \Omega_4 = \mathrm{diag}\{0.0400\ 0.0400\}, \\
&\Omega_5 = \mathrm{diag}\{0\ 0\}, \Omega_6 = \mathrm{diag}\{0\ 0\}, \\
&\mathrm{d} = 1, \mathrm{k} = 2
\end{aligned}
\tag{30}
$$

Using the MATLAB LMI toolbox[21], if feasible solution is given by

$$
\begin{aligned}
&\mathcal{V} = \begin{bmatrix} 22.9072 & -2.0584 \\ -2.0584 & 2.3343 \end{bmatrix}, \Psi_1 = \begin{bmatrix} 1.1216 & -0.1313 \\ -0.1313 & 0.0793 \end{bmatrix} \\
&\Psi_2 = \begin{bmatrix} 5.0361 & -1.7454 \\ -1.7454 & 1.5951 \end{bmatrix}, \Psi_3 = \begin{bmatrix} 2.3918 & -0.2478 \\ -0.2478 & 0.0879 \end{bmatrix} \\
&\Psi_4 = \begin{bmatrix} 0.0083 & 0.0004 \\ 0.0004 & 0.0000 \end{bmatrix}, \Psi_5 = \begin{bmatrix} 0.0052 & 0.0001 \\ 0.0001 & 0.0000 \end{bmatrix} \\
&\text{Ч}_1 = [-6.9163 \quad 1.1508], \text{Ч}_2 = [-0.1704 \quad -0.2222], \\
&\text{Ч}_3 = [3.4948 \quad -7.4609], \text{Ч}_4 = [-0.5521 \quad 0.1886], \\
&\text{Ч}_5 = [0.0372 \quad -0.0002], \text{Ч}_6 = [0 \quad 0], \\
&\varepsilon = 0.9032, \curlywedge = 0.0122
\end{aligned}
\tag{31}
$$

According to Theorem 2, for this system, the optimal guaranteed cost controllers are.

$$
\begin{aligned}
&\text{Ꝓ}_1 = [-0.2798 \quad 0.2463], \text{Ꝓ}_2 = [-0.0174 \quad -0.1105], \\
&\text{Ꝓ}_3 = [-0.1462 \quad -3.3251], \text{Ꝓ}_4 = [-0.0183 \quad 0.0647], \\
&\text{Ꝓ}_5 = [0.0018 \quad 0.0014], \text{Ꝓ}_6 = [0 \quad 0],
\end{aligned}
\tag{32}
$$

The corresponding cost function has the least upper bound

$$\omega^* = 0.0331 \tag{33}$$

5 Conclusions

In this paper, a technique has been presented to find an optimal guaranteed cost controller for an uncertain and shift delayed FM first model via memory state feedback. Under the LMI framework. For selecting the optimal guaranteed cost controllers, the problem has been developed into a convex optimization problem. The guaranteed cost controllers thus obtained minimizes the upper bound on the closed-loop cost function. Finally, the applicability of the proposed techniques has been demonstrated on two industrial processes.

References

1. Fornasini, E., Marchesini, G.: Doubly-indexed dynamical systems: state-space models and structural properties. Math. Syst. Theory **12**(1), 59–72 (1978). https://doi.org/10.1007/BF01776566
2. Dhawan, A., Kar, H.: LMI-based criterion for the robust guaranteed cost control of 2-D systems described by the Fornasini-Marchesini second model. Signal Process. **87**(3), 479–488 (2007). https://doi.org/10.1016/j.sigpro.2006.06.002
3. Tiwari, M., Dhawan, A.: Robust suboptimal guaranteed cost control for 2-D discrete systems described by Fornasini-Marchesini first model. J. Signal Inf. Process. **03**(02), 252–258 (2012). https://doi.org/10.4236/jsip.2012.32034
4. Tandon, A., Dhawan, A., Tiwari, M.: Optimal guaranteed cost control of uncertain 2-D discrete state-delayed systems described by the Roesser model via memory state feedback. Trans. Inst. Meas. Control **41**(1), 285–294 (2019). https://doi.org/10.1177/0142331218754623
5. Ye, S., Zou, Y., Wang, W., Yao, J.: Delay-dependent stability analysis for two-dimensional discrete systems with shift delays by the general models. In: 2008 10th Int. Conf. Control. Autom. Robot. Vision, ICARCV 2008, December, pp. 973–978 (2008). https://doi.org/10.1109/ICARCV.2008.4795650
6. Xu, J., Yu, L.: Delay-dependent H ∞ control for 2-D discrete state delay systems in the second FM model. Multidimension. Syst. Signal Process. **20**(4), 333–349 (2009). https://doi.org/10.1007/s11045-008-0074-z
7. Tandon, A., Dhawan, A.: An LMI approach to non-fragile robust optimal guaranteed cost control of uncertain 2-D discrete systems with both state and input delays. Trans. Inst. Meas. Control **40**(3), 785–804 (2018). https://doi.org/10.1177/0142331216667476
8. Dhawan, A.: Non-fragile controller design for 2-D discrete uncertain systems described by the Roesser model. J. Signal Inf. Process. **03**(02), 248–251 (2012). https://doi.org/10.4236/jsip.2012.32033
9. Dhawan, A., Kar, H.: An LMI approach to robust optimal guaranteed cost control of 2-D discrete systems described by the Roesser model. Signal Process. **90**(9), 2648–2654 (2010). https://doi.org/10.1016/j.sigpro.2010.03.008

10. Wu, L., Lam, J., Yao, X., Xiong, J.: Robust guaranteed cost control of discrete-time networked control systems. Optim. Control Appl. Methods **32**(1), 95–112 (2011). https://doi.org/10.1002/oca.932
11. Ye, D., Yang, G.H.: Reliable guaranteed cost control for linear state delayed systems with adaptive memory state feedback controllers. Asian J. Control **10**(6), 678–686 (2008). https://doi.org/10.1002/asjc.68
12. Wei, Z., Liu, X.: Finite-time state and output feedback guaranteed cost control of discrete-time stochastic systems. In: Proc. 32nd Chinese Control Decis. Conf. CCDC 2020, vol. 0, pp. 4513–4519 (2020). https://doi.org/10.1109/CCDC49329.2020.9164120
13. Negi, R., Kar, H., Purwar, S.: Stability analysis of 2D discrete linear system described by the Fornasini-Marchesini second model with actuator saturation. ISRN Comput. Math. **2012**, 1–9 (2012). https://doi.org/10.5402/2012/847178
14. Vidyarthi, A., Tiwari, M.: LMI approach to optimal guaranteed cost control for uncertain 2-D discrete shift delayed systems described, September 2017. https://doi.org/10.5281/zenodo.161066
15. Bolajraf, M.: LP Conditions for stability and stabilization of positive 2D discrete state-delayed roesser models. Int. J. Control. Autom. Syst. **16**(6), 2814–2821 (2018). https://doi.org/10.1007/s12555-017-0464-9
16. Hernández-Galván, B.L., López-Rentería, J.A., Aguirre-Hernández, B., Fernández-Anaya, G.: Robust stability in discrete control systems via linear controllers with single and delayed time. Math. Probl. Eng. **2018** (2018). https://doi.org/10.1155/2018/3674628
17. Ye, S., Wang, W., Zou, Y.: Robust guaranteed cost control for a class of two-dimensional discrete systems with shift-delays. Multidimens. Syst. Signal Process. **20**(3), 297–307 (2009). https://doi.org/10.1007/s11045-008-0063-2
18. Tiwari, M., Dhawan, A.: An LMI approach to optimal guaranteed cost control of uncertain 2-D discrete shift-delayed systems via memory state feedback. Circuits Syst. Signal Process. **31**(5), 1745–1764 (2012). https://doi.org/10.1007/s00034-012-9410-5
19. Wan, K.: Convergence analysis of iterative learning control for two classes of 2-D linear discrete Fornasini-Marchesini model. Complexity **2020** (2020). https://doi.org/10.1155/2020/6843730
20. Tiwari, M., Dhawan, A.: Optimal guaranteed cost control of uncertain 2D discrete systems with both shift delays and input delays via memory state feedback. Trans. Inst. Meas. Control **35**(4), 491–502 (2013). https://doi.org/10.1177/0142331212453375
21. Boyd, S., El Ghaoui, L., Feron, E., Balakrishnan, V.: Some Standard Problems Involving LMIs (1994)

Survey: Advancement in Materials and Trends in Various Fields of 3D Printing

Om Maheshwari, Anant Singhal, Vaibhav Pachaulee, Mayank Trehan, and Pushpendra Kumar Gupta

Abstract 3D printing technology, also referred to as rapid prototyping or additive manufacturing, creates physical objects from a CAD model by successive addition or layer by layer addition of materials. It is a fast-emerging technology and nowadays, widely used in the world. This fast-emerging technology is increasingly used for the mass as well as customized production in the field of agriculture, bio medical, automotive industry, bio-printing and electronics industry.

This paper presents a review on various advancement and development in 3D printing technology. The research papers considered for this survey were primarily focused on advancement of existing 3D printing techniques, thermo mechanical properties and cost effectiveness of materials used in 3D printing. Superior techniques of scanning and modeling the object to be printed were also reviewed. This paper also includes a survey regarding specific application, the best alternative material in terms of cost and durability and the technique which would be the most suitable for the specific requirement. Eventually this survey wraps up with a summary of the recent advancements in three-dimensional scanning, the associated technologies which are not yet fully implemented, but will be an inseparable part of the 3D Printing scenario in the near future.

Keywords 3D printing · 3D scanning · Bio-printing · Stretchable electronics

O. Maheshwari
Department of Electrical Engineering, Indian Institute of Technology Gandhinagar, Gandhinagar 382355, India

A. Singhal
Department of Electrical Engineering, Indian Institute of Technology Jodhpur, Jodhpur 342037, India

V. Pachaulee (✉)
Department of Physics, Indian Institute of Technology Bombay, Mumbai 400076, India
e-mail: vaibhavpachaulee1999@gmail.com

M. Trehan · P. K. Gupta
Department of Electronics and Communication Engineering, College of Technology Pantnagar, Pantnagar 263145, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_37

1 Introduction

3D printing finds its application in almost every sector; hence it becomes necessary to weigh different techniques and the materials used in 3D printing in terms of cost, speed of manufacturing and thermo-mechanical properties. Moreover, the pros and cons of materials and techniques vary with area of applications. Figure 1 shows the taxonomy of the review. A brief overview of these materials and techniques in various areas of applications of 3D printing are given below.

1.1 Medical Applications

Medical applications for 3D printing are expanding very fast and are expected to revolutionize the health care sector, printing complicated structural geometries such as the morphological architecture of the human bones, which cannot be produced by milling or any other conventional machining process alone. Also, printing craniomaxillofacial scaffolds to cure defects induced from trauma, manufacturing bone augmentation using alloplastic materials (titanium) [1, 2]. 3D printing allows for production of implants with controlled porosity, similar to natural bone.

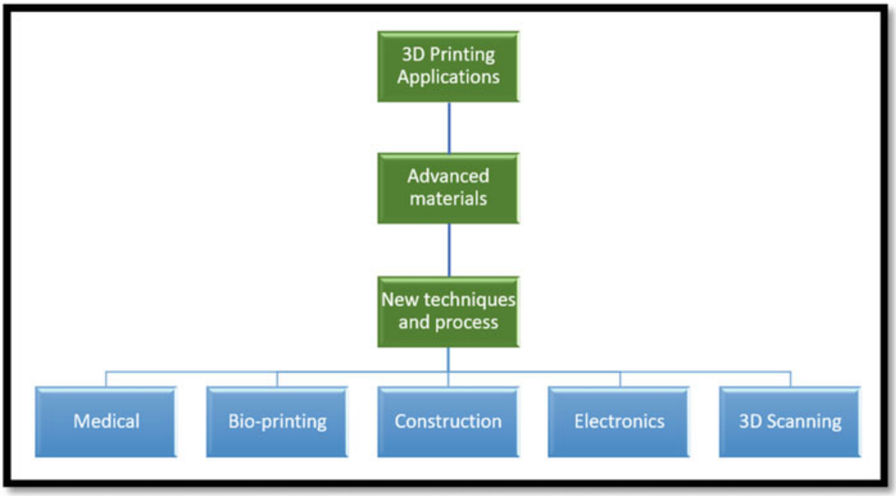


Fig. 1 Categorization of the applications reviewed in the survey

1.2 Bio-printing

Bio-printing is an umbrella term, used for various 3D printing technologies. This term is specifically used to produce human cells by bio-materials, organs mimicking cell dense tissue networks either individually or in a serially connected tandem like fashion [3–8]. Printing is executed by layer-by-layer deposition method.

1.3 Architectural and Construction Applications

3D printing assists in faster and more accurate construction of complex items. The process leads to minimum wastage and overall expense in carrying out the construction process [9, 10]. 3D printing assisted construction practices are more immune to harsh or dangerous environments which are not suitable for human workers. Construction in space is such an example where 3D printing proves more apt as compared to traditional construction procedure.

1.4 Electronic Components and Sensors

Using lead-based alloys having low melting point for 3D printing circuit boards can result in very high bonding strength between solder and PCB. The size of circuit boards can be reduced by about 34%. The number of required steps was also reduced as analyzed Kim Min-Saeng [11] et al. The new research in the electronics field of 3D printing for manufacturing stretchable electronics and sensors is extensive for the development of wearable electronics and soft robotics.

1.5 3D Scanning

3D scanning is a major advancement in modeling techniques of the prototype to be printed. Scanning of the object beforehand is especially beneficial in digitizing a physical object for mass production. The scanning process speeds up the prototyping design procedure and also allows the user to model with actual material on digital platform rather than build from scratch in computer software.

2 Literature Survey

2.1 Medical Applications

Majority of 3D printing applications in medical practices are in fields of dentistry and orthopedics. Printing of complex geometries is a wide area of 3D printing in medical applications as shown in Fig. 2.

Andrew Dawood [1] et al., analyzed the spectrum of techniques and materials used in 3D printing applications in dentistry. This research paper analyzed resins, photopolymers, metal alloys and thermoplastics. The associated techniques were also discussed and while resin proved to be most cost-effective material. Selective laser sintering (SLS) technique was suggested as the technique best suited to fabricate a finely detailed scaffold.

Dodziuk Helena [12] presented the generally "unheard" monopoly of 3D printed manufacturing practices in the industry of producing hearing aids. The paper reveals that more than ten million hearing aids were in circulation worldwide as of 2013. The technique of choice seems to be Stereolithography (SLA) as more than 95% of the production is executed using this technique.

Del Junco [13] et al., examined the implementation of 3D printing in manufacturing of surgical training models. The growing scarcity of healthy organ implants requires the surgeon to be very precise in the surgical procedure. A test surgery can be performed on a 3D printed part of the vascular structures of the patient. The

Fig. 2 3D applications in biomedical sciences [15]

patient specific model is manufactured by using silicone. These practices have been successfully applied in liver transplants and removal of renal malignancies.

Nicholas Herbert [2] et al. developed a prosthetic foot using simple and efficient 3D printing techniques. The patient treated by this method had a severe amputation but reported comfort and mobility after a few weeks of implant procedure. The foot had a titanium scaffold with silicone and hydroxyapatite providing flexibility and movement support.

Cacco Tommaso [14] et al., successfully implemented 3D printed full face snorkel masks and compared its performance with the standardized PPE kit equipment. The use of 3D printing in rapidly producing protective equipment and devices in times of emergency has proved to be an efficient strategy in quenching the ever-present demand of essential protective gear.

2.2 Bio-printing

Bio-printing refers more or less to print cell or tissue structures. It holds an upper hand over traditional methods to create 3D scaffolds.

Helena N Chia [3] et al., investigated the direct bio-printing technique which seems to witness a recent surge in application. The technique works by extruding a viscous liquid into a liquid medium of matching density. The nozzle deposits the extruded bio material in form of dots or strands to create the bespoke tissue structure. The extruded material used was gelatin and the liquid medium was a doubly ionized calcium ion reservoir. This process can be carried out at room temperature and leads to well-arranged distribution of cells.

Andrade [4] et al., investigated the robot assisted deposition of bioactive ceramic scaffolds. These ceramic structures supported the cell growth and were flexible as well as strong, mimicking the natural human cartilage structures. Calcium carbonate and Calcium phosphate dibasic were processed to create a gelatinous paste. The printing was done by a Computer Aided RoboCoaster. The scaffolds showed an observable compressive strength of order 2–12 MPa which is comparable to human bones. Rapid growth in cell formation over the scaffold was noticed on cyclically applying shear stresses to the scaffold.

Narayanan Lokesh Karthik [5] et al., examined the bio-printing of a bio ink consisting the three main components viz. poly lactic acid (PLA) nano-fiber, Alginate hydrogel. A bone structure resembling human knee was printed and cell growth on the scaffold was observed for a period of 8 weeks. Nano fibers increased cell proliferation by 28.5% by the end of first week.

Koch Loather [6] et al., worked on laser printing of skin cells and human stem cells. The technique used is Laser Induced Forward Transfer (LIFT). The printed bio structures of fibroblasts and stem cells were studied after being processed through LIFT process, the transfer process has about 90% success rate, where a successful transfer indicating that the lifespan of cell was not affected after the LIFT process.

Pallab Datta [7] et al., presented astonishing insights into the development of carcinoma microenvironments. The printing of the three dimensional dynamically evolving microenvironment of cancer affected tissue aids in developing remedial measures for the seemingly incurable ailment of cancer. Fabrication of ‘cancer on a chip’ environments like that of glioblastoma which targets the spine and the associated nervous system help quickly analyze the drug response of such conditions if they unfortunately happen in a human system.

McCormack Andrew [8] et al., discuss the leaps that tissue fabrication by 3D printing can make if tissue printing is directly carried out in suspension baths of suitable medium compatible with the type of tissue being printed. The technique is compared with the traditional bio printing carried out in the air on the printing surface.

2.3 Architectural and Construction Applications

3D Printing has paved way in the building and construction industry and has brought the ease in the construction process by using existing construction framework and techniques along with the adoption of new 3D printing process. Figure 3 lists some of the advantages of 3D printed construction methods compared to conventional methods. In 2019, Shanghai opened the longest 3D printed bridge in the world [16]. 3D printing in construction reduces the operating costs, and makes the operations time, labor, and resource efficient. Concrete material to be used for printing has to be flowable with low plasticity, viscosity and moderate yield stress for the pumpability. At the same time the material must exhibit plastic viscosity and high yield stress to cease the flow immediately after extrusion from the printer for the better layer by layer buildability.

Gibbons Gregory [17] et al., investigated the RPHC (Rapid-Hardening Portland Cement) for 3D printability using additive layer manufacturing. A printed result has properties, comparable to Plaster of Paris based components. Water curing of the structures reduced porosity and minimized distinction between the layers.

Malaeb Ziena [18] et al., studied and optimized a mix for 3D concrete printing which consisted of cement 125 g, sands 80 g, fine aggregates 160 g with w/c ratio 0.39. For optimum results, accelerator 1 ml and retarder 0.625 ml were added to the mix. Approximate compressive strength of 42 MPa was observed the studied the mixture.

Xia Ming and Sanjayan Jay [19] demonstrated a low waste, environment friendly powder-based 3D printing material. Slag based geo-polymer with silicate activator and fine sand along with powder-based polymer. The results suggested that geo-polymer-based material could replace conventional powder-based 3D printing materials. With post processing, compressive strength of about 16 MPa was observed.

Venkatesh Naidu Nerella and Viktor Mechtcherine [20] carried out work on printable mixture found suitable for most concrete 3D printing applications, referred as 3M3. It is composed of cement, fly ash, micro silica suspensions, sand 0.06–0.2,

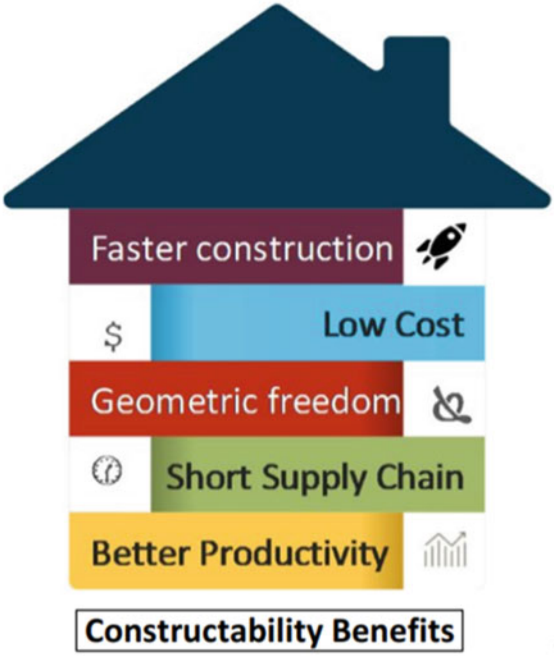


Fig. 3 Advantages of 3D printing in construction applications [9]

sand 0–1, sand 0–2, water produce heavy duty construction quality prototypes in a reasonably short span of time.

Integration of steel and fiber reinforcement along with cement is under development. A blend of Portland cement with basalt fibers and glass has been studied by Hambach M [21] et al. Esoda Agusti-Juan FM [22] et al., researched upon frameworks made of metal wires are used as permanent support and a mesh mold for concrete using digital fabrication.

Yifan Pan [10] et al., reviewed the cost benefits of 3D printing. Adopting 3D printing in construction industry reportedly reduces production costs up to 60%. The energy required constitutes only 0.2% of total costs. Usage of recycled materials in additive manufacturing reduces material costs, which is the major part of total costs. Using additive manufacturing for unique constructions is highly cost effective compared to traditional construction practices.

2.4 Electronic Components and Sensors

Complex 3D geometries can be created by combining electronically functional materials with additive manufacturing and direct-write, which can be used to create multilayer circuit boards, 3D antennas and embedded electronics.

Lu Yanfeng [23] et al., proposed a direct-write/cure and projection micro stereo lithography technique for designing 3D structural electronics. A dispersion of conductive photopolymers with carbon nano-tubes in photopolymer solution was used to make embedded conductive paths or wires. This process allowed the introduction of conductive paths over irregular and 3D structures rather than planes in conventional PCBs, thus allowed layering of more complex conductive paths in the electronic components.

Flowers Patrick [24] et al., examined dual material fused filament to fabricate electronic components using conductive thermoplastic filaments. Copper, Graphene, and Carbon Black were used as the conductive filaments under study. Their resistivity was found to be 0.014, 0.78, and 12 Ω/cm. It can be deduced from the result that resistors with varying values up to 3 orders of magnitude can be manufactured with these materials. It was observed that the copper-based filament was 500 times more flexible than the counterparts with little change in resistivity, whereas the Graphene and Carbon filaments were brittle and could fracture with ease.

The work demonstrated the ability of 3D printing electronic circuits and components with fused filaments at lower cost, higher accuracy and high efficiency, with printing on demand but with no performance loss as compared to conventional PCBs.

Stretchable electronics is a new field of consumer electronics supporting the wearable electronics and soft robotics. Stretchable 3D printed skins can be mounted onto irregular surfaces using flexible and stretchable electronic materials.

Muth Joseph [25] et al., presented a new technique of embedded 3D printing to fabricate strain sensors for stretchable electronics. The method uses visco-elastic ink, which acts as a resistive element for sensing. Filler and reservoir fluid are used as the monolithic part. Carbon black particles suspended in silicone oil was used in the research as the ink for sensing element. The filler fluid is a mixture of Ecoflex and Platinum Silicone cure retarder with silicone thinner.

Vatani Morteza [26] et al., presented a hybrid process of Direct Print/Cure and Projection based Stereo-lithography to fabricate stretchable tactile sensor. Photopolymer resin was used as stretchable elastic body of the sensor and piezo-resistive sensing element based on multiwall carbon nano-tubes is suspended in the resin. The materials used were photo curable and stretchable. The proposed sensors along with suitable signal processing algorithms were able to identify the location of the applied forces.

Soft robotics is an emerging field of nature inspired automation which uses non-conventional modes of actuation with flexible materials. Ela Sachyani Keneth [27] et al., reviewed the 3D printing application in soft robotics. Hydrogels are a class of soft, stretchable polymers which are highly applicable in soft robotics. Hydrogels can actuate signals in response to aqueous environments. They are water compatible. Yuting Dong [28] et al., studied the various properties and stimuli response of hydrogels to be implemented in soft robotics. Alginate, PNIPAm, Agarose are some commonly used hydrogels.

2.5 3D Scanning

Low-cost 3D scanners can be embedded into the 3D printers which enable us to scan the models and create the 3D catalogue of the objects without having to create them from scratch in the modeling software. 3D scanners speed up the design procedure and enable easy mass production. 3D scanners have capability of acquiring self-registered and high-resolution shape and color data. The scanners produce good result even on a complex test bed.

Claudio Rocchini [29] et al., had constructed and demonstrated a low-cost 3D scanner in their research. In this model, a projector was used to project a light beam with different patterns (triangle) known as structured light on surface of the test object. Cameras mounted on two sides simultaneously scan the object. The modeling is done by the detection of the structured light pattern on the object. The distortions in the light patterns were sensed to mark the curves in the object. The acquired image and video data is processed for shape and color reconstruction.

Kazunori Ohno [30] et al., demonstrated a 3D scanner using LASER scanning and mapping and created a map using a scan matching algorithm and odometry data (to use data from sensors and estimate change in position over time), which led to the development of the TK scanner that measures uniform and dense 3D shape which consists of Pan Tilt base and 2D LRF which rotates in yaw angle and pitch angle. Thus, it takes shorter duration to do one complete scan and the scanning density could be increased or decreased by changing the number of points scanned at a time.

Christoph Schmalz [31] et al., demonstrated 3D scanning using structured light in the endoscopy The system consists of a slide projector and a catadioptric camera. The light source is connected with a glass fiber and a spherical mirror is used for a wide field of view. The structured light consists of 15 rings with each ring having a different color. The wide-angle lens catadioptric camera finds the spacing to the side of the endoscope. Another camera that has a narrow angle lens looks for the area directly in front of the camera and provides guidance instructions. A bright Illumination for the enhancement of surface visuals is provided by two micro-LEDs in front of the scanner then all the scan points and color information is collected and calibrated to create the model using pinhole and pinhole mirror tube model.

Jeremy Straub and Scott Kerlin [32] developed and demonstrated a large-scale instant 3D Scanner that uses 50 Raspberry Pi cameras, 3 Adjustable power supplies, 50 8 GB SD cards, 2 48-port Ethernet switches, 1 File Server, 1 User console workstation and with a geometric hexagon frame comprising of PVC pipes. Each camera set at the points of each layer of hexagon, where each layer collects images from

Fig.4 **a** FreeScan X7 **b** HandySCAN 700

Table 1 Comparison table of HandySCAN 700 and FreeScan X7 [33]

	Manufacturer's data		Quality parameter according to VDI/VDE 2634		
	Range [m]	Accuracy [mm]	PS [mm]	PF [mm]	SD [mm]
Handy SCAN 700	0.30–4.00	0.02 + 0.06/m	0.20	1.08	0.07
FreeScan X7	0.10–8.00	0.02 + 0.06/m	0.10	1.01	0.23

different perspective. It integrates them together to finally produce a 3D design output with the help of Autodesk ReCap 360 or Agisoft's PhotoScan.

In the recent years, a plethora of professional 3D scanning devices are available but after thorough analysis only two of the devices provided some astonishing results. The first handheld device is FreeScan X7 (shown in Fig. 4(a)), the scanning hardware is directly connected to a computing device called Shining3D and a calibration board) and the second is HandySCAN 700 (shown in Fig. 4(b)), the working principle of both these scanners is quite similar which is based on the photogrammetric stereo system [33]. A brief comparison of accuracy between HandySCAN 700 and FreeScan X7 is shown below in Table 1 where PS is Probing Error Size, PF is Probing Error Shape and SD is Sphere-spacing Error.

3 Results and Discussion

Analysis of aforementioned research papers highlights a few points in each of the reviewed sectors. These are listed in the paragraphs:

In Medical applications of 3D printing, the use of technique and materials tend to be highly objective. The major usage of 3D techniques is in Orthodontics and

Dentistry. For orthodontic applications light cured resin seems to be a cost-effective alternative, the scaffolds can be manufactured relatively faster [1]. Metal fused alloys prove to be stronger and more bio compatible but are not practical cost wise.

In Bio-printing, conventional practices include the uses of hydroxyapatite and hydrogel infused gelatin like bio inks. The direct write printing technique seems to be an effective method of handling such type of material. The LIFT (Laser Induced Forward Transfer) technique shows considerable promise as a printing technique for stable cell transfer [6]. LIFT seems to be a future ready option for printing of tissue structures mimicking human cellular constructs.

The survey in the construction and architecture sector reveals many quality records of digital construction. Additive Layer Manufacturing of cement composites including Rapid Hardening Portland Cement, Cement mixtures with retarder and accelerator, geo-polymers and 3M3 mix [20]. Recent studies reveal 3M3 to be optimum cement composite for large scale and heavy-duty construction application. 3M3 has fine grained composition with high flexural and compressive strength and high workability. A comparison of compressive strength between 3M3 and other composites is shown in Fig. 5. Although powder based geo-polymers are environment friendly building composites, their properties are not as per with the conventional cement like materials in the industry [19].

3D Printing of electronics and sensors has paved way for new possibilities for complex 3D circuitry and soft robotics. Copper filament due to its flexibility and conductivity stood at top in circuit printing. Direct write/cure technique of 3D printing has resulted in more functional and complex conductive paths on irregular surfaces [23]. Carbon nano-tubes for conductive and sensing applications along with post

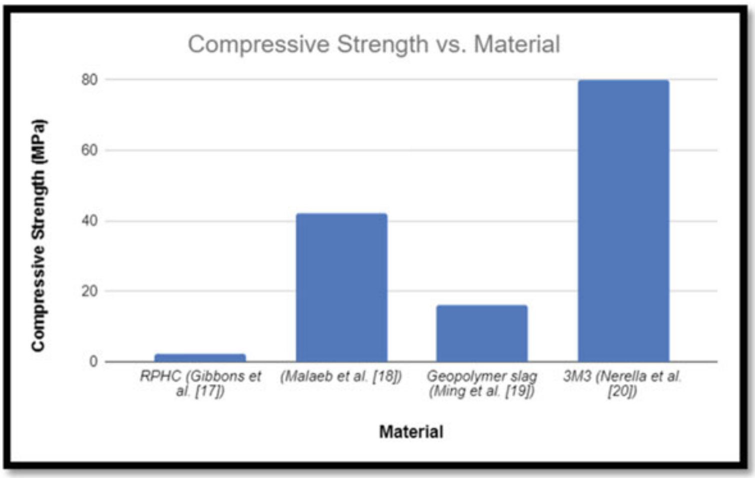


Fig. 5 Compressive strengths of various 3D printing composites in construction application [17–20]

processing algorithms have made the stretchable electronics and wearable a reality [26].

Modern 3D scanners show promising results even for convoluted shapes of test objects. The ability to control the precision of a TK scanner makes its use flexible [30]. 3D endoscopy scanners are now able to create full scale 3D replica of internal organs [31]. The scanners are inept in scanning non solid objects inside the organs. The problem however is being worked upon. Laser scanning seems to be an overall balanced option for most of the applications. 3D scanners can be easily implemented in a 3D printer with only slight hardware changes. Ongoing research will further increase the accuracy of 3D scanners.

4 Conclusion and Future Scope

Requirement of customized manufactured objects is an inseparable part of almost every modern-day sector. A considerable amount of research has been carried out in the biomedical field related to resin prototyping, artificial organs manufacturing and implants manufacturing. Stereolithography (SLA) has been proved to be the most suited techniques for hearing aids and artificial body parts manufacturing. In the field of architecture this technology brought the ease of construction by reducing the operation cost and construction time. This technology has the potential to reduce the operation and to increase the efficiency of labour in the field of construction work by saving the time. This technique has unfolded the new dimensions in the modern industry after successful manufacturing of complex 3D structures used to create multilayer circuit boards, 3D antennas and PCBs. Acquiring high-resolution shape and colour data for 3D printing of objects is prime important in this fastest growing technology. For this, 3d laser scanner has been embedded into a 3D printer.

The major industries using 3D printing technologies have been discussed in this paper and it seems revealed that the technological advancements in the last decade has exponentially increased the influence of 3D printing in modern day industries. The next big leap forward in 3D printing is in implementing machine learning technologies and algorithms with the existing 3D printing technologies. This will make 3D printing smarter and even more suitable for future industries. The scale of 3D printing usage in the education sector will grow exponentially. Use of metal and metal alloys in 3D printing techniques is one of the fastest growing aspects of 3D printing. This shows that the involvement of 3D printing will become quintessential in future industries as well as daily lives. 3D printing evolves every day and it is not wrong to say that the manufacturing and designing industries of the near future will heavily rely on said technologies.

References

1. Dawood, A., Marti, B., Sauret-Jackson, V., Darwood, A.: 3D printing in dentistry. Br. Dental J. **219**, 521–529 (2015). https://doi.org/10.1038/sj.bdj.2015.914
2. Herbert, N., Simpson, D., Spence, W.D., Ion, W.: A preliminary investigation into the development of 3-D printing of prosthetic sockets. J. Rehabil. Res. Dev. **42**(2), 141–146 (2005). https://doi.org/10.1682/JRRD.2004.08.0134
3. Chia, H.N., Wu, B.M.: Recent advances in 3D printing of biomaterials. J. Biol. Eng. **9**(4) (2015). https://doi.org/10.1186/s13036-015-0001-4
4. Andrade, S., Abdalla, A., Montufar, E., Corté, L., Vanegas, P.: Fabrication of 3D bioactive ceramic scaffolds by robocasting. In: Braidot, A., Hadad, A. (eds.) VI Latin American Congress on Biomedical Engineering CLAIB 2014, vol. 49, pp. 167–169. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-13117-7_44
5. Narayanan, L.K., Huebner, P., Fisher, M.B., Spang, J.T., Starly, B., Shirwaiker, R.A.: 3D-bioprinting of polyactic acid (PLA) nanofibers-alginate hydrogel bioink containing human adipose_derived stem cells. ACS Biomater. Sci. Eng. **2**(10), 1732–1742 (2016). https://doi.org/10.1021/acsbiomaterials.6b00196
6. Koch, L., Kuhn, S., Sorg, H., Gruene, M., Schlie, S., Gaebel, R., Polchow, B., Reimers, K., Stoelting, S., Ma, N., Vogt, P.M., Steinhoff, G., Chichkov, B.: Laser printing of skin cells and human stem cells. Tissue Eng. Part C Methods **16**(5), 847–854 (2010). https://doi.org/10.1089/ten.tec.2009.0397
7. Datta, P., Dey, M., Ataie, Z., et al.: 3D bioprinting for reconstituting the cancer microenvironment 4 (2020)
8. McCormack, A., Highley, C.B., Leslie, N.R., Melchels, F.P.W.: 3D printing in suspension baths: keeping the promises of bioprinting afloat. Trends Biotechnol. **38**(6) (2020)
9. El-Sayegh, S., Romdhane, L., Manjikian, S.: A critical review of 3D printing in construction: benefits, challenges, and risks. Archiv. Civ. Mech. Eng **20**, 34 (2020). https://doi.org/10.1007/s43452-020-00038-w
10. Pan, Y., Zhang, Y., Zhang, D., et al.: 3D printing in construction: state of the art and applications. Int. J. A Manuf. Technol. **115**, 1329–1348 (2021). https://doi.org/10.1007/s00170-021-07213-0
11. Kim, M.S., Chu, W.S., Kim, Y.S., Avila, A.P.G., Ahn, S.H.: Direct metal printing of 3D electrical circuit using rapid prototyping. Int. J. Precis. Eng. Manuf. **10**, 147–150 (2009). https://doi.org/10.1007/s12541-009-0106-0
12. Dodziuk, H.: Application of 3D printing in healthcare. Pol. J. Thoracic Cardiovasc. Surg. **13**(3), 283–293 (2016). https://doi.org/10.5114/kitp.2016.62625
13. Del Junco, M., Okhunov, Z., Yoon, R., Khanipour, R., Juncal, S., Abedi, G., Lusch, A., Landman, J.: Development and initial porcine and cadaver experience with three-dimensional printing of endoscopic and laparoscopic equipment. J. Endourol. **29**(1), 58–62 (2015). https://doi.org/10.1089/end.2014.0280
14. Cacco, T., Fragale, M., Sampieri, C., et al.: Modified full-face snorkeling mask for thoracic surgery and otolaryngology surgical use: comfort and usability assessment during the COVID-19 pandemic **112**(2), 107–114 (2021)
15. Ali, A., Ahmad, U., Akhtar, J.: 3D printing in pharmaceutical sector: an overview, pharmaceutical formulation design - recent practices, Usama Ahmad and Juber Akhtar. IntechOpen (2020). https://doi.org/10.5772/intechopen.90738
16. https://edition.cnn.com/style/article/shanghai-3d-printed-bridge-scli-intl/index.html
17. Gibbons, G.J., Williams, R., Purnell, P., Farahi, E.: 3D printing of cement composites. Adv. Appl. Ceramics **109**, 287–290 (2010). https://doi.org/10.1179/174367509X12472364600878
18. Malaeb, Z., Hachem, H., Tourbah, A., Maalouf, T., Zarwi, N., Hamzeh, F.: 3D concrete printing: machine and mix design. Int. J. Civil Eng. Technol. **6**, 14–22 (2015)
19. Xia, M., Sanjayan, J.: Method of formulating geo-polymer for 3D printing for construction applications. Mater. Des. **110**, 382–390 (2016). https://doi.org/10.1016/j.matdes.2016.07.136

20. Nerella, V.N., Mechtcherine, V.: Studying the printability of fresh concrete for framework-free concrete onsite 3D printing technology (CONPrint3D). 3D Concr. Print. Technol. **16**, 333–347 (2019). https://doi.org/10.1016/b978-0-12-815481-6.00016-6
21. Hambach, M., Volkmer, D.: Properties of 3D-printed fiber reinforced Portland cement paste. Cem. Concr. Compos. **79**, 62–70 (2017). https://doi.org/10.1016/j.cemconcomp.2017.02.001
22. Isolda Agusti-Juan, F.M., Hack, N., Wangler, T.: Potential benefits of digital fabrication for complex structures: environmental assessment of a robotically fabricated concrete wall. J. Clean Prod. **154**, 330–340 (2017). https://doi.org/10.1016/j.jclepro.2017.04.002
23. Lu, Y., Vatani, M., Choi, J.W.: Direct-write/cure conductive polymer nanocomposites for 3D structural electronics. J. Mech. Sci. Technol. **27**, 2929–2934 (2013). https://doi.org/10.1007/s12206-013-0805-4
24. Flowers, P.F., Reyes, C., Ye, S., Kim, M.J., Wiley, B.: 3D printing electronic components and circuits with conductive thermoplastic filament. Addit. Manuf. **18**, 156–163 (2017). https://doi.org/10.1016/j.addma.2017.10.002
25. Muth, J.T., Vogt, D.M., Truby, R.L., Menguc, Y., Kolesky, D.B., Wood, R.J., Lewis, J.A.: Embedded 3D printing of strain sensors within highly stretchable elastomers. Adv. Mater. **26**, 6307–6312 (2014). https://doi.org/10.1002/adma.201400334
26. Vatani, M., Lu, Y., Engeberg, E.D, Choi, J.W.: Combined 3D printing technologies and material for fabrication of tactile sensors. Int. J. Precision Eng. Manuf. **16**, 1375–1383 (2015). https://doi.org/10.1007/s12541-015-0181-3
27. Keneth, E.S., Kamyshny, A., Totaro, M., Beccai, L., Magdassi, S.: 3D printing materials for soft robotics. Adv. Matter. **33** (2021). https://doi.org/10.1002/adma.202003387
28. Dong, Y., Wang, S., Ke, Y., Ding, L., Zeng, X., Magdassi, S., Long, Y.: Adv. Mater. Technol. **5** (2020)
29. Rocchini, C., Cignoni, P., Montani, C., Pingi, P., Scopigno, R.: A low cost 3D Scanner based on structured light. Eurographics **20**(3) (2001). https://doi.org/10.1111/1467-8659.00522
30. Ohno, K., Kawahara, T., Tadokoro, S.: Development of 3D laser scanner for measuring uniform and dense 3D shapes of static objects in dynamic environment. In: Proceedings of the 2008 IEEE International Conference on Robotics and Biomimetics, pp. 2161–2167 (2009). https://doi.org/10.1109/ROBIO.2009.4913337
31. Schmalz, C., Forster, F., Schick, A., Angelopoulou, E.: An endoscopic 3D scanner based on structured light. Med. Image Anal. **16**, 1063–1072 (2012). https://doi.org/10.1016/j.media.2012.04.001
32. Straub, J., Kerlin, S.: Development of a large, low-cost, instant 3D scanner. Technologies **2**, 76–95 (2014). https://doi.org/10.3390/technologies2020076
33. Kersten, T., Przybilla, H.J., Lindstaedt, M., Tschirschwitz, F., Misgaiski-Hass, M.: Comparative geometrical investigations of hand-held scanning systems. ISPRS Int. Arch. Photogram. Remote Sens. Spat. Inf. Sci. **XLI-B5**, 507–514 (2016). https://doi.org/10.5194/isprsarchives-XLI-B5-507-2016

Polychromatic Behavior of Reflectance and Field Performance of Graphene Coated SPR Sensor

S. K. Jaiswal and J. B. Maurya

Abstract In this article, conventional and graphene coated surface plasmon resonance (SPR) sensor based on prism coupled Kretschmann configuration are investigated using wavelength interrogation (polychromatic). Considered structure is simulated by finite element method. The performance defining parameters e.g. minimum reflection intensity, beam width, resonance wavelengths of reflectance curve and penetration depth are studied. It is observed that resonance wavelength shifts towards the higher wavelength *i.e.*, from 0.654 μm (for conventional SPR sensor) to 0.751 μm (graphene coated SPR sensor with ten number of graphene layer) by increasing the number of graphene layer. Further, the reflected field intensity increases with the number of graphene layer which signifies less absorption of incident light at prism-metal interface. It is found that addition of graphene makes reflectance curves broader and shallower which eventually increases beam width and minimum reflection intensity. Further, the beam width and minimum reflectance intensity increases monotonically by increasing the number of graphene layer *i.e.*, from 0.105 μm (for conventional SPR sensor) to 0.238 μm (for graphene coated SPR sensor with ten number of graphene layer), and from 0.000015 a. u. (for conventional SPR sensor) to 0.212320 a. u. (for graphene coated SPR sensor with ten number of graphene layer) respectively. Penetration depth (PD) increases with the wavelength and addition of graphene layer has negligible effect on the penetration depth. The obtained PD at 0.650 μm and 0.750 is 205.73 and 237.85 respectively.

Keywords Field distribution · Finite element method · Graphene · Penetration depth · SPR · Wavelength interrogation

S. K. Jaiswal · J. B. Maurya (✉)
Department of Electronics and Communication Engineering, National Institute of Technology Patna, Patna 800005, Bihar, India
e-mail: jbm.ec@nitp.ac.in

S. K. Jaiswal
e-mail: Shailendraj.phd20.ec@nitp.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_38

1 Introduction

The surface plasmon resonance (SPR) sensor is an important tool for the detection of biomolecules and chemical species. This technique is very high sensitive and provides label-free sensing and real-time probing of the presence of molecule in the sensing medium [1]. The basic structure of SPR sensor consists of coupler (prism, waveguide and grating coupler), metal layer (gold, silver, copper, aluminium, sodium, and indium) and sensing medium (sample) [2]. Gold (Au) is preferred among all above mentioned metal because of its highest stability against corrosion and oxidation) [2] Further, gold is highly absorptive in the optical range of wavelength, which makes it suitable to generate surface plasmon at its surface. However, due to its high absorptive nature its thickness must be optimized in order to control over radiation damping and internal damping [2, 3]. The performance parameters of SPR sensor are sensitivity, Beam width (BW) detection accuracy (DA), quality factor, selectivity, limit of detection, propagation length (PL), and penetration depth (PD) [4]. These parameters are optimized by altering different constructional parameters (coupler, metal, addition of coating layer, and sensing medium) of the basic structure of the SPR sensor [1]. To optimize the sensitivity of the SPR sensor, coating of 2-dimensional (2D) nanomaterial viz. graphene, dichalcogenides, meta-material, silicone, MXene, antimonene etc. can be used due to high surface volume ratio and higher affinity towards molecules [5]. Charge carrier mobility of graphene is 106 $cm^2\ V^{-1}\ s^{-1}$ and work function of graphene is 4.5 eV [6]. It has high surface to volume ratio and high binding or adsorption affinity towards biomolecules [7]. Graphene attachés biomolecule with the help of π-stacking property between graphene and carbon rings presented in the biomolecules. Compared to the gold surface, the graphene adsorbs biomolecules more strongly and stably [8]. Graphene is found to absorb a significant fraction (2.3%) of incident white light [9].

From the literature it is found that many researchers have studied the effect of absorbing coating on metal layer in SPR sensors structures to find effect on various performance parameters depending on their reflectance curve of the SPR [2, 10, 11]. But, effect on the field distributions, effect of graphene coating on SPR reflectance curve using wavelength interrogation, and variation of PD are still unexplored. Therefore, in this article it is aimed to explore the field distribution of p-polarized polychromatic incident wave and their PD in the sensing medium. It is worth noting here that PD gives the information of interaction volume of fields with the biomolecules present in the sensing medium.

2 Design Consideration and Theoretical Model

The mathematical calculation of reflection intensity using Fresnel's coefficients with the help of transfer matrix method for N layer model SPR sensor is given by S. Zeng et al. [6]. In this SPR sensor structure, we have used wave optics module of

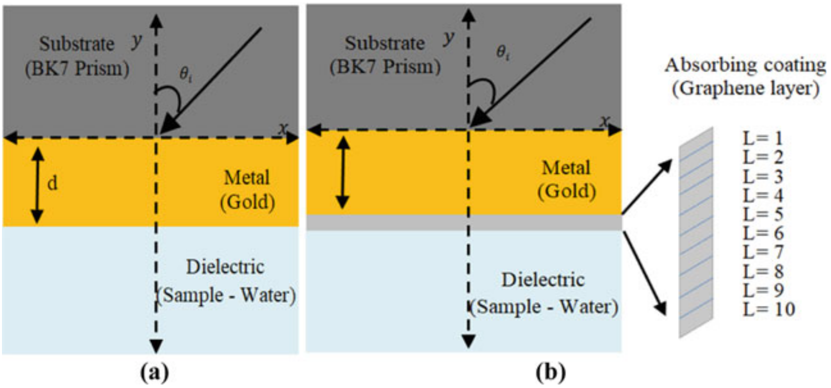


Fig. 1 Schematic diagram of considered SPR sensor structure: **a** Conventional SPR sensor **b** Graphene coated SPR sensor

COMSOL Multiphysics software based on finite element method to simulate the SPR sensor structure consist of N layer prism coupled kretschmann configuration. The 2D structure is made in the *xy* plane in the COMSOL and simulated by considering extra fine mesh. The considered structure shown in Fig. 1(a) consist of three layer system *i.e.*, prism (BK7 glass), metal (gold), and sensing medium (water) and Fig. 1(b) consist of four layer system *i.e.*, prism (BK7 glass), metal (gold), absorbing layer (graphene layer) and sensing medium (water). The thickness of metal layer (d) and graphene layer are 50 nm and (L × 0.34) nm, respectively, Where L is number of graphene layer whereas prism and sensing medium are considered as semi-infinite. Refractive indices of prism (n_1) [6], gold (n_2) [6], and graphene (n_3) [12], water (n_4) [13], are given in Eqs. (1)–(4), respectively, here λ is operating wavelength in µm;

$$n_1^2 - 1 = \frac{1.03961212\lambda^2}{\lambda^2 - 0.00600069867} + \frac{0.231792344\lambda^2}{\lambda^2 - 0.0200179144} + \frac{1.01046945\lambda^2}{\lambda^2 - 103.560653} \quad (1)$$

$$n_2^2 = (\mathcal{E}_r + \mathcal{E}_i) = \left(1 - \frac{\lambda^2\lambda_c}{\lambda_p^2(\lambda_c + i\lambda)}\right) \quad (2)$$

where, $\lambda_c = 8.9342$ and $\lambda_p = 0.16826$ are collision and plasma wavelength in µm.

$$n_3 = \left(3 + i\frac{c_1}{3}\right), \text{ with } c_1 = 5.446\,\mu\text{m}^{-1} \quad (3)$$

$$n_4^2 - 1 = \frac{0.75831\lambda^2}{\lambda^2 - 0.01007} + \frac{0.08495\lambda^2}{\lambda^2 - 8.91377} \quad (4)$$

The operating principle of conventional SPR sensor has been reported by many researcher [3, 14–16]. For the graphene coated SPR sensor, the work function of

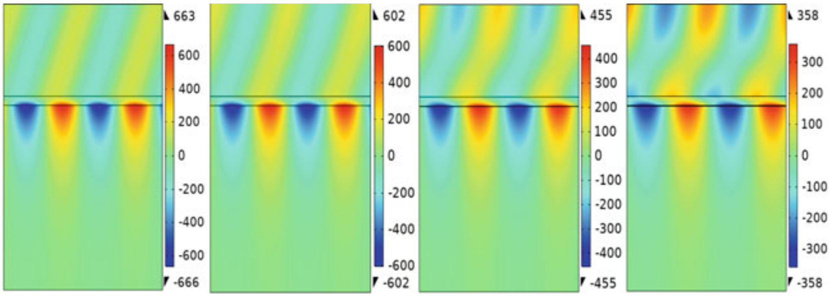
Au (5.54 eV) is higher than 2D nanomaterial graphene (4.5 eV), successful transfer of electrons from graphene layers to gold film will occur under optical excitation. This arrangement leads to a larger electric field enhancement at the metal-sensing interface thereby resulting in a higher sensitivity to the target analytes/molecule [6]. the We have analyzed the effect of graphene coating in the field distribution of incident light, penetration depth, and change in resonance wavelength (λspr) using wavelength interrogation from 300 to 1200 nm in prism coupled kretschmann configuration SPR sensor with a constant angle of incident of incident light *i.e.*, $\theta_i = 70°$.

3 Result and Discussions

The variation of field distribution of $\Re\ (H_Z)$ for a conventional SPR sensor (L = 0) and graphene coated SPR sensor (L = 1 to 10) using wavelength interrogation for p-polarized incident light has been studied. Further, the effect of graphene coating is analyzed on the SPR reflectance curve and PD has been investigated for the number of graphene layer 1 to 10.

3.1 Field Distribution

The field distribution of $\Re(H_Z)$ at resonance wavelength for conventional SPR sensor and graphene coated SPRsensor (L = 1, 5 and 10) are shown in Fig. 2. In this figure, field distribution of $\Re(H_Z)$ for conventional SPRsensor at resonance wavelength gives the information about the incident light is fully absorbed at prism-metal interface and very small amount of light is reflected back in prism. In contrast, in graphene coated SPRsensor, the intensity of reflected light is increases if number of graphene layer increases from L = 1 to 10, which shows that the absorption of the incident light at prism-metal interface decreases. Further, one more interesting thing is also notifying that the maximum value of field intensity of $\Re(H_Z)$ decreases at metal-graphene interface with increase in the number of layer from L = 1 to 10 ingraphene coated SPRsensor. The numerical value of field intensity at metal-sensing medium interface ($\text{Interface}_{\text{m-s}}$), metal-graphene interface ($\text{Interface}_{\text{m-gr}}$), and graphene-sensing medium ($\text{Interface}_{\text{gr-s}}$) for conventional and graphene coated SPRsensor is listed in Table 1. It is also noted that the maximum field intensity is high for the conventional SPR sensor at the metal-sensing medium interface after that field decrease exponentially normal to the interface in the sensing medium but in the graphene coated SPR sensor the distribution of field intensity is different from the conventional SPR sensor. In graphene coated SPR sensor maximum field intensity is highest at L = 1 and lowest at L = 10 at both of the interfaces *i.e.*, metal-graphene and graphene-sensing medium and field intensity decay more rapidly comparison to the field decay in sensing medium. The distance at which field decay more rapidly is defined as the



(a) L = 0 **(b) L = 1** **(c) L = 5** **(d) L = 10**

Fig. 2 Variation of field intensity (z-component of magnetic field, $\Re(H_Z)$) at different number of graphene layers

Table 1 List of variation in field intensity at interfaces with increase in number of graphene layer

L	$\Re(H_Z)$ at interface$_{m\text{-}s}$ (A/m)	$\Re(H_Z)$ at interface$_{m\text{-}gr}$ (A/m)	$\Re(H_Z)$ at interface$_{gr\text{-}s}$ (A/m)	λspr (μm)	BW (μm)	R_{min} (a.u.)
0	653.96	–	–	0.654	0.105	0.000015
1	–	602.04	597.91	0.664	0.116	0.008047
2	–	555.14	547.13	0.674	0.129	0.025710
3	–	517.20	507.02	0.683	0.141	0.047999
4	–	483.94	471.57	0.693	0.155	0.072335
5	-	455.38	441.18	0.703	0.168	0.097285
6	–	431.28	415.48	0.712	0.182	0.121930
7	–	409.39	392.29	0.722	0.195	0.145930
8	–	390.53	372.24	0.731	0.209	0.169200
9	–	373.15	353.91	0.741	0.223	0.191270
10	–	357.51	337.47	0.751	0.238	0.212320

cumulative thickness of graphene layer (see inset of Fig. 3). Thereafter, field decay exponentially in sensing medium as similar to conventional SPR shown in Fig. 3.

3.2 Reflectance Curve

Reflectance curves are plotted using COMSOL and shown in Fig. 4. It is a function of operating wavelength. The BW is defined as either addition of half of the difference between maximum reflection intensity (R_{max}) and minimum reflection intensity (R_{min}) and minimum reflection intensity or difference between maximum reflection intensity and half of the difference between maximum reflection intensity and minimum reflection intensity, see Eq. (2) of Ref. [17]. The obtained resonance

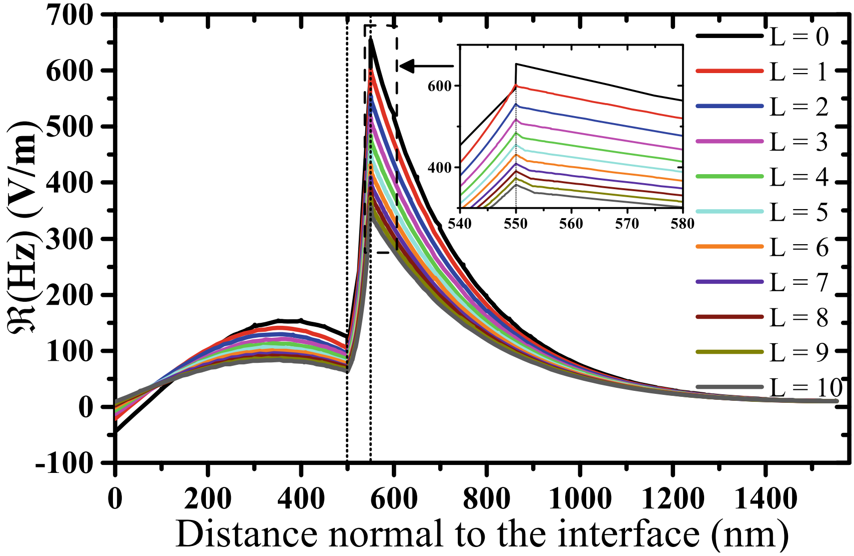


Fig. 3 Field distribution of magnetic field intensity normal to the interface at different number of graphene layers

wavelength, R_{min} and BW are listed in Table 1. The obtained reflectance curves show that an increase in the number of graphene layers increases shift in the resonance wavelength, R_{min} and BW with respect to conventional SPR sensor, which is also shown in Fig. 5. This observed effect can be attributed to the change in thickness and increased absorption of SPW field by graphene layer. This absorption increases the damping of SPW oscillation in the graphene layer. This damping shallows and broadens the reflectance curves [10]. This broadening increases the BW which eventually decreases the detection accuracy ($\propto$ BW^{-1}). Further, shallowing of reflectance curves increases the R_{min} which signifies the loss of SPs. BW and R_{min} increases with the number of graphene layers. The detection accuracy of conventional SPR sensor is more than the graphene coated SPR sensor. Although BW and R_{min} increases with the number of graphene layer which degrades the sensor performance, the shift in resonance angle increases with the number of graphene layers which signifies the improvement in the performance of sensor. In addition, graphene increase the sensitivity of the sensor by providing higher surface area and higher affinity towards biomolecules having carbon rings. Further, increased number of graphene layer improves the adhesion which gives efficient adsorption of biomolecules. Therefore, on increasing the number of graphene layer from 1 to 10, exists a tradeoff between sensitivity and detection accuracy.

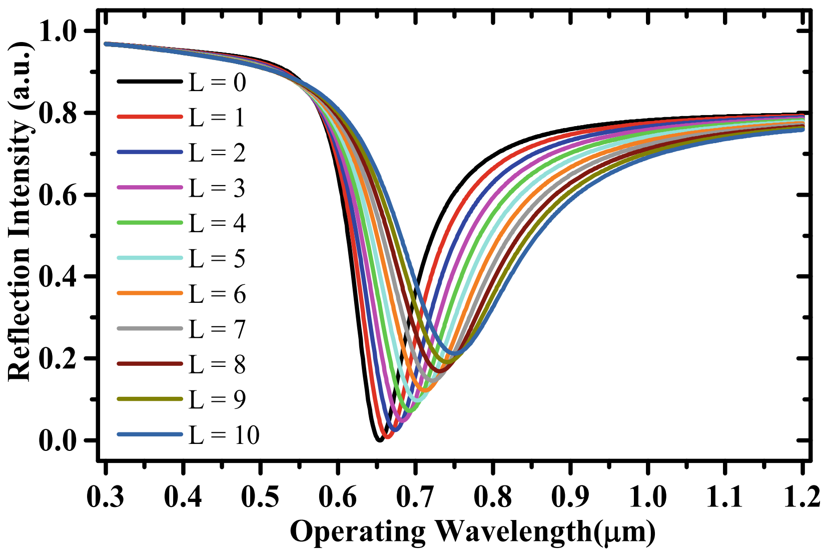


Fig. 4 Reflectance curve as a function of operating wavelength

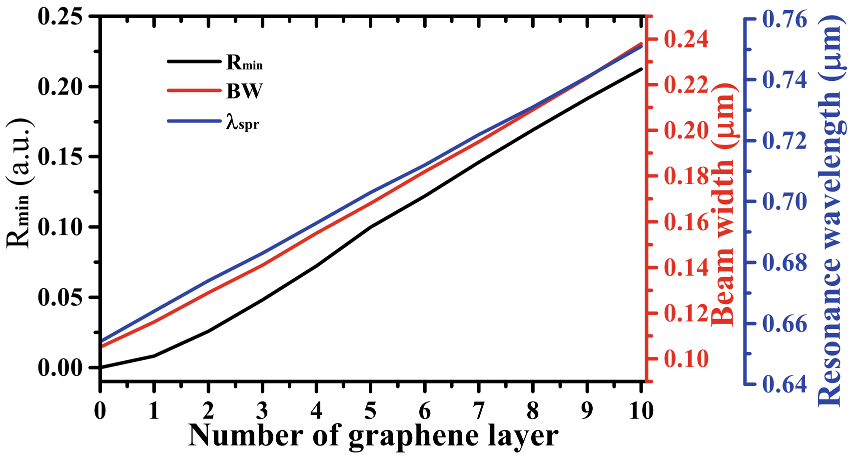


Fig. 5 Variation of R_{min}, resonance wavelength, and beam width with the number of graphene layer

3.3 Penetration Depth

In conventional SPR sensor, SPW propagates at metal-sensing medium interface in *x*-direction and its intensity decreases in normal to the interface as well as along the interface. The PD is defined as distance travelled by SPW normal to the metal-sensing medium interface from maximum at the metal-sensing medium interface to

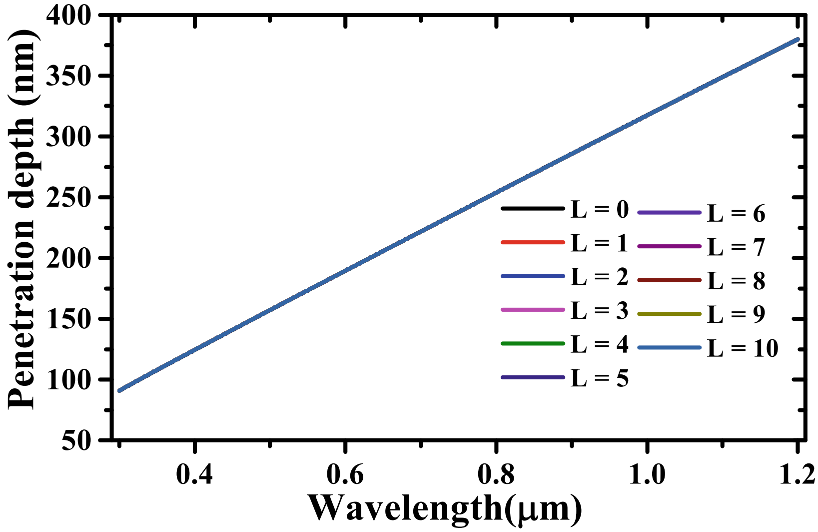


Fig. 6 Variation of penetration depth due to increase in number of graphene layer

the distance at which its intensity decays to 1/e of maximum field [15]. However, in the case of graphene coated SPR sensor, the PD is defined as distance travelled by SPW normal to the graphene-sensing medium interface from maximum at the graphene-sensing medium interface to distance at which intensity decays 1/e of the maximum field. L. Laplatine et al. reported the variation of PD with increasing the operating wavelength [18]. In considered conventional SPR sensor, the PD linearly increases with the operating wavelength which is shown in Fig. 6. In contrast, PD is independent on number of graphene *i.e.,* the PD is nearly same for conventional and graphene coated SPR sensor but field intensity varies with the increase in number of graphene layer. This variation in field intensity is shown in Fig. 3 and listed in Table 1.

4 Conclusion

Polychromatic investigation has been employed for considered conventional and graphene coated SPR sensor structure. The resonance wavelength, beam width, and Rmin increase monotonically with the increase in the number of graphene layer from 0.654 μm (for conventional SPR sensor) to 0.751 μm (graphene coated SPR sensor with ten number of graphene layer), from 0.105 μm (for conventional SPR sensor) to 0.238 μm (for graphene coated SPR sensor with ten number of graphene layer), and from 0.000015 a. u. (for conventional SPR sensor) to 0.212320 a. u. (for graphene coated SPR sensor with ten number of graphene layer) respectively. Sensitivity and detection accuracy show a tradeoff behavior. The penetration depth of the both considered structure is almost same but the field distribution in conventional SPR sensor is completely different from graphene coated SPR sensor. The obtained penetration depth at 0.650 μm and 0.750 μm is 205.73 nm and 237.85 nm respectively and the magnetic field at metal-graphene is 602.04 A/m for monolayer of graphene and 357.51 A/m for ten number of graphene layer. This studies shows single layer of graphene has best performance but fabrication of uniform single layer is quite typical therefore more number of graphene layer has been considered which degrade the performance of SPR sensor.

References

1. Rahman, M.M., Rana, M.M., Anower, M.S., et al.: An SPR-based optical biosensor with shared plasmonic materials and optimized graphene layer: an angular interrogation approach. In: 2020 IEEE Region 10 Symposium, TENSYMP 2020. Institute of Electrical and Electronics Engineers Inc., pp. 1680–1683 (2020)
2. Maurya, J.B., Prajapati, Y.K.: A comparative study of different metal and prism in the surface plasmon resonance biosensor having MoS2-graphene. Opt. Quantum Electron. **48**, 1–12 (2016)
3. Pockrand, I.: Surface plasma oscillations at silver surfaces with thin transparent and absorbing coatings. Surf. Sci. **72**, 577–588 (1978)
4. Sharma, A.K., Jha, R., Gupta, B.D.: Fiber-optic sensors based on surface plasmon resonance: a comprehensive review. IEEE Sens. J. **7**, 1118–1129 (2007)
5. Roh, S., Chung, T., Lee, B.: Overview of the characteristics of micro- and nano-structured surface plasmon resonance sensors. Sensors (Basel) **11**, 1565–1588 (2011)
6. Zeng, S., Hu, S., Xia, J., et al.: Graphene-MoS2 hybrid nanostructures enhanced surface plasmon resonance biosensors. Sens. Actuators B Chem. **207**, 801–810 (2015)
7. Bhavsar, K., Prabhu, R.: Investigations on sensitivity enhancement of SPR biosensor using tunable wavelength and graphene layers. In: IOP Conference Series: Materials Science and Engineering (2019)
8. Song, B., Li, D., Qi, W., et al.: Graphene on Au(111): a highly conductive material with excellent adsorption properties for high-resolution bio/nanodetection and identification. Chemphyschem **11**, 585–589 (2010). https://doi.org/10.1002/CPHC.200900743
9. Wu, L., Chu, H.S., Koh, W.S., et al.: Highly sensitive graphene biosensors based on surface plasmon resonance (2010)
10. Maharana, P.K., Padhy, P., Jha, R.: On the field enhancement and performance of an ultra-stable SPR biosensor based on graphene. IEEE Photonics Technol. Lett. **25**, 2156–2159 (2013)
11. Maurya, J.B., Prajapati, Y.K.: Experimental demonstration of DNA hybridization using graphene based plasmonic sensor chip. J. Light Technol. **38**, 5191–5198 (2020)
12. Bruna, M., Borini, S.: Optical constants of graphene layers in the visible range. Appl. Phys. Lett. **94**, 031901 (2009)
13. Kedenburg, S., Vieweg, M., Gissibl, T., et al.: Nonlinear optical materials; (260.2030) Dispersion; (300.1030) Ab-sorption. (C) 2 (2012)
14. Homola, J.: Electromagnetic theory of surface plasmons. In: Springer Ser Chem Sens Biosens., pp. 3–44 Springer, Heidelberg (2006)
15. Raether, H.: Surface plasmons on smooth surfaces. In: Surface Plasmons on Smooth and Rough Surfaces and on Gratings, pp. 4–39. Springer, Heidelberg (1988)
16. Homola, J.: Surface plasmon resonance sensors for detection of chemical and biological species. Chem. Rev. **108**, 462–493 (2008)

17. Maurya, J.B., Prajapati, Y.K.: A novel method to calculate beam width of SPR reflectance curve: a comparative analysis. IEEE Sensors Lett. **1**, 1–4 (2017)
18. Laplatine, L., Leroy, L., Calemczuk, R., et al.: Spatial resolution in prism-based surface plasmon resonance microscopy. Opt. Express **22**, 22771 (2014)

A Low Power High Speed 15T FinFET-GDI Based Hybrid Full Adder Using 18 nm Technology

Parthiv Bhau and Vijay Savani

Abstract Over the years, adder has become an essential component in any digital system, especially in modern processor architectures. Several different adders have been designed and proposed by researchers to perform the task of addition. In this paper, a low-power high speed 15T Fin Field Effect Transistor (FinFET)-Gate Diffusion Input (GDI) based Hybrid Full Adder (HFA) is proposed. The design is implemented at 1, 0.9, and 0.8 V power supply using 18 nm FinFET technology at 27 °C temperature with the help of the Cadence Virtuoso tool. The various performance parameters like average power dissipation, maximum propagation delay, and Power Delay Product (PDP) of the proposed adder are compared with existing adder topologies like Conventional Complementary Metal Oxide Semiconductor (CCMOS-Mirror) adder, Complementary Pass Logic (CPL) adder, Double Pass Transistor Logic (DPL) adder, Swing Restored Complementary Pass Logic (SRCPL) adder, Transmission Gate adder (TGA), Transmission Function adder (TFA), and 10T Gate Diffusion Input (GDI) adder in terms of average power dissipation, maximum propagation delay and PDP. It has been observed that the proposed adder has the lowest PDP amongst all adders discussed in this paper for 1, 0.9, and 0.8 V power supply. The Process Corner and the Monte Carlo analysis are also performed for the proposed adder at a supply voltage of 1 V.

Keywords Fin Field Effect Transistor (FinFET) · Gate Diffusion Input (GDI) · Very Large Scale Integrated Circuits (VLSI) · Arithmetic and Logic Unit (ALU) · Complementary Metal Oxide Semiconductor (CMOS)

1 Introduction

In the modern world, fast computation with the most negligible power dissipation becomes necessary for Very Large Scale Integrated circuits (VLSI). The adders are a

P. Bhau (✉) · V. Savani
Electronics and Communication Engineering Department, Institute of Technology, Nirma University, Ahmedabad, India
e-mail: 20ftphde40@nirmauni.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_39

crucial component of the Arithmetic Logic Unit (ALU) of various microprocessors and digital signal processors [1]. The full adder serves as a fundamental building block for Multipliers, Subtractors, and Multi-bit adders [2]. The improved performance of full adders benefits the overall system's performance because adders are one of the key element in the critical paths of the system [3]. Over the years, the designers have tried to optimize the performance parameters of the adders [4]. Robustness and low power consumption are the significant benefit parameters of the static adders compared to the dynamic adders; hence, they are more preferred [5].

1.1 Literature Review of Adder Architectures

The Conventional Complementary Metal Oxide Semiconductor (CCMOS-Mirror) adder requires total 28 transistors, providing full output voltage swing with good driving capabilities, but it has more power consumption and more propagation delay due to large input capacitance than Hybrid adders as stated by Basireddy HR et al. [6].

Hasan M et al. reveal that Complementary Pass Logic (CPL) adder requires total 32 nMOS transistors. It provides good output voltage swing, but due to the large transistor count and high switching activity, there is more power dissipation in the circuit [1]. Transmission Gate adders (TGA) has some drawbacks like degradation due to lack of driving capability as stated by Goel S et al. [3] but it gives inherently low power consumption as it is having a lower transistor count of 20 as declared by Kadu CP et al. [4]. Transmission Function adder (TFA) has also the problem of degradation due to the lack of driving capability as stated by Goel S et al. [3] but with 16 transistors it has a small transistor stack height and due to that it has less power consumption and better speed than the other adders, as discussed by Wairya S et al. [7].

Double Pass Transistor Logic (DPL) adder built with 28 transistors gives less propagation delay as stated by Hasan M et al. [1]. It reduces power dissipation, but it occupies more area due to more pMOS transistors discussed by Wairya S et al. [8]. Yadav M et al. reveal that Swing Restored Complementary Pass Logic (SRCPL) adder is CPL with swing restoration transistor with 26 transistors [9]. The propagation delay of SRCPL adder is low whereas it has high power consumption, and due to that, the PDP is increased as stated by Hasan M et al. [1]. Lee PM et al. suggested that The 10T Gate Diffusion Input (GDI)-based XNOR adder has less power dissipation due to less transistor count of 10 but its driving capability is degraded [10].

1.2 Basics of FinFET Architecture

A Fin Field Effect Transistor (FinFET) is a multigate device. It is a non-planner or 3D transistor [11]. The gate terminal is wrapped all over the channel which gives terrific

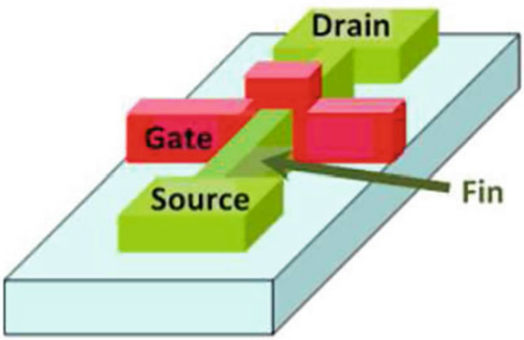


Fig. 1 FinFET structure [11]

control. Fin is a channel between source and drain terminals. Flexibility is obtained when the Fin size is smaller so that multiple Fins can be fabricated which ultimately occupies more area. If the Fin size is larger then it can cause unstable structure but it occupies less silicon [11]. One of the advantages is that the FinFET is having a high drive current which gives a higher speed compared to the Complementary Metal Oxide Semiconductor (CMOS) device [11]. Another advantage is that the oxide thickness of FinFET is lower as compared to CMOS devices which ultimately reduces sub-threshold leakage current [12]. Various researchers proposed different architectures of FinFETs. Figure 1 shows the basic FinFET architecture.

The rest of the paper is organized as follows: Sect. 2 gives an overview of the Gate Diffusion Input Scheme which is responsible for low power consumption. Section 3 describes various performance parameters of the adder. In the next Sect. 4, the working of the proposed full adder design is explained with the schematic. Simulation results with test bench schematic are presented in Sect. 5. Finally, conclusion is described in Sect. 6.

2 Gate Diffusion Input (GDI) Scheme Overview

Figure 2 shows Gate Diffusion Input (GDI) Scheme, which is a low power alternative to normal gate driven approach due to its reduced voltage swing at the output [13]. GDI cell has three inputs and its body terminals are biased as shown in Fig. 2 [14]. Many complex functions can be realized using GDI techniques by utilizing only two transistors [14]. Table 1 shows various function implementations using GDI Technique.

Table 1 shows how different Boolean functions like AND, OR, NOT, MUX etc. which can be realized by connecting different inputs to the N, P, and G terminals of the GDI cell.

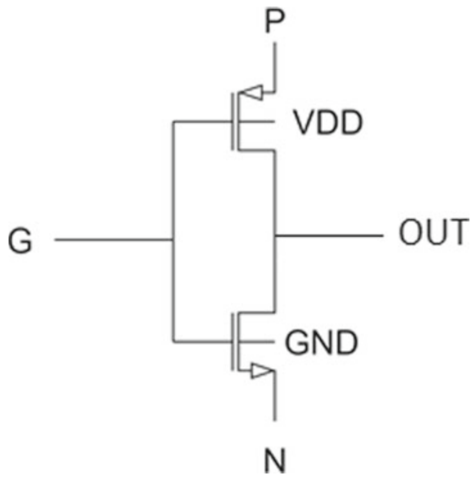


Fig. 2 Basic GDI cell [13]

Table 1 Function implementation using GDI technique [13]

N	P	G	OUT	Function
B	0	A	AB	AND
1	B	A	A+B	OR
0	1	A	A’	NOT
C	B	A	A’B+AC	MUX

3 Performance Parameters of the Adder

To improve the performance of adder circuits, the optimization of device parameters and modification in architecture is necessary. Designers aim to develop the adder with less power dissipation, less area, and less propagation delay [15]. The total power dissipated in the design is given by Eq. 1 as below. Equation 2 gives the Power Delay Product (PDP) which is a Figure of Merit (FoM).

$$P_{total} = P_{static} + P_{switching} + P_{shortcircuit} \tag{1}$$

$$Figure\ of\ Merit\ (FoM) = Average(P_{total}) * Maximum\ Propagation\ Delay \tag{2}$$

P_{static} is the static power dissipated where switching activity is not present in the circuit [16]. $P_{switching}$ is the dynamic power dissipated during transitions at the gate inputs, depending on the switching activity factor [16]. $P_{shortcircuit}$ is the short circuit power dissipation that occurs because of short circuit current [16]. The propagation delay is the time difference between the 50% amplitude value of the input and the output [16]. The propagation delay depends on the input transition time and output load [16]. The PDP becomes a Figure of Merit (FoM) for adder circuits defined by taking the product of average total power dissipation and maximum propagation delay and its value should be as low as possible.

4 Proposed Full Adder Design

Figure 3 shows the proposed Low Power High Speed 15T FinFET-GDI based Hybrid Full Adder. The proposed adder schematic has 4 blocks. First block used a 9T (9-Transistor) based XOR gate which is followed by second block of an inverter which generates XNOR functionality. Two 2:1 GDI-based Multiplexers are used as block 3 and 4 to generate Sum and C_{out} outputs with low power consumption as compared to other adders described in the literature. The proposed adder does not have full swing operation; however, when Sum and C_{out} outputs of the proposed adder are passed through a buffer (cascade pair of inverters), full output voltage swing is achieved.

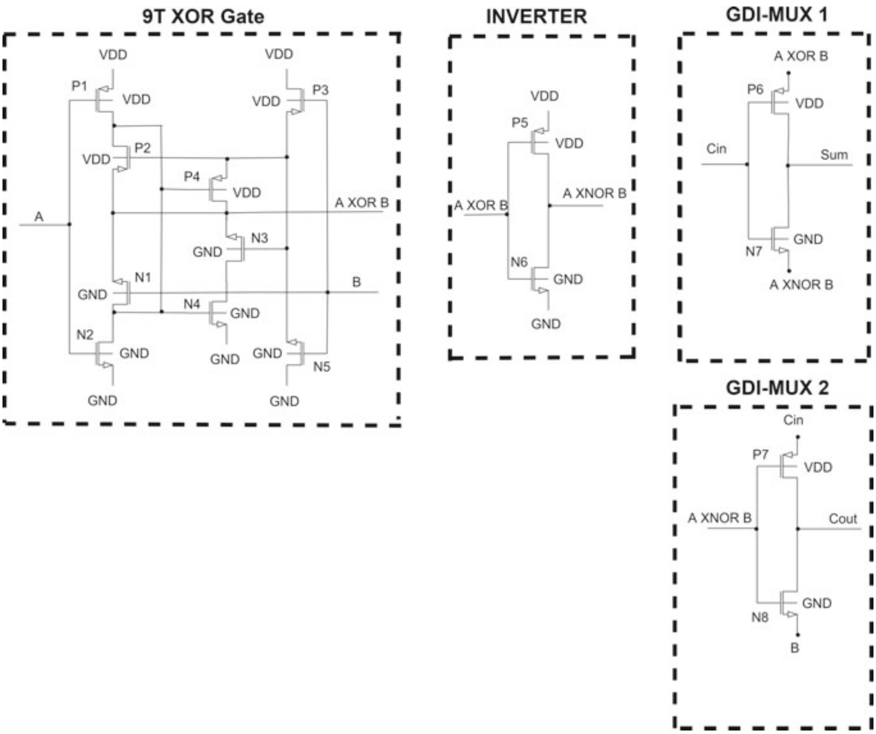


Fig. 3 Schematic of low power high speed 15T FinFET-GDI based hybrid full adder

The proposed adder uses 15 transistors. The logical equations for Sum and C_{out} are given by Eqs. 3 and 4 as below for the proposed adder.

$$Sum = (A \oplus B) \cdot C_{in}^{'} + (A \odot B) \cdot C_{in} \tag{3}$$

$$C_{out} = (A \oplus B) \cdot C_{in} + (A \odot B) \cdot B \tag{4}$$

5 Simulation and Results

Figure 4 Shows the Test bench setup for simulation purposes for all mentioned adders in this paper.

Figure 5 shows the functional simulation of the proposed Low Power High Speed 15T FinFET-GDI based Hybrid Full Adder using 57 input transitions in the simulation to measure maximum propagation delay, estimate the fair amount of power dissipation and check the correct functionality which ultimately is a good input test pattern. The stop time of the transient analysis is set as 58 ns. The period, rise time, and fall time of input pulses A, B, and C_{in} is set as 1 ns, 1 ps, and 1 ps respectively. All adders are designed with 18 nm technology FinFET devices using Cadence Virtuoso tool. All adders are simulated with 18 nm technology FinFET devices at 27 °C temperature using the Cadence Spectre simulator tool.

The FinFET device parameters are as shown in Table 2. Here Fin Pitch represents the sum of Fin width and space between Fins. The number of Fins per finger indicates the width of each poly finger in integer Fin pitches. The multiplier indicates the number of parallel devices.

Influence of supply voltage variation on performance parameters (i.e. average power, maximum propagation delay, and PDP) are simulated at different supply voltages (i.e. 1, 0.9, and 0.8 V) for all 8-different adder architectures.

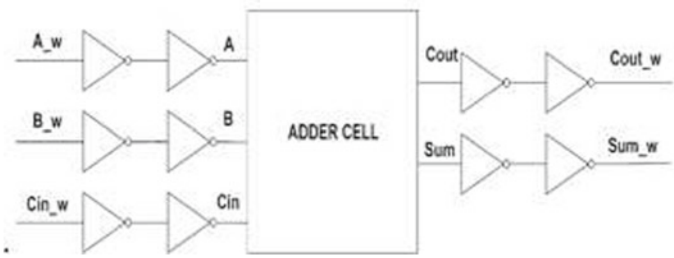


Fig. 4 Test bench setup for simulation

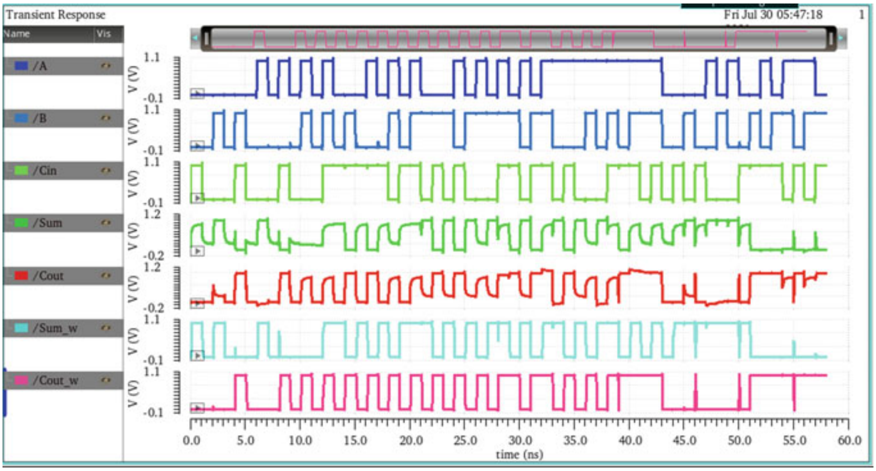


Fig. 5 Simulation of proposed full adder

Table 2 FinFET device parameters

Parameter type	Parameter
Library	cds_ff_mpt 1.0
Cell name	P1–P7: p1lvt & N1–N8: n1lvt
Fin pitch	48 nm
No. of Fins per finger	p1lvt: 2, n1lvt: 2
No. of Fins per finger (For Test Bench Inverters)	p1lvt: 4, n1lvt: 2
Poly pitch	poly 86
Drawn gate length	18 nm
No. of fingers	1
Multiplier	1

Figure 6 shows the simulation results of average power dissipation for all 8 different adders. The observed values of average power dissipation for the proposed adder are 2.5, 1.99, and 1.54 μW at different input supply voltages like 1V, 0.9V, and 0.8V, respectively. The simulation results of maximum propagation delay for all 8 different adders are shown Fig. 7. The observed values of maximum propagation delay for the proposed adder are 88, 98, and 112 ps at different input supply voltages like 1V, 0.9V, and 0.8V, respectively. The simulation results of the PDP for all 8 different adders are presented and compared in the Fig. 8. The observed values of PDP for the proposed adder are 220, 194.73, and 172.82 aJ at different input supply voltages like 1V, 0.9V, and 0.8V, respectively.

The results of Kumar PM et al. [12] are comparable with the proposed adder. However, it is to be mentioned that the technology, the supply voltage, and the

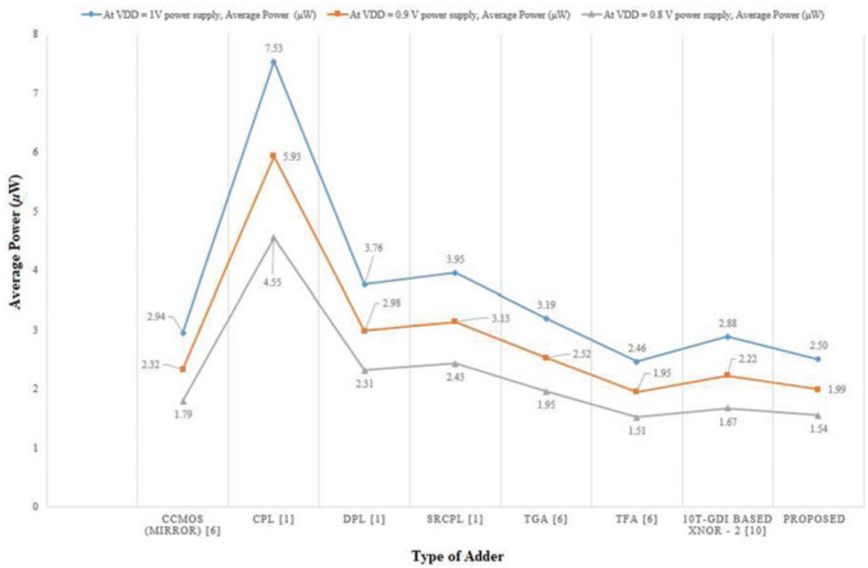


Fig. 6 Comparative analysis of average power dissipation for various adders

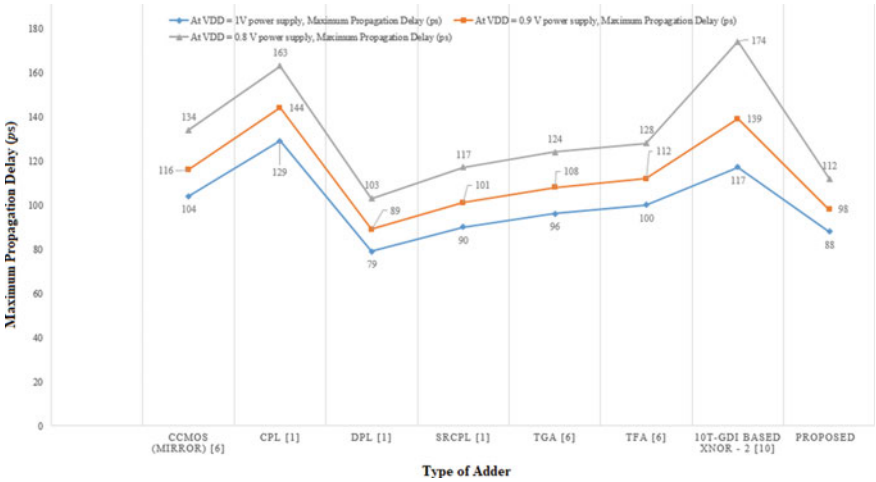


Fig. 7 Comparative analysis of maximum propagation delay for various adders

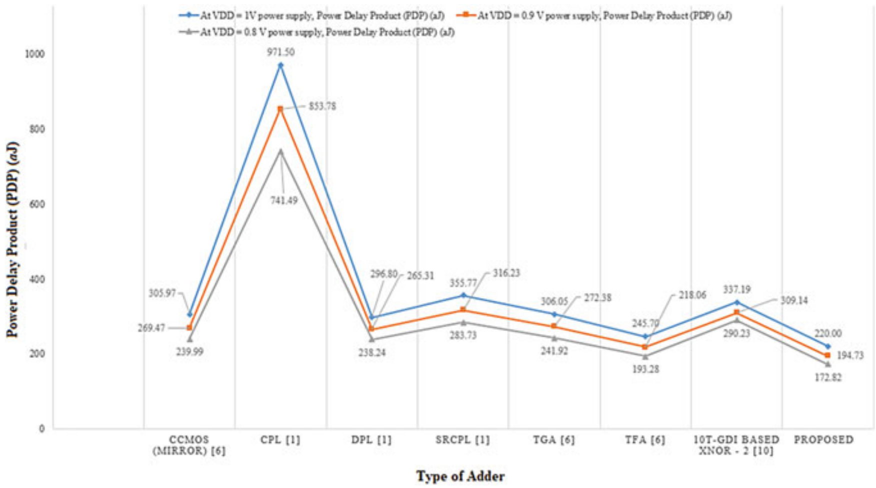


Fig. 8 Comparative analysis of power delay product for various adders

Table 3 Different process corner analysis of proposed adder at 1 V power supply

Parameters	FF	FS	SF	SS	TT
Average power dissipation (μW)	3.03	2.62	2.44	2.27	2.5
Maximum propagation delay (ps)	76.11	1912	89.32	2938	88.67

simulation methodology are different in Kumar PM et al. [12]. Kumar PM et al. [12] has not used 57 input transitions for simulation methodology to get the results.

Table 3 shows the results for influence of process corner variation on the performance parameters i.e. average power and maximum propagation delay at supply voltage of 1 V. During the simulation for Process Corner analysis, the FinFET parameters are changed to have effect of Slow or Fast device. Process corners analysis for Fast N-type FinFET—Fast P-type FinFET (FF), Fast N-type FinFET—Slow P-type FinFET (FS), Slow N-type FinFET—Fast P-type FinFET (SF), Slow N-type FinFET-Slow P-type FinFET (SS), and Typical N-type FinFET-Typical P-type FinFET (TT) is given. The result supports that there is a significant change in the maximum propagation delay for FS and SS Process Corners as compared to other Process Corners.

Process Corner analysis simulates extreme cases i.e. minimum and maximum values of each parameter but in the real fabrication process, Monte Carlo analysis is required because it is based on realistic statistical distributions. The Monte Carlo analysis is used to evaluate mismatch and process variation. Monte Carlo analysis is useful to give a prediction about random variations of the process parameters of the design. In Monte Carlo analysis maximum value, minimum value and the average value which is called as mean is also generated. Apart from that median and standard deviation is also calculated.

After sorting the data from lowest to highest, the median is the ‘middle’ value or the 50^{th} Percentile meaning that 50% of the results from the simulation are less than the median. If there is an even number of data points, then the median is the average of the middle two points. Standard deviation is a standard measure of the spread in data values. If the value of standard deviation is small then there is high confidence in the estimated value.

The Monte Carlo Simulations are carried out for 200 samples for average power and maximum propagation delay given the process variation. The Simulation is performed for the proposed Hybrid adder is at the Supply Voltage of 1 V and the results of average power and maximum propagation delay are enumerated in Table 4. The statistical variations are also shown in terms of minimum, maximum, mean, median, and standard deviation values in Table 4 for the proposed adder.

Figures 9 and 10 shows the Monte Carlo analysis result of the proposed Hybrid adder for average power dissipation and maximum propagation delay using 18 nm FinFET Technology at the supply voltage of 1 V for 200 samples which shows the process variability as an outcome.

Table 4 Monte Carlo analysis of proposed adder

Parameters	Min	Max	Mean	Median	Std. deviation
Average power dissipation (μW)	2.43	2.76	2.50	2.49	0.05
Maximum propagation delay (ps)	86.11	1915	107	88.73	181.9

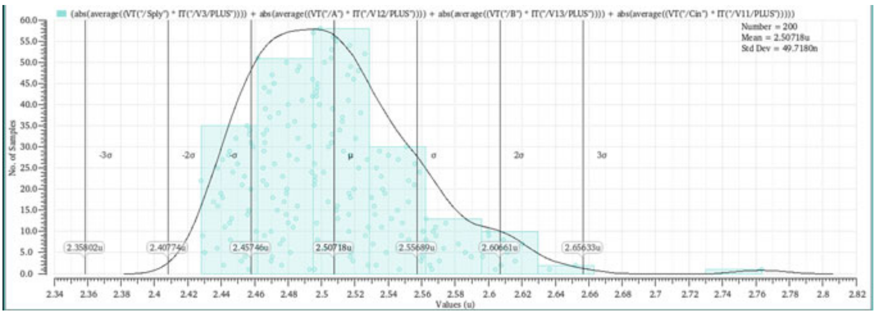


Fig. 9 Monte Carlo analysis for average power dissipation of proposed adder

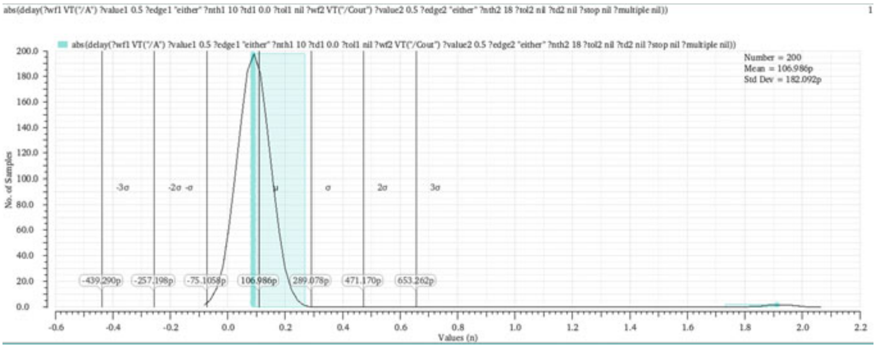


Fig. 10 Monte Carlo analysis for maximum propagation delay of proposed adder

6 Conclusion

The proposed 15T FinFET-GDI based Hybrid adder has high speed and low power dissipation. Simulations were carried out by changing the supply voltage and observing the maximum propagation delay, average power dissipation, and the PDP for eight different adders including well known adder architectures along with proposed. It has been observed that for different power supply voltages 1, 0.9, and 0.8 V, the proposed adder gives the minimum PDP amongst all eight adders. The measured average power dissipation of the proposed adder are 2.5 μW with supply voltage of 1 V, 1.99 μW at 0.9 V power supply, and 1.54 μW at 0.8 V power supply. The measured maximum propagation delay of the proposed adder are 88 ps at 1 V power supply, 98 ps at 0.9 V power supply, and 112 ps at 0.8 V power supply. The observed PDP for the proposed adder is 220 aJ, 194.73 aJ, and 172.82 aJ with the supply voltage of 1 V, 0.9 V, and 0.8 V respectively. The Process Corner analysis and the Monte Carlo analysis are also being carried out and their results are populated in Table 3 and Table 4 respectively. In summary, the proposed Hybrid Adder offer lower Figure of Merit (FoM) as compared to other adders.

References

1. Hasan, M., Hossein, M.J., Saha, U.K., Tarif, M.S.: Overview and comparative performance analysis of various full adder cells in 90 nm technology. In: 4th International Conference on Computing Communication and Automation (ICCCA), pp. 1–6 (2018). https://doi.org/10.1109/CCAA.2018.8777684
2. Mewada, M., Zaveri, M.: A low-power high-speed hybrid full adder. In: 2016 20th International Symposium on VLSI Design and Test (VDAT), pp. 1–2 (2016). https://doi.org/10.1109/ISVDAT.2016.8064900
3. Goel, S., Kumar, A., Bayoumi, M.A.: Design of robust, energy-efficient full adders for deep-submicrometer design using hybrid-CMOS logic style. IEEE Trans. Very Large Scale Integr. (VLSI) Syst. **14**, 1309–1321 (2006). https://doi.org/10.1109/TVLSI.2006.887807

4. Kadu, C.P., Sharma, M.: Area-efficient high-speed hybrid 1-bit full adder circuit using modified XNOR gate. In: 2017 International Conference on Information, Communication, Instrumentation and Control (ICICIC), pp. 1–5 (2017). https://doi.org/10.1109/ICOMICON.2017.8279064
5. Theja, M.N., Balakumaran, T.: Energy efficient low power high speed full adder design using hybrid logic. In: 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT), pp. 1–8 (2016). https://doi.org/10.1109/ICCPCT.2016.7530209
6. Basireddy, H.R., Challa, K., Nikoubin, T.: Hybrid logical effort for hybrid logic style full adders in multistage structures. IEEE Trans. Very Large Scale Integr. (VLSI) Syst. **27**, 1138–1147 (2019). https://doi.org/10.1109/TVLSI.2018.2889833
7. Wairya, S., Nagaria, R.K., Tiwari, S.: Performance analysis of high speed hybrid CMOS full adder circuits for low voltage VLSI design. J. VLSI Des. **2012** (2012). https://doi.org/10.1155/2012/173079
8. Wairya, S., Singh, G., Vishant, Nagaria, R.K., Tiwari, S.: Design analysis of XOR (4T) based low voltage CMOS full adder circuit. In: 2011 Nirma University International Conference on Engineering, pp. 1–7 (2011). https://doi.org/10.1109/NUiConE.2011.6153275
9. Yadav, M., Gupta, A., Rai, S.: A study and comparative analysis of low power hybrid-CMOS 1-bit full adders in deep-submicron technology. In: 2017 4th International Conference on Power, Control Embedded Systems (ICPCES), pp. 1–7 (2017). https://doi.org/10.1109/ICPCES.2017.8117618
10. Lee, P.M., Hsu, C.H., Hung, Y.H.: Novel 10-T full adders realized by GDI structure. In: 2007 International Symposium on Integrated Circuits, pp. 115–118 (2007). https://doi.org/10.1109/ISICIR.2007.4441810
11. Khadir, M., Chaitanya, K., Sushma, S., Preethi, V., Vijay, V.: Design of carry select adder based on a compact carry look ahead unit using 18nm FinFet technology. J. Crit. Rev. **7**, 1164–1171 (2020). https://doi.org/10.31838/jcr.07.06.210
12. Pavan Kumar, M., Dharmendra, M., Sai, R., Shrity, M., Vamsi, P.: An efficient full adder using FinFet technology. Int. J. Manage. Technol. Eng. **9**, 4754–4761 (2019)
13. Shoba, M., Nakkeeran, R.: GDI based full adders for energy efficient arithmetic applications. Eng. Sci. Technol. Int. J. **19**, 485–496 (2016). https://doi.org/10.1016/j.jestch.2015.09.006
14. Sanapala, K., Sakthivel, R.: Ultra-low-voltage GDI-based hybrid full adder design for area and energy-efficient computing systems. IET Circ. **13**, 465–470 (2019). https://doi.org/10.1049/iet-cds.2018.5559
15. Hasan, M., Zaman, H.U., Hossain, M., Biswas, P., Islam, S.: Gate diffusion input technique based full swing and scalable 1-bit hybrid full adder for high performance applications. Eng. Sci. Technol. Int. J. **23**, 1364–1373 (2020). https://doi.org/10.1016/j.jestch.2020.05.008
16. Harish, B., Sivani, K., Rukmini, M.: Performance comparison of various CMOS full adders. In: International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), pp. 3789–3792 (2017). https://doi.org/10.1109/ICECDS.2017.8390172

Sensitivity Enhancement of SPR Sensor Based on $Ti_3C_2T_x$ (MXene) with Composite Layers of TiO_2-SiO_2

Rajeev Kumar, Maneesh Kumar Singh, Sarika Pal, Narendra Pal, and Y. K. Prajapati

Abstract In this seminal, the surface plasmon resonance sensor (SPR) sensor based on $Ti_3C_2T_x$ (MXene) with a composite layer of TiO_2-SiO_2 is theoretically presented. The $Ti_3C_2T_x$ as biomolecular recognition element (BRE) layer enhances the sensitivity of the proposed sensor. The other two important parameter such as detection accuracy (DA) and figure of merit (FoM) have been also deliberated. Firstly, the thicknesses of TiO_2 and SiO_2 are optimized with a monolayer of gold and $Ti_3C_2T_x$ layers for calculation of the sensitivity. The maximum sensitivity achieved is 276°/RIU. The variation in the sensitivity, Rmin, DA, and FoM w.r.t the RI of the sensing medium, is also analyzed. To evaluate the penetration depth for the proposed sensor, we have evaluated the transverse electromagnetic field distribution too. The penetration depth of 197 nm is obtained for the proposed SPR sensor.

Keywords Detection accuracy · Electromagnetic field distribution · FoM · Sensitivity · Sensor · Surface plasmon resonance

R. Kumar · M. K. Singh (✉) · S. Pal
Department of ECE, National Institute of Technology Uttarakhand, Pauri Garhwal, Srinagar 246174, Uttarakhand, India
e-mail: maneeshsingh.phd19@nituk.ac.in

R. Kumar
e-mail: rajeevkumar.phd2020@nituk.ac.in

S. Pal
e-mail: sarikapal@nituk.ac.in

N. Pal
Department of ECE, Rajkiya Engineering College, Kannauj 209732, U.P., India

Y. K. Prajapati
Department of ECE, Motilal Nehru National Institute of Technology Allahabad, Paryagraj 211004, U.P., India
e-mail: yogendrapra@mnnit.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_40

1 Introduction

In biomedical field, a number of biosensors have been developed based on PCR [1], ellipsometry [2], quartz crystal microbalance [3], electrochemistry [4], piezoelectric, spectroscopy, interferometry, colorimetry [5], fluorescence resonance energy transfer [6], surface plasmon resonance, etc. Among the various methods, optical biosensor has attracted the attention of researchers. Surface plasmon resonance (SPR) technique among optical biosensors is a powerful tools for the detection and investigation of DNA hybridization, protein, antigen [7], and virus [8] in real-time with high sensitivity and high-speed response [9]. In a specific application such as detection of DNA hybridization [9, 10], miRNA [11], SPR biosensor has attracted the attention of researchers because of their attractive sensing capability, compactness, robustness, high precision, and reliability. The most widely used SPR biosensor works on the principle of attenuated total reflection (ATR). Surface plasmon polaritons (SPPs) are electron density wave (collective oscillation of free electron gas density) propagating along the interface of dielectric (positive permittivity) and metal (negative permittivity) [12]. Surface plasmon resonance phenomenon achieved when evanescent wave vector of transverse magnetic (TM) polarized light matches with wave vector of SPP wave [13]. At resonance condition, dip in the reflection intensity is observed at SPR angle [14]. Various methods are proposed to improve or enhance the sensitivity of biosensors such as a colloidal gold nanoparticle, nano slits, bimetallic SPR active metal, hybrid layer structure, and 2D material [11]. 2D materials open a new window for researchers working in the field of material science, optoelectronics, and biosensing [15].

Numeral emerging 2D materials such as graphene, black phosphorous (BP), transition dichalcogenides (TMDCs), metal oxides and antimonene has attracted significant interest due to their electrical, and optical properties to enhance the sensitivity of biosensors [16]. Most of the 2D material has certain drawbacks such as weak interaction with biomolecules or poor chemical stability. Recently explored 2D material, mixed metal carbides (MXenes) has unique properties such as large surface area, strong carrier confinement, incredible stability and hydrophillicity much better than other 2D materials [17]. In 2018, Wu et al. [18] compared the sensitivity of MXene based SPR sensor on using different metals, the highest sensitivity of 160°/RIU is achieved with Au-Mxene configuration. In 2019, Xu et al. [19] demonstrated highest sensitivityof198°/RIU for MXene and transition metal dichalcogenide based SPR sensor. Recently, in 2020, Pal et al. [20] used the BlueP/MoS_2 heterostructure with MXene and low refractive prism CaF_2 for enhancing the sensitivity up to 203°/RIU. Recently in 2021, Pandey et al. [21], theoretically investigated a SPR sensor based on heterojunction (MoO_3-$Ti_3C_2T_x$), which obtained the highest sensitivity (268.14°/RIU). For the further enhancement of the performance of the SPR sensor, the adherence composite layer of lower RI over the high RI material is used over the prism. The SiO_2 is used over the TiO_2 to make a composite layer. The light-trapping near the interface at the composite layer is enhanced due to the plasmon effect that enhances the surface plasmons (SPs), resulting in larger resonance angle shift, which leads to advance the sensitivity of the SPR sensor [22]. In this paper, we proposed the 6-layer SPR sensor with constituent layers of TiO_2/SiO_2/Au/$Ti_3C_2T_x$/SM over the low RI (BaF_2) prism. The proposed SPR sensor is theoretically and numerically analyzed at operating wavelength of 633 nm for RI range (1.330–1.355) of sensing medium.

This paper is prepared as follows. In Sect. 2, The sensor design consideration and the performance parameters are presented. In Sect. 3, the results are discussed under numerous sub-sections. To conclude the work presented here, finally the conclusion and required references are presented at the end.

2 Proposed Sensor Design

2.1 Design Consideration and Modeling

The schematic of the 6-layer proposed SPR sensor (Configuration: Prism BaF_2-TiO_2-SiO_2-Au-$Ti_3C_2T_x$-SM) is shown in Fig. 1. The first layer is the BaF_2 glass prism and its RI is $n_1 = 1.4733$ [21]. The second layer is TiO_2 and its RI $n_2 = 2.5837$ [23]. The third layer is SiO_2 and its RI is $n_3 = 1.4570$ [23]. The fourth layer is Au (gold) and its RI (n_4) is calculated from Eq. (1) as per the Drude model [10]. The fifth layer is $Ti_3C_2T_x$ and its RI is $n_5 = 2.38 + 1.33i$ [17]. The sixth layer is the sensing medium having RI of 1.33, which may vary up to 1.335 on adsorption of biomolecules. The thickness of Au and $Ti_3C_2T_x$ considered are 47 and, 0.993 nm with an optimized thickness of TiO_2 and SiO_2 (65–65 nm).

$$n = (\varepsilon_r + \varepsilon_i)^{\frac{1}{2}} = \left(1 - \frac{\lambda^2\lambda_c}{\lambda_p^2(\lambda_c + i\lambda)}\right)^{\frac{1}{2}} \tag{1}$$

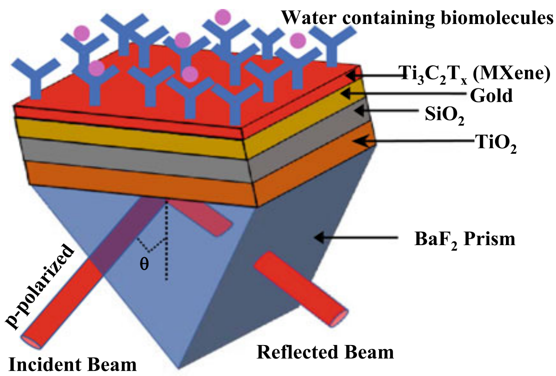


Fig. 1 Proposed SPR sensor

where collision and plasma wavelengths for Au are $\lambda_c = 8.9342 \times 10^{-6}$ m and $\lambda_p = 1.6826 \times 10^{-7}$ m respectively.

The reflectance of the p-polarized incident light is calculated using the transfer matrix method (TMM) as precise modeling with zero approximation [11]. The N-layer modeling is used for the calculation and systematic investigation of the reflectivity of the reflected light in the proposed SPR sensor. The mathematical modeling is described in our previous papers [11].

2.2 Performance Parameters of the SPR Sensor

Sensitivity is the ratio of a difference between two resonance angles ($\Delta\theta_{res} = \theta_2 - \theta_1$) to the RI shift of sensing layer ($\Delta n_s = 0.005$)

$$S = \frac{\Delta\theta_{\mathrm{Res}}}{\Delta n_s}(^{\circ}/\mathrm{RIU}) \tag{2}$$

FWHM is change of resonance angles ($\theta_2 - \theta_1$) at 50% reflection intensity and measures the angular width of the SPR curve [24].

$$\mathrm{FWHM} = (\theta_2 - \theta_1) \tag{3}$$

DA is inversely proportional to the FWHM and given as:

$$DA = \frac{1}{FWHM}(1/^{\circ}) \tag{4}$$

FoM is multiplication of the sensitivity and detection accuracy.

$$FoM = S \times DA(1/\mathrm{RIU}) \tag{5}$$

3 Discussion and Analysis of Numerical Results

First, we have optimized the thickness of the TiO_2-SiO_2 composite layer. The thickness of the TiO_2-SiO_2 has been optimized corresponding to reflectance and sensitivity with 47 nm of Au and monolayer of $Ti_3C_2T_x$ at 1.33 values of n_s, shown in Fig. 2. In Fig. 2(a), it is clearly observed that the minimum reflectivity, Rmin. is reached at 65 nm thickness of both TiO_2 and SiO_2 layers. Similarly, at 65 nm of TiO_2 and SiO_2, the maximum sensitivity 276°/RIU is obtained, shown in Fig. 2(b). Thus, the optimized thickness of TiO_2 and SiO_2 are 65 nm.

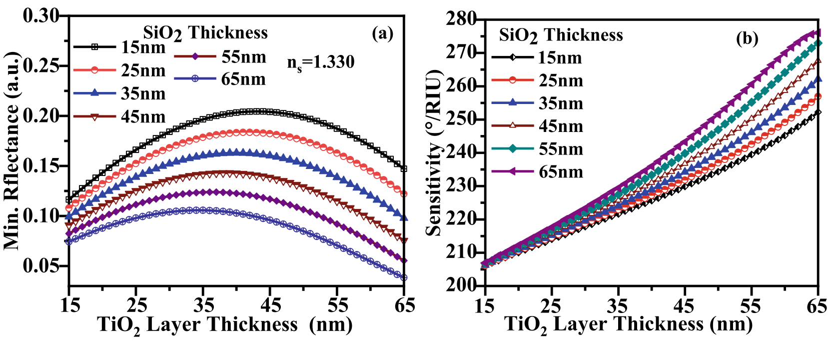


Fig. 2 TiO_2 and SiO_2 thickness optimization of in reference to **a** Min. reflectance and **b** sensitivity for Au (47 nm) and monolayerTi3C2Tx

The performance parameters of the sensor are calculated from the SPR characteristic curve illustrated in Fig. 3(a). The resonance angle shift, $\theta_{res.}$ due to alteration in RI of the sensing medium from 1.33 to 1.335, shown in Fig. 3(b). The performance parameters has been intended from SPR curves with $\Delta n_s = 0.005$. The sensitivity, FWHM, detection accuracy, and FoM are 276°/RIU, 12.02°, 0.083/° & 23/RIU, respectively. The Fig. 3(b) shows that the variation (shifting) in SPR curve in reference to alteration in sensing medium RI. From the Fig. 3(b) a linear variation in resonance angle is detected which varies from 80.2089° to 81.5891° corresponding to alteration the RI of sensing medium (1.330–1.335).

The performance parameter's analysis corresponding to the variation in RI of the sensing medium is presented, in Fig. 4. The maximum sensitivity 281.87°/RIU is achieved at 1.330 RI of the sensing medium then further it decreases from 281.87°/RIU to 276°/RIU with the alteration in RI of sensing medium (1.33–1.335), as per Fig. 4(a). In Fig. 4(a), the minimum reflectance increases from 3.84×10^{-2} to

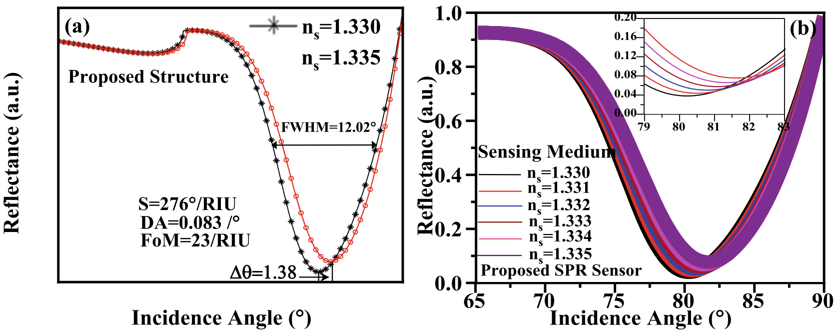


Fig. 3 **a** SPR reflectance curves of proposed SPR **b** Reflectance curve at various thickness of SiO_2 layer with 65 nm thickness of TiO_2 and monolayer of Ti3C2Tx and Au (47 nm)

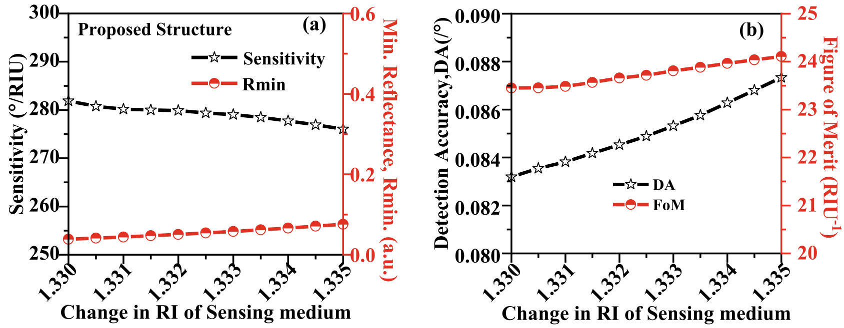


Fig. 4 **a** Sensitivity, Min. reflectance vs RI of sensing medium **b** Detection accuracy, FoM vs RI of sensing medium

7.95×10^{-2} a.u. with the increasing the RI (1.330 to 1.335) of the sensing medium due to increase in energy loss.

Figure 4(b) plot shows, the variation of detection accuracy and FoM with a slight modification in RI of the sensing medium. The FWHMs of the proposed SPR sensor are in decreasing manner, so its reciprocal, detection accuracy are increased from 0.0831/oto 0.0873/o. The overall performance of the SPR sensor is defined by FoM that is the product of sensitivity and detection accuracy. The FoM variations are from 23.450/RIU to 24.109/RIU corresponding to alteration in RI of sensing medium (1.330–1.335), respectively.

From Fig. 5, we can easily analyze the effect of the Mxene and composite layers on the SPR sensor. The Mxene based conventional SPR curves indicated by red colour demonstrates the more shifting of resonance angle as compared to conventional SPR due to improved adsorption of bio-molecules by the MXene, and change in resonance angle 0.967° is obtained. High real RI of TiO_2 gives large change in resonance angle up to 1.231° as compared to SiO_2, as shown in Fig. 5. For further upgrading the sensitivity of the sensor, we can introduce composite layers of TiO_2-SiO_2 which give the much larger change in resonance angle (1.38°) as compared to conventional SPR and SPR sensor without TiO_2 and SiO_2 layers.

Table 1 shows that the comparison between the proposed SPR sensor with the conventional SPR sensor as well as the proposed sensor with and without TiO_2 and SiO_2. From the Table 2, it is interpreted that the composite layer enhances the sensitivity as compare to conventional and rest of proposed SPR sensors due TiO_2 and SiO_2 enhancing the light trapping that enhance the plasmonic effect. The variation in the performance of the proposed SPR sensor corresponding to thickness variation in SiO_2 from 15 to 65 nm.

The COMSOL Multiphysics software is used for the analysis of the electric field enhancement by the finite element analysis (FEA) methods. The 1-D electric field distribution for the proposed SPR sensor within each layer is revealed in the Fig. 6(a). It is clear that, the TiO_2 and SiO_2 layer help to enhance the electric field at the interface of metal and composite layer and 2D material MXene also enhanced the electric

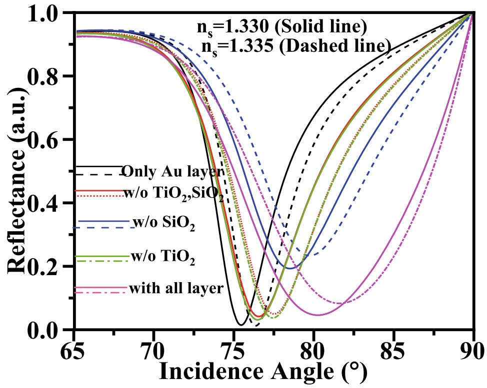


Fig. 5 SPR curves for the conventional SPR and proposed SPR with and without composite layer at ns = 1.330 and 1.335

Table 1 Performance parameters analysis with and without MXene, TiO_2 and SiO_2 layers

Structure layer	Δθ (deg.)	FWHM (1.33)	DA (deg^{-1})	S (°/RIU)	FoM (RIU^{-1})
Only Au layer	0.9	4.9	0.204	180	36.72
w/o TiO_2, SiO_2	0.967	6.73	0.148	193	28.56
w/o SiO_2	1.231	6.91	0.144	246	35.42
w/o TiO_2	0.955	6.83	0.146	191	27.88
With all layer	1.38	12.02	0.083	276	22.9

Table 2 Performance parameters with increasing the thickness of SiO_2 layer at 65 nm of TiO_2, monolayer of Au (47 nm) and $Ti_3C_2T_x$ (0.993 nm)

SiO_2 thickness	Δθ (deg.)	FWHM (1.33)	DA (deg^{-1})	S (°/RIU)	FoM (RIU^{-1})
15	1.26	7.92	0.126	252	31.75
25	1.28	8.66	0.115	257	29.55
35	1.31	9.46	0.105	262	27.51
45	1.34	10.31	0.096	268	25.72
55	1.37	11.17	0.089	273	24.29
65	1.38	12.02	0.083	276	22.9

field distribution interface of metal and MXene. The penetration depth of electric field in sensing medium is calculated from this 1-D electric field distribution. The penetration depth is 197 nm in the sensing medium for the proposed SPR configuration. The penetration depth of the SPR sensor is important parameter to build SPR with the molecule sized specific that immobilized at the surface. The range of penetration depth of conventional SPR is near to 100–250 nm with visible range of wavelength. The electric field intensity is plotted by the COMSOL Multiphysics

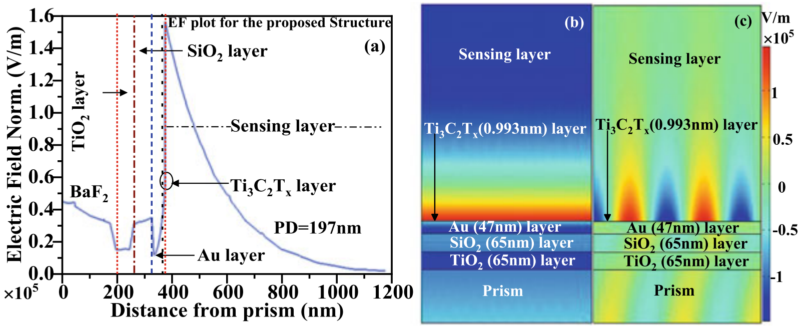


Fig. 6 **a** Electric filed distribution as a function of distance BaF2 to analyte **b** and **c** Field enhancement and SPs distribution at resonance state

Table 3 Comparison table of SPR sensors utilizing $Ti_3C_2T_x$, MXene

Structure	Sensitivity	References
BK7-Au-$Ti_3C_2T_x$	160	Wu et al. [18]
BK7-Au-TMD-Au-$Ti_3C_2T_x$	198	Xu et al. [19]
Prism-Metal-BlueP/MoS_2- $Ti_3C_2T_x$	203	Pal et al. [20]
SiO_2-Ag-MoO_3-$Ti_3C_2T_x$	268.14	A. K. Pandey [21]
BaF_2-TiO_2-SiO_2-Au-$Ti_3C_2T_x$	276	This work

as shown in Fig. 6(b and c). The electric field intensity is an important role in the SPR sensor. Figure 6(b and c) represents TM field distribution and SPs propagation respectively at various interface of the proposed sensor [BaF_2-TiO_2(65 nm)–SiO_2(65 nm)–Au(47 nm)-$Ti_3C_2T_x$(0.993 nm)-SM] w.r.t the normal distance from BaF_2 to sensing medium at resonance condition. The electric field enhancement in the y component at resonance angle is plotted in Fig. 6(b). The X–Y plane is considered as the TM-polarized incidence light. The mode coupling is responsible for E_y improvement at resonance state at 633 nm wavelength.

Table 3 is presented to compare proposed sensor with recent works for achieving maximum sensitivity. The work proposed here is capable to conquer the maximum sensitivity at 633 nm wavelength for ns = 1.330 and Δns = 0.005.

4 Conclusion

The proposed work is theoretically presented of an SPR sensor using a composite layer (TiO_2-SiO_2) with a monolayer of $Ti_3C_2T_x$ and Au. After adding the composite layer, the sensitivity is improved. It demonstrates the great sensitivity of 276°/RIU as compared to existing SPR sensor. The performance parameter is calculated at optimized TiO_2 and SiO_2 layer thicknesses at 65 nm with 47 nm of Au thickness.

It is further analyzed that on using the $Ti_3C_2T_x$ layer over Au layer, the sensitivity is enhanced. The penetration depth is 197 nm has been analyzed by the normalized electric field distribution. The proposed SPR sensor can be used for the field of medical and environmental science.

References

1. Sun, Y., Cai, H., Qiao, X., Wang, X.: High-performance polarization control modulated surface plasmon resonance sensor based on monolayer graphene/Au-NPs architecture for detection of DNA hybridization. Meas. Sci. Technol. **30**(12), 125701 (2019)
2. Arwin, H., Poksinski, M., Johansen, K.: Total internal reflection ellipsometry: principles and applications. Appl. Opt. **43**, 3028 (2004)
3. Caruso, F., Rodda, E., Furlong, D.N., Niikura, K., Okahata, Y.: Quartz crystal microbalance study of DNA immobilization and hybridization for nucleic Acid sensor development. Anal. Chem. **69**, 2043–2049 (1997)
4. Kukkar, M., Mohanta, G.C., Tuteja, S.K., Kumar, P., Bhadwal, A.S., Samaddar, P., Kim, K.-H., Deep, A.: A comprehensive review on nano-molybdenum disulfide/DNA interfaces as emerging biosensing platforms. Biosens. Bioelectron. **107**, 244–258 (2018)
5. Bayraç, C., Eyidoğan, F., AvniÖktem, H.: DNA aptamer based colorimetric detection platform for Salmonella Enteritidis. Biosens. Bioelectron. **98**, 22–28 (2017)
6. Xie, N., Huang, J., Yang, X., He, X., Liu, J., Huang, J., Fang, H., Wang, K.: Scallop-inspired DNA nanomachine: a ratio metric nano thermometer for intracellular temperature sensing. Anal. Chem. **89**, 12115–12122 (2017)
7. Basak, C., Hosain, M.K., Sazzad, A.A.: Design and simulation of a high sensitive surface plasmon resonance biosensor for detection of biomolecules. Sensing Imaging **21**(1), 2 (2020)
8. Miyoshi, H., Suehiro, N., Tomoo, K., Muto, S., Takahashi, T., Tsukamoto, T., Ohmori, T., Natsuaki, T.: Binding analyses for the interaction between plant virus genome-linked protein (VPg) and plant translational initiation factors. Biochimie **88**(3–4), 329–340 (2006)
9. Hossain, M., Rana, M.: DNA hybridization detection based on resonance frequency readout in graphene on Au SPR biosensor. J. Sens. (2016)
10. Singh, M.K., Pal, S., Prajapati, Y.K., Saini, J.P.: Highly sensitive antimonene based SPR biosensor for miRNA detection. Mater. Today Proc. **28**, 1776–1780 (2020)
11. Singh, M.K., Pal, S., Verma, A., Prajapati, Y.K., Saini, J.P.: Highly sensitive antimonene-coated black phosphorous-based surface plasmon-resonance biosensor for DNA hybridization: design and numerical analysis. J. Nanophotonics **14**(4), 046015 (2020)
12. Hasib, M.H.H., Nur, J.N., Rizal, C., Shushama, K.N.: Improved transition metal dichalcogenides-based surface plasmon resonance biosensors. Condensed Matter. **4**(2), 49 (2019)
13. Zhu, J., Ke, Y., Dai, J., You, Q., Wu, L., Li, J., Guo, J., Xiang, Y., Dai, X.: Topological insulator overlayer to enhance the sensitivity and detection limit of surface plasmon resonance sensor. Nanophotonics (2019)
14. Moznuzzaman, M., Islam, M.R., Hossain, M.B., Mehedi, I.M.: Modeling of highly improved SPR sensor for formalin detection. Results Phys. **16**, 102874 (2020)
15. Pal, S., Verma, A., Prajapati, Y.K., Saini, J.P.: Sensitive detection using heterostructure of black phosphorus, transition metal di-chalcogenides and MXene in SPR sensor. Appl. Phys. A **126**(10) (2020). https://doi.org/10.1007/s00339-020-03998-1
16. Xue, T., Liang, W., Li, Y., Sun, Y., Xiang, Y., Zhang, Y., Bao, Q.: Ultrasensitive detection of miRNA with an antimonene-based surface plasmon resonance sensor. Nature Commun. **10**(1) (2019). https://doi.org/10.1038/s41467-018-07947-8
17. Kumar, R., Pal, S., Prajapati, Y.K., Saini, J.P.: Sensitivity enhancement of MXene based SPR sensor using silicon: theoretical analysis. SILICON **13**(6), 1887–1894 (2021)

18. Wu, L., You, Q., Shan, Y., Gan, S., Zhao, Y., Dai, X., Xiang, Y.: Few-layer $Ti_3C_2T_x$ MXene: a promising surface plasmon resonance biosensing material to enhance the sensitivity. Sens. Actuators B Chem. **277**, 210–215 (2018)
19. Xu, Y., Ang, Y.S., Wu, L., Ag, L.K.: High sensitivity surface plasmon resonance sensor based on two-dimensional MXene and transition metal dichalcogenide: a theoretical study. Nanomaterials **9**(2), 165 (2019)
20. Pal, S., Pal, N., Prajapati, Y.K., Saini, J.P.: Sensitivity analysis of surface Plasmon resonance biosensor based on Heterostructure of 2D BlueP/MoS_2 and MXene. In: Inamuddin, B.R., Ahamed, M.I., Asiri, A.M. (eds.) Layered 2D Advanced Materials and Their Allied Applications, pp. 103–129. Wiley (2020) https://doi.org/10.1002/9781119655190.ch5
21. Pandey, A.K., Hashemi, M.: Plasmonic sensor based on molybdenum trioxide-MXene heterojunction for refractive index sensing. Arab. J. Sci. Eng. 1–6 (2021)
22. Xu, J., Xiao, X., Stepanov, A.L., Ren, F., Wu, W., Cai, G., Zhang, S., Dai, Z., Mei, F., Jiang, C.: Efficiency enhancements in Ag nanoparticles-SiO_2–TiO_2 sandwiched structure via plasmonic effect-enhanced light capturing. Nanoscale Res. Lett. **8**, 73 (2013)
23. Maurya, J.B., Prajapati, Y.K., Singh, V., Saini, J.P.: Sensitivity enhancement of surface plasmon resonance sensor based on graphene–MoS_2 hybrid structure with TiO_2–SiO_2 composite layer. Appl. Phys. A **121**(2), 525–533 (2015)
24. Maurya, J.B., Prajapati, Y.K.: A novel method to calculate beam width of SPR reflectance curve: a comparative analysis. IEEE Sens. Lett. **1**(4), 1–4 (2017)

Performance Analysis of LiF and Graphene Based Long-Range SPR Sensor

Vipin Kumar Verma, Sarika Pal, Alka Verma, and J. P. Saini

Abstract In this paper we have proposed a novel long-range surface plasmon resonance sensor (LRSPR) based on LiF and graphene having single metallic layer. Angular interrogation method is used for the detection of biomolecules. Imaging sensitivity for conventional SPR sensor and LRSPR sensor has been calculated and compared. Also, the imaging sensitivity is analyzed with increasing layers of graphene. In our result imaging sensitivity of LRSPR sensor with single layer of graphene is found to be 867 RIU^{-1} while for the conventional SPR is 308 RIU^{-1}, which shows around three times enhancement. The detection accuracy of LRSPR with single layer of graphene is found to be 5.94 Deg^{-1}, which also shows around three times improvement than that of conventional SPR.

Keywords Graphene · Long Range Surface Plasmon Resonance Sensor (LRSPR) · Detection accuracy · Imaging sensitivity

1 Introduction

Surface plasmon resonance (SPR) sensors are those plasmonic sensors which are linked with use of surface plasmons (SPs) generated by using the interface of two different materials, having opposite nature of permittivity. To analyze the broad

V. K. Verma (✉)
KIET Group of Institutions, Delhi NCR, Ghaziabad 201206, U.P., India
e-mail: vipin.phd20@nituk.ac.in; vipinkumarverma060@gmail.com

V. K. Verma · S. Pal
National Institute of Technology Uttarakhand, Garhwal 246174, U.K., India
e-mail: sarikapal@nituk.ac.in; narensarru@gmail.com

A. Verma
Institute of Engineering and Rural Technology, Allahabad, Prayagraj 211001, U.P., India
e-mail: alkapra25@iert.ac.in; alkapra25@gmail.com

J. P. Saini
Netaji Subhas University of Technology, New Delhi 110078, India
e-mail: jps@nsut.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_41

variety of intermolecular interactions like receptor ligand contacts, SPR sensors have become a very important tool for the analysis. [1]. After comparing with conventional analysis methods used in sensing of intermolecular interactions, technology using SPR sensors can be categorized with high accuracy and sensitivity, label free technology, a technology in which interference can be avoided, no need to purify the sample to be analysed. These properties make this technology to use it widely for inspecting biochemistry medicine water, assuring the quality of food and monitoring the environment. In contemporary time, with the advancement of technology it has also raised its requirements for the better accuracy and sensitivity and upgrading of the bioassay methods. Therefore, it is very difficult to meet these kinds of properties with the conventional SPR, and the new modes of SPR sensors like long-range SPR (LRSPR), has been studied. LRSPR is basically different than that of conventional SPR in only the kind of propagation of plasmon, that takes place in LRSPR. It can be differentiated with the conventional SPR structure wise that it uses a dielectric buffer layer between metal and the substrate. It provides weaker confinement into the materials and hence allows to interact deeper with the sensing medium, which increases the penetration depth of electromagnetic field in the case of LRSPR and also strengthen the field intensities generated. This makes it to have better sensitivity and to reduce the full width half maximum (FWHM). Narrower FWHM leads to have greater detection accuracy, which was not possible with the conventional SPR. Apart from it there are many more choices for metal to use in the design of LRSPR other than gold and silver [2]. It is also observed that 2D transition metal dichalcogenide (TMDC) materials can also be a potential candidate for SPR sensing as the sensitivity with these materials increases with the number of TMDC layers instead of decreasing, which happens with the graphene [3]. For DNA detection a new structure of LRSPR is suggested which uses cytop layer between prism and gold layers. The suggested structure is Prism (SF10)/cytop/Au/water, which increases the sensitivity than that of a conventional SPR sensor [4]. SPR sensor used to deal with the sensing of toxicity involved with living cells shows lower sensitivity values. Using LRSPR sensor shows relevant improvement in the sensitivity in sensing the morphological changes in cells, due to deeper penetration of evanescent field. LRSPR shows 50% higher sensitivity than that of SPR sensor [5]. Ag is also a potential metal which shows narrower angular width of SPR curves and thus improves the detection accuracy. However, to prevent the oxidation of Ag layer, bi-metal layer combinations may be used by coating Ag with thin Au layer. It improves the performance of SPR sensor which was present when Au coating is not done. Hence, double metal (Ag/Au) layers can also be used for improved biosensing [6].

It is numerically demonstrated to use black phosphorus coated with graphene on the metal surface to detect the hybridization of DNA. The suggested SPR sensor uses Prism-Au-BP-Graphene-PBS Solution as its structure in Kretschmann configuration. The suggested SPR sensor shows sensitivity (125°/RIU), detection accuracy (0.95) and quality factor (13.62 RIU^{-1}) for complementary DNA detection [7]. Discussion on detection of hybridization of DNA with SF10 Prism-Au-ZnO-Graphene-PBS solution structure of SPR sensor. In the evaluation process the methodology used is angular interrogation and the wavelength of operation is 633 nm. The proposed sensor exhibits the sensitivity-141.9°/RIU, DA-0.64 Degree^{-1} and FoM-9.14 RIU^{-1} at 0 eV chemical potential of graphene at the room temperature. If the chemical potential of graphene is changed to 1.25 eV, then the proposed SPR sensor shows the maximum value of sensitivity of 156.33°/RIU [8]. The most intriguing property of TMDCs layers useful for sensing is their adsorption capabilities better to graphene and their rate of light adsorption which may be tuned by changing their layer numbers. Number of researchers have tried TMDCs along with other 2D materials to enhance the SPR sensor performance [9].

LRPSR sensors have become quite popular in biosensing, chemical sensing, biochemistry and food quality etc. measurements, due to its increased interaction with sensing medium, which leads the higher sensitivity [1–9]. In LRSPR sensors the reflectance curve is narrower than that of a conventional SPR sensor that is why the detection accuracy and FOM gets increased. Different dielectric buffer layers can be used to fabricate the LRSPR sensor [10]. In Kretschmann configuration the base of prism is coated with a metal. When a parallel polarized light is incident on it and the wave vector of incident light matches the SP wave vector, then maximum surface polaritons are generated, which is the fundamental for SPR sensing [11]. The drawback associated with the angle interrogation method is not allowing the simultaneous inspection of many biomolecular interactions. To monitor the parallel molecular interaction imaging SPR sensors have been suggested [12–17]. One more disadvantage of conventional SPR sensor is its broader reflectance curve, due to which the detection accuracy is poor. LRSPR imaging sensor is an effective way to improve the detection accuracy and sensitivity [18]. When a metallic layer is sandwiched between two dielectrics having approximately same refractive indices, then plasmon wave propagates along the metallic layer. In this type of structure, the evanescent waves penetrate deeper into the sensing medium leading to give narrower reflectance curve than that of a conventional SPR sensor [19–23]. Hence, we proposed here a LRSPR imaging sensor (Prism(2S2G)/LIF/Ag/Graphene/Sensing Medium) for biomolecule sensing. In most of the literature part Teflon and cytop are used as dielectric to generate surface plasmon. The drawbacks of using these dielectrics are with respect to high temperature operation, stability and cost. Therefore, in this work a try has been made to analyze the performance of SPR sensor which uses LiF (refractive index 1.391) as dielectric material.

2 Proposed Structure and Theoretical Modelling

The suggested structures for both the conventional and LRSPR sensor is shown in Fig. 1. Figure 1a shows the conventional structure of SPR sensor (Prism(2S2G)/Ag/Graphene/Sensing Medium) in which the structure is of 4 layers while the Fig. 1b shows the LRSPR structure (Prism(2s2G)/LiF/Ag/Graphene/Sensing Medium), which consists of 5 layers. In LRSPR structure LiF layer is sandwiched between prism and Ag layer, and which was absent in the conventional SPR. The configuration used for the analysis is

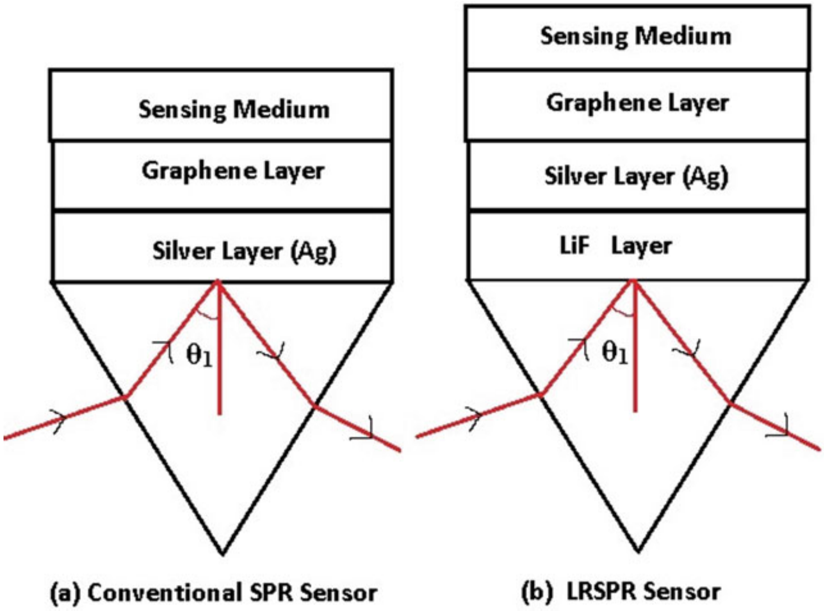


Fig. 1 Proposed structure of SPR and LRSPR sensor

Kretschmann configuration and the methodology used is angular interrogation. Refractive index for the sensing medium is taken as 1.33 and the wavelength at which all the parameters are analyzed is 633 nm. The refractive index of Ag metal can be found by using the Drude model [7].

$$n_{Ag} = \left(1 - \frac{\boldsymbol{\lambda}^2\boldsymbol{\lambda}_c}{\boldsymbol{\lambda}_c^2(\boldsymbol{\lambda}_c + \mathbf{i}\lambda)}\right)^{\frac{1}{2}} \tag{1}$$

where, $\boldsymbol{\lambda}_c$ and $\boldsymbol{\lambda}_p$ are representing collision and plasma wavelength respectively and $\boldsymbol{\lambda}$ is the wavelength of operation. The value of $\boldsymbol{\lambda}_c$ and $\boldsymbol{\lambda}_p$ for the Ag layers are 1.7614 $\times$ 10^{-5} m and 1.4541 $\times$ 10^{-7} m respectively at 633 nm. Graphene leads to easy attachment of biomolecules by forming pi bonds with the biomolecules. Hence the biomolecular interaction is enhanced by using the graphene layer on Ag layer, The thickness of monolayer graphene is 0.34 nm. Refractive index of graphene layer is calculated by using the following relation [8].

$$n_{Gra} = 3 + i\left(\frac{C_1}{3}\lambda\right) \tag{2}$$

Sensing medium refractive index rises after attachment of biomolecules on sensor surface. In this work the sensing medium with wide refractive index range (1.33–1.38) is chosen for analysis. To obtain the reflectivity of the reflected light is calculated by using the transfer matrix method developed for N-layers. It is efficient method without any approximations. Readers may find complete detail regarding mathematical modeling for reflectivity calculation from references [24].

3 Performance Parameters Evaluated for SPR/LRSPR

I. Sensitivity: Sensitivity is defined as the ratio of shift in resonance angle to shift of RI of sensing medium.

$$\text{sensitivity} = \frac{\Delta\theta_{\text{res}}}{\Delta n_s}\left({}^{\circ}/\text{RIU}\right) \tag{3}$$

II. Full width half maximum (FWHM): It is defined as the difference in the angles or angular width corresponding to 50% reflectivity in the reflection curve.

$$\text{FWHM} = \Delta\theta_{0.5}(\text{Degrees}) \tag{4}$$

III. Detection Accuracy (DA): Detection accuracy is defined as the inverse of FWHM with the unit degree inverse.

$$DA = \frac{1}{FWHM}\left(1/^{\circ}\right) \tag{5}$$

IV. Imaging Sensitivity: It is the ratio of the reflectivity shift to the RI shift of the sensing medium for a fixed angle.

$$\text{Imaging Sensitivity} = \frac{dR_p}{dn_s} \tag{6}$$

V. Figure of Merit (FOM): Figure of merit of a SPR sensor is defined as the ratio of sensitivity to FWHM.

$$\text{FOM} = \frac{\text{Sensitivity}}{\text{FWHM}} \tag{7}$$

4 Discussions on Simulation Results

According the structure mentioned, we defined the refractive indices of each layer in modeling section. Thereafter, optimization of LiF layer is done. The optimized thickness of LiF layer comes out to be 968 μm. Once we fixed the thickness of LiF layer the next optimization is done for the Ag layer thickness for both the conventional and LR SPR sensors. The Fig. 2a shows the variation reflectance for conventional SPR whereas Fig. 2b shows the variation of reflectance for LRSPR sensor. The optimization is done on the basis of minimum value of reflectance as it indicates the maximum coupling of SPs. For conventional SPR sensor, the optimized thickness of Ag layer is 48 nm and for the LRSPR sensor it is 30 nm, where the minimum reflectance is obtained as shown in Fig. 2a and b. For all the optimizations and calculations, we have used MATLAB software.

Figure 3a shows the angular sensitivity variation with the increasing number of graphene layer thickness. As shown in the Fig. 3a, the angular sensitivity is almost constant with and without the graphene layer and the value of angular sensitivity is equal to 40 deg/RIU as the number of graphene layers are increased. Hence only single graphene layer is enough for the conventional SPR sensor. Figure 3b shows the variation of detection accuracy with the increasing number of graphene layers. As per Fig. 3b the value of detection accuracy is small and it also decreases as the number of graphene layers are increased. The reason behind this is the damping effect of SPs on increasing the number of graphene layers, which leads to broader SPR curve and which in turn decreases the detection accuracy of SPR sensor.

To overcome the low detection accuracy shown by conventional SPR sensor, we tried to examine the performance of a LRSPR sensor by using LiF layer in structure between the prism and Ag layer. Figure 4a shows the variation of reflectance of conventional and LRSPR sensor, which implies that after using LiF layer in the structure the reflectance curve becomes narrower than that of conventional SPR sensor, which leads to higher detection accuracy. Figure 4b shows the variation of imaging sensitivity of conventional and LRSPR sensor.

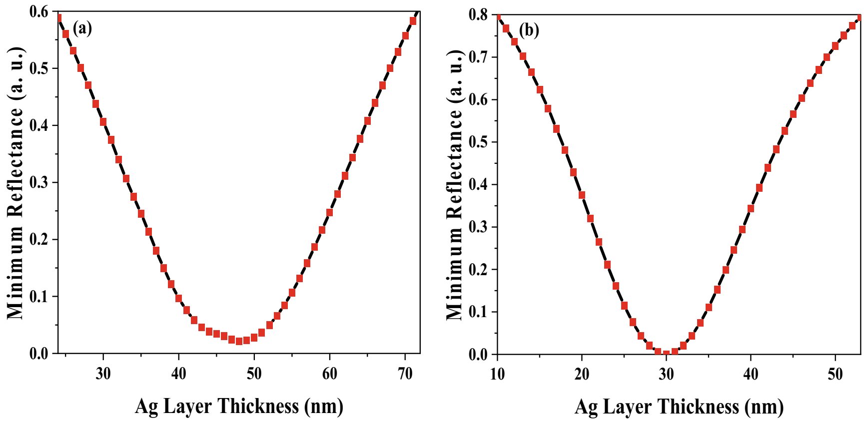


Fig. 2 Optimizing the Ag layer thickness in terms of minimum reflectance (Rmin.) for conventional SPR sensor (**a**) and LRSPR sensor (**b**)

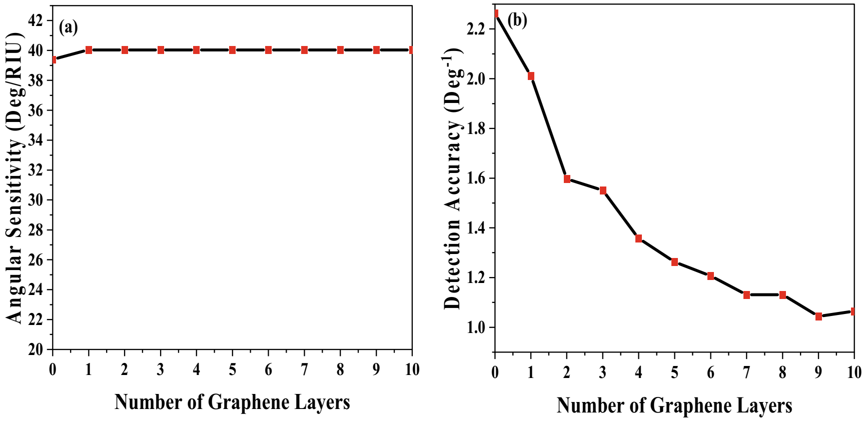


Fig. 3 Angular sensitivity (**a**) and detection accuracy (**b**) of conventional SPR sensor

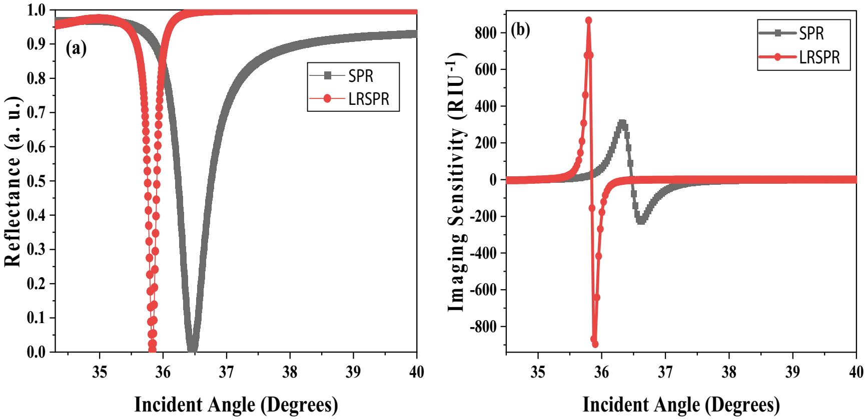


Fig. 4 SPR curves (**a**) and imaging sensitivity of conventional and long range SPR sensors (**b**)

As the Fig. 4b implies that the imaging sensitivity of LRSPR sensor is much more improved than that of conventional SPR sensor. These imaging sensitivities are plotted for the optimized values of thickness of Ag layer in both the cases of conventional and LRSPR sensors. The thickness of Ag layer for conventional SPR sensor is 48 nm while for the LRSPR is 30 nm. Thickness of graphene monolayer is 0.34 nm for both the curves. The maximum value of imaging sensitivity of LRSPR is 866.91 RIU^{-1} for LRSPR and 308.74 RIU^{-1} for conventional SPR sensor, which shows the improvement.

In Fig. 5a and b the variation of normalized electric field for the proposed structure at different interfaces of conventional and LRSPR sensor is shown. The field enhancement in sensing medium leads to the higher performance of SPR sensor. It is obvious from the results that in case of LRSPR sensor the field distribution is better than that of conventional SPR sensor. That is reason why the sensitivity of LRSPR sensor is better than the conventional SPR sensor. In Fig. 5c and d the 1D field plots for the conventional and LRSPR sensor is shown with the distance from the prism. The penetration depth evaluated for the conventional SPR sensor is 222.32 nm and for the LRSPR sensor is 261.6 nm. Hence, the interaction between the evanescent field and the biomolecules is more in case of LRSPR, which is the basic cause of enhanced sensitivity. Also, there is rise in electric field at the LiF interface and again there is rise of electric field at graphene interface after that there is an exponential decay of electric field in the sensing medium. All these field simulations are done with COMSOL Multiphysics software.

Figure 6a shows the variation of imaging sensitivity of proposed LRSPR with number of graphene layers, which decreases as the number of graphene layers are increased. It can be explained by the increasing the non-zero imaginary part of refractive index with the increasing numbers of graphene layers, which leads the damping of plasmons. In Fig. 6b resonance angle variation with the number of graphene layers is shown. It shows that the resonance angle increases with the number of graphene layers. Figure 6c shows that the minimum reflectance first decreases with single layer of graphene and on adding more layers of graphene it increases, which leads to broader SPR curve.

In Fig. 7a variation of detection accuracy of long range SPR sensor is plotted against the number of graphene layers. As from the plot it is evident that LRSPR shows improvement in terms of detection accuracy than that of conventional SPR sensor. However, the detection accuracy decreases with the increasing number of graphene layers. It can be explained with the broad SPR curves as the number of

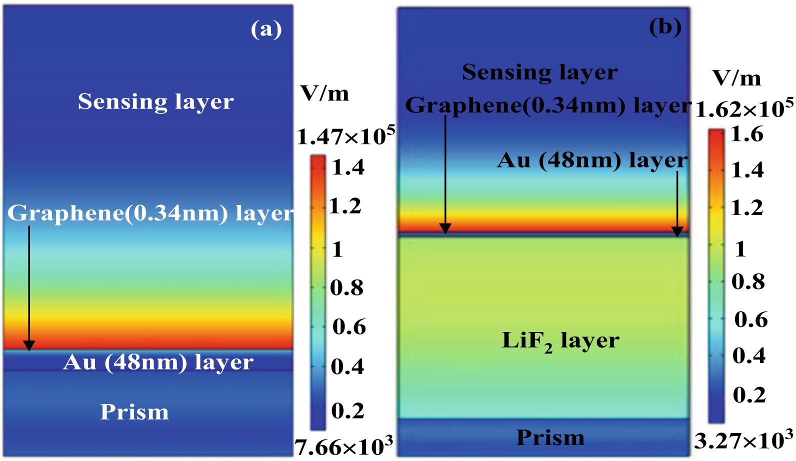


Fig. 5 **a-b** Variation of electric field at different interfaces of **a** conventional SPR and **b** LRSPR sensor. **c-d** Variation of electric field with the distance from the prism to analyte **c** conventional SPR and **d** LRSPR sensor

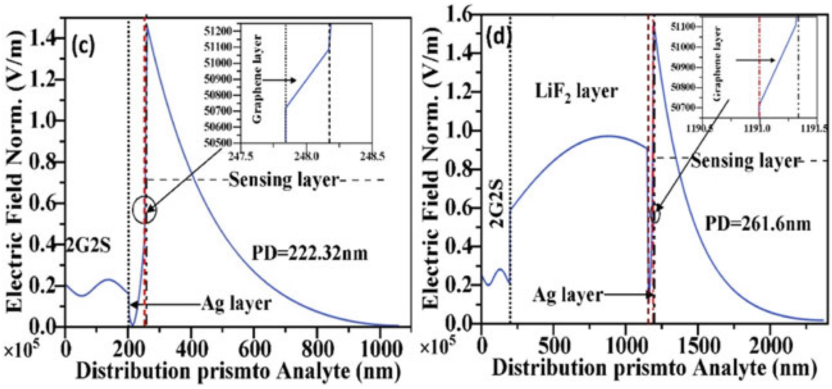


Fig. 5 (continued)

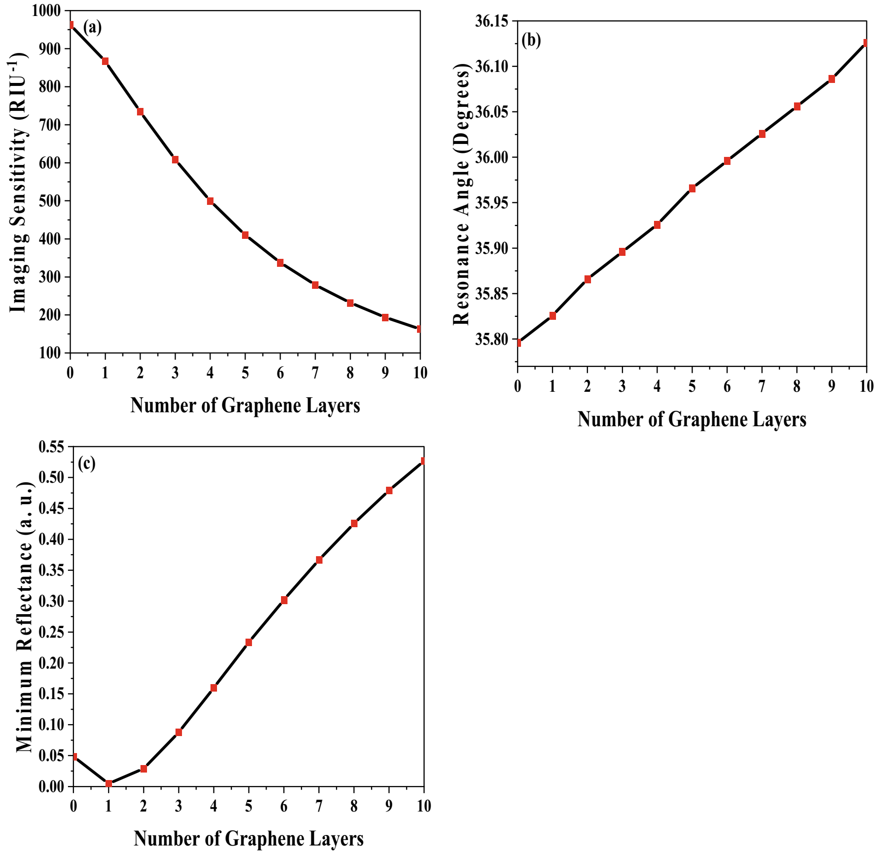


Fig. 6 Variation of imaging sensitivity (**a**), resonance angle (**b**) and minimum reflectance (**c**) with number of graphene layers for LRSPR

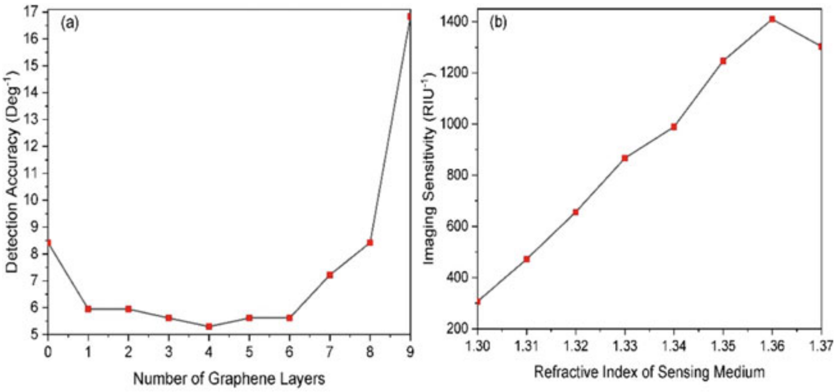


Fig. 7 Variation of detection accuracy with number of graphene layers (**a**) and imaging sensitivity with the refractive index of the sensing medium (**b**) for LRSPR

graphene layers are increased. In Fig. 7b imaging sensitivity variation with the refractive index of the sensing medium is shown. Firstly, the imaging sensitivity increases with increasing value of refractive index of the sensing medium up to 1.36, after that it starts decreasing. The maximum value of imaging sensitivity is approximately 1400 RIU^{-1} at 1.36 refractive index of the sensing medium.

5 Conclusion

The proposed structure of LRSPR shows the improvement in detection accuracy and imaging sensitivity than that of the conventional SPR sensor. The maximum value of imaging sensitivity at the refractive index of sensing medium is 866.91 RIU^{-1}. As the imaging sensitivity decreases with the number of graphene layers so only one graphene layer is sufficient for this structure. It is necessary to prevent the silver layer from oxidation and for efficient binding of biomolecules with carbon–carbon pi-stacking interactions.

References

1. Robelek, R.: Surface plasmon resonance sensors in cell biology: basics and application. Bioanal. Rev. **1**, 57–72 (2009). https://doi.org/10.1007/s12566-009-0005-y
2. Jing, J.-Y., Wang, W., Zhao, W.-M., Wang, B.-T.: Long-range surface plasmon resonance and its sensing applications: a review. Opt. Lasers Eng. **112**, 103–118 (2019)
3. Xu, Y., Hsieh, C.-Y., Wu, L., Ang, L.K.: Ultrasensitive and highly accurate long-range surface plasmon resonance biosensors based on two-dimensional transition metal dichalcogenides. arXiv:1709.08813v1 [physics.app-ph], 26 September 2017

4. Wark, A.W., Hye Lee, J., Corn, R.M.: Long-range surface plasmon resonance imaging for bioaffinity sensors. Anal. Chem. **77**, 3904–3907 (2005)
5. Chabot, V., Miron, Y., Grandbois, M., Charettea, P.G.: Long range surface plasmon resonance for increased sensitivity in living cell biosensing through greater probing depth. Sens. Actuators B **174**, 94–101 (2012)
6. Choi, S.H., Kyung, M.N.: Investigation on an application of silver substrates for sensitive surface plasmon resonance imaging detection. J. Opt. Soc. Am. A **27**(10), 2229 (2010)
7. Pal, S., Verma, A., Raikwar, S., Prajapati, Y.K., Saini, J.P.: Detection of DNA hybridization using graphene-coated black phosphorus surface plasmon resonance sensor. Appl. Phys. A **124**(5), 394 (2018)
8. Pal, S., Prajapati, Y.K., Saini, J.P.: Influence of graphene's chemical potential on SPR biosensor using ZNO for DNA hybridization. Opt. Rev. **27**(1), 57–64 (2020)
9. Pal, S., Verma, A., Saini, J.P., Prajapati, Y.K.: Sensitivity enhancement using silicon-black phosphorus-TDMC coated surface plasmon resonance biosensor. IET Optoelectron. **13**(4), 196–201 (2019)
10. Vala, M., Etheridge, S., Roach, J.A., Homola, J.: Long-range surface plasmons for sensitive detection of bacterial analytes. Sens. Actuators B Chem. **139**(1), 59–63 (2009)
11. Kretschmann, E., Rather, H.: Zeitschrift für Naturforschung A **23**, 2135 (1968)
12. Lee, H.J., Nedelkov, D., Corn, R.M.: Analyt. Chem. **78**, 6504 (2006)
13. Choi, S.H., Byun, K.M.: J. Opt. Soc. Am. A **27**, 2229 (2010)
14. Choi, S.H., Kim, Y.L., Byun, K.M.: Opt. Express **19**, 458 (2011)
15. Wong, C.L., Olivo, M.: Plasmonics **9**, 809 (2014)
16. Zeng, Y., Hu, R., Wang, L., Gu, D., He, J., Wu, S.Y., Ho, H.P., Li, X., Qu, J., Gao, B.Z.: Nanophotonics **6**, 1017 (2017)
17. Wark, A.W., Lee, H.J., Corn, R.M.: Anal. Chem. **77**, 3904 (2005)
18. Sarid, D.: Phys. Rev. Lett. **47**, 1927 (1981)
19. Alek, J.D., Kasry, A., Knoll, W.: Plasmonics **2**, 97 (2007)
20. Slavık, R., Homola, J.: Sens. Actuators B Chem. **123**, 10 (2007). 16
21. Nenninger, G., Tobiska, P., Homola, J., Yee, S.: Sens. Actuators B Chem. **74**, 145 (2001)
22. Chabot, V., Miron, Y., Grandbois, M., Charette, P.G.: Sens. Actuators B Chem. **174**, 94 (2012)
23. Berini, P.: Adv. Opt. Photonics **1**, 484 (2009)
24. Pal, S., Verma, A., Prajapati, Y.K., Saini, J.P.: Influence of black phosphorous on performance of surface plasmon resonance biosensor. Opt. Quant. Electron. **49**, 403 (2017)

Ultra Low Power Fully Differential Double Recycling Current Mirror OTA for Biomedical Applications

Nikhil Deo and Tripurari Sharan

Abstract This paper presents a low-voltage low-power fully differential bulk-driven operational transconductance amplifier (OTA). The input core of this OTA is based on the double recycling structure where input PMOS are bulk-driven. The flipped-voltage-follower adaptive biased input differential pairs operating in class AB mode ensures dynamic current enhancement along with increased slew rate. This OTA also incorporates partial-positive feedback structure at the current mirror load thus enhancing the effective transconductance. Upon simulation, the proposed OTA provided DC gain of 81.4 dB, unity gain frequency of 27.2 kHz, and phase margin of 71°. It utilizes ±0.25 V of dual power supply and dissipates 45 nW. These results validate its effectiveness as for low-voltage, low-power applications for low-frequency signal processing. Process corner simulations were also performed to evaluate the robustness of the proposed OTA's low-frequency voltage gain, unity gain frequency, and phase margin, against process variations. A SIMO biquad filter is also realized using the proposed OTA. This OTA is designed and simulated using UMC 180 nm standard CMOS process technology.

Keywords Low-power · Low-voltage · Subthreshold · Fully differential OTA

1 Introduction

With the rising demand of small portable electronic systems, such as medical implant devices, IoT sensor nodes, wearable devices, and other small portable systems [1], plenty of effort is made to design analog and mixed-signal circuits that can operate with low voltage of threshold voltage order of CMOS devices and ultra-low power budget (<50 nW) while providing acceptable performance [2].

Operational transconductance amplifier (OTA) is an important part of many analog and mixed-signal circuits [3], biomedical devices [4, 5], analog filter, and operational

N. Deo (✉) · T. Sharan
Department of Electronics and Communication Engineering, North Eastern Regional Institute of Science and Technology, Nirjuli 791109, Arunachal Pradesh, India
e-mail: nikh.do@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_42

amplifiers [6]. These OTAs often make a considerable part of mixed-signal IC's total power budget [7]. Thus, designing OTAs operating at low-voltage (LV) and low-power (LP) is important in many applications.

At low-supply voltage, output signal swing might degrade, as a minimum drain-source voltage needs to be maintained for proper operation of CMOS circuits. In such situation, often times bulk-driven (BD) technique has been preferred over the conventional gate-driven (GD) technique [8]. To limit the power consumed by an OTA to few tens of nano-watt, driving the MOS transistors in sub-threshold region instead of strong inversion region is preferred, as high g_m/Id ratio and thus high degree of energy efficiency can be obtained [8].

There are various types of OTA topologies, such as current mirror, folded cascode, telescopic etc., the decision to choose a particular topology depends on the application requirements. For an OTA which is required to operate at LV supply while consuming ultra-low-power, single stage current mirror OTA topology is a suitable candidate. Single-stage design is power efficient as compared to multistage design while current mirror OTA can provide very good rail-to-rail output swing [3, 9]. Thus, a LV-LP OTA can be successfully realized by a bulk-driven single stage current mirror OTA operating in subthreshold region. Figure 1 shows a conventional BD-current mirror OTA, as such the gain provided by this OTA in deep-submicron CMOS technology range from 20 to 40 dB, which insufficient for many applications [3, 9]. To improve the performance of BD-CM OTA a solution has been reported in [10], where the gain was increased by reducing the quiescent output current with the help of two additional voltage controlled current sources incorporated in the OTA which were controlled by an extra pair of differential pair with diode connected load. OTA designed using this technique is often referred to as recycling OTA structure [7].

To further enhance the performance of recycling OTA, a nested-current-mirror (NCM) single stage amplifier based on double recycling structure, was proposed in [11, 12], which achieved improved gain, unity gain frequency (UGF), and slew-rate (SR). For this NCM amplifier to be used for LV LP application there have to be some modification, from gate-driven it has to made bulk-driven, and also operating

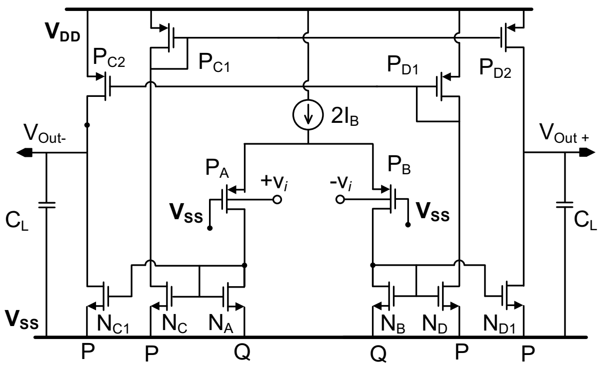


Fig. 1 A conventional fully differential current mirror OTA

in subthreshold instead of strong inversion. Gate-driven NCM amplifier in [12] operating in strong inversion region need to be modified to make them compatible for LV LP applications. Figure 2 shows a single stage BD-fully differential implementation of double recycling CM-OTA based on [12]. Doing so does make it suitable for LV LP application, now it can operate at low voltage and also provide good output swing, but still there is room for improvement.

The bulk-transconductance (g_{mb}) of a MOS transistor is substantially small as compared to gate-transconductance (g_m), generally g_{mb} is around 20 to 30% of g_m [6, 8], which results in bulk-driven OTA (BD-OTA) transconductance being lower than its gate-driven counterpart. Reduced effective transconductance also eventually lowers the gain, UGF, and SR of BD-OTA.

The remaining parts of this paper is organized in next four sections. Section 2 presents conventional and fully differential balanced current mirror OTA structures, Section 3 depicts simulation results of proposed OTA to demonstrate its performance features and its comparison with few existing recent works, and Sect. 4 presents its application for low-frequency analog signal filtering. Finally, Sect. 5 concludes this work.

2 Proposed OTA Structure

A conventional fully differential current mirror OTA is displayed in Fig. 1. The circuit schematic of the proposed OTA is depicted in Fig. 2. The proposed OTA which is a low-voltage low-power fully-differential double recycling current mirror OTA, has being referred to as LFD-DRCM OTA in this paper. As pointed out in previous section a conventional BD-OTA operating in (subthreshold region), provides low

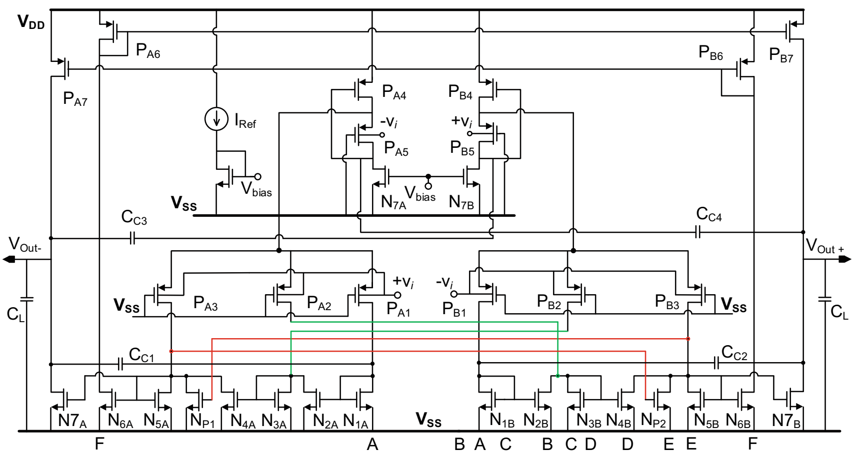


Fig. 2 Proposed fully differential low-voltage low-power (LFD-DRCM) OTA

transconductance, gain, UGF and slew rate, so some improvement in circuit structure is required. The core of the proposed OTA is based on the double recycling structure [11, 12], where compared to Fig. 1 the input differential pair (P_A–P_B) of Fig. 1 is split into three number of differential pairs, namely P_{A1}–P_{B1}, P_{A2}–P_{B2}, and P_{A3}–P_{B3} (see Fig. 2). Similarly the current mirror (N_A–N_C) and (N_B–N_D) in Fig. 1 is partitioned into multiple cur-rent mirrors bearing different current mirror ratios (N_{A1}–N_{A2}, N_{A3}–N_{A4}, N_{A5}–N_{A6}) and (N_{B1}–N_{B2}, N_{B3}–N_{B4}, N_{B5}–N_{B6}) with current mirror ratios as A, B, C, D, E, and F, respectively and their outputs are combined in a sequential manner that results in improved effective transconductance and output resistance, which in turn results in improvement in gain, slew rate, and UFG as well.

To further improve the performance, instead of the constant current biasing as provided in Fig. 1, an adaptively biased class AB differential input based on flipped voltage follower (FVF) is utilized which improved the current drivability of the input core [13–15].

In [16], constant current shunting is done for the outermost current mirror pair of the recycling OTA to achieve higher performance. In our design, instead of using constant biasing for the shunting transistor pairs at the outermost current mirror pair (N_{A5}–N_{A6} and N_{B5}–N_{B6}), we have adopted the cross coupled self-biasing for the shunting transistor pairs, this technique is referred to as partial-positive feedback (PPF) [17]. Partial positive feedback, when employed with care, so as to not make OTA unstable, has shown to increase the transconductance of an OTA [17–22]. Thus, in the LFD-DRCM OTA, a PPF circuit is placed across the last current mirror load. Though, the fully differential version of OTA requires a CMFB network to set the DC reference signal at its both of the output nodes. However, for the symmetric or balanced version of fully differential current mirror OTA the DC reference voltage at output nodes is pre-defined. So, the CMFB network is not mandatory for balanced current mirror fully differential OTA and it is not used in this design [22].

2.1 Differential Open-Loop Voltage Gain of LFD-DRCM

For circuit analysis, the half-circuit of the input portion of proposed OTA is shown in Fig. 3. The bulk-voltage of P_{A5} is $-v_i$ and the bulk-voltage of P_{A1}, P_{A2}, and P_{A3} is $+v_i$.

The small signal model of half-circuit in Fig. 3 is shown in Fig. 4 depicts, where drain of all the PMOS in the signal path are short circuited.

In Fig. 4, G_{mbe} and G_{me} denote the total equivalent bulk transconductance and gate transconductance of P_{A1}, P_{A2}, and P_{A3}, also the input PMOS are in parallel having same gate, bulk, and source voltages of V_{SS}, v_i, and v_x. Further, they provide an equivalent degeneration resistor, $Rs \approx \frac{1}{(g_{mA1}+g_{mA2}+g_{mA3})}$ while G_{me} is given by (1) and G_{mbe} are by (2).

$$G_{me} = g_{mA1} + g_{mA2} + g_{mA3} \tag{1}$$

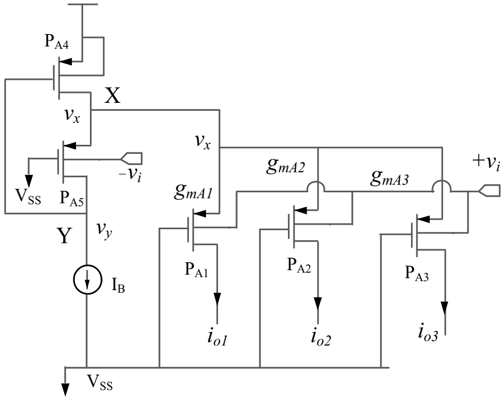


Fig. 3 Schematic of half-circuit of differential input pair and FVF

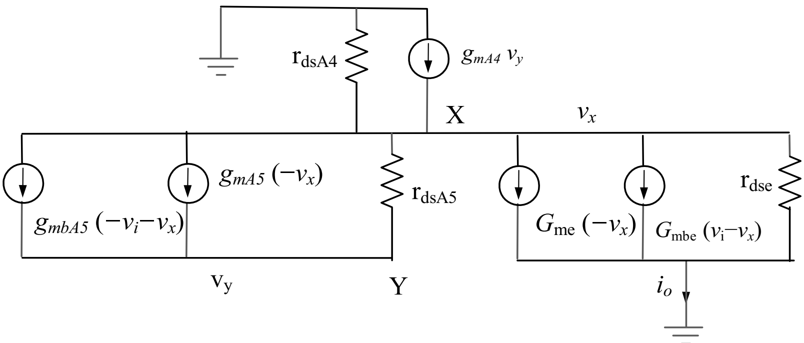


Fig. 4 Small signal ac model of Fig. 3

$$G_{mbe} = g_{mbA1} + g_{mbA2} + g_{mbA3} \tag{2}$$

In [19] it is shown that equivalent transconductanceof adaptive FVF and P_{A1}, P_{A2}, and P_{A3} can be obtained as (3)

$$G'_m = \frac{i_0}{v_i} = \frac{g_{mA4}v_y}{v_i} = \left[G_{mbe} + G_{me}\left(n_p - 1\right)\right] = 2G_{mbe} \tag{3}$$

n_p denotes the input PMOS subthreshold slope, while i_o is the gross output signal current, i_o where $i_o = (i_{o1} + i_{o2} + i_{o3})$.

Next, taking in the account the effect of resistance R_S, as reported in [19, 22], the transconductance gain of (3) now becomes G_m as shown in (4).

$$G_m = \frac{2G_{mbe}}{\left(1 + n_p g_{mA5} R_s\right)} \tag{4}$$

Further, LFD-DRCM has PPF circuit across at the third current mirror load transistors N_{5A} and N_{5B}, which will enhances the transconductance of that branch by factor of $1/(1-K)$ [18, 20], here K represents $(W/L)_{NP1}/(W/L)_{N5A}$. To avoid instability K is kept less than 1, here it is taken as 0.91. The effective small-signal transconductance (G_{meff}) of LFD-DRCM is represented by (5)

$$G_{meff(LFD-DRCM)} = 2\left[\frac{2}{\left(1 + n_p g_{mA5} R_s\right)} \frac{F}{E} \frac{1}{(1-K)} \left[\frac{D}{C}\left(g_{mbPB2} + \frac{B}{A} g_{mbPA1}\right) + g_{mbPA3}\right]\right] \quad (5)$$

The output resistance of proposed OTA is given by the parallel connection of small signal output resistance r_{oPB7} and r_{oN7B} of P_{B7} and N_{7B}, as depicted by (6).

$$R_{out} = r_{oPB7} || r_{oN7B} \quad (6)$$

The overall open-loop gain A_V is given by (7)

$$\begin{aligned} Av(LFD - DRCM) &= G_{meff} \times R_{out} \\ &= 2\left[\tfrac{2}{(1+n_p g_{mA5} R_s)} \tfrac{F}{E} \tfrac{1}{(1-K)} \left[\tfrac{D}{C}\left(g_{mbPB2} + \tfrac{B}{A} g_{mbPA1}\right) + g_{mbPA3}\right]\right][r_{oPB7} || r_{oN7B}] \end{aligned} \quad (7)$$

3 Simulation Results

The LFD-DRCM was simulated using UMC 180 nm CMOS process technology with rail-to-rail power supply of 0.5 V (±0.25 V dual power supply). The dual power supply eliminates the need of common-mode voltage (V_{CM}), as it becomes zero in this case. The bias current was set to 13 nA for proposed OTA. Table 1 outlines the size of transistors used. To maintain the stability and phase margin, the compensation capacitors used were $C_{C1} = C_{C2} = 0.4$ pF and $C_{C3} = C_{C4} = 0.6$ pF.

3.1 Simulation of Open Loop AC Response

For the fully differential LFD-DRCM OTA, the AC response simulations has been done for a differential load capacitor of 15 pF. Figure 5 depicts the open loop differential gain and phase response of the proposed OTA. A DC gain of 81.4 dB, unity gain frequency of 27.3 kHz, and PM of 71° is achieved while dissipating 45 nW of power.

Table 1 Transistors aspect ratios

Device	W/L (μm/μm)	Device	W/L (μm/μm)
P_{A1}, P_{B1}	3.6/0.6	N_{1A}, N_{1B}	2.4/1.2
P_{A2}, P_{B2}	7.2/0.6	N_{2A}, N_{2B}	3.6/1.2
P_{A3}, P_{B3}	8.6/0.6	N_{4A}, N_{4B}	3.6/1.2
P_{A6}, P_{B6}, P_{A7}, P_{B7}	9.8/0.6	N_{3A}, N_{3B}	1.2/1.2
P_{A4}, P_{B4}	20/1	N_{5A}, N_{5B}	2.4/1.2
P_{A5}, P_{B5}	10/0.5	N_{6A}, N_{6B}	6.0/1.2
N_{P1}, N_{P2}	2.2/1.2	N_{7A}, N_{7B}	6.0/1.2
N_B	2/1	N_{8A}, N_{8B}	2/1

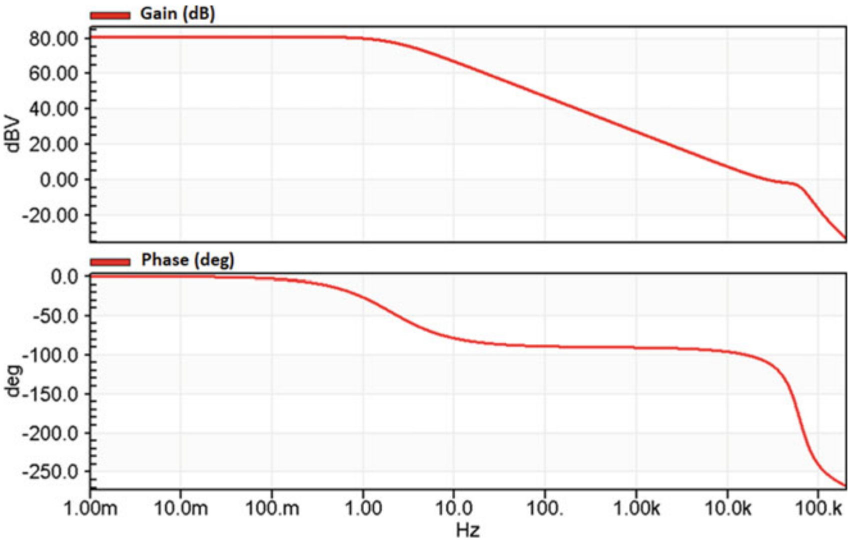


Fig. 5 Simulated differential gain and phase AC response

3.2 Unity Gain Response

A widely use method to assess various distortions introduced by OTA is to look at the Unity gain response of the OTA, where OTA is configured to act like a voltage follower. Figure 6 shows the differential unity gain configuration. It can provides information about noise and non-linearity generated by an OTA. Under unity gain the pulse transient response, DC sweep response, and total harmonic distortion (THD) analysis have been performed.

The large signal response is obtained for a square input signal of 400 mV peak-to-peak voltage and 1 kHz signal frequency. The large signal transient response is displayed in Fig. 7, the output response do not have any peak overshoots or

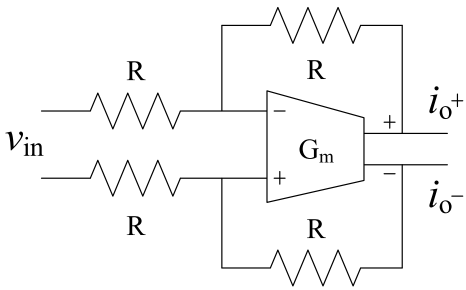


Fig. 6 Differential unity gain circuit

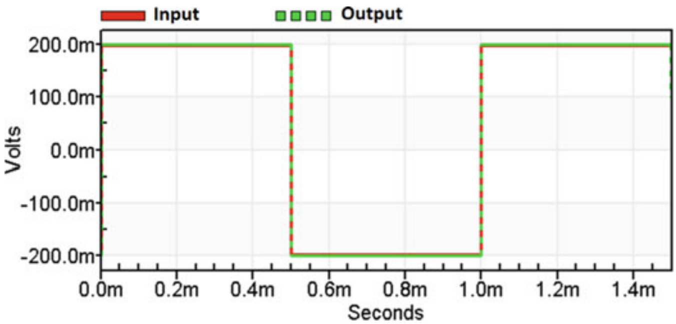


Fig. 7 Large single response to a square wave input of ±200 mV

oscillations thus producing good large signal response with an average SR of 750 V/ms.

Sinusoidal transient signal response (Fig. 8), has also been performed for this OTA by providing input sinusoidal signal of 1 kHz frequency having a peak-to-peak voltage of 1 V. To obtain the total harmonic distortion (THD) a parametric sweep is done by varying peak amplitude from 50 to 500 mV. The THD remain in between −47.4 to −48.8 dB, as depicted in Fig. 9, and for the rail-to-rail input signal swing THD obtained is −48.5 dB.

3.3 Simulation Demonstrating Effect of PVT Variations

Three factors that are considered important in terms of performance variation process voltage and temperature (PVT) variations. The process variation represents the deviation in this semiconductor fabrication process. All the five process corner were considered to analyze the effect of process variation upon gain, phase and UGF. Figure 10 shows the gain and phase corner response under five process corners namely

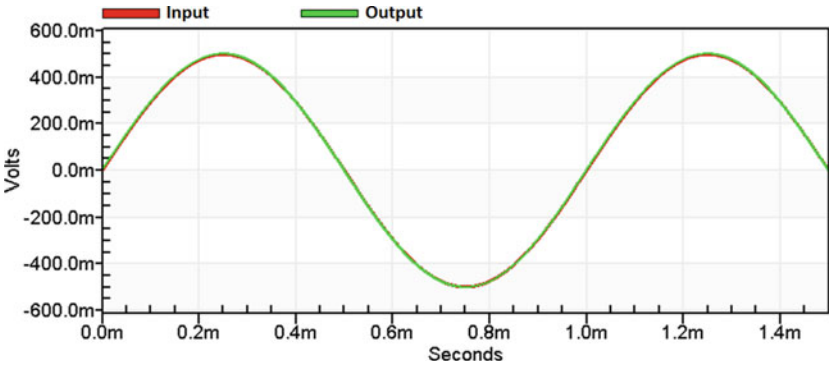


Fig. 8 Sinusoidal Transient response

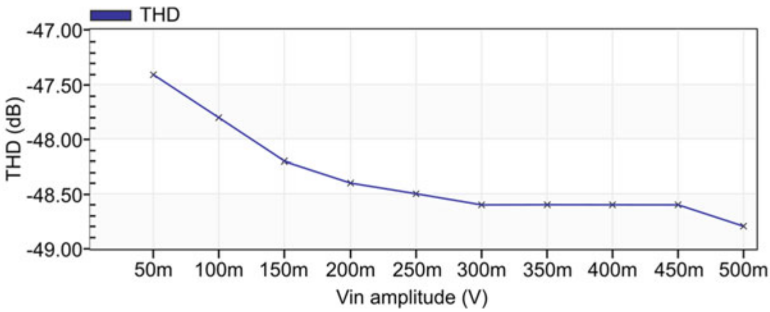


Fig. 9 THD for sinusoidal input of 1 kHz and amplitude varying from 50 to 500 mV

typical-typical (TT), slow-slow (SS), fast-fast (FF), slow-fast (SF), and fast-slow (FS).

Table 2 displays DC gain, phase margin and UGF obtained under five-process corners, the table also shows the effect of temperature and power supply variations on DC gain, PM, and UGF.

Voltage and temperature variation also need to be kept in consideration, so to assess their effect on DC gain, PM, and UGF the power supply voltage variation considered is ±10% of power supply voltage, whereas two temperature conditions are considered, one is low temperature of −10 °C and another high temperature of +80 °C.

Proposed OTA provided 99.6 dB common-mode rejection ratio (CMRR) in (single ended mode) at 10 Hz frequency, CMRR above 70 dB is preferred for good CM signal rejection. The input-referred noise (*inoise*) obtained at 1 kHz signal frequency is 1.01 $\mu V/\sqrt{Hz}$ for the proposed OTA.

Two commonly used Figure of Merits (FoM) shown in (8) and (9), have been used to examine OTA performance with respect to some of the other recently proposed low-power low-voltage OTA's. The FoM_S in (8) evaluates the small signal performance

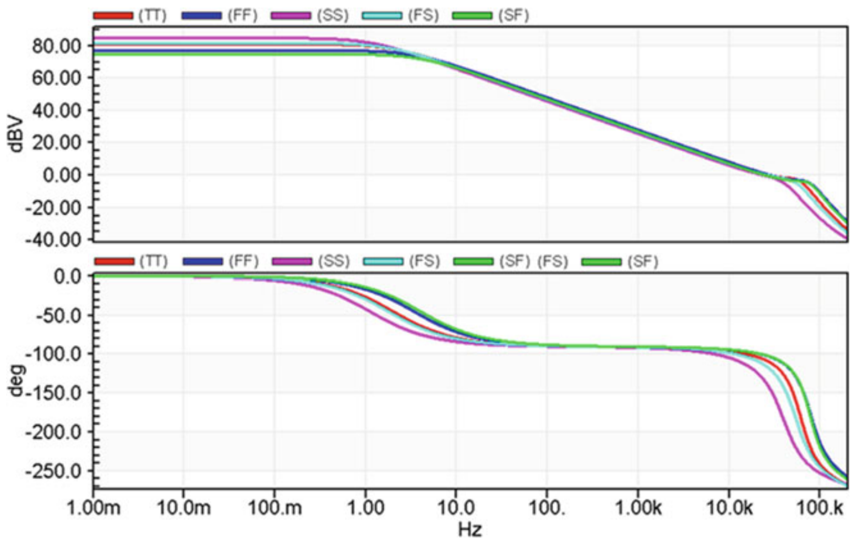


Fig. 10 Gain, phase response at five corner process

Table 2 Performance under five corner, small voltage variation, and temperature variation

Parameters	Five process corner					Power supply level		Temperature	
	TT	FF	SS	FS	SF	(V_{DD} + $\lvert V_{SS}\rvert$) − 10%	(V_{DD} + $\lvert V_{SS}\rvert$) + 10%	−10 °C	+80 °C
DC gain (dB)	81.4	77.1	84.8	81.6	74.9	80.4	81.1	83.5	73.4
UGF (kHz)	27.2	28.0	23.8	26.4	24.9	26.1	28.3	28.8	27.4
Phase margin (°)	71	78	48	64	77	69	71	62	76

of the amplifier, while FoM_L in (9) examines the large signal performance, both FoM also take into account the load capacitance and power consumed.

$$FoM_s = \frac{Gain(dB)UGF(kHz)C_L(pF)}{Power(nW)} \tag{8}$$

$$FoM_L = \frac{SR_{avg}(V/ms)C_L(pF)}{Power(nW)} \tag{9}$$

Table 3 shows the overall performance of the proposed OTA. Figure 11 depicts the layout of the proposed OTA covering an area of 8,075 μm^2.

Table 4 lists and compares the LFD-DRCM with few recent LV LP OTAs [23–25]. From Table 4, it can be observed that the proposed OTA provides considerably higher open loop gain, and lower power consumption as compared to other OTAs referred in this Table. The proposed OTA also achieved highest performance in terms of FoM_L and *FoMs*.

Table 3 Overall performance of the proposed OTA

Parameters	Proposed OTA
Technology node (μm)	0.18
Capacitance Load (pF)	15
Supply voltage	±0.25
DC gain (dB)	81.4
PM (deg)	71
UGF (kHz)	27.2
Power dissipation (nW)	45
CMRR (dB)(@10 Hz) (Single Ended)	99.6 dB
Average slew rate (V/ms)	750
THD (dB) (400Vpp @1 K)	−48.6
Input referred noise at 1 kHz (μV/√Hz)	1.01

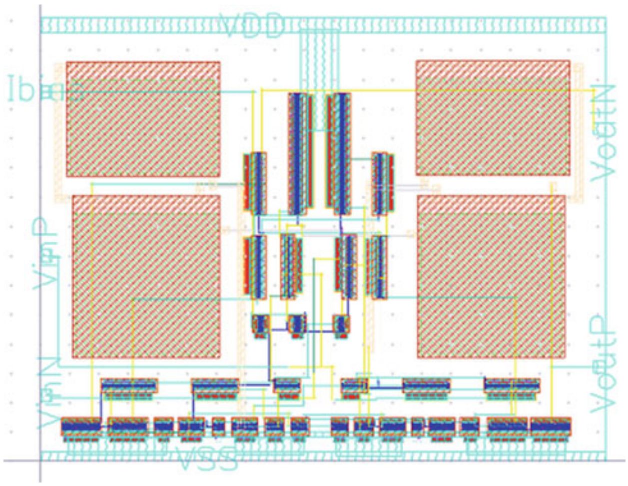


Fig. 11 Layout of proposed OTA

4 FD-OTA Based MISO Filter

A fully differential OTA (FD-OTA) based Multiple-input Single-output (MISO) type biquadratic filter structure is described in [25]. It is further used to validate application of the FD-OTA recently reported in [26]. This MISO bi-quadratic filter is designed using two numbers of proposed FD-OTA and two capacitors (see Fig. 12) to validate the usefulness of this work.

The Fig. 12 presents multiple inputs, say 4-inputs named V_1 through V_4 and two output nodes which are named as V_{o1} and V_{o2}. The routine nodal analysis can

Table 4 Comparison with few other recent OTAs

Parameters	This work	[23]		[24]	[22]
Power supply (V)	±0.25	0.4	0.6	0.6	0.5
Year	2021	2020	2020	2019	2017
Technology (nm)	180	130	130	180	180
Capacitive Load (pF)	15	6	6	15	15
DC gain (dB)	81.4	33	43.9	71	70.4
PM (deg)	71	90	88.8	74	54
UGF (kHz)	27.2	33.14	112.9	18.2	9.24
Power (nW)	45	430	1800	144	64
CMRR (dB)	99.6 @10 Hz			201 @10 Hz	106 @1 Hz
SR, avg (V/ms)	750	181	845	6.6	967
THD (dB)	−48.6*	–	–	–	– 52.4
Input referred noise (μV/$\sqrt{}$Hz)	1 @ 1 kHz	–	–	0.25@ 0.1 Hz	2.5 @ 1 kHz
FoM_S	738.02	15.25	16.52	134.2	152.06
FoM_L	250	2.52	2.81	0.69	226.64
Area (μm)^2	8075	–	–	16,000	–

* Simulation done with 1 kHz sinusoidal signal of 250 mV peak amplitude

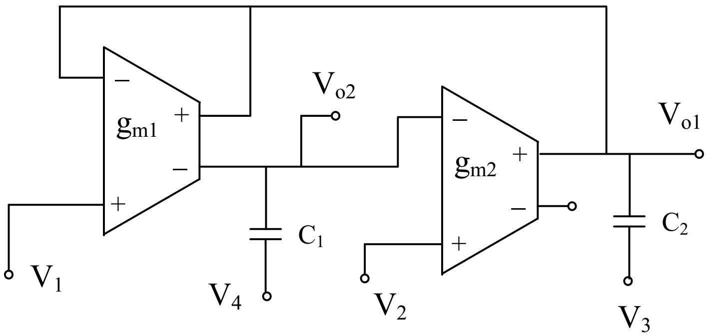


Fig. 12 Two-FD-OTA based MISO biquadratic filter

be performed to achieve the transfer functions related to these two output nodes as shown in (10) and (11).

$$V_{o1} = \frac{s^2 C_1 C_2 V_3 + sC_1(g_{m1}V_1 + g_{m2}V_2 - g_{m2}V_4) + g_{m1}g_{m2}V_1}{s^2 C_1 C_2 + sC_1 g_{m1} + g_{m1}g_{m2}} \tag{10}$$

$$V_{o2} = \frac{s^2 C_1 C_2 V_4 + sg_{m1}(C_1 V_4 + C_2 V_3 - C_2 V_1) + g_{m1}g_{m2}V_2}{s^2 C_1 C_2 + sC_1 g_{m1} + g_{m1}g_{m2}} \tag{11}$$

Design statements of this Biquadratic filter can be described in the following statements [25]:

(1) If $V_1 = V_2 = V_4 = 0$ V and input signal is connected to the input node V_3 a non-inverting high-pass (HP) and band-pass (BP) filter responses is obtained at V_{o1} and V_{o2} nodes, respectively.
(2) If $V_1 = V_3 = V_4 = 0$ V and input signal is connected to the input node V_2 a non-inverting BP and low-pass (LP) filter response are realized on V_{o1} & V_{o2} nodes, respectively.
(3) If $V_1 = V_2 = V_3 = 0$ V and input signal is connected to the input node V_4 an inverting BP filter responses are realized at V_{o1} node.
(4) If $V_2 = V_3 = 0$ V while $g_{m1} = g_{m2}$, $C_1 = C_2$ and $V_1 = V_4 =$ input signal voltage a non-inverting LP response is realized on V_{o1} and non-inverting HP is generated on V_{o2} node.
(5) If $V_2 = 0$ V while $g_{m1} = g_{m2}$, $C_1 = C_2$ and $V_1 = V_3 = V_4 =$ input signal voltage a non-inverting notch response is realized at V_{o1}.
(6) If $V_2 = 0$ V while $g_{m2} = 2\ g_{m1}$, $C_1 = C_2$ and $V_1 = V_3 = V_4 =$ input signal voltage a non-inverting all-pass response is realized at V_{o1}.
(7) If $V_2 = V_3 = V_4 = 0$ V and input signal is connected to the input node V_1, an inverting type of BP filter response is obtained at V_{o2}.
(8) If $V_3 = 0$ V and $C_1 = C_2$ and $V_1 = V_2 = V_4 =$ input signal voltage a non-inverting notch response is realized at V_{o2}.
(9) If $V_3 = 0$ V and $C_2 = 2C_1$ and $V_1 = V_2 = V_4 =$ input signal voltage a non-inverting all-pass filter response is realized at V_{o2}.

To confirm the usability of the designed FD-OTA, high-pass filter response has been realized as per the statement-1 and low-pass as well as band-pass filter functions are realized as per the statement-2 and notch filter response has been simulated as per the statement-8.

The comparison of (10), and (11) with second order generic standard transfer function shown in (12), the angular central/notch frequency (ω_o) is given by (13) while the quality factor (Q) can be expressed as (14).

$$T(s) = \frac{Vo(s)}{Vin(s)} = \frac{s^2 + a_1 s + \omega_o^2}{s^2 + (\omega_o/Q)s + \omega_o^2} \tag{12}$$

where V_o and V_{in} are output and input voltage, and a_1 is the filter coefficient which determines the pass band gain (A_o) of band-pass response as $A_o = a_1\ Q/\omega_o$.

$$\omega_o = \sqrt{\frac{g_{m1} g_{m2}}{C_1 C_2}} \tag{13}$$

$$Q = \sqrt{\frac{C_2 g_{m2}}{C_1 g_{m1}}} \tag{14}$$

Thus, its quality factor can be adjusted by choosing the ratio of C_2 and C_1, for a selected value of g_{m1} and g_{m2}.

To simulate the LP, HP, BP, and notch filter response the C_1 and C_2 have been chosen to be 100 pF and realized filter functions are shown in Figs. 13 and 14.

The LP filter configuration provided −3 dB corner frequency of 4.3 kHz whereas the HP filter configuration has realized −3 dB corner frequency of 3.8 kHz dissipating a total power of 90 nW, since two-OTAs are used in each circuit. The low-pass filter sinusoidal response for a sine wave input of 200 mV amplitude with frequency ranging from 10 to 100 Hz offered THD of −54 to −37 dB, while high pass filter sinusoidal response for amplitude and frequency of 0.2 V and 10 kHz offered THD of −40 dB. The band-pass response offered the lower and upper corner frequencies of 2.2 kHz and 7.32 kHz, respectively as depicted in Fig. 13. Its bandwidth (BW) is 5.21 kHz, central frequency $f_o = 3.9$ kHz, and quality factor $Q = f_o / BW = 0.74$.

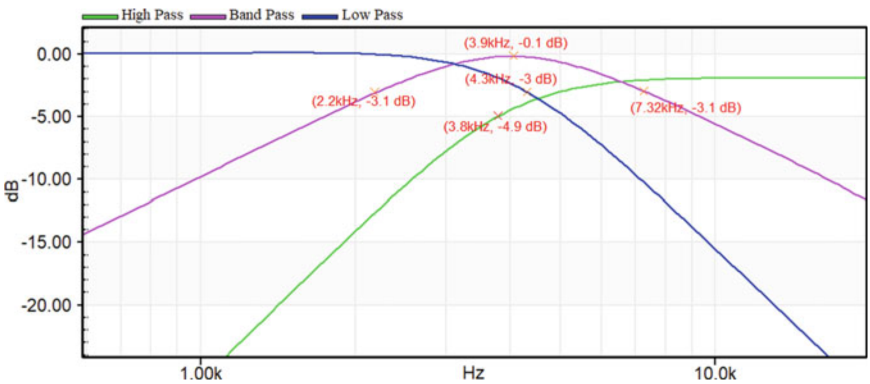


Fig. 13 Low-pass, high-pass and band-pass filter responses

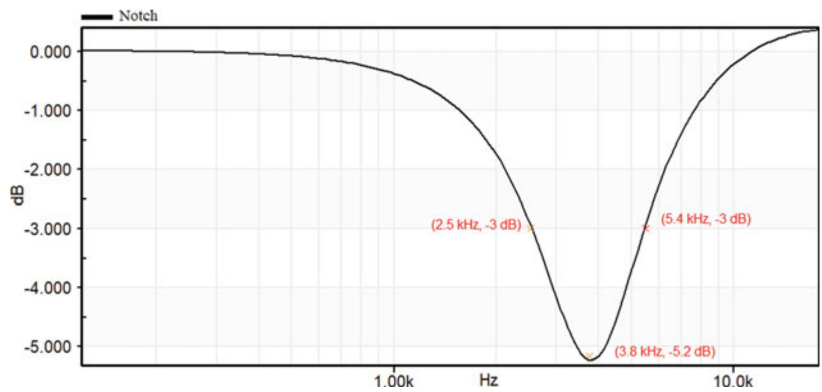


Fig. 14 Notch-filter response

The Fig. 14 shows the notch response, which is realized by setting the $C_1 = C_2 =$ 100 pF and offered the notched frequency $f_N = 3.8$ kHz, notch-bandwidth of 2.9 kHz and quality factor $Q = f_N/BW = 3.8/2.9 = 1.31$. The notch filter is used to remove out narrow band of un-desired interferences.

5 Conclusion

This paper has presented a LV LP fully differential OTA, which is bulk-driven and operates in sub-threshold region. To achieve improved characteristics, in this work double recycling input core with adaptive biasing utilizing FVFs is implemented. It also incorporates PPF at the final current mirror load. The simulation results obtained are DC gain of 81.4 dB, PM of 71°, SR of 750 V/ms, UGF of 27.2 kHz while dissipating substantially less power of 45 nW. The proposed LFD-DRCM OTA in comparison to OTAs in Table 4 gives high FoMs and FoM_L. Finally, a MISO Biquadratic Filter is also implemented to obtain LP, HP, BP and notch responses usable for bio-signals and sub-audio range of analog signal filtering.

References

1. Cabrera-Bernal, E., Pennisi, S., Grasso, A.D., Torralba, A., Carvajal, R.G.: 0.7-V three-stage class-AB CMOS operational transconductance amplifier. IEEE Trans. Circ. Syst. I Regular Pap. **63**(11), 1807–1815 (2016)
2. Blalock, B.J., Allen, P.E., Rincon-Mora, G.A.: Designing 1-V op amps using standard digital CMOS technology. IEEE Trans. Circ. Syst. II Analog Digit. Sig. Process. **45**(7), 769–780 (1998)
3. Yao, L., Steyaert, M., Sansen, W.: A 1-V 140-μW 88 dB audio sigma-delta modulator in 90-nm CMOS. IEEE J. Solid State Circ. **39**(11), 1809–1818 (2004)
4. Harrison, R.R., Charles, C.: A low-power low-noise CMOS amplifier for neural recording applications. IEEE J. Solid State Circ. **38**(6), 958–965 (2003)
5. Sharan, T., Bhadauria, V.: Sub-threshold, cascode compensated, bulk-driven OTAs with enhanced gain and phase-margin. Microelectron. J. **54**, 150–165 (2016)
6. Kulej, T.: 0.5-V bulk-driven OTA and its applications. Int. J. Circ. Theory Appl. **43**(2), 187–204 (2015)
7. Assaad, R.S., Silva-Martinez, J.: The recycling folded cascode: a general enhancement of the folded cascode amplifier. IEEE J. Solid State Circ. **44**(9), 2535–2542 (2009)
8. Ferreira, L.H.C., Sonkusale, S.R.: A 60-dB gain OTA operating at 0.25-V power supply in 130-nm digital CMOS process. IEEE Trans. Circ. Syst. I Regular Pap. **61**(6), 1609–1617 (2014)
9. Yao, L., Steyaert, M., Sansen, W.: A 0.8-V, 8-μW, CMOS OTA with 50-dB gain and 1.2-MHz GBW in 18-pF load. In: Proceedings of 29th European Solid-State Circuits Conference, Estoril, Portugal, pp. 297–300 (2004). https://doi.org/10.1109/esscirc.2003.1257131
10. Roh, J.: High-gain class-AB OTA with low quiescent current. Analog Integr. Circ. Sig. Process **47**(2), 225–228 (2006)
11. Yan, Z., Mak, P.I., Martins, R.P.: Double recycling technique for folded-cascode OTA. Analog Integr. Circ. Sig. Process **71**(1), 137–141 (2012)

12. Yan, Z., Mak, P.I., Law, M.-K., Martins, R.P., Maloberti, F.: Nested-current-mirror rail-to-rail-output single-stage amplifier with enhancements of DC gain, GBW and slew rate. IEEE J. Solid State Circ. **50**(10), 2353–2366 (2015)
13. Carvajal, R.G., Angulo, J.R., Martin, A.J.L., Torralba, A., Galan, J.A.G., Carlosena, A., Chavero, F.M.: The flipped voltage follower: a useful cell for low-voltage, low-power circuit design. IEEE Trans. Circ. Syst. I Regular Pap. **52**(7), 1276–1291 (2005)
14. Yodtean, A., Thanchayanont, A.: Sub 1–V highly linear low power class AB bulk driven tunable CMOS transconductor. Analog Integr. Circ. Sig. Process **75**(3), 383–397 (2013)
15. Lv, X., Zhao, X., Wang, Y., Jia, D.: Super class AB-AB bulk-driven folded cascode OTA. Integr. VLSI J. **63**, 196–203 (2018)
16. Attili, B., Mahmoud, S.A.: Optimizing the performance of a low power – area efficient OTA design that is based on hybrid current shunting technique. In: 2018 International SoC Design Conference (ISOCC), pp. 95–96 (2018)
17. Wang, R., Harjani, R.: Partial positive feedback for gain enhancement of low-power CMOS OTAs. Analog Integr. Circ. Sig. Process **8**(1), 21–35 (1995)
18. Carrillo, J.M., Torelli, G., Pérez-Aloe, R., Carrillo, J.F.D.: 1-V rail-to-rail CMOS OpAmp with improved bulk-driven input stage. IEEE J. Solid State Circ. **42**, 508–517 (2007)
19. Deo, N., Sharan, T., Dubey, T.: Subthreshold biased enhanced bulk-driven double recycling current mirror OTA. Analog Integr. Circ. Sig. Process **105**, 229–242 (2020)
20. Sharan, T., Bhadauria, V.: Ultra low-power rail-to-rail linear sub-threshold bulk-driven transconductor. In: 2014 International Conference on Power, Control and Embedded Systems (ICPCES), pp. 1–6 (2014)
21. Sharan, T., Bhadauria, V.: Ultra low-power bulk and gate-driven, class AB, sub-threshold transconductor. In: 2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN), pp. 921–926 (2015)
22. Sharan, T., Bhadauria, V.: Fully differential, bulk-driven, class AB, sub-threshold OTA with enhanced slew rates and gain. J. Circ. Syst. Comput. **26**(01), 1750001 (2017)
23. Centurelli, F., Fava, A., Olivieri, M., Tommasino, P., Trifiletti, A.: A low-voltage class-AB OTA exploiting adaptive biasing. AEU Int. J. Electron. Commun. **122**, 153282 (2020)
24. Ghaemnia, A., Hashemipour, O.: An ultra-low power high gain CMOS OTA for biomedical applications. Analog Integr. Circ. Sig. Process **99**(3), 529–537 (2019)
25. Horng, J.-W.: Voltage-mode universal biquadratic filter using two OTAs. Active Passive Elec. Comp. **27**, 85–89 (2004)
26. Mahendra, M., Kumari, S., Gupta, M.: DTMOS based low power adaptively biased fully differential transconductance amplifier with enhanced slew-rate and its filter application. IETE J. Res. 1–20. Online version on 24 May 2021

Design of CFOA-Based Electronically-Tunable Band Pass Filter Suitable for Filtering of EEG Signals

Shanvi Srivastava

Abstract In this paper, a Current Feedback Operational Amplifier (CFOA) based electronically-tunable bandpass filter suitable for filtering electroencephalogram (EEG) signals has been presented. The EEG contains information about the brain; hence its sub-band decomposition with different frequency bands such as alpha, beta, delta, theta, and gamma provide useful information for analyzing many brain diseases. The EEG signals are low-frequency signals, and therefore, the bandpass filter used for them has large capacitance values. Various capacitor multiplier circuits are used to eliminate the need for large valued capacitors, and the best circuit is used for the final results. Electronic tunability is achieved by using voltage control resistances in place of regular resistors. The proposed circuit has three controlling voltages, and each of them changes either frequency or bandwidth or the gain. The simulation results are included to confirm the workability of the proposed circuit using Pspice.

Keywords Current Feedback Operational Amplifier (CFOA) · Grounded capacitance multiplier · Electronically-tunable bandpass filter · Voltage controlled resistor · EEG signals

1 Introduction

Filters designed with a very low-frequency range are widely used in biomedical fields. They have a wide range of applications, like filtering electrocardiogram (ECG) and electroencephalography (EEG) signals. Typically, the EEG signal ranges from 0.1 to 100 Hz. The electroencephalogram (EEG) contains information about the brain; hence the sub-band decomposition of EEG is used for analyzing many brain diseases [15]. This sub-band decomposition extracts brain waves with different frequency bands such as alpha, beta, delta, theta and gamma from EEG signal to get more information from it. Five bands of brain waves, Delta (1–4 Hz), Theta

S. Srivastava (✉)
ECE Division NSUT, Delhi 110078, India
e-mail: shanvi265@gmail.com; shanvi.sp19@nsut.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_43

(4–7 Hz), Alpha (8–12 Hz), Sigma (12 to 16 Hz), and Beta (13–40 Hz), constitute the basic EEG signals [13, 14]. The Bandpass filter used for EEG signals, therefore, has a low cutoff frequency, i.e., the value of capacitances used will be very high. Large value capacitors require a large chip area. Capacitance multipliers are used to obtain a large-valued equivalent capacitor from a small-valued capacitor. Many different capacitance multipliers circuits [2–9] were studied, and finally, three different types of grounded capacitance multipliers were used in the proposed Bandpass filter. Their simulation results are compared, and the most suitable one is used for the final implementation of the bandpass filter. The Current Feedback Operational Amplifier (CFOA) is used to implement the Bandpass filter as it requires no component matching, higher speed, higher slew rate, and gain-bandwidth independence. The proposed circuit implements a bandpass filter suitable for filtering EEG signals which is also electronically tunable. Electronic tunability is achieved by using voltage control impedances (resistance in this case) in place of regular resistors. The gain-bandwidth independence of CFOA helps us in tuning bandwidth and gain independently. The proposed circuit has three controlling voltages, and each of them changes either frequency or bandwidth or the gain.

Grounded capacitance multipliers using different active elements have been presented in the literature. The ground capacitance multiplier [2] presents a topology suitable for both positive and negative capacitance multiplication and uses only a single CFOA, two resistors, and the reference capacitor [2]. It has fewer parasitic components and occupies a smaller chip area [2]. The circuit proposed in [3] includes two CFOA, one ground capacitor, and three resistors. Passive component matching constraints are not required for this circuit. The Capacitance multiplier circuit presented in [4] employs two CFOA, two resistors, and a single capacitor. This circuit attains a high multiplication factor at the cost of the lower component spread, whereas in most of the grounded capacitance multipliers higher multiplication factor is achieved by the larger component spread. Paper [5] presents a modified CFOA which is used to realize capacitance multiplier circuits. The presented circuits need no component-matching conditions and/or cancellation constraints and therefore are easy to implement [5]. The drawback of this circuit is that it uses MCFOA, which is not readily available as packaged IC. Paper [8] also presents new realizations of grounded negative capacitance. The capacitance multiplier circuits presented in [5] and [7] are realized using three AD844 [10], which in turn will increase the total number of CFOA used in our proposed tunable bandpass filter.

There are numerous circuits for grounded capacitance multiplier circuits using CFOA. A brief comparison of the workability of the three capacitance multiplier circuits [2–4] used in the proposed Bandpass filter is presented in Table 2, and the best-suited is used in the final circuit (Fig. 3).

2 Basic Circuit

A bandpass filter is an electronic device that allows a specified range of frequencies to pass through it and rejects other frequencies by providing a large attenuation for the stopband frequencies. Active Bandpass filters require an external power supply and have active components such as transistors. The central frequency, gain, and bandwidth of the active bandpass filter can be changed. CFOA based bandpass filter is depicted in Fig. 1. The circuit comprises three sections: V to I converter, RLC resonator, and amplifier circuit.

A straightforward analysis of the circuit reveals that its transfer function is

$$T(s)\Big|_{BP} = \frac{V_{out}}{V_{in}} = \left[1 + \frac{R_5}{R_4}\right]\left[\frac{{}^{S}\!/C_2R_0}{s^2 + \frac{s}{R_3C_2} + \frac{1}{R_1R_2C_1C_2}}\right] \tag{1}$$

The equation for the second-order transfer function of the standard Bandpass filter is

$$T(s)\frac{ho\left(\frac{\omega_0}{Q_0}\right)s}{s^2 + \frac{\omega_0}{Q_0}s + \omega_0} \tag{2}$$

Comparing Eqs. (2) and (3):

$$\omega_0 = \frac{1}{\sqrt{R_1R_2C_1C_2}} \tag{3}$$

$$Bandwidth = \frac{\omega_0}{Q_0} = \frac{1}{R_3C_2} \tag{4}$$

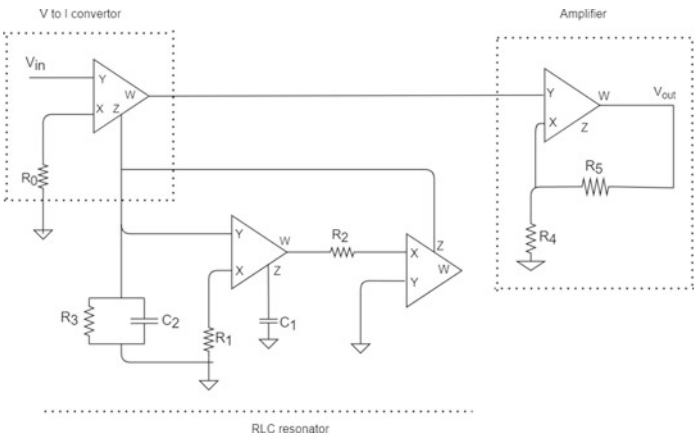


Fig. 1 Circuit for band pass filter

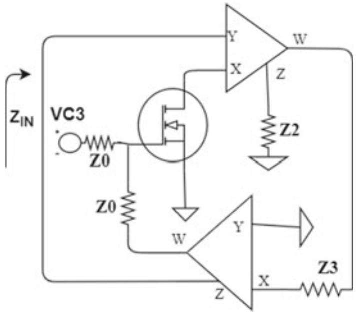


Fig. 2 Voltage-controlled resistance

$$Gain = h_0 = \frac{R_3}{R_0} \tag{5}$$

In Eq. (4) bandwidth of the filter is inversely proportional to C_2, so after replacing C_2 with a capacitance multiplier, the error in that becomes significant, BUT from Eq. (3), ω_o (central frequency) is inversely dependent on the square root of R_1 R_2 C_1 C_2, so the error percentage comparatively decreases if C_1 C_2 in ω_o is replaced with capacitance multiplier.

From Eqs. (3), (4), and (5), we observe that if resistors R_0, R_1 and R_3 are made electronically tunable, it will change the ω_0, bandwidth and gain of the circuit independently.

To make the Bandpass filter electronically tunable, grounded voltage control resistance is used in place of a general resistor. A linear voltage-controlled (VC) grounded impedance circuit [1] is shown in Fig. 2. R_0, R_1 and R_3 are replaced with the grounded voltage-controlled resistor. The gate voltage of MOSFET is made equal to $(V_C + V_1)/2$ and is obtained by using the w-terminal of the second CFOA, as shown in the figure.

The nonlinear term in r_{DS} is effectively cancelled. The overall circuit realizes a linear voltage-controlled the grounded impedance of the value

$$Z_{in}(s) = r_{DS}\frac{Z_2}{Z_3} \tag{6}$$

This voltage-controlled resistance is simulated in a simple voltage divider circuit, and its equivalent resistance at a particular applied DC voltage is observed. This table provides us with an idea about the range of voltages to be applied and the resistance values which will be obtained.

From the above Table 1, it is clear that as the ratio of Z_2/Z_3 increases, the spread in the equivalent resistance increases, i.e. have a larger range of resistance.

Capacitance C_1 and C_2 are replaced by capacitance multiplier circuits presented in [2–4], respectively, and the resultant bandpass filter is simulated using Pspice. Supply voltages are specified with the final bandpass filter simulation result. The

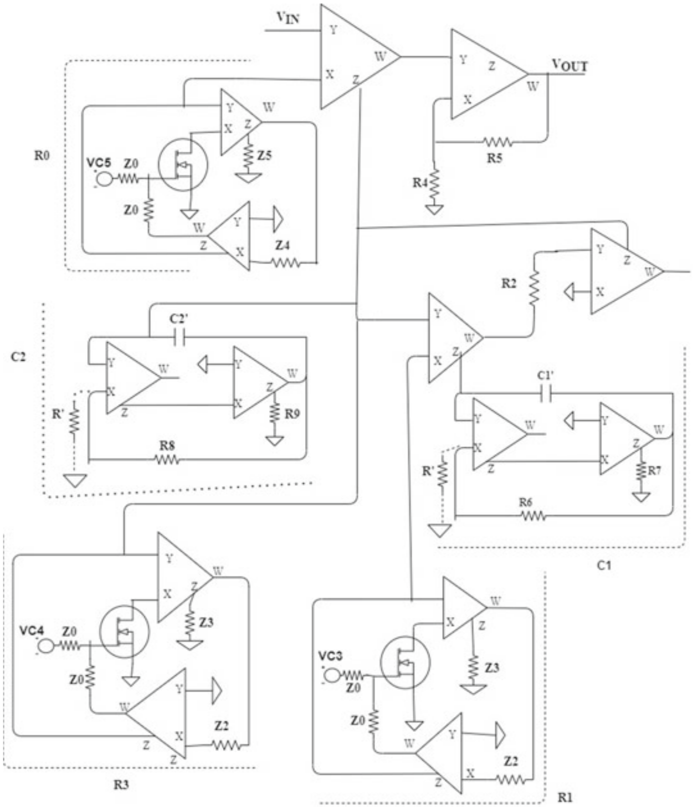


Fig. 3 Final proposed circuit for the bandpass filter

Table 1 Applied voltage and its equivalent resistance

Applied DC voltage (in V)	For $Z_2 = 200$ kΩ and $Z_3 = 10$ kΩ (k)	For $Z_2 = 250$ kΩ and $Z_3 = 5$ kΩ (k)
8	80	190
7	98	235
6	130	305
5	185	435
4	300	665

error percentage in Central frequency and bandwidth for all the three circuits is presented in Table 2.

To reduce the value of capacitances to the lowest value, a high multiplication factor is preferred. By comparing the results for $k = 100$, the best option for capacitance multiplier is [4]. It has less error percentage for ω_o and bandwidth, which is further reduced by using compensating resistor.

Table 2 Comparison of performance of Band pass filter with different capacitance multiplier

Gain	$R_1(\Omega)$ $R_2(\Omega)$ C_1' C_2'				Error in [2]	Error in [3]	Error in [4]
5	10 k	10 k	90.909 nF	45.454 nF	$f_o = 18.22\%$ $BW = 3.38\%$	$f_o = 9.38\%$ $BW = 2.31\%$	$f_o = 3.47\%$ $BW = 7.97\%$
22	10 k	10 k	90.909 nF	45.454 nF	$f_o = 6.77\%$ $BW = 62.60\%$	$f_o = 0.497\%$ $BW = 41.3\%$	$f_o = 66.66\%$ $BW = 75.08\%$
22	100 k	10 k	9.0909 nF	45.454 nF	$f_o = 11.802\%$ $BW = 228.9\%$	$f_o = 3.7\%$ $BW = 158.42\%$	$f_o = 20.73\%$ $BW = 31.88\%$
100	10 k	10 k	10 nF	10 nF	NA (component spread will be very high)	$f_o = 3.47\%$ $BW = 182.91\%$	$f_o = 97.6\%$ $BW = 107.4\%$ $f_o = 43.65\%$ $BW = 41.65\%$ (with R')

3 Proposed Circuit

In the final proposed circuit Fig. 3 C1 and C2 are replaced by a grounded capacitance multiplier circuit [4], and resistors R_0, R_1, and R_3 are replaced by voltage-controlled resistance [1]. The grounded capacitance multiplier circuit also uses compensating resistor R' for improved performance. It contains a total of fourteen CFOAs, three MOSFETs, two capacitance, and twelve resistors.

Without considering any non-idealities, the transfer function of the proposed bandpass filter in Fig. 3 is evaluated as

$$\boldsymbol{T.F.} = \frac{\boldsymbol{V_{out}}}{\boldsymbol{V_{in}}} = \left[1 + \frac{\boldsymbol{R_5}}{\boldsymbol{R_4}}\right]\left[\frac{\dfrac{s}{k_2 C_2' \frac{2k}{(v_{C5}-2V_T)} \frac{Z_4}{Z_5}}}{s^2 + \dfrac{s}{\frac{2k}{(v_{C4}-2V_T)} \frac{Z_2}{Z_3} \cdot k_2 C_2' + \frac{1}{\frac{2k}{(V_{C3}-2V_T)} \frac{Z_2}{Z_3} \cdot R_2 k_1 C_1' k_2 C_2'}}}\right] \tag{7}$$

V_T (threshold voltage) and k are the same for R_1, R_0, and R_3 as the same MOSFETs are used for all three of them.

$$\text{Where } k = \mu_n C_{ox} \frac{W}{L},\; k_1 = \frac{1}{1 - \frac{R_7}{R_6}},\; k_2 = \frac{1}{1 - \frac{R_9}{R_8}}$$

Comparing Eqs. (2) and (5) we get

$$\omega_0 = \frac{1}{\sqrt{\frac{2k}{(v_{C3}-2V_T)} \frac{Z_2}{Z_3} \cdot R_2 k_2 C_2' k_1 C_1'}} \tag{8}$$

$$Bandwidth = \frac{\omega_0}{Q_0} = \frac{1}{\frac{2k}{(v_{C4}-2V_T)} \frac{Z_2}{Z_3} k_2 C_2'} \tag{9}$$

$$Gain = h_0 = \frac{(v_{C5} - 2V_T)\,\frac{Z_2}{Z_3}}{(v_{C4} - 2V_T)\,\frac{Z_4}{Z_5}} \tag{10}$$

There are three control voltages. VC_3 (R_1) is used to change the frequency (ω_o), VC_5 (R_0) is used to change the gain, and VC_4 (R_3) is used to change the bandwidth.

4 Simulation Result

The proposed electronically tunable Bandpass filter depicted in Fig. 3 is simulated for possible ranges of EEG signals using the PSpice program in which the AD844 [11] model is utilized. Supply voltages are selected as ±12 V. Values of some parasitic impedances of the AD844 are $R_X \cong 50\ \Omega$, $R_Z \cong 3\ \mathrm{M}\Omega$ and $C_Z \cong 5.5$ pF. A 40 Hz 50 mV sine wave signal with zero DC amplitude and peak to peak value of 20 mV is applied as the input of the bandpass filter. Furthermore, $R_2 = 100$ k $R_4 = 50$ k, $R_5 = 200$ kΩ (i.e. gain is 4 + 1 = 5) $C_1 = 450$ pF and $C_{eq} = 90.45$ nF (k = 201) $C_2 = 0.25$ nF and $C_{eq} = 50.25$ nF (k = 201) $Z_2 = 250$ kΩ, $Z_3 = 5$ k $Z_4 = Z_5 = 100$ kΩ For R_1, R_3, R_0 values are from Table 2.

From Table 2, it is clear that as we reduce the applied DC voltage at the voltage-controlled resistor, the value of equivalent resistance increases. From Eq. (8), we conclude that as the value of resistance increases as ω_0 decreases. Similarly, from Eq. (9), we conclude that the bandwidth of the proposed circuit is inversely proportional to the value of resistance or is directly proportional to the applied DC voltage. An Eq. (5) states that the gain of the proposed circuit depends on both R_0 and R_3, and both of them are variable. Initially, the voltage across VC_4 (R_3) is applied as per the required bandwidth, and then the voltage across VC_5 (R_0) is adjusted to get the desired gain.

[1] Beta (13 to 30 Hz): For applied voltages Vcc3 = 4.6 V, Vcc4 = 9 V and Vcc_5 = 6 V. We get Central freq = 19.099 Hz, Upper freq = 30.01 Hz, Lower freq = 12.09 Hz, Bandwidth = 17.92 Hz (Fig. 4).

[2] Sigma (12 to 16 Hz): For applied voltages Vcc_3 = 6.5, Vcc4 = 3.5 and Vcc_5 = 6 V. We get Central freq = 14.028 Hz, Upper freq = 16.42 Hz, Lower freq = 12.07 Hz, Bandwidth = 4.35 Hz (Fig. 5).

[3] Alpha (8 to 12) Hz: For applied voltages Vcc_3 = 4.7, Vcc_4 = 3.5 and Vcc_5 = 6 V. We get Central freq = 9.977 Hz, Upper freq = 12.42 Hz, Lower freq = 8.008 Hz, bandwidth = 4.412 Hz (Fig. 6).

[4] Theta (4 to 7) Hz: For applied voltages Vcc_3 = 3.25, Vcc_4 = 4 and Vcc_5 = 6 V. We get Central freq = 5.21 Hz, Upper freq = 7.21 Hz, Lower freq = 3.76 Hz, Bandwidth = 3.76 Hz (Fig. 7).

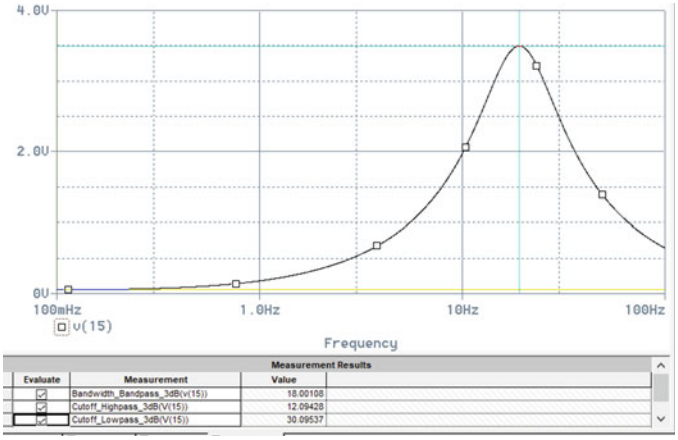


Fig. 4 Simulation result for beta wave

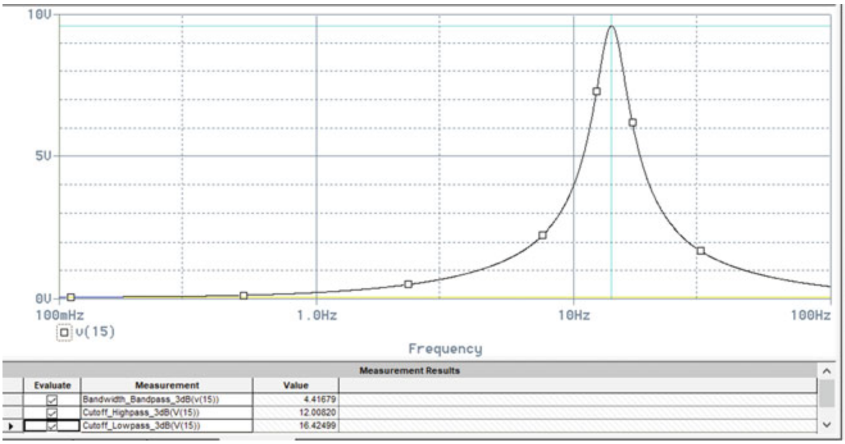


Fig. 5 Simulation result for Sigma wave

[5] Delta (1 to 4) Hz: For applied voltages $Vcc_3 = 2.9$ V, $Vcc_4 = 3.3$ V, and $Vcc_5 = 6$ V. We get Central freq = 2.20 Hz, Upper freq = 4.411 Hz, Lower freq = 1.08 Hz, Bandwidth = 3.331 Hz (Fig. 8).

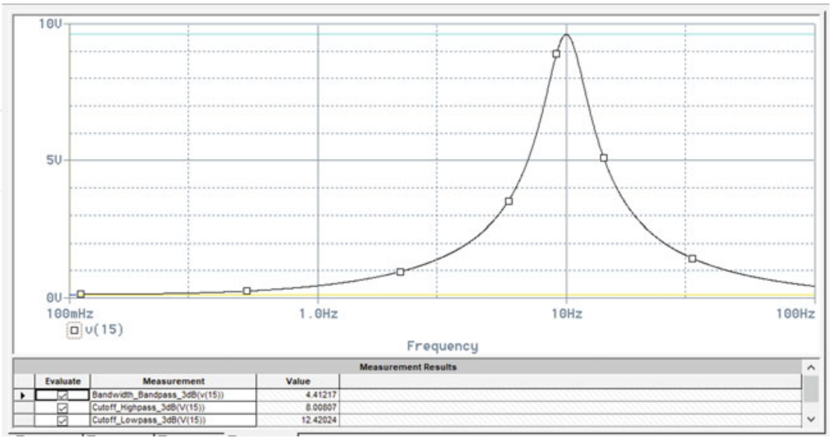


Fig. 6 Simulation result for Alpha wave

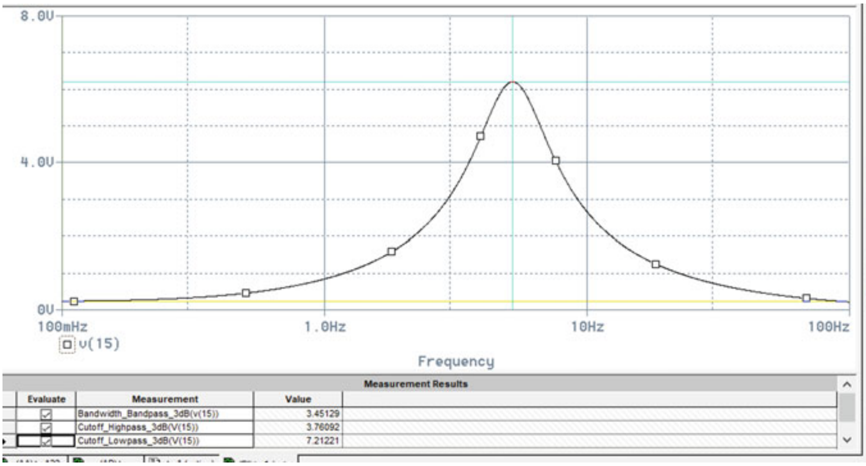


Fig. 7 Simulation result for Theta wave

5 Conclusion

In this paper, an electronically tunable bandpass filter suitable for filtering EEG signals is presented, which uses fourteen CFOAs, three MOSFETs, two capacitance, and twelve resistances. The proposed circuit is successfully implemented for the desired range of frequencies with adequate bandwidth for each of the EEG signals, such as alpha, beta, delta, theta and gamma. Various capacitor multiplier circuits are used to eliminate the need for large valued capacitors, and the best circuit is used for the final results. The circuit is made electronically tunable by using a grounded voltage control resistor, and three control voltages are provided, each for changing

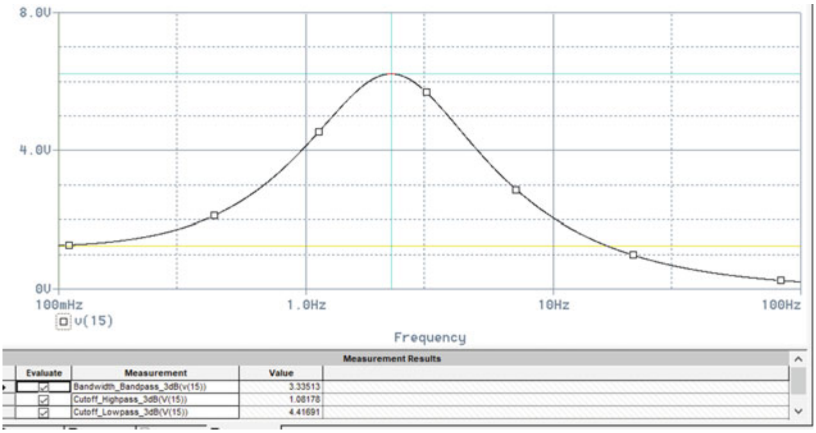


Fig. 8 Simulation result for Delta wave

the gain, central frequency and bandwidth of the circuit. The workability of the proposed circuit is verified using Pspice simulation. Future work can be done on reducing the number of CFOAs used, which will make the circuit more suitable for IC implementation.

Acknowledgements I wholeheartedly express my indebtedness and profound gratitude to my guide Professor Raj Senani for his simulating guidance and profuse assistance at every juncture of my endeavours and for the fruitful completion of my research work. I will always cherish my association with him with great pleasure.

References

1. Senani, R.: Realization of a class of analog signal processing/signal generation circuits: novel configurations using current. Frequenz **52**(9-J) (1998)
2. Keskin, A.U.: Grounded capacitance multipliers with electronic tuning possibility using single current feedback amplifier. U.S. Patent 10,382,011, 13 August 2019
3. Dogan, M., Yuce, E.: A new CFOA based grounded capacitance multiplier. AEU Int. J. Electron. Commun. **115**, 153034 (2020)
4. Verma, R., Pandey, N., Pandey, R.: Novel CFOA based capacitance multiplier and its application. AEU Int. J. Electron. Commun. **107**, 192–198 (2019)
5. Yuce, E., Minaei, S.: A modified CFOA and its applications to simulated inductors, capacitance multipliers, and analog filters. IEEE Trans. Circuits Syst. I Regul. Pap. **55**(1), 266–275 (2008)
6. Senani, R., Bhaskar, D.R., Singh, A.K., Singh, V.K.: Current Feedback Operational Amplifiers and their Applications. Springer, New York (2013). https://doi.org/10.1007/978-1-4614-5188-4
7. Yesil, A., Yuce, E., Minaei, S.: Grounded capacitance multipliers based on active elements. AEU Int. J. Electron. Commun. **79**, 243–249 (2017)
8. Lahiri, A., Gupta, M.: Realizations of grounded negative capacitance using CFOAs. Circ. Syst. Sig. Process. **30**(1), 143–155 (2011)
9. Dogan, M., Yuce, E.: CFOA based a new grounded inductor simulator and its applications. Microelectron. J. **90**, 297–305 (2019)

10. AD844 60 MHz Monolithic Op amp. http://www.analog.com/static/imported-files/data_sheets/AD844.pdf. Accessed 21 July 2021
11. http://www.spicemodel.com/models/ad/AD844.html. Accessed 21 July 2021
12. Soliman, A.M.: Applications of the current feedback operational amplifiers. Analog Integr. Circ. Sig. Process. **11**(3), 265–302 (1996)
13. Qian, X., Xu, Y.P., Li, X.: A CMOS continuous-time low-pass notch filter for EEG systems. Analog Integr. Circ. Sig. Process. **44**(3), 231–238 (2005)
14. Nayak, C.S., Anilkumar, A.C.: EEG normal waveforms. In: StatPearls. StatPearls Publishing, Treasure Island (FL) (2020)
15. Anatomy of Brain, Brain parts and function of each part, Encyclopedia

Design and Analysis of Non-uniform Transmission Line Based Dual-Band Bandpass Filter

Mridul Gupta, Mayank Kansal, Shriram Thyagarajan, Prajjwal Singh Chauhan, and Dharmendra Kumar Upadhyay

Abstract A novel methodology for designing a microwave filter has been implemented. The designed filter is a wideband dual-band bandpass filter which has been constructed in digital domain utilizing the chain-scattering parameters for low-pass parallel coupled lines (LP-PCL), serial transmission lines and shunt connected short as well as open circuited stubs. Nature inspired metaheuristic algorithms have been utilized to approximate the magnitude response of the desired transfer function in reference to the magnitude response of ideal transfer function. The dual-band bandpass filter designs are proposed for the 2.1 and 5.2 GHz band. The configuration obtained is to be then simulated in the ADS environment utilizing RT Duroid 5880 substrate. The proposed microstrip structure is capable of covering various microwave applications, most important and common being, the WLAN operating range.

Keywords Dual-band bandpass filter · Metaheuristic optimization · Microstrip · Transmission line elements

1 Introduction

Non-uniform transmission lines have been used to design microwave dual-band bandpass filter with more precise and accurate response using Z-domain chain scattering matrices than the existing models. Specifically, the filter is constructed utilizing a synthesis method for determining the adequacy of the design technique proposed.

A novel perspective to make use of the well-developed digital signal processing techniques and optimization algorithms for designing and constructing microwave

M. Gupta (✉)
Department of Electronics and Communication Engineering, Graphic Era Deemed to be University, Dehradun 248002, India
e-mail: mri.gupta@gmail.com

M. Kansal · S. Thyagarajan · P. S. Chauhan · D. K. Upadhyay
Department of Electronics and Communication Engineering, Netaji Subhas University of Technology, Sector 3, Dwarka, New Delhi 110078, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_44

frequency range filters happened to be stated by Chang et al. [1]. Using the chain scattering matrices of various line sections [2], at first a prototype configuration will be proposed. The response of the prototype will be approximated to the ideal response as desired. The desired response will be achieved by adjusting the impedances of the stub using Harris Hawks algorithm for optimisation which was also used by Bao et. al for multilevel thresholding segmentation of colour image [3]. The system is synthesized using method of moments to determine the validity and usefulness of the design technique.

Standard methods for designing microwave filters generally commence with the prototype made for lumped element circuits. These elements are then transformed using Richard's transformation to stubs of same electric length [4–7]. Recently, several filters have been designed using heuristic optimization approach. Oraizi et al. [8] constructed a microwave LPF with the help of Particle Swarm Optimization, Mahata et al. [9] gave the design for fractional order digital integrators utilizing Colliding Bodies Optimization technique. Several microwave filters were designed and implemented using such algorithms by Gupta et al. [10–13].

A bandpass filter (BPF) allows frequencies to pass through within a range of onterest to pass while rejecting (attenuating) unwanted frequency components. Butterworth filter is a form of signal processing filter which is constructed to obtain a frequency response as flat as achievable in the passband. Thus, for constructing a filter with maximally flat passband, a butterworth prototype of the filter is used to approximate the filter design.

Numerous ways to design dual-band BPFs have been formulated in recent years [14–16]. The optimum method of designing a dual-band BPF is by assembling two singular band filters having dissimilar passbands [17]. One more efficient technique is to cascade a bandstop filter (BSF) and a BPF for obtaining dual-band performance [18]. A technique of designing a selectivity-enhanced Stepped impedance resonator (SIR) dual-band BPF having two open-circuited stubs is implemented in [19]. For minimizing circuit area, SIR frequencies are often utilized to form the second passband [20–23]. A dual-band BPF utilizing transversal filtering for obtaining wide stopband suppression has been implemented in [24]. In [25], a dual-band differential BPF filter has been implemented through coupling of two similar stepped impedance Ring loaded resonators (RLR). Stepped impedance ring loaded resonator was utilized by Weng et al. [26] for implementing a dual-band BPF. Cascaded structure of transversal filters and loading stubs were used to implement dual-band BPFs in [27]. Wide bandwidth using two end short-circuited SIR was achieved in the design [28]. Using the same technology, a dual-wideband has been implemented [29]. Wang et al. [30] made use of bridged-T coils was also identified to obtain dual-band BPF design with reduced size.

Bandpass filters are widely used in optics such as LIDARS, lasers, etc. These filters are applicable in WLAN, sonar, control instruments, various medical, and seismology applications. This paper is organized four sections which includes the introduction discussed in Sect. 1. Desing of dual-band BPF is explained in Sect. 2 which is followed by Sect. 3 comprising the simulations results for the proposed design. Finally, the conclusion is included as Sect. 4.

2 Design of Dual-Band BPF

Several dual-band BPFs have been designed using cascaded stuctures of wideband BPF and bandstop filter. Similar approach has been used to design dual-band BPF using serial lines, low-pass parallel coupled lines (LP-PCL) and shunt connected short as well as open circuited stubs [18]. Highly accurate and wideband design has been obtained using the proposed methodology.

2.1 *Two-Port Network*

There are many parameters that can be defined for the two port network namely Y-parameters, Z-parameters, ABCD parameters, S- or Scattering parameters etc. The microwave two port network is described in respect of S parameters. This is due to the fact that for a microwave system, it is impossible to measure voltage, current and phase accurately but power can be measured accurately. S-parameters are converted to chain scattering parameters or T-parameters since for a network which is made up by cascading multiple two port networks in series, the T-matrix for the overall network can be obtained by simply multiplying the individual T-matrices [1].

Given in Fig. 1 is a two-port network. Here, u(1) is the wave incident and v(1) is the wave reflected at the first port, while u(2) is the wave incident and v(2) is the wave reflected at the second port. The aforementioned waves are linked to each other by chain scattering matrices (or T-matrices) T_{ab}, a,b = 1, 2 of the network as shown:

$$\begin{pmatrix} u(1) \\ v(1) \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} u(2) \\ v(2) \end{pmatrix} \tag{1}$$

For each of the line element under consideration, all the finite sections are assumed to have equal electrical length, i.e., $\beta l = \omega\tau$, where ω is the angular frequency and τ is propagation delay for each line element. With this, the T-parameters are converted to discrete-time domain by substituting:

$$z^{-1} = exp(-2j\beta l) \tag{2}$$

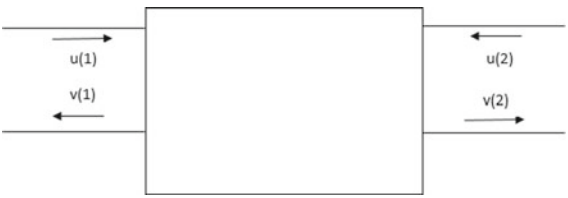


Fig. 1 Two port network device

T-matrices for the transmission line elements to be used for the proposed filter configuration are as follows:

A. *Serial transmission line* [18]:

$$T_S = \frac{1}{z^{-1/2}\left(1-\Gamma^2\right)}\begin{bmatrix} 1-\Gamma^2 z^{-1} & -\left(\Gamma - \Gamma z^{-1}\right) \\ \Gamma - \Gamma z^{-1} & -\Gamma^2 + z^{-1} \end{bmatrix} \tag{3}$$

where $z = e^{jw}$, $\Gamma = (Z_1 - Z_0)/(Z_1 + Z_0)$. Here, reference characteristic impedance is Z_0 (Fig. 2).

B. *Shunt short stub* [18]:

$$T_{SSS} = \frac{1}{1-z^{-1}}\begin{bmatrix} (1+a)-(1-a)z^{-1} & a+az^{-1} \\ -a-az^{-1} & (1-a)-(1+a)z^{-1} \end{bmatrix} \tag{4}$$

where $a = Z_0/2Z_2$ (Figs. 3, 4 and 5).

III. *Shunt open stub* [18]:

$$T_{SOS} = \frac{1}{1+z^{-1}}\begin{bmatrix} (1+a)+(1-a)z^{-1} & a-az^{-1} \\ -a+az^{-1} & (1-a)+(1+a)z^{-1} \end{bmatrix} \tag{5}$$

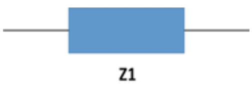


Fig. 2 Serial transmission line configuration

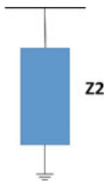


Fig. 3 Shunt short stub configuration

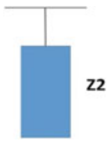


Fig. 4 Shunt open stub configuration

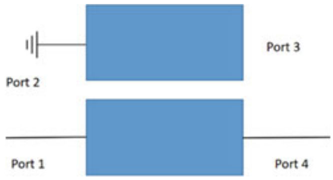


Fig. 5 LP-PCL configuration

IV. *Low-pass parallel coupled lines* [18]:

$$T_{LP-PCL} = \frac{1}{8\left(b_1^2 - b_2^2\right)Z_0 z^{-\frac{1}{2}}\left(1 - z^{-1}\right)}\begin{bmatrix} p & -s \\ r & \left(4q^2 - rs\right)/p \end{bmatrix} \tag{6}$$

where $b_1 = \mathrm{Z_{ev}} + \mathrm{Z_{ov}}$, $b_2 = \mathrm{Z_{ev}} - \mathrm{Z_{ov}}$,
$\mathrm{p} = [4\,\mathrm{b}_1^2\mathrm{Z}_0 - 2\mathrm{b}_2^2\mathrm{Z}_0 + \mathrm{b}_1(\mathrm{b}_1^2 - \mathrm{b}_2^2) + 4\,\mathrm{b}_1\mathrm{Z}_0^2] + (8\mathrm{b}_1^2\mathrm{Z}_0 - 12\mathrm{b}_2^2\mathrm{Z}_0)\,\mathrm{z}^{-1}$
$+[4\,\mathrm{b}_1^2\mathrm{Z}_0 - 2\mathrm{b}_2^2\mathrm{Z}_0 - \mathrm{b}_1(\mathrm{b}_1^2 - \mathrm{b}_2^2) - 4\,\mathrm{b}_1\mathrm{Z}_0^2]\,\mathrm{z}^{-1}$,
$q = 4\left(b_1^2 - b_2^2\right)Z_0 z^{-\frac{1}{2}}\left(1 + z^{-1}\right)$,
$r = -2\mathrm{b}_2^2\mathrm{Z}_0 + \mathrm{b}_1\left(\mathrm{b}_1^2 - \mathrm{b}_2^2\right) - 4\,\mathrm{b}_1\mathrm{Z}_0^2\big] + 4\,\mathrm{b}_2^2\mathrm{Z}_0\,\mathrm{z}^{-1}$
$+\left[-2\mathrm{b}_2^2\mathrm{Z}_0 - \mathrm{b}_1\left(\mathrm{b}_1^2 - \mathrm{b}_2^2\right) + 4\,\mathrm{b}_1\mathrm{Z}_0^2\right]\mathrm{z}^{-2}$,
$s = [2\mathrm{b}_2^2\mathrm{Z}_0 + \mathrm{b}_1(\mathrm{b}_1^2 - \mathrm{b}_2^2) - 4\,\mathrm{b}_1\mathrm{Z}_0^2] - 4\,\mathrm{b}_2^2\mathrm{Z}_0\,\mathrm{z}^{-1}$
$+[2\mathrm{b}_2^2\mathrm{Z}_0 - \mathrm{b}_1(\mathrm{b}_1^2 - \mathrm{b}_2^2) + 4\,\mathrm{b}_1\mathrm{Z}_0^2]\,\mathrm{z}^{-2}$.

2.2 Location of Zeros for Multi Section Stubs

Equations (3)–(6) mention T-matrices for unit line, low-pass parallel coupled lines (LP-PCL) and shunt connected short as well as open circuited stubs in the Z-domain. Shunt connected open circuited stub provides one zero at $z = -1$ ($\Omega = \pi$) and shunt connected short circuited stub provides one zero at $= 1(\Omega = 0)$. Zero at $\Omega = \pi$ and $\Omega = 0$, improves the quality of the parameters for the stopband, while also assuring that at normalising frequency, a transmission zero is obtained. This ensures a sharp rejection at the normalising frequency which is in the stop band. Despite the fact that the serial transmission line element provides a zero at $z = 0$, the zero obtained is not found to be feasible. This predicament occurs as the delay in time is represented by $z^{-1/2}$ and for all frequencies its magnitude is one. The LP-PCL also contributes one zero at $z = -1$ which indicates that it may also be replaced by an open stub but it provides a higher stopband attenuation rate than the latter.

2.3 Design Method

The comprehensive chain scattering parameter for a cascaded sequence of multiple elements which are unit-line, low-pass parallel coupled lines and shunt connected short as well as open circuited stubs is obtained by successive multiplication of each element's chain scattering matrix, i.e.,

$$\begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} = \prod_{i=1}^{N} \begin{pmatrix} T_{11}^{(i)} & T_{12}^{(i)} \\ T_{21}^{(i)} & T_{22}^{(i)} \end{pmatrix} \tag{7}$$

where total number of stubs is denoted by N, while $T_{11}^{(i)}$, $T_{12}^{(i)}$, $T_{21}^{(i)}$, $T_{22}^{(i)}$ are the T parameters of the i^{th} stub element.

Assuming that the network comprises of P serial lines, Q short stubs and R open stubs and S LP-PCL, the T_{11} of the overall cascaded network can be given as:

$$T_{11}, \text{network (z)} = \frac{\sum_{\text{i=1}}^{\text{Q+R+P+2S}} g_i z^{-1}}{\prod_{p=1}^{P} (1 - \Gamma^2) z^{-\frac{p}{2}} \prod_{r=1}^{R} (1 + z^{-1}) \prod_{q=1}^{Q} (1 - z^{-1}) \prod_{s=1}^{S} (1 + z^{-1}) z^{-\frac{s}{2}}} \tag{8}$$

where g_i is imaginary and obtained from the characteristic impedance valuess of all the line elements.

After this, for the network of comprehensive chain scattering matrix, T_{11} is calculated and transmission coefficient is obtained as:

$$H_p(z) = \frac{1}{T_{11overall}(z)} \tag{9}$$

For the line elements which form the structure, impedances are obtained through a Least Square (LS) based error function which is given by:

$$E(\omega) = \sum_{\omega} \left[\left| H_p(\omega) \right| - |H_i(\omega)| \right]^2 \tag{10}$$

Here, $H_p(\omega)$ represents the transmission coefficient and $H_i(\omega)$ represents the ideal magnitude response.

A metaheuristic optimization algorithm named Colliding Bodies Optimization (CBO) is utilized for minimizing the error function. High level of accuracy in the magnitude response and fast convergence of rate is shown by CBO as compaired with other widely used available metaheuristic algoriths [13].

3 Simulation Results

A third order Butterworth bandpass filter is used as a prototype at normalizing frequency of 10 GHz based on which a network consisting of one shunt connected open circuit stub, one shunt connected short circuit stub and three serial lines is obtained as shown in Fig. 6. Simulated magnitude response for the proposed bandpass filter is shown in Fig. 7.

Afterwards, a third order Butterworth lowpass filter is used as a prototype at normalizing frequency of 4 GHz. When we unfold the magnitude response due to property of symmetricity the overall response for the desired 10 GHz band maps to a bandstop filter based on which a network configuration consisting of one shunt connected open circuit stub, two serial lines and one LP-PCL is obtained as shown as in Fig. 8 and the simulated response for this configuration is shown in Fig. 9.

These proposed designs are then simulated in the ADS environment. The substrate used is Rogers RT Duroid 5880 which has a relative permittivity of 2.2, substrate thickness of 20 mil and a loss tangent value of 0.0009. Equal length of $l = \lambda_o/4$ is

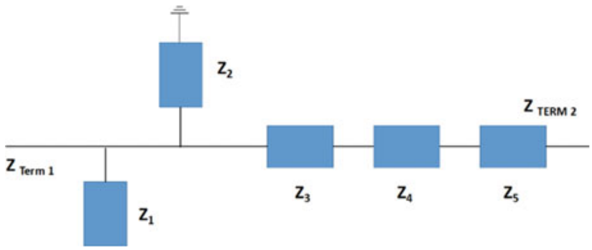


Fig. 6 Configuration of proposed bandpass filter

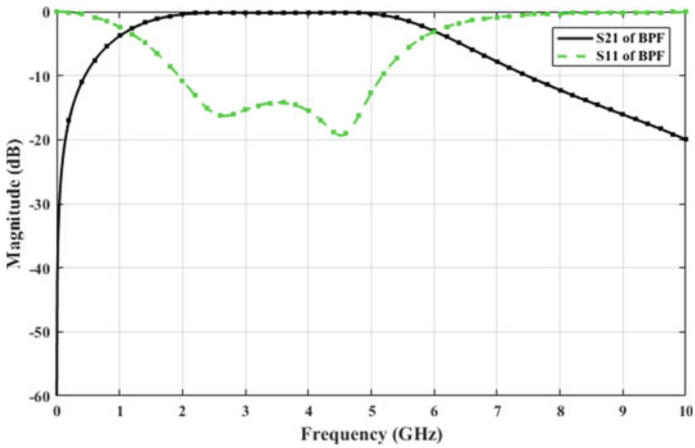


Fig. 7 Magnitude response of proposed bandpass filter

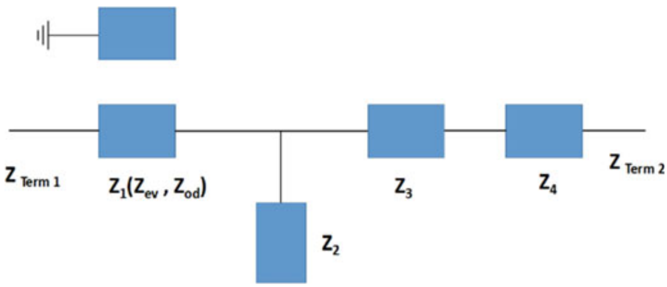


Fig. 8 Configuration of proposed bandstop filter

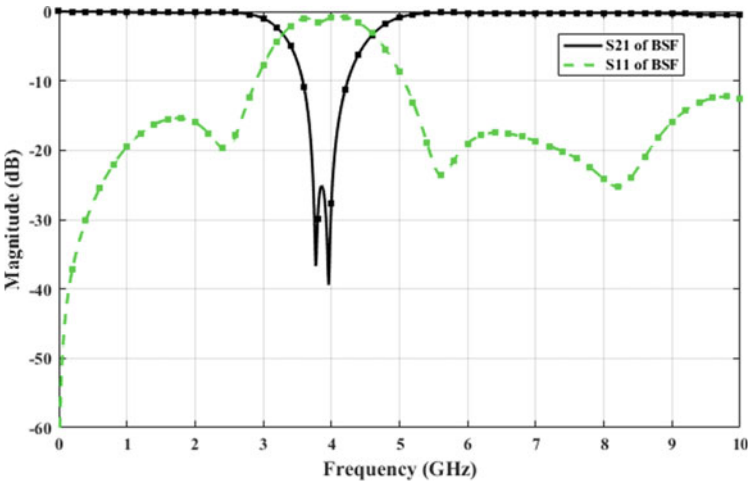


Fig. 9 Configuration of the proposed bandstop filter

considered for each line element, wherein λ_o denotes wavelength at the normalizing frequencies stated above. For the proposed designs, equivalent widths and lengths of the impedance values are calculated with the help of line calculator and further optimized in ADS using CBO algorithm and are enumerated in Tables 1 and 2.

The dual-band bandpass filter is obtained by cascading the bandpass filter and the bandstop filter designs obtained above. Figure 10 gives layout of proposed microstrip dual-band bandpass filter. The stubs have been bent at 90 degrees to make the design more compact. The dimensions of the filter are $W_1 = 1.526$ mm, $W_2 = 3.162$ mm, $W_3 = 0.288$ mm, $W_4 = 1.4$ mm, $W_5 = 3.819$ mm, $W_6 = 1.0004$ mm, $W_7 = 0.338$ mm, $W_8 = 1.4$ mm, $W_9 = 2.2$ mm, $L_1 = 4.51$ mm, $L_2 = 4.881$ mm, $L_3 = 5.882$ mm, $L_4 = 4.127$ mm, $L_5 = 4.341$ mm, $L_6 = 5.299$ mm, $L_7 = 14.959$ mm, $L_8 = 7.158$ mm, $L_9 = 7.496$ mm, $L_{10} = 6.7$ mm, $L_{11} = 6.493$ mm, $L_{12} = 8.917$ mm, $L_{13} = 6.778$ mm. 0.2 mm of diameter is used to create via holes. Total size of proposed filter is 58.3×15.8 mm. Simulated response for the proposed filter is shown in Fig. 11.

Table 1 Widths and lengths of line elements for the proposed bandpass configuration

Line element	Impedance (ohms)	Width (mm)	Length (mm)
Z_0	50	1.526	5.466
Z_1	29.791	3.162	5.332
Z_2	114.41	0.288	5.738
Z_3	53.203	1.385	5.484
Z_4	25.718	3.819	5.299
Z_5	61.805	1.083	5.529

Table 2 Widths and lengths of line elements for the proposed bandstop configuration

Line element	Impedance (ohms)	Width (mm)	Length (mm)
Serial Transmission Line	$Z_3 = 53.534$	1.369	13.757
	$Z_4 = 39.638$	2.139	13.555
Shunt Open Stub	$Z_2 = 107.768$	0.338	14.315
Coupled Line	$Z_{ev} = 77.735$ $Z_{od} = 45.71$	1.004	14.457
		Spacing between coupled line = 0.2 mm	

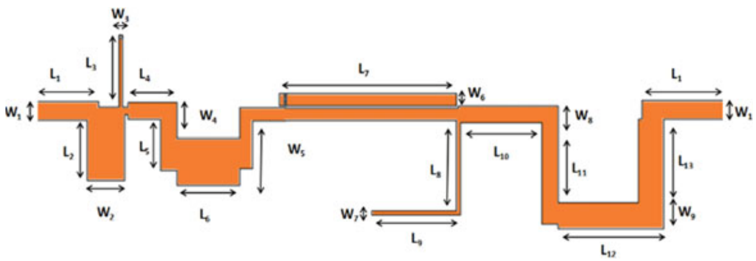


Fig. 10 Microstrip layout of proposed dual-band bandpass filter

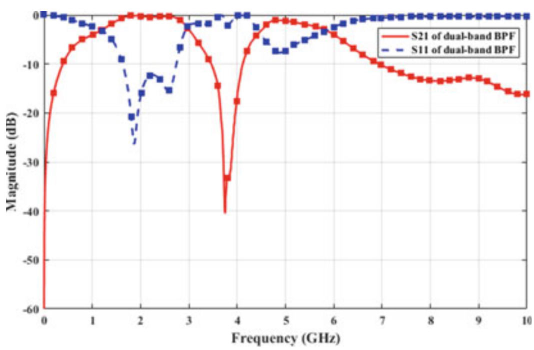


Fig. 11 Magnitude response of proposed dual-band bandpass filter

Table 3 Performance of proposed microstrip dual-band bandpass filter

Passbands	Centre frequency (GHz)	Band (3 dB) (GHz)	Bandwidth (GHz)	3 dB Fractional BW (%)	Insertion loss (dB)	Return loss (dB)
Passband 1	2.1	1.215–2.98	1.765	84	0.13	26.3
Passband 2	5.16	4.501–5.833	1.332	25.81	1.1	7.7

Performance parameters of the designed filter are shown in the Table 3. Furthermore, the comparison of proposed filter in this work with other few recently proposed dual-band bandpass filter based on various performance parameters is shown in the Table 4.

It can be observed from Table 4 that the filter design proposed in this work outperforms many other previously acclaimed dual-band bandpass filters. The size of the proposed design is also comparable and the design is easy to fabricate as it is single-layered.

Table 4 Comparison of proposed dual-band bandpass filter with its counterparts

Filter	Centre frequency (GHz)	3 dB Fractional BW (%)	Insertion loss (dB)	Return loss (dB)	Size ($\lambda_o * \lambda_o$)
Designed in [23]	2.82/3.21	5.2/5.1	1.9/1.7	21.6/16.1	2.76 * 1.3
Designed in [24]	3.78/4.82	11.3/10.6	1.38/1.82	14/33	0.16 * 0.31
Designed in [25]	1.57/2.38	9.9/6.5	1.21/1.95	19/24	–
Designed in [26]	2.6/5.8	10.4/3.6	1.1/2.15	>20	0.26 * 0.34
Designed in [27]	2.4/4	8/39	1.4/1	–	0.48 * 0.9
Designed in [30]	2.3/5.25	54/20	0.8/0.8	>20	0.3 * 0.3
Designed in [32]	2.4/5.2	51.9/23.3	0.3/0.7	22.1/20.8	0.28 * 0.2
Proposed dual-band bandpass filter	2.1/5.16	84/25.81	0.13/1.1	26.3/7.7	0.6 * 0.16

4 Conclusion

An efficient and realizable design of a dual-band bandpass filter having a wide passband has been provided in this work. The discrete-time (Z-domain) transfer function is obtained using the T-matrices. CBO algorithm has been used to minimize the Least Square based error function. After running multiple iterations, the optimum impedance values are achieved. The proposed design is obtained by implementing a cascaded structure comprising of a bandpass filter and a band-stop filter. Simulated response of the proposed design has wide bandwidths at 2.1/5.16 GHz which makes it suitable for various applications.

References

1. Chang, D.C., Hsue, C.W.: Design and implementation of filters using transfer functions in the Z domain. IEEE Trans. Microw. Theory Tech. **49**(5), 979–985 (2001)
2. Hsue, C.W., Tsai, Y.H., Hsu, C.C., Wu, C.Y.: Sharp rejection low-pass filter using three-section stub and Z-transform technique. IET Microwaves Antennas Propag. **4**(9), 1240–1246 (2010)
3. Bao, X., Jia, H., Lang, C.: A novel hybrid harris hawks optimization for color image multilevel thresholding segmentation. IEEE Access **7**, 76529–76546 (2019)
4. Maity, B.: Stepped impedance low pass filter using microstrip line for C-band wireless communication. In: 2016 International Conference on Computer Communication and Informatics (ICCCI), pp. 1–4. IEEE, January 2016
5. Jadhav, S.A., Misal, S.B., Mishra, A., Murugkar, A.: Designing of stepped impedance butterworth and chebyshev filters for wireless Communication. In: 2017 IEEE Applied Electromagnetics Conference (AEMC), pp. 1–2. IEEE, December 2017
6. Chen, F.C., Li, R.S., Chu, Q.X.: Ultra-wide stopband low-pass filter using multiple transmission zeros. IEEE Access **5**, 6437–6443 (2017)
7. Hayati, M., Abbasi, H., Shama, F.: Microstrip lowpass filter with ultrawide stopband and sharp roll-off. Arab. J. Sci. Eng. **39**(8), 6249–6253 (2014)
8. Oraizi, H., Esfahlan, M.S., Forati, E.: Design of stepped-impedance low pass filters with impedance matching by the particle swarm optimization and conjugate gradient method. In: 2009 European Conference on Circuit Theory and Design, pp. 639–642. IEEE, August 2009
9. Mahata, S., Saha, S.K., Kar, R., Mandal, D.: Optimal design of wideband infinite impulse response fractional order digital integrators using colliding bodies optimisation algorithm. IET Sig. Process. **10**(9), 1135–1156 (2016)
10. Gupta, M., Kansal, M., Thyagarajan, S., Chauhan, P.S., Upadhyay, D.K.: Design of an optimal microstrip butterworth low-pass filter using colliding bodies optimization. In: Dutta, D., Kar, H., Kumar, C., Bhadauria, V. (eds.) Advances in VLSI, Communication, and Signal Processing. LNEE, vol. 587, pp. 925–934. Springer, Singapore (2020). https://doi.org/10.1007/978-981-32-9775-3_82
11. Gupta, M., Upadhyay, D.K.: Design and implementation of fractional-order microwave differentiator. IET Microwaves Antennas Propag. **12**(8), 1375–1381 (2018)
12. Gupta, M., Upadhyay, D.K.: Design and implementation of second-order microwave integrator. Int. J. Electron. **107**(1), 125–140 (2020)
13. Gupta, M., Upadhyay, D.K.: Design of a coupled-line microstrip butterworth low pass filter. In: 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN), pp. 939–943. IEEE, March 2019

14. Velez, P., Bonache, J., Martın, F.: Dual-band balanced bandpass filter with common-mode suppression based on electrically small planar resonators. IEEE Microwave Wirel. Compon. Lett. **26**(1), 16–18 (2015)
15. Zhu, H., Abbosh, A.M.: Single-and dual-band bandpass filters using coupled stepped-impedance resonators with embedded coupled-lines. IEEE Microwave Wirel. Compon. Lett. **26**(9), 675–677 (2016)
16. Xu, J., Chen, Z.Y., Cai, Q.H.: Design of miniaturized dual-band low-pass–bandpass and bandpass filters. IEEE Trans. Compon. Packag. Manuf. Technol. **8**(1), 132–139 (2017)
17. Chen, C.Y., Hsu, C.Y.: A simple and effective method for microstrip dual-band filters design. IEEE Microwave Wirel. Compon. Lett. **16**(5), 246–248 (2006)
18. Tsai, L.C., Hsue, C.W.: Dual-band bandpass filters using equal-length coupled-serial-shunted lines and Z-transform technique. IEEE Trans. Microw. Theory Tech. **52**(4), 1111–1117 (2004)
19. Kuo, J.T., Shih, E.: Microstrip stepped impedance resonator bandpass filter with an extended optimal rejection bandwidth. IEEE Trans. Microw. Theory Tech. **51**(5), 1554–1559 (2003)
20. Chang, S.F., Jeng, Y.H., Chen, J.L.: Dual-band step-impedance bandpass filter for multimode wireless LANs. Electron. Lett. **40**(1), 38–39 (2004)
21. Zhang, Y.P., Sun, M.: Dual-band microstrip bandpass filter using stepped-impedance resonators with new coupling schemes. IEEE Trans. Microw. Theory Tech. **54**(10), 3779–3785 (2006)
22. Weng, M.H., Wu, H.W., Su, Y.K.: Compact and low loss dual-band bandpass filter using pseudo-interdigital stepped impedance resonators for WLANs. IEEE Microwave Wirel. Compon. Lett. **17**(3), 187–189 (2007)
23. Gómez-García, R., Muñoz-Ferreras, J.M., Feng, W., Psychogiou, D.: Balanced symmetrical quasi-reflectionless single-and dual-band bandpass planar filters. IEEE Microwave Wirel. Compon. Lett. **28**(9), 798–800 (2018)
24. Wang, L.T., Xiong, Y., Gong, L., Zhang, M., Li, H., Zhao, X.J.: Design of dual-band bandpass filter with multiple transmission zeros using transversal signal interaction concepts. IEEE Microwave Wirel. Compon. Lett. **29**(1), 32–34 (2018)
25. Gómez-García, R., Yang, L., Muñoz-Ferreras, J.M., Psychogiou, D.: Selectivity-enhancement technique for stepped-impedance-resonator dual-passband filters. IEEE Microwave Wirel. Compon. Lett. **29**(7), 453–455 (2019)
26. Ren, B., Liu, H., Ma, Z., Ohira, M., Wen, P., Wang, X., Guan, X.: Compact dual-band differential bandpass filter using quadruple-mode stepped-impedance square ring loaded resonators. IEEE Access **6**, 21850–21858 (2018)
27. Weng, M.H., Lan, S.W., Chang, S.J., Yang, R.Y.: Design of dual-band bandpass filter with simultaneous narrow-and wide-bandwidth and a wide stopband. IEEE Access **7**, 147694–147703 (2019)
28. Sánchez-Soriano, M.Á., Gómez-García, R.: Sharp-rejection wide-band dual-band bandpass planar filters with broadly-separated passbands. IEEE Microwave Wirel. Compon. Lett. **25**(2), 97–99 (2014)
29. Xu, J., Ji, Y.X., Miao, C., Wu, W.: Compact single-/dual-wideband BPF using stubs loaded SIR (SsLSIR). IEEE Microwave Wirel. Compon. Lett. **23**(7), 338–340 (2013)
30. Chin, K.S., Yeh, J.H.: Dual-wideband bandpass filter using short-circuited stepped-impedance resonators. IEEE Microwave Wirel. Compon. Lett. **19**(3), 155–157 (2009)
31. Wang, Y.R., Lin, Y.S.: Size-reduction of dual-wideband bandpass filters using bridged-T coils. IEEE Access **7**, 42836–42845 (2019)
32. Liang, G.Z., Chen, F.C.: A compact dual-wideband bandpass filter based on open-/short-circuited stubs. IEEE Access **8**, 20488–20492 (2020)

Simulation and Modeling of Novel TFETs

Puja Singh and Santosh Kumar Gupta

Abstract In this paper, we have proposed an Extended Source Dielectric Modulated Split Gate Tunnel Field Effect Transistor (ES DM SG TFET) based biosensor for label free detection. In the proposed structure, the source has been extended towards channel and cavity has been introduced between split gates. Gate region has been splitted into four parts: Two towards top gate side and two towards bottom gate side. Drain current and sensitivity of bio-molecules vary due to immobilization of bio-molecules into cavity region. The source extension results in improved drain current. The thickness of cavity and length of extended source have been varied from 7 to 14 nm and 2 to 10 nm, respectively for different dielectric constants. The sensitivity has been found to increase with dielectric constant of biomolecule. Further, the ES DM SG TFET based biosensor provides better I_{ON}/I_{OFF} current ratio and sensitivity as compared to dielectric modulated split gate TFET.

Keywords Extended source (ES) · Split gate (SG) · Dielectric modulation (DM) · TFET · Band to band tunneling (BTBT) · Biosensor · Sensitivity

1 Introduction

Recently, Biosensor has grown in popularity among researcher for label free detection of bio-molecules due to its diverse variety of applications in medical, agriculture, environmental fields and many more. Bio-molecules [1] like glucose, protein, streptavidin, DNA etc. are used for label free detection. FET [2] based biosensors have been studied due to its various advantages such as better sensitivity for level free detections, low cost, and higher scalability. The first FET based biosensor Ion sensitive Field Effect Transistor (ISFET) [3] was proposed by P. Bergveld in 1970. In this gate electrode region was replaced by ionic solution, reference electrode and

P. Singh (✉) · S. K. Gupta
Department of ECE, MNNIT Allahabad, Prayagraj 211004, India
e-mail: kmpujasingh221@gmail.com

S. K. Gupta
e-mail: skg@mnnit.ac.in

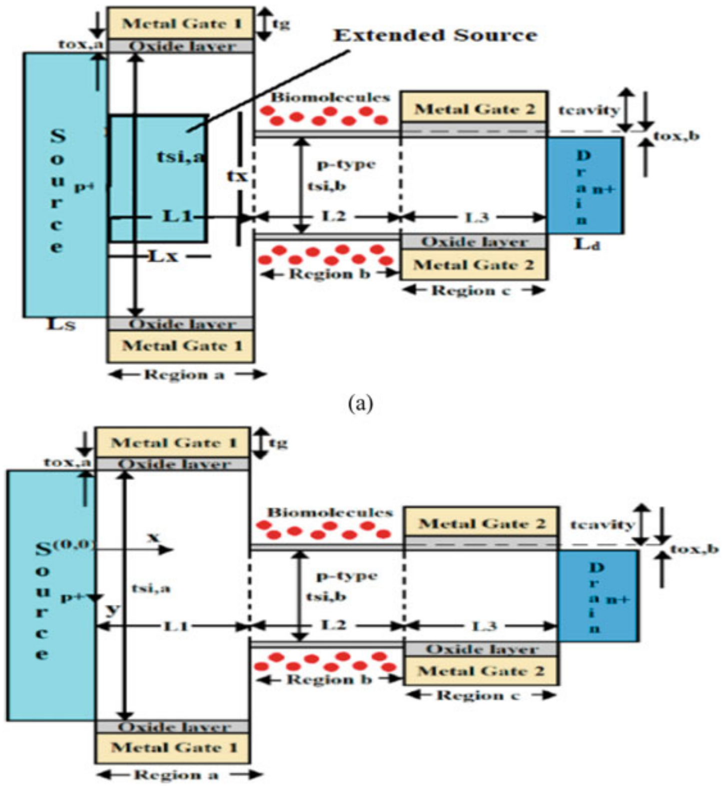
A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_45

ion-sensitive membrane. The electrical characteristics change due to charged particles present between the gate regions. It shows admirable sensitivity performance for charged bio-molecules but unable to sense the neutral bio-molecules. Later, a dielectric modulated FET [4] based biosensor had been introduced to sense neutral as well as charged bio-species. The FET has drawbacks such as short channel effects (SCEs), and to overcome this, a Dielectric modulated MOSFET based biosensor was introduced. The dielectric modulated transistors have a cavity under the gate for immobilization of bio-targets and due to this, the electrical characteristic of device get changed. The changes in parameters of transistors are key factor for the detection of a bio-species. However, MOSFETs also have some drawback, i.e., higher subthreshold slope (SS), lower I_{ON}/I_{OFF} current ratio, high leakage current etc. Hence, tunnel field effect transistor (TFET) [5] based biosensors were introduced because TFETs wok on band to band tunneling (BTBT) phenomenon at source and channel junction and due to BTBT, TFETs [6–8] can exhibit lower subthreshold slope that results in high sensitivity. DM TFET gives commendable results in the field of bio-sensing and various researches [9–12] are going on to improve the performance of DM TFETs.

Split gate TFET (SG TFET) in bio-sensing application was introduced by Saha et al. [13] in which device channel is left exposed in between the split gate. This SG TFET offered higher sensitivity and better ON current. In the present work, to enhance the performance of split gate TFET, an extended source dielectric modulated split gate tunnel field effect transistor (ES DM SG TFET) has been proposed. The proposed structure has been discussed in Sect. 2. The results are discussed in Sect. 3. Finally, Sect. 4 summarizes the findings.

2 Device Structure

Figures 1(a) and (b) show the schematic diagrams of extended source dielectric modulated split gate (ES DM SG) TFET and SG TFET which have been used as bio-sensing devices. It detects the presence of bio-targets depend on the permittivity andcharge of bio-molecules. In this device, the gate structure has been splitted into four parts that resembles the split gate configuration. The total length of channel has been splitted into 3 parts: region a, b and c of lengths L1, L2 and L3, respectively. The thickness of source is $t_{si,a}$ and the thickness of drain is $t_{si,b.}$ nanogaps at the source side are used to control the coupling between gate electrode and channel regions and sense the bio-molecules. So Higher source region has been used to increase the controllability of gate. A part of source has been extended towards channel region in the proposed structure to increase the tunneling area which results in an increased current. The cavity towards drain side has been used to reduce the ambipolar conduction and sense the biomolecule. So thinner width of drain in device helps to decrease in the leakage current and ambipolar current. The length of source is L_S and drain length is L_d. The thickness of the gate oxide below the gates 1 & 2 of regions a & c is denoted by $t_{ox,a}$, while thickness of oxide layer below bio-targets of region b is denoted by $t_{ox,b}$. The thickness of extended source is t_x and extended



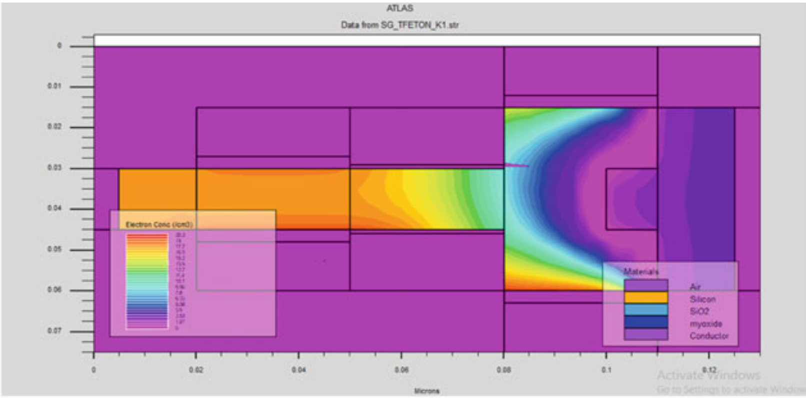
(b)

Fig. 1 Schematic diagram of **a** ES DM SG TFET **b** SG TFET

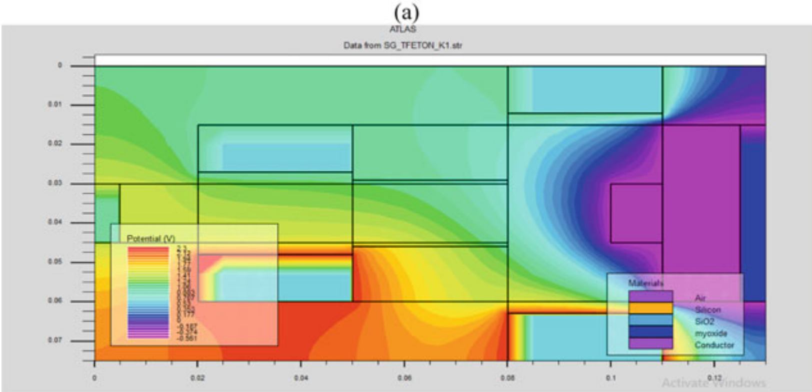
region length towards channel region is L_x. The thickness of cavity region is denoted by t_{cavity}. Thickness of metal gates 1 & 2 in the symmetric configuration is implied by t_g and work function by φ_m. The doping of source and drain regions are p-type and n-type, respectively, while the channel is doped lightly with p-type impurities. Figure 2(a) and (b) show the contour plots of electron concentration and potential variation of ES DM SG TFET respectively.

3 Results

This section investigates the electrical characteristics of ES DM SG TFET by varying the dielectric constant of the cavity, extended source length and bio targets thickness. For simulation, in Silvaco TCAD [14] tool BTBT model is activated. In the simulations, the geometries and biasing conditions similar to [16] are initially used and the parameters are calibrated with the experimental data accordingly for double



(a)



(b)

Fig. 2 ON state contour plots for **a** electron concentration, and **b** potential variation of ES DM SG TFET for Vds = 1 V

gate TFET as shown in Fig. 3. The parameters taken for the simulation of ES DM SG TFET based biosensor has been tabulated in Table 1.

The presence of bio-molecules in the cavity region is introduced with dielectric constant K > 1 during simulation. Neutral charged bio-molecules for various bio target thickness and permittivity are simulated to see the effect on drain current. The length of extended source has also been varied to better understand its effect on various device parameters, viz. ON current, subthreshold slope (SS) and threshold voltage (V_{th}).

Figure 4 represents the variation of drain current w.r.t. gate voltage of ES DM SG TFET and DM SG TFET based biosensors for various dielectric constants. In this, it clearly shows that drain current increases when source is extended towards channel, however, there is a slight increase in leakage current. The extended source increases the tunneling region between source and channel, and hence, the drain current has been increased. OFF current, ON current, SS, transconductance and V_{th} of devices have been found and listed in Table 2. It can be inferred from the results that for ES

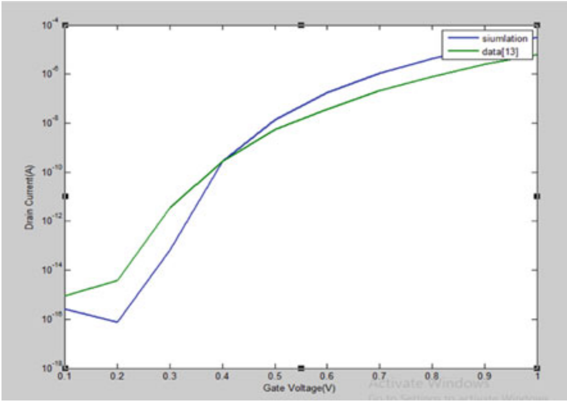


Fig. 3 Calibration of tunneling model with [16] for DG TFET

Table 1 Parameters of ES DM SG TFET

Parameters	Values	Parameters	Values
$N_{ch}(cm^{-3})$	1×10^{17}	$t_{si,a}$ (nm)	45
N_s (cm^{-3})	5×10^{20}	$t_{si,b}$ (nm)	15
N_d (cm^{-3})	1×10^{18}	Φ_m (eV)	4.3
t_g (nm)	12	t_x (nm)	15
t_{cavity} (nm)	14	L1, L2, L3 (nm)	30
$t_{ox,a}$ (nm)	3	L_x (nm)	10
$t_{ox,b}$ (nm)	1	L_S, L_d (nm)	20

DM SG TFET device, the SS and threshold voltage is lower and I_{ON}/I_{OFF} ratio is higher as compared to SG TFET for K = 20.

Figures 5, 6, 7 and 8 shows the effect of dielectric constant on drain current for an extended source length Lx. It can be observed from Fig. 5 for extended source length Lx = 2 nm with fully filled cavity. The difference in drain current w.r.t. gate voltage is shown for neutral bio-molecules with different dielectric constants (K = 5, 10 and 20) w.r.t. empty cavity i.e. K = 1. Again in Fig. 6 for Lx = 5 nm, variations of drain current w.r,t. gate voltage is shown. Further increase in the length of extended source to Lx = 10 nm the drain current for K = 5, 10 and 20 differs from K = 1 as illustrated in Fig. 7. Therefore, it can be concluded from the figures that raising the dielectric constant of the bio-species tend to increase the drain current w.r.t. gate voltage. Figure 8 represents the variation of drain current of biosensor w.r.t. gate voltage at constant dielectric K = 5. From Fig. 8, it is observed that the drain current increases with increase in source length Lx.

The characteristic of the ES DM SG TFET established biosensor to immobilization of neutral bio-molecules is shown in Fig. 9 through drain current variation with gate voltage at drain bias Vds = 1 V, extended source length Lx = 10 nm and K = 5 for different heights of the bio-molecules. It can be observed that by varying t_{bio} drain current also changes w.r.t. gate voltage.

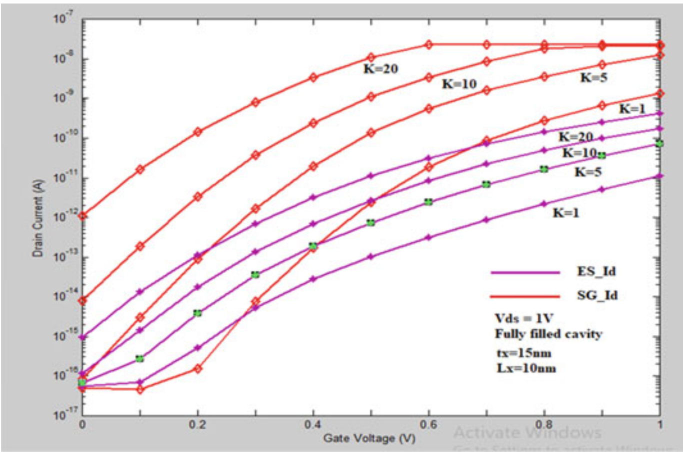


Fig. 4 Dielectric constant impact on I_D of SG TFET and ES DM SG TFET based biosensor for extended source length Lx = 10 nm

Table 2 Various output parameters of ES TFET and SG TFET

Parameters	I_{ON} (μA)	I_{OFF} (fA)	SS (mV/dec)	I_{ON}/I_{OFF} ($\times 10^6$)	V_{th} (V)	g_m (mho)
ES DM SG TFET	0.0129	0.0842	64.9068	153.104	0.2771	5.78471×10^{-8}
DM SG TFET	7.25×10^{-5}	0.0657	86.5744	1.10339	0.2974	3.5795×10^{-10}

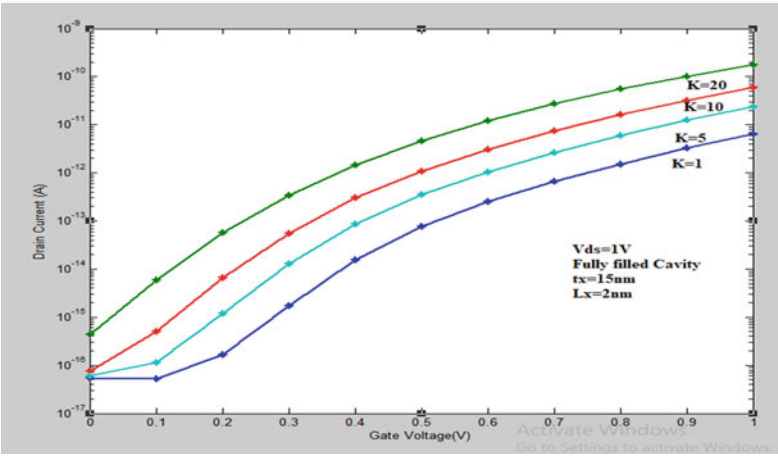


Fig. 5 Dielectric constant influence on I_D of ES DM SG TFET based biosensor for extended source length Lx = 2 nm

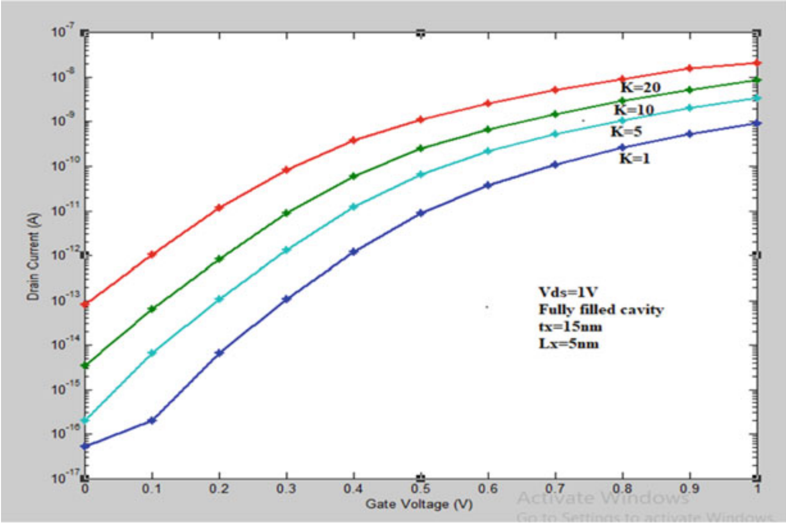


Fig. 6 Dielectric constant influence on I_D of ES DM SG TFET based biosensor for extended source length Lx = 5 nm

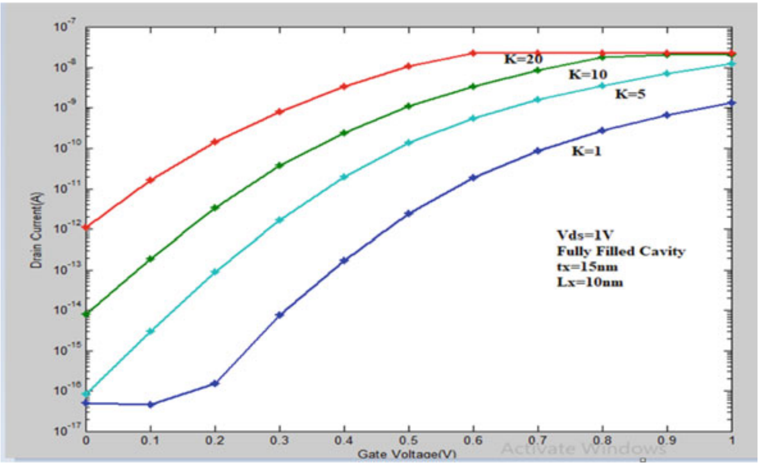


Fig. 7 Dielectric constant influence on I_D of ES DM SG TFET based biosensor for extended source length Lx = 10 nm

Figure 10 represents the change in surface potential along channel length of ES DM SG TFET based biosensor for different dielectric constants. In this it clearly noticed that the surface potential increases by extending source towards channel. An increase in surface potential increases the inversion charge density and more energy band bending. The higher energy band bending results in shortened tunneling

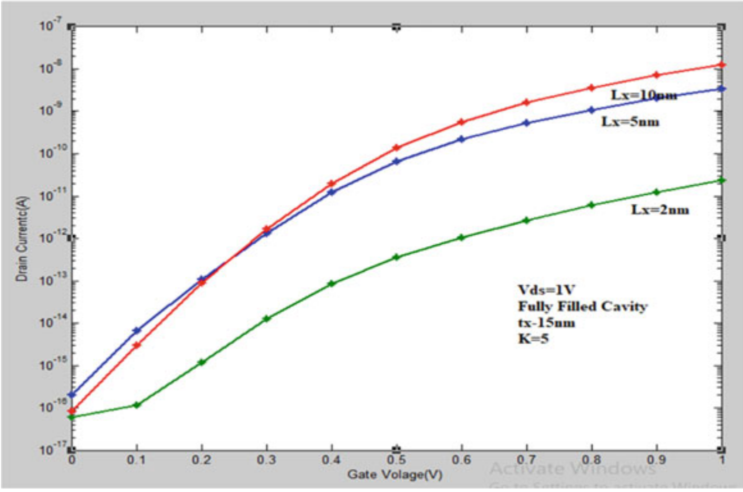


Fig. 8 Effect of variation of extended source length on I_D of ES DM SG TFET based biosensor for neutral bio-molecules K = 5

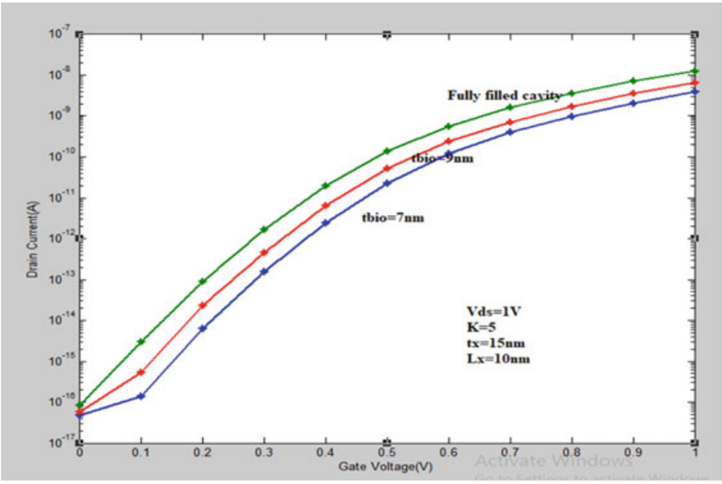


Fig. 9 Influence of t_{bio} on I_D of ES DM SG TFET based biosensor for extended source length Lx = 10 nm and dielectric constant K = 5

path is the main cause of increased tunneling across source channel interface. Also, extending the source into the channel increases the tunneling area, thus, the drain current increases.

Transconductance g_m is expressed by first order differentiation of drain current w.r.t. gate source voltage as shown in Eq. (1). For faster switching speed gm should be high.

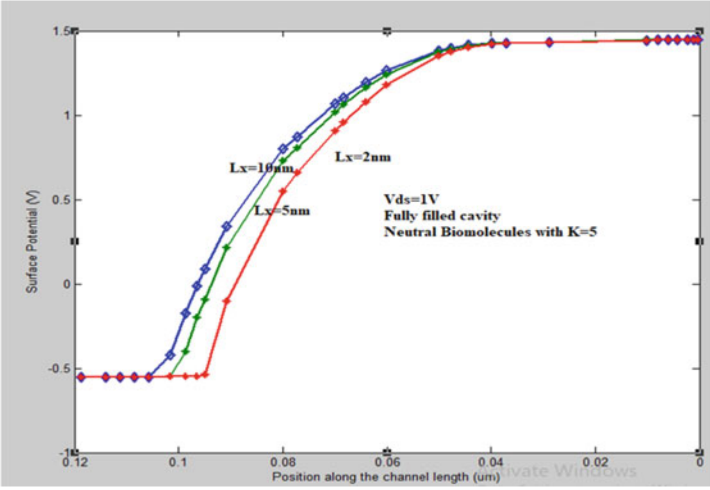


Fig. 10 Surface potential along channel of ES DM SG TFET based biosensor for constant dielectric constant K = 5

$$g_m = \frac{dI_d}{dV_{gs}} \tag{1}$$

Figures 11, 12 and 13 show the influence of dielectric constant (K = 1, 5, 10 and 20) on transconductance with fully filled cavity having constant extended source lengths of 2, 5 and 10 nm, respectively. So, it can be concluded from the Figs. 11, 12 and 13 that raising the dielectric constant of the bio-species tend to increase the transconductance w.r.t. gate voltage.

Figure 14 shows the variation of transconductance of biosensor w.r.t. gate voltage for different source extensions into channel at constant dielectric K = 5. From Fig. 14 it is clearly observed that by increasing source length Lx, tranconductance also increases.

Transconductance generation factor (TGF) is very useful parameter for device efficiency and the trade-off between operating frequency and power dissipation. TGF should be high for low power circuit operation with high speed. The mathematical equation of TGF is given by Eq. (2).

$$TGF = \frac{g_m}{I_d} \tag{2}$$

The influence of TGF having extended source lengths Lx = 2, 5, and 10 nm are shown in Figs.15, 16 and 17 for neutral bio-molecules with different dielectric constants K = 1, 5, 10 and 20. It can be concluded from the results that raising the dielectric constant of the bio-species tend to increase the TGF w.r.t. gate voltage. Figure 18 shows the variation of TGF of biosensor w.r.t. gate voltage at constant

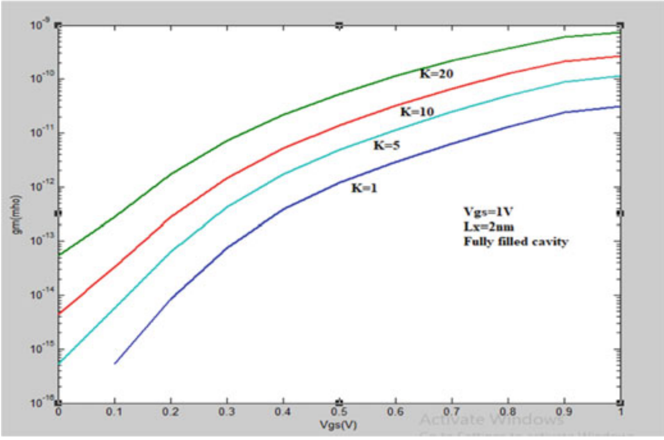


Fig. 11 Influence of dielectric constant on transconductance of ES DM SG TFET based biosensor for expended source length Lx = 2 nm

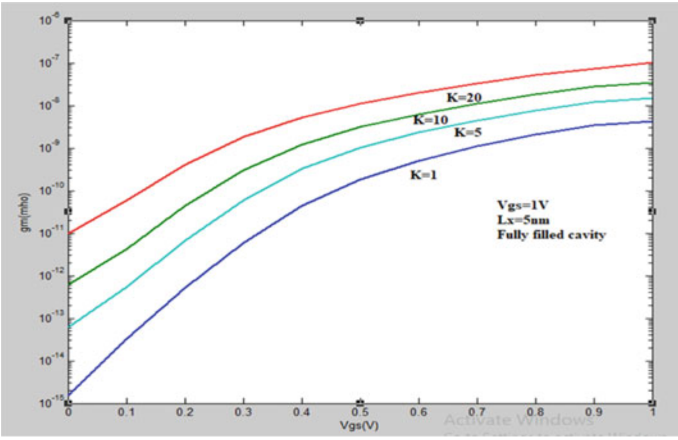


Fig. 12 Influence of dielectric constant on transconductance of ES DM SG TFET based biosensor for expended source length Lx = 5 nm

dielectric K = 5 for different extended source lengths Lx = 2, 5, and 10 nm. From Fig. 18 it is observed that TGF increases by increasing source length Lx.

The characteristics of the ES DM SG TFET based biosensor to immobilization of neutral bio-molecules are demonstrated in Figs. 19 and 20 through transconductance and TGF variations, respectively, w.r.t. gate voltage with Vds = 1 V and different heights of the bio-molecules (t_{bio}). It can be observed that by varying t_{bio} transconductance and TGF also changes w.r.t. gate voltage.

Table 3 shows threshold voltage shift [15] for various dielectric constants of bio-molecules w.r.t. K = 1 for different extended source lengths. Sensitivity parameter

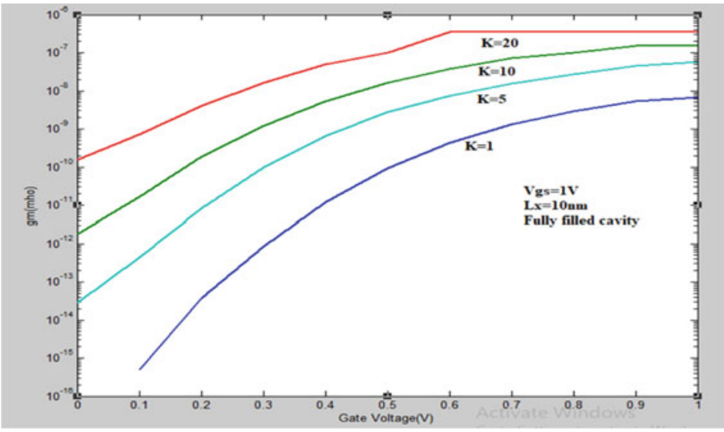


Fig. 13 Influence of dielectric constant on transconductance of ES DM SG TFET based biosensor for expended source length Lx = 10 nm

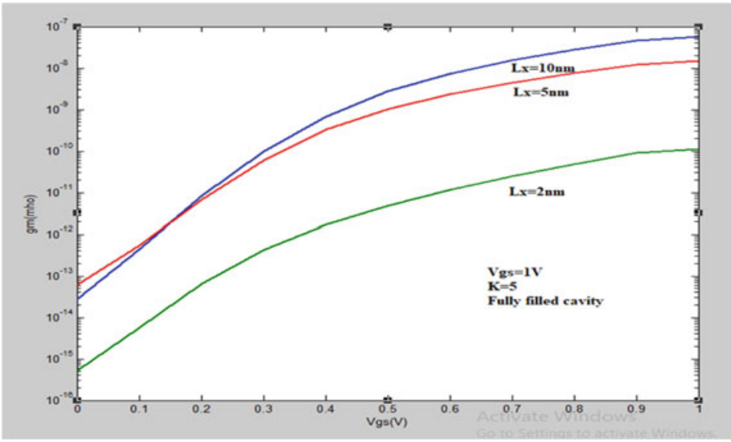


Fig. 14 Variation of extended source length variation on transconductance curve of ES DM SG TFET based biosensor for neutral bio-molecules K = 5

of threshold voltage $\left(\Delta V_{th\_SS}\right)$ is defined as

$$\Delta V_{th\_SS} = V_{th\_air} - V_{th\_bio} \tag{3}$$

where ΔV_{th_SS} is the threshold voltage shift, V_{th_air} and V_{th_bio} denote the threshold voltage at the absence of bio-molecules, and when the cavity is occupied by biomolecules, respectively. The increase in the dielectric constant of bio-molecules makes the device ON very quickly. At higher values of K, lower threshold voltage

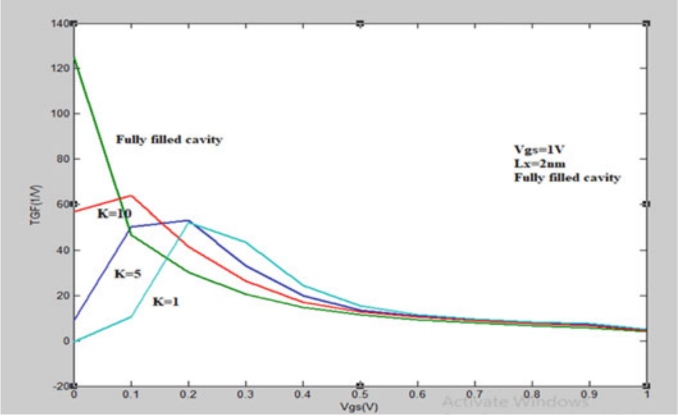


Fig. 15 Impact of dielectric constant on TGF of ES DM SG TFET based biosensor for extended source length Lx = 2 nm

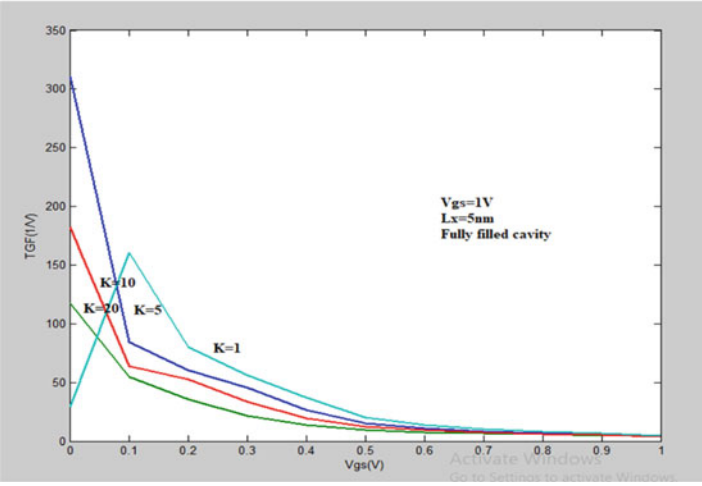


Fig. 16 Impact of dielectric constant on TGF of ES DM SG TFET based biosensor for extended source length Lx = 5 nm

is obtained. Therefore, as the value of K is increased, the sensitivity of threshold voltage also increases that makes device ON early.

It is observed in Table 3 as values of K increases, sensitivity also starts increasing. In the proposed device, maximum threshold voltage sensitivity is obtained for extended source length Lx = 10 nm and dielectric constant K = 20.

Table 4 shows the variation in drain current sensitivity at different dielectric constants and different extended source lengths. Higher sensitivity is preferred for biosensor based devices because for higher probability of sensing, sensitivity factor

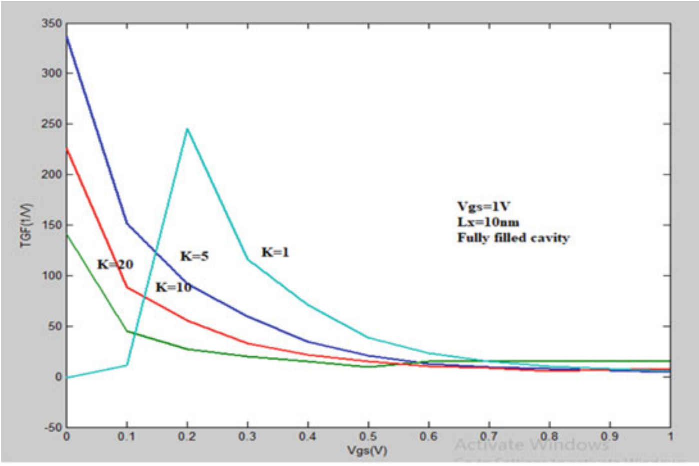


Fig. 17 Impact of dielectric constant on TGF of ES DM SG TFET based biosensor for extended source length Lx = 10 nm

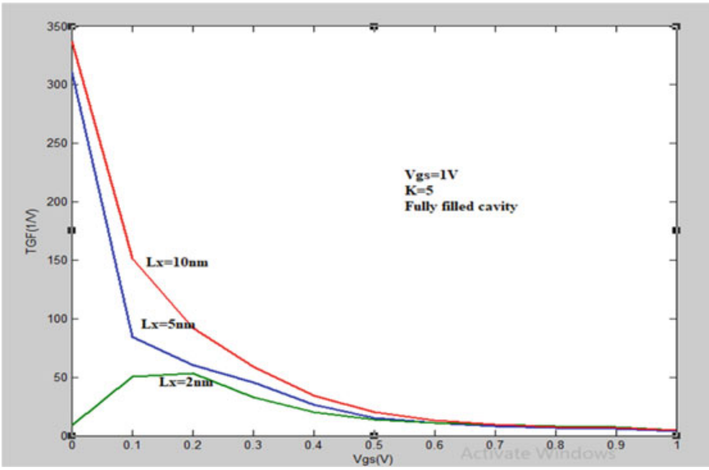


Fig. 18 Variation of extended source length variation on TGF of ES DM SG TFET based biosensor for neutral bio-molecules K = 5

should be high. Drain current sensitivity $\left(I_{d\_SS}\right)$ is calculated by

$$I_{d\_SS} = \frac{I_{d\_bio} - I_{d\_air}}{I_{d\_air}} \tag{4}$$

From Table 4 it is clear that with increase in dielectric constant, the drain current sensitivity also starts increasing. However, it is observed that with increase in

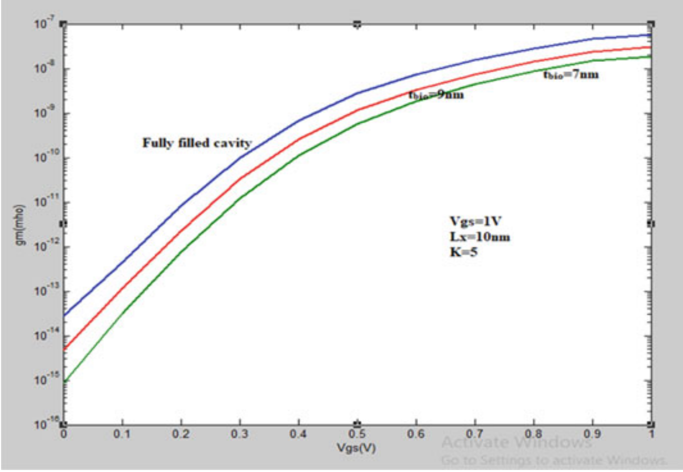


Fig. 19 Impact of t_{bio} on transconductance of ES DM SG TFET based biosensor for extended source length Lx = 10 nm and dielectric constant K = 5

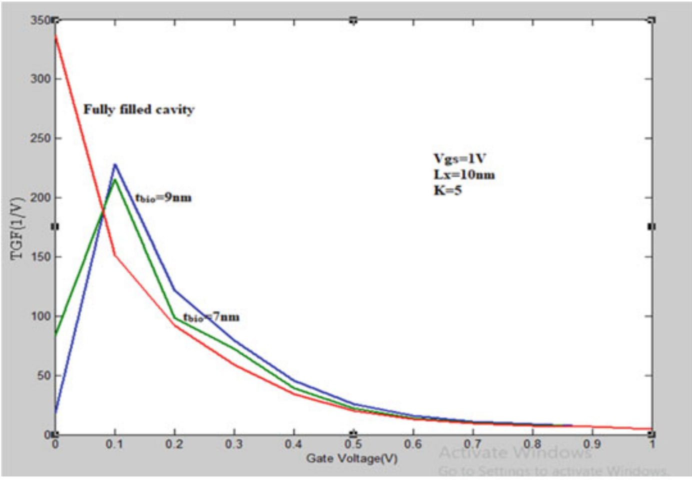


Fig. 20 Impact of t_{bio} on TGF of ES DM SG TFET based biosensor for extended source length Lx = 10 nm and dielectric constant K = 5

Table 3 Threshold voltage sensitivity at different extended source length

Parameters	Lx = 2 nm	Lx = 5 nm	Lx = 10 nm
K = 5	0.006099	0.016832	0.024208
K = 10	0.017584	0.029219	0.118378
K = 20	0.037823	0.018795	0.293909

Table 4 Variation of drain current sensitivity

Parameters	Lx = 2 nm	Lx = 5 nm	Lx = 10 nm
K = 5	2.7104	2.67713	8.49812
K = 10	8.3071	8.020306	14.8499
K = 20	26.414	21.01259	15.9744

Table 5 Variation of Subthreshold slope sensitivity

Parameters	Lx = 2 nm	Lx = 5 nm	Lx = 10 nm
K = 5	– 0.00061633	– 0.0034192	– 0.098479
K = 10	0.070952202	– 0.1780784	– 0.231643
K = 20	0.087796472	– 0.3655758	– 0.437197

extended source, sensitivity is not as much high as for Lx = 2 nm. Therefore, there is a tradeoff between sensitivity and low power dissipation.

Table 5 shows variation in the average subthreshold slope sensitivity as a function of dielectric constants and extended source lengths. By increasing the dielectric constant inside cavity, the tunneling junction capacitance becomes higher that results in an increase of leakage current. Formula for average value of subthreshold slope sensitivity ($SS_S(avg)$) is calculated using

$$SS_S(avg) = \frac{SS_{air} - SS_{bio}}{SS_{air}} \tag{5}$$

From Table 5 it is clearly observed that by increasing K values, subthreshold slope sensitivity also get increased.

4 Conclusion

In this work, an ES DM SG TFET has been analyzed for label free bio-sensing. The influence of dielectric constant of bio-molecules and extended source length into channel on the surface potential and drain current, which are considered as important sensitivity parameters of the detection device are investigated. By varying length of extended source it is observed that drain current, surface potential, transconductance and transconductance generation factor increases. It has been found that the extended source dielectric modulated split gate TFET based biosensor gives an increase in $I_{\mathrm{ON}}/I_{\mathrm{OFF}}$ current ratio by two orders of magnitude, subthreshold slope decreases by 25%, higher On-current (I_{ON}) and higher sensitivity factors as compared to split gate TFETs. Depending on the comparative analysis w.r.t. conventional TFET and SG TFET, it can be predicted that the ES DM SG TFET device can be a better device for future label free bio-sensing applications.

References

1. Chen, X., et al.: Electrical nanogap devices for biosensing. MaterialsToday **13**(11), 28–41 (2010). https://doi.org/10.1016/S1369-7021(10)70201-7
2. Bulyha, A., Heitzinger, C., Mauser, N.J.: Bio-sensors: modelling and simulation of biologically sensitive field-effect-transistors. ERCIM News, vol. 04 (2011)
3. Bergveld, P.: Development of an ion-sensitive solid-state device for neurophysiological measurements. IEEE Trans. Biomed. Eng. **17**(1), 70–71 (1970). https://doi.org/10.1109/TBME.1970.4502688
4. Im, H., Huang, X.-J., Gu, B., Choi, Y.-K.: A dielectric-modulated field effect transistor for biosensing. Nat. Nanotechnol. **2**(7), 430–434 (2007)
5. Narang, R., Saxena, M., Gupta, M.: Comparative analysis of dielectric-modulated FET and TFET-based biosensor. IEEE Trans. Nanotechnol. **14**(3), 427–435 (2015). https://doi.org/10.1109/TNANO.2015.2396899
6. Seabaugh, A.C., Zhang, Q.: Low-Voltage Tunnel Transistors for beyond CMOS Logic. Proc. IEEE **98**(12), 2095–2110 (2010)
7. Saurabh, S., Kumar, M.J.: Estimation and compensation of process induced variations in nanoscale tunnel field effect transistor (TFETs) for improved reliability. IEEE Trans. Device Mater. Reliab. **10**, 390–395 (2010)
8. Saurabh, S., Jagadesh Kumar, M.: Fundamental of Tunnel Field Effect Transistors. John and Wiley and Sons Ltd. Published 2017
9. Anand, S., Singh, A., Intekhab Amin, S., Thool, A.S.: Design and performance analysis of dielectrically modulated doping-less tunnel FET-based label free biosensor. IEEE Sens. J. **19**(12), 4369–4374 (2019)
10. Reddy, N.N., Panda, D.K.: Simulation study of dielectric modulated dual material gate TFET based biosensor by considering ambipolar conduction. Silicon **13**, 4545–4551 (2021). https://doi.org/10.1007/s12633-020-00784-9
11. Nagendra Reddy, N., Panda, D.K.: A comprehensive review on tunnel field-effect transistor (TFET) based biosensors: recent advances and future prospects on device structure and sensitivity. Springer, Cham (2020). https://doi.org/10.1007/s12633-020-00657-1
12. Masurkar, N., Varma, S., Arava, L.M.R.: Supported and suspended 2D material-based FET biosensors. Electrochem **1**(3), 260–277 (2020)
13. Saha, P., Sarkar, S.K.: Drain current characterization of dielectric modulated split gate TFET for bio-sensing application. Mater. Sci. Semiconductor Process. **124**, 105598 (2021)
14. ATLAS: Device simulation software. Silvaco Int, Santa Clara (2012)
15. Anam, A., Anand, S., Intekhab Amin, S.: Design and performance analysis of tunnel field effect transistor with buried strained $Si_{1-x}Ge_x$ source structure based biosensor for sensitivity enhancement. IEEE Sens. J. **20**(22), 13178–13185 (2020)
16. Bala, S., Khosla, M.: Design and simulation of nanoscale double-gate TFET/tunnel CNTFET. J. Semiconductors **39**(4), 044001 (2018)

Modeling and Investigation of Novel Two-Port UWB-MIMO Antenna with Enhanced Isolation

Aditya Kumar Singh, Ajay Kumar Dwivedi, Nagesh Kallollu Narayanaswamy, Vivek Singh, and R. S. Yadav

Abstract In this communication, a low-profile two-port MIMO antenna is proposed and investigated for bandwidth and gain enhancement. Low mutual coupling (≤–31 dB) and ultra-wideband (UWB) (6.39–13.46 GHz) are two prominent attributes of the proposed configuration. The diversity performance of the proposed antenna is also calculated in terms of ECC, DG, and TARC. DGS is utilized for the decoupling mechanism. Simulation and optimization are carried out by ANSYS HFSS 18 electromagnetic solver.

Keywords MIMO · Diversity gain · Envelope correlation coefficient · Defected ground structure

1 Introduction

Through the use of advanced antenna systems, wireless communication network coverage, capacity, throughput, and customer experience have improved in recent years. Beamforming, mm-wave communication, and multiple-input multiple-output (MIMO) antenna systems are among the most advanced antenna systems [1–3]. In addition to increasing throughput and system capacity, MIMO systems may also reduce fading effects, reduce tapping susceptibility, and increase system coverage by enhancing signal quality. Microstrip antennas provide several benefits over other antenna designs, including ease of manufacture, cheap cost, compact size, multi-band capabilities, support for both linear and circular polarizations, and the ability to be placed on a solid surface. Surface waves are the most common cause of mutual coupling in microstrip MIMO antennas. In MIMO, there is a trade-off between

A. K. Singh · R. S. Yadav
Department of Electronics and Communication Engineering, University of Allahabad, Prayagraj, India

A. K. Dwivedi · N. K. Narayanaswamy · V. Singh (✉)
Department of Electronics and Communication Engineering, Nagarjuna College of Engineering and Technology, Bengaluru, India
e-mail: vivek.10singh@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_46

coupling and size. For MIMO communication, a small antenna with excellent isolation is recommended due to the limited amount of area available [4, 5].

Several approaches for reducing mutual coupling between radiating components in MIMO antenna systems have been described, including the use of defective ground structures, diversity techniques, neutralization lines, partially reflecting surface, and orthogonal orientation-based designs [6]. Ultra-Wideband (UWB) technology has certain advantages, including data transfer at high speeds, economicalcost, and ease of production. However, such a major area of technology suffers from multipath fading in real implementations [7]. To address this issue, MIMO technology was developed. When UWB and MIMO technologies are coupled with space multipath and parallel transmission of several signals, considerable multiplexing and diversity benefits can be achieved. X-band has broad applications such as satellite, mobile, radiolocation, etc., and the Ku/K band has been used in radar and satellite television broadcasting. Several designs addressing the above applications, as well as their parametric study, have been published [8, 9].

In the current work, we propose a defected ground structure (DGS) loaded two-port MIMO antenna with optimized dimensions for bandwidth and gain enhancement for X-band applications. The proposed design has an ultra-wideband of 6.39–13.46 GHz (71.23%) and a rectangular stepped-slot structure in the ground plane to obtain immenseisolation between the antenna elements. Multiple-input-multiple-output antenna is modeled and simulated with the help of electromagnetic tool HFSS and evolution of antenna design, results and conclusion for the proposed work is conferred in foregoing sections.

2 Design Methodologies and Geometrical Specifications

The antenna top view (green) and bottom view (orange) are illustrated in Fig. 1. Dimensions are given I n millimeters. FR4 substrate is used for the suggested design. The thickness of the substrate is 1.6 mm, and the loss tangent is 0.02.

Step-wise evolution path to obtain the proposed design and corresponding scattering parameters plots are mentioned in Fig. 2 and Fig. 3. Ant-1 is obtained by designing two symmetrical structures of deformed hexagon shape by keeping them apart from each other at an optimized distance of D (cf. Figure 1). Ant-2 is achieved by introducing two parallel asymmetrical rectangles A (12×3 mm^2) and B (20×3 mm^2) in the ground plane of Ant-1. Antenna-3 (proposed) is obtained by incorporating rectangle C (20×3 mm^2) in the middleware portion and by cuttingslot of size (4×1mm^2) in-plane B by Antenna 2. Ant-1 has a narrow single band with isolation greater than 15 dB. By the introduction of DGS of rectangular shape, the resonant frequency has been shifted to a higher frequency range with wider bandwidth and with improved isolation (cf. Ant-2 and Fig. 3).Further introduction of rectangular

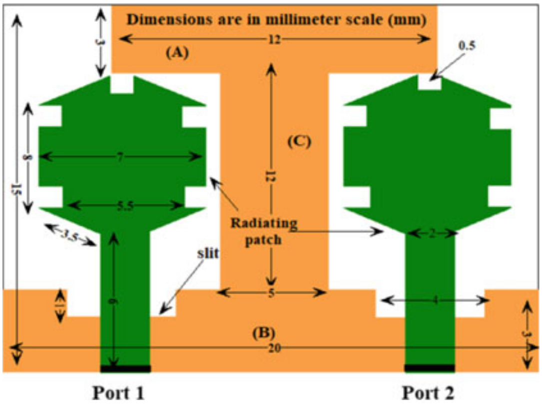


Fig. 1 Antenna's frontal (green) and rear (orange) views

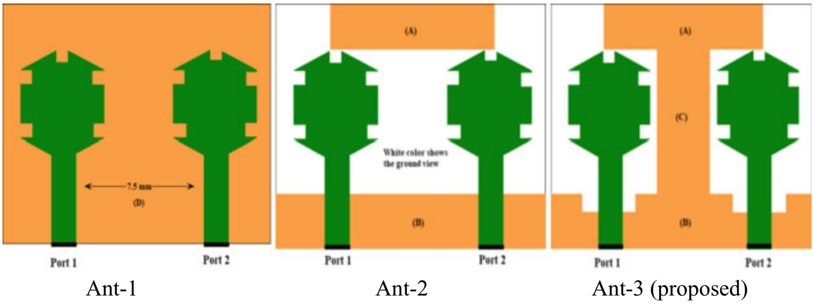


Fig. 2 Step-wise evolution of the Ant-1, Ant-2 and Ant-3 (Proposed)

shape design C on DGS result in proposed antenna Ant-3 which finally achieves ultra-wideband characteristics with isolation greater than 30 dB.

3 Results and Discussion

The proposed antenna is investigated in terms of returnloss, gain, radiation efficiency, radiation patterns, mutual coupling between ports, ECC, DG, and TARC. Table-1, represents the data of Antenna 1, Antenna 2 and Antenna 3 (proposed) discretely at both the ports, i.e. port-1, and port-2. The Table 1 depicts the operating band and its percentage impedance bandwidth, resonant frequency (in GHz) of the concerned band, reflection coefficient (in dB), and peak gain (in dBi).

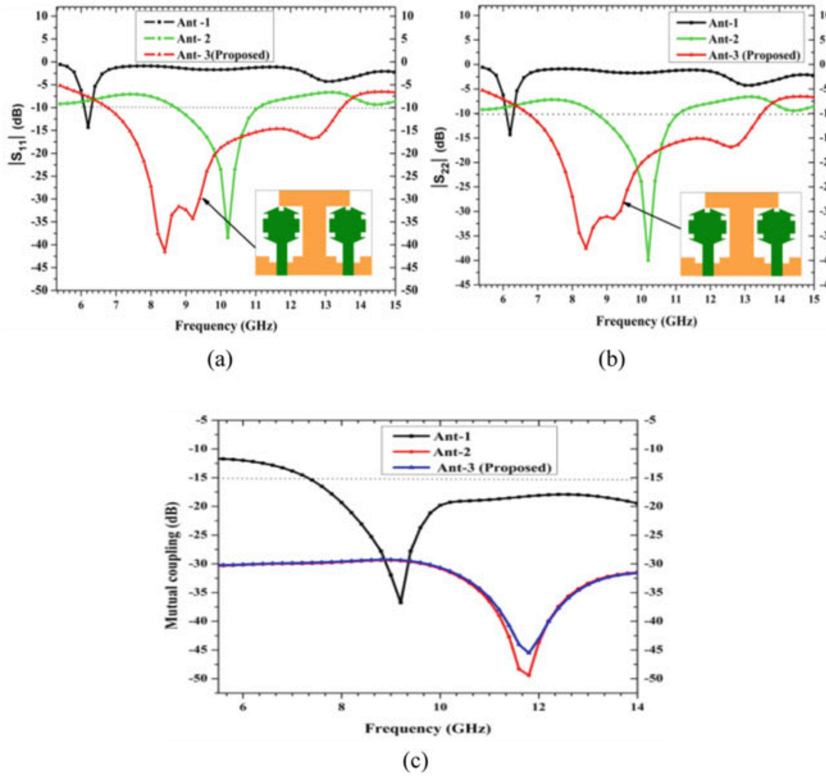
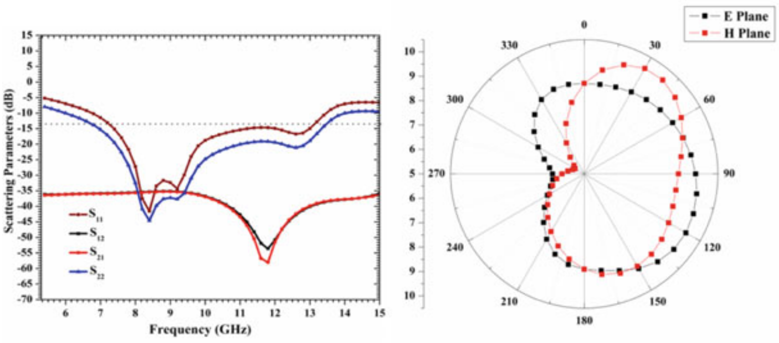


Fig. 3 **a** $|S_{11}|$, **b** $|S_{22}|$ and **c** Mutual coupling (between the ports) of the (Ant 1, Ant 2, Ant 3)

Table 1 ANT 1, ANT 2 and ANT 3 port characteristics

Antenna	PortNo	Operating band (GHz)/Impedance bandwidth (in%)	Resonant frequency (GHz)/Reflection coefficients (dB)	Peak gain (dBi)	Isolation/Mutual coupling (dB)
ANT-1	Port-1 (S_{11}) Port-2 (S_{22})	6.02–6.31/4.70 6.01–6.31/4.69	6.18/–14.18 6.18/–14.19	1.21 1.21	≤–13
ANT-2	Port-1 (S_{11}) Port-2 (S_{22})	8.67–11.05/24.34 8.36–10.56/23.51	10.17/–39.81 10.17/–39.82	2.64 2.63	≤–31
ANT-3	Port-1 (S_{11}) Port-2 (S_{22})	6.39–13.46/71.23 6.40–13.46/70.36	8.38/–41.38 8.38/–39.26	3.89 3.90	≤–31



(a) (b)

Fig. 4 **a** Scattering parameters plot for the proposed Antenna 3 **b** Radiation patterns at 8.38 GHz (E and H-plane)

3.1 Scattering Parameters and Radiation Patterns

Figure 4(a) depicts the simulated mutual coupling (S_{21} and S_{12}) and return loss (S_{11} and S_{22}) between the antenna components for both ports. MIMO arrangement with a single ultra-wideband from 6.39–13.46 (71.23%) for $|S_{11}| < 10$ dB and $|S_{12}| < 30$ dB is observed.

The simulated far-field radiation pattern (dB) at all resonating frequencies is presented in Fig. 4(b). Figure 4(b) illustrates that the antenna is ideally suitable for Omni-directional radiation patterns.

3.2 Radiation Efficiency and Gain

Figure 5(a) and (b) show the simulated gain (dBi) and radiation efficiency curve respectively.For the resonating bandwidth of 6.39–13.46 GHz, the gain varies from 1.56 to 3.89 dBi. The radiation efficiency of the MIMO antenna is >91% for all working bands of (6.39–13.46 GHz). Figure 5(b) demonstrates the simulated gain of Ant1, Ant 2 and Ant 3. From the perusal of Fig. 5(b), it can be observed that the Ant-3 has the highest gain compared to Ant-1 and Ant-2.

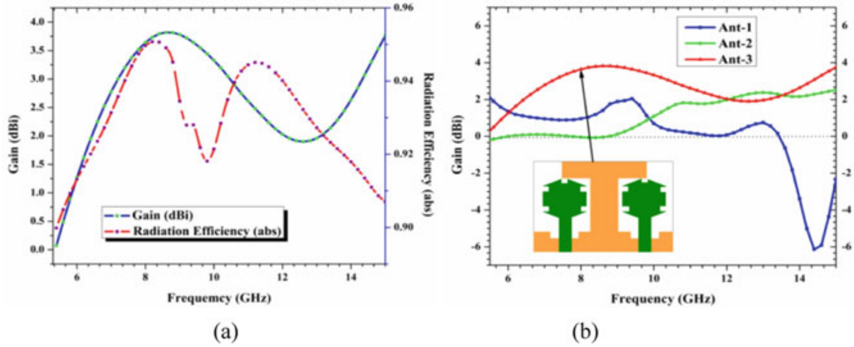


Fig. 5 **a** Plot of the Ant 3's gain (dBi) and radiation efficiency (abs) **b** Comparative observation of gain of Ant 1, Ant 2 and Ant 3

3.3 Electric Field Distribution

Figure 6 depicts surface current patterns with and without decoupling stubs at a resonant frequency of 8.38 GHz. Only port 2 is excited to monitor the surface current distribution. Figure 6 shows significant mutual coupling in the absence of a decoupling stub. This is due to the MIMO antenna's coupling current flowing from port 2 to port 1. The insertion of a decoupling stub focuses most of the current on port 2 to the right of the decoupling stub, resulting in high isolation on the other port.

4 MIMO Diversity Performances

To evaluate the capacity of a MIMO-UWB antenna design, three essential characteristics must be investigated: envelope correlation coefficient (ECC), total active

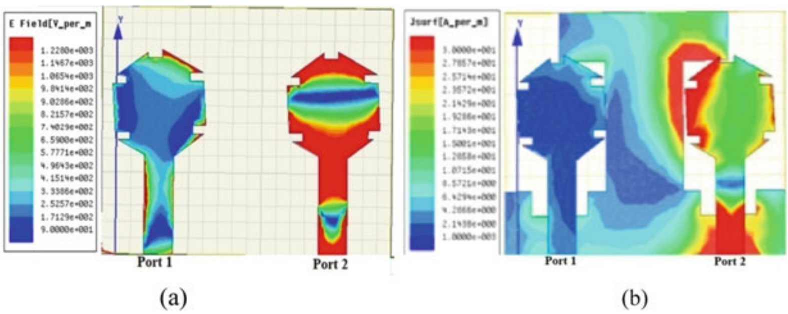


Fig. 6 Surface current distribution at 8.38 GHz, **a** Without using decoupling stub **b** With decoupling stub

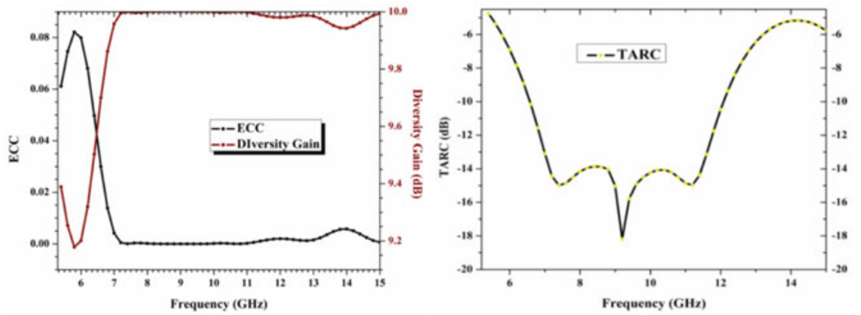


Fig. 7 **a** ECC/DG plot of the suggested setup. **b** TARC of the considered antenna (Ant-3)

reflection coefficient (TARC), and diversity gain (DG) features. The permissible standards limitations are ECC < 0.5, TARC < 10, and DG near 10 dB [10].

$$\mathrm{ECC} = \frac{|S_{11} * S_{12} * S_{21} * S_{22}|}{\left(1 - |S_{11}^2| - |S_{21}^2\right)\left(1 - |S_{22}^2| - |S_{12}^2\right)}. \tag{1}$$

$$\text{Diversity Gain} = 10 \times \sqrt{1 - (\mathrm{ECC})^2} \tag{2}$$

$$\mathrm{TARC} = \frac{\sqrt{(S_{11} + S_{12})^2 + (S_{22} + S_{21})^2}}{\sqrt{2}} \tag{3}$$

The value of DG for the proposed antenna is greater than 9.92 dB which legitimizes its diversity performance for multiple-input-multiple-output applications (cf. Fig. 7a).The total active reflection coefficientis the major element to better symbolize the diversity indices for the 2-element antenna system (MIMO). Equation 3 calculates the TARC, which is utilized in multiple-input multiple-output (MIMO) antenna systems [11].

4.1 Comparative Analysis with Recent Antennas

To justify the novelty of the proposed configuration it is necessary to do a comparison study between the suggested MIMO antenna and the previously published MIMO antennas in Table 2 for the same operational bandwidthin terms of antennas dimensions, the material used, impedance bandwidth, radiation efficiency, isolation, ECC, peak gain.

Table 2 Comparative study between the proposed MIMO antenna operating band and impedance bandwidth

Antennas dimensions (mm^3)	Material used	Operating band (GHz)/Impedance bandwidth (in%)	Radiation efficiency (%)	Isolation (dB)	ECC	Max gain (dBi)
50 × 100 × 1.6 [12]	FR4(εr = 4.4, tanδ = 0.02)	0.803–0.823/28.15	NR	− 17	<0.2	2.4
25 × 30 × 1.6 [13]	FR4(εr = 4.4, tanδ = 0.02)	7.4–11.8/45	95	− 26	<0.2	4.2
16 × 20 × 1.6 [14]	FR4(εr = 4.3, tanδ = 0.02)	6.5–15.8/63.4	NR	− 15	<0.1	3.22
75.7 × 58.35 × 0.76 [15]	FR4(εr = 3.5, tanδ = 0.08)	3.3–12/113	NR	NR	<0.1	3.60
27 × 21 × 0.6 [16]	FR4(εr = 4.4, tanδ = 0.02)	5.8-5.3, 7.30–7.66/4.81	79	− 22.5	NR	9
30 × 25 × 0.8 [17]	FR4(εr = 4.3, tanδ = 0.02)	7.2–9.6/28.57	NR	− 17	<0.2	4
48 × 48 × 0.8 [18]	FR4(εr = 4.4, tanδ = 0.02)	2.5–12/71	NR	− 18	<0.05	2.61
20 × 15 × 1.6 (proposed)	FR4(εr = 4.4, tanδ = 0.02)	6.39–13.46/71.23	91	− 30	<0.1	3.89

5 Conclusions

This paper presents anew and small MIMO antenna with a unique decoupling structure stub, increased bandwidth of 71.23%, and excellent isolation (<30 dB). The antenna is as small as 20 × 15 mm^2, and the isolation is enhanced by the use of a decoupling stub. The gain ranges from 1.56 to 3.89 dB, with a complete agreement between theoretical and simulated findings. The peak gain, ECC, TARC, and DG values demonstrate that the suggested antenna is suitable for a wide range of MIMO-UWB wireless applications.

References

1. Garg, R., Bhartia, P., Bahl, I., Ittipiboon, A.: Microstrip Antenna Design Handbook. Artech House, Norwood (2001)

2. Jehangir, S.S., Sharawi, M.: S: A miniaturized UWB biplanar Yagi-like MIMO antenna system. IEEE Antennas Wirel. Propag. Lett. **16**(1), 2320–2323 (2017)
3. Dwivedi, A.K., Sharma, A., Tripathi, P.N., Singh, A.K.: Dual Placed CDRA-Based MIMO Antenna for Wi-f/WLAN Applications. Springer, Singapore (2020). https://doi.org/10.1007/978-9815-6840-4
4. Dwivedi, A.K., Sharma, A., Singh, A.K., Singh, V.: Circularly polarized two port MIMO cylindrical DRA for 5G applications. In: 2020 International Conference on UK-China Emerging Technologies (UCET), pp. 1–4 (2020)
5. Wang, F., Duan, Z., Wang, X., Zhou, Q., Gong, Y.: High isolation millimeter-wave wideband MIMO antenna for 5G communication. Int. J. Antennas Propag. **10**(1), 1–12 (2019)
6. Tathababu, A., Vaddinuri, R.: Compact two-port mimo antenna with high isolation using parasitic reflectors for UWB, X and Ku band applications. Progress Electromagn. Res. C **102**(8), 63–77 (2020)
7. Singh, V., Mishra, B., Singh, R.: A compact and wide band microstrip patch antenna for X-band applications. In: Second International Conference in Computing and Communication Engineering, pp. 296–300 (2015)
8. Mishra, B.: An ultra-compact triple band antenna for X/Ku/K band applications. Microw Opt. Technol. Lett. **61**(9), 1857–1862 (2019)
9. Biswas, K., Chakraborty, U.: Textile multiple input multiple output antenna for X-band and ku-band uplink-downlink applications. In: 2020 National Conference on Emerging Trends on Sustainable Technology and Engineering Applications, vol. 13, no. 9, pp. 1–4 (2020)
10. Votis, C., Tatsis, G., Kostarakis, P.: Envelope correlation parameter measurements in a MIMO antenna array configuration. Int. J. Commun. Network Syst. Sci. **03**(04), 350–354 (2010)
11. Chae, S.H., Oh, S., Park, S.O.: Analysis of mutual coupling, correlations, and TARC in WiBro MIMO array antenna. IEEE Antennas Wirel. Propag. Lett. **6**, 122–125 (2007)
12. Mohammad, S.S., Ahmed, B.N., Khan, M.U., Daniel, N.A.: A dual-element dual-band MIMO antenna system with enhanced isolation for mobile terminals. IEEE Antennas Wirel. Propag. Lett. **11**, 1006–1009 (2012)
13. Singh, A., Kumar, A., Kanaujia, B.K.: A compact low-prole high isolation MIMO antenna For X-band applications. Wirel. Pers. Commun. (2021)
14. Biswas, K., Chakraborty, U.: Textile multiple input multiple output antenna for X-band and ku-band uplink-downlink applications. In: 2020 National Conference on Emerging Trends on Sustainable Technology and Engineering Applications (NCETSTEA), vol. 19, no. 2, pp. 1–4 (2020)
15. Pahadsingh, S., Sahu, S.: A two port UWB-dual narrowband antenna for cognitive radios. Microw. Opt. Technol. Lett **58**, 1973–1978 (2016)
16. Dkiouak, A., Zakriti, A., Ouahabi, M.E.: Design of a compact dual-band MIMO antenna with high isolation for WLAN and X-band satellite by using orthogonal polarization. J. Electromagn. Waves Appl. **34**(9), 1254–1267 (2019)
17. Mohanty, A., Behera, B.R., Nasimuddin, N.: Hybridmetasurface loaded tri-port compact antenna with gain enhancement and pattern diversity. Int. J. RF Microw Comput. Aided Eng. **31**(11), 22795 (2021)
18. Gao, P., He, S., Wei, X., Xu, Z., Wang, N., Zheng, Y.: Compact printed UWB diversity slot antenna with 5.5-GHz band-notched characteristics. IEEE Antennas Wirel. Propag. Lett **13**(1), 376–379 (2014)

Performance Enhancement of SiGe-Based Junctionless Tri-Gate (JL-TG) FinFETs Using Hetero-High-K Gate Oxide Material

Himanshi Awasthi, Deepak Sigroha, Vikrant Varshney, Manish Kumar Rai, Sanjeev Rai, and Abhinav Gupta

Abstract This paper endeavors to utilize the combined advantages of Hetero-High-K gate stack material on SiGe fin with Junctionless triple gate structure and proposes three devices with different oxide materials and placement. The potential distribution, electrostatic, analog, and RF parameters of the proposed devices are analyzed and compared. Crucial FOMs like on/off current ratio, DIBL, Transconductance, Early voltage, cut-off frequency, GFP, TFP, GTFP, etc., are discussed.

Keywords Junctionless multigate · SiGe · DHGO · THGO · QHGO · Analog/RF performance

1 Introduction

The increasingly high demand for high-performance and low-power electronics has been the driving force for the miniaturization of device sizes. It has led the feature size

H. Awasthi
Department of Electronics and Communication Engineering, Kanpur Institute of Technology, Kanpur, Uttar Pradesh, India

D. Sigroha · A. Gupta (✉)
Department of Electronics Engineering, Rajkiya Engineering College Sonbhadra, Sonbhadra, Uttar Pradesh, India
e-mail: abhinavkit87@gmail.com

V. Varshney
Department of Electronics and Communication Engineering, Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh, India
e-mail: vikrant.vershney@miet.ac.in

M. K. Rai · S. Rai
Department of Electronics and Communication Engineering, M.N.N.I.T. Allahabad, Prayagraj, Uttar Pradesh, India
e-mail: mrai@mnnit.ac.in

S. Rai
e-mail: srai@mnnit.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_47

down to the nanometer range. However, due to the severe short channel effects (SCEs) observed in MOSFETs' scaling, scientists have been working on advanced techniques to control the device leakage current. The Multigate (MG) device structures [1–6] are becoming potential candidates due to excellent controllability over the channel region. A Multigate (MG) device structure with the same doping concentration in the source, drain, and channel region has been introduced by *Colinge et al.* [1] through the junctionless method. It comprises an ultrathin semiconductor film with a gate stack to control its resistance and modulate the current flowing through it. When the MOS structure length is downscaled below 100 nm, the oxide thickness also needs to be reduced. However, the SiO_2 width is diminished below 3 nm; after that, leakage current increased dramatically due to tunneling, leading to high power consumption and reduced device usability [7]. Some other researcher's works have been reviewed in the scholarly literature on the high-K gate oxide-based tri-gate FinFET performances [8–13].

Bera et al. [8] explained that replacing SiO_2 with a high-K oxide would enable gate capacitance improvement resulting in the attenuation of the energy bands due to fringing fields, resulting in reduced OFF-state current. The electrostatic characteristics are enhanced by using various high-k gate oxides, studied by *Atan et al.* [9]. *Das et al.* [10] examined the impact of different gate oxides on tri-gate FinFET with Heterojunction structure. The numerical-based comprehensive studies have been executed using high-K spacer oxide in tri-gate Junctionless (HKS TG-JL) FinFET by *Bousari et al.* [11]. The impact of high-K dielectric on analog/RF performance of junctionless tri-gate FinFET has been examined by *Tayal et al.* [12]. *Bha et al.* [13] studied 10 nm tri-gate FinFET with a high-K dielectric spacer in the underlap region.

This present article has designed three different Junctionless Tri-gate (JL-TG) FinFET structures using the gate oxide engineering technique. The various structures are identified as Double hetero gate oxide (DHGO), Triple hetero gate oxide (THGO), and Quadruple hetero gate oxide (QHGO), each with a different combination of oxide materials. These Junctionless tri-gate (JL-TG) FinFET structures have been analyzed for electrostatics and analog/RF performance characteristics, where we have chosen the SiGe material in this paper because the SiGe material has low energy bandgap and high mobility compared to the Silicon material."

2 Device Structure and Simulation Framework

The Junctionless tri-gate (JL-TG) FinFET is a device with three gates around the channel, and it provides excellent control over the channel region. The authors have proposed the SiGe-based tri-gate (JL-TG) FinFET with different structures by using oxide-engineering techniques. DHGO, THGO, and QHGO based junctionless Tri-gate (JL-TG) FinFETs are presented in Fig. 1. The SiGe material has been used as the substrate because it has higher mobility compared to silicon. The physical design parameters of SiGe-based Tri-gate (JL-TG) FinFET are summarized in Table

1. The same doping concentration dopes the source-channel-drain region, and the work function of all three device gates is maintained based on the constant threshold voltage. In the DHGO TG-JL FinFET, SiO_2 ($\kappa = 3.9$) layer of thickness 0.5 nm is used over the SiGe film because of the lattice compatibility of the materials.

Over the SiO_2 layer, a layer of HfO_2 ($\kappa = 22$) with a thickness of 1 nm is used to increase gate control. The expression can be used to calculate the effective oxide thickness (EOT):

$$EOT = t_{high-\kappa} \frac{\kappa_{high-\kappa}}{\kappa_{SiO_2}} \quad (1)$$

With SiO_2 and HfO_2, the total effective oxide thickness is approximately 0.67 nm in the THGO TG-JL FinFET, SiO_2 ($\kappa = 3.9$) layer of thickness 0.5 nm is used. Above this layer, two oxides, HfO_2 ($\kappa = 22$) and Al_2O_3 ($\kappa = 9$) have been used. The total

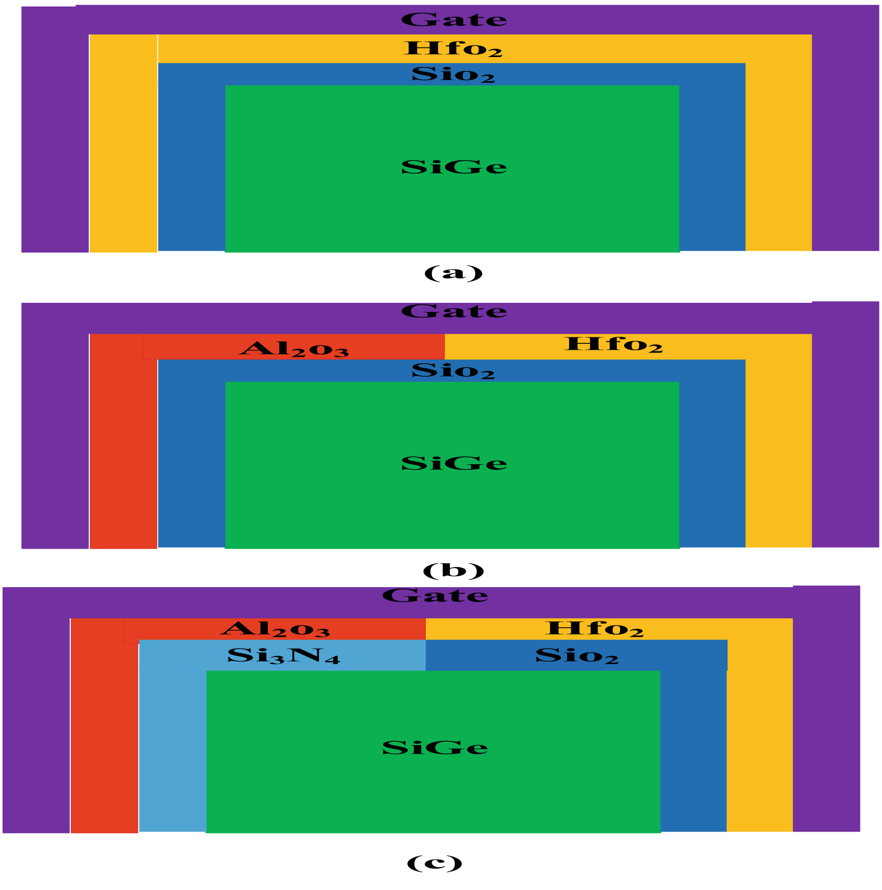


Fig. 1 Cross-section view of **a** DHGO, **b** THGO, and **c** QHGO TG-JL FinFETs

Table 1 Physical parameters of JL-TG FinFET

Parameters	DHGO	THGO	QHGO
Channel Length	14 nm	14 nm	14 nm
Fin Height	5 nm	5 nm	5 nm
Fin Width	5 nm	5 nm	5 nm
Source/Drain Length	10 nm	10 nm	10 nm
Doping Concentration	2×10^{19} cm^{-3}	2×10^{19} cm^{-3}	2×10^{19} cm^{-3}
Gate Metal Work-function	5.01 eV	5.03 eV	5.0 eV

EOT for SiO_2 and Al_2O_3 layers is 0.933 nm, and on the other hand, the SiO_2 and HfO_2 layers EOT is 0.67 nm. In the QHGO TG-JL FinFET, SiO_2 ($\kappa = 3.9$) and Si_3N_4 ($\kappa = 7.5$) layers of 0.5 nm thickness have been used and divided along XY-plane. Over these layers, HfO_2 ($\kappa = 22$) and Al_2O_3 ($\kappa = 9$) layers of 1 nm thickness have been used and divided along the XY-plane. The total effective oxide thickness (EOT) for Si_3N_4 and Al_2O_3 layers is 1.02 nm, and SiO_2 and HfO_2 layers EOT is 0.67 nm. The following numerical Models have been initiated in the simulation at room temperature (300 K) [14]. The concentration and field-dependent mobility model, Shockley–Read–Hall recombination model, band-to-band Tunneling model, and bandgap narrowing model. According to the ref [15], the quantum confinement does not occur at the 5 nm channel thickness in junctionless devices.

3 Results and Discussion

This section highlights the electrostatic characteristics and analog performance of all high-K gate stack-based Junctionless Tri-gate (JL-TG) FinFETs. The potential distribution of all devices is shown in Fig. 2 along the channel length. The plot exposes that the potential variation is found minimum at the center of the channel. However, the minimum potential of QHGO-based JL-TG FinFET is higher compared to other devices. Hence, the QHGO-based JL-TG FinFET is becoming on-state at the low power supply. Figure 3 shows the variation of drain current (I_{DS}) and Transconductance (g_m) against the V_{GS} at $V_{DS} = 1$ V of all devices. The figure shows that the drain current (I_{DS}) significantly increases owing to the high effective dielectric constant of the QHGO device compared to other devices. However, the transconductance (g_m) of QHGO is found more elevated than other devices. It shows that the optimum bias point and RF performance. All proposed devices' electrostatic characteristics are shown in Table 2, where we found excellent electrostatics characteristics in DHGO devices. Figure 4 presents the variation of early voltage and intrinsic gain against the V_{GS} at $V_{DS} = 1.0$ V. In this figure, the early voltage increases with an increase in V_{GS}, but the intrinsic gain decreases with V_{GS}. The DHGO device shows a high intrinsic gain; however, the QHGO has shown high early voltage. The early voltage and intrinsic gain are calculated as [16, 17].

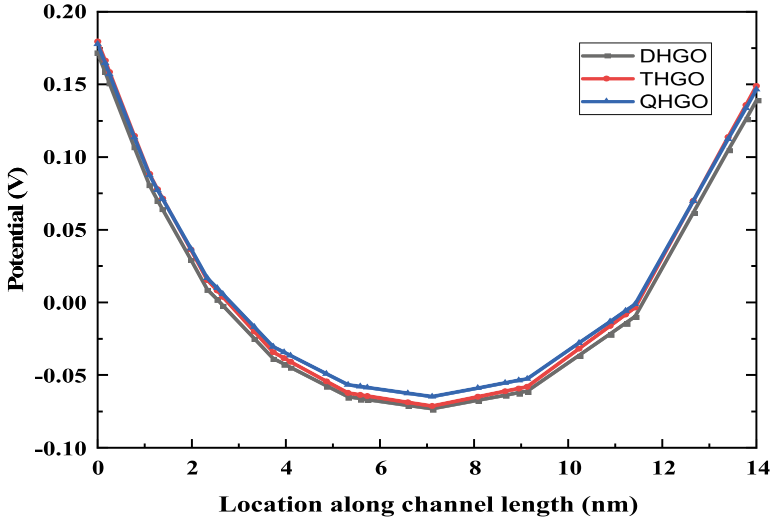


Fig. 2 Cut-lines potential distribution along channel length of all proposed JL-TG FinFETs

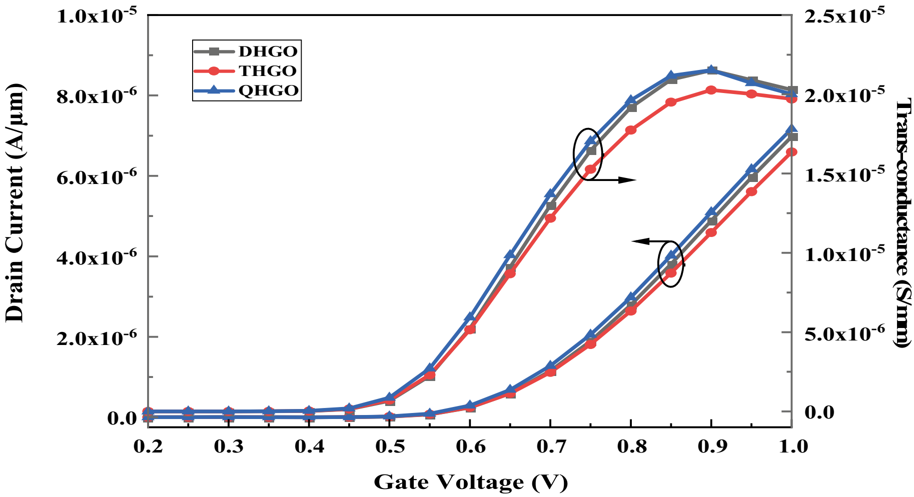


Fig. 3 Drain current (I_D) and Trans-conductance (g_m) variations against Gate Voltage (V_{GS}) for all proposed JL-TG FinFETs at $V_{DS} = 1.0$ V

The fast-emerging RF wireless communication market demands devices with excellent RF performance and low standby power requirements. RF figures of merit (FOMs) like capacitances (C_{gs} and C_{gd}), cut-off frequency (f_T), gain frequency product (GFP), trans-conductance frequency product (TFP), gain trans-conductance frequency product (GTFP) have been extracted as [16, 17] and discussed for all

Table 2 Electrostatics performance of all devices

Parameters	DHGO	THGO	QHGO
Subthreshold swing(mV/dec)	61.6573	61.9009	61.6786
I_{ON}/I_{OFF} ratio	7.635×10^9	5.821×10^9	5.460×10^9
I_{ON} (A)	6.975×10^{-6}	6.596×10^{-6}	7.169×10^{-6}
I_{OFF} (A)	9.136×10^{-16}	1.133×10^{-15}	1.313×10^{-15}
DIBL (mV/V)	20.372	22.820	17.485

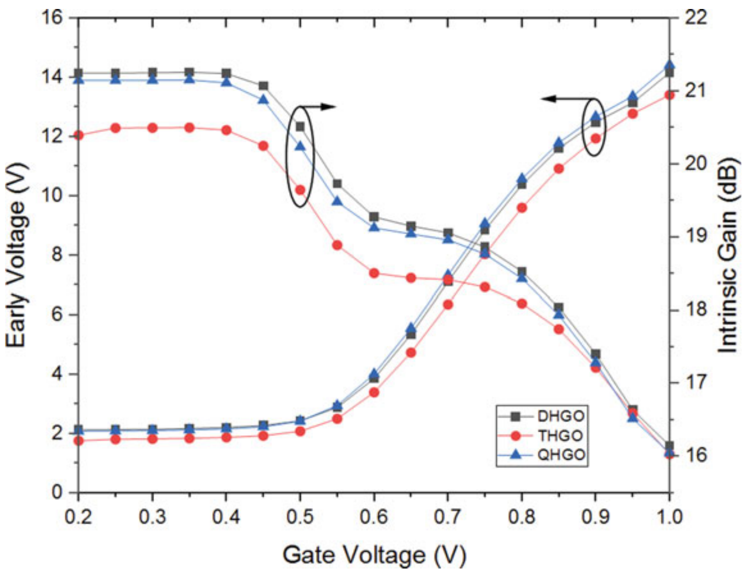


Fig. 4 Early voltage and intrinsic gain variation against gate voltage (V_{GS}) at $V_{DS} = 1.0$ V

proposed devices. Figure 5 shows gate capacitances (C_{gs} and C_{gd}) as a function of gate voltage. C_{gs} for all devices increase around the threshold voltage and then decrease shortly after reaching a peak value. C_{gd} continues to grow till the threshold voltage and then falls and rises dramatically again. Figure 6 illustrates the variation of GFP and TFP against the V_{GS}. The plot reveals that the higher GFP and TFP in the DHGO device than other devices, GFP and TFP play an essential role in Op-Amps' high-frequency applications [16].

Figure 7 shows the cut-off frequency and GTFP against the V_{GS}. We observed the lowest cut-off frequency and highest GTFP in the DHGO device in this plot compared to other proposed devices. The cut-off frequency is the most critical parameter in RF performance analysis as it reflects intrinsic delay associated with electron transit time from source to drain. The GTFP peak helps to determine the gain and speed of devices.

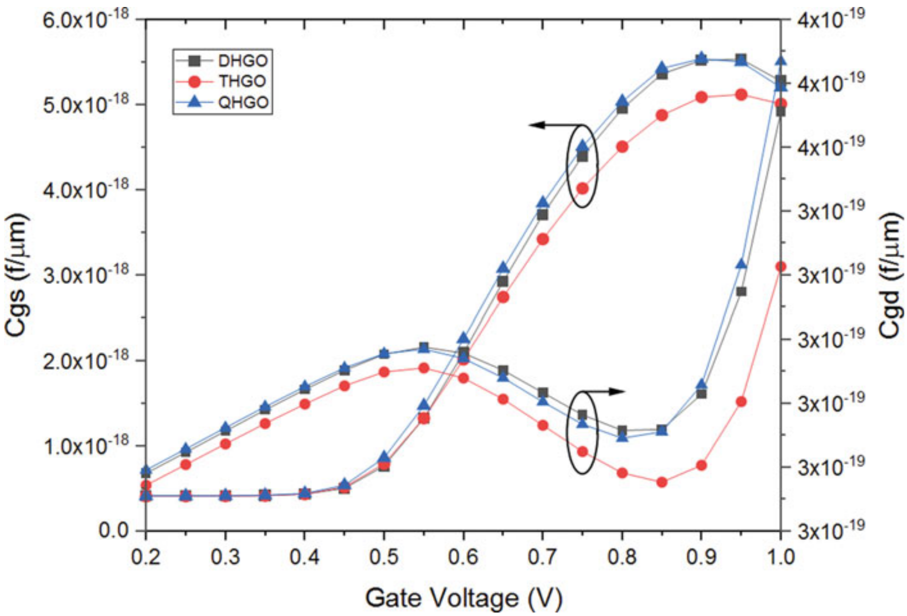


Fig. 5 The C_{gs} and C_{gd} variation against V_{GS} at $V_{DS} = 1.0$ V

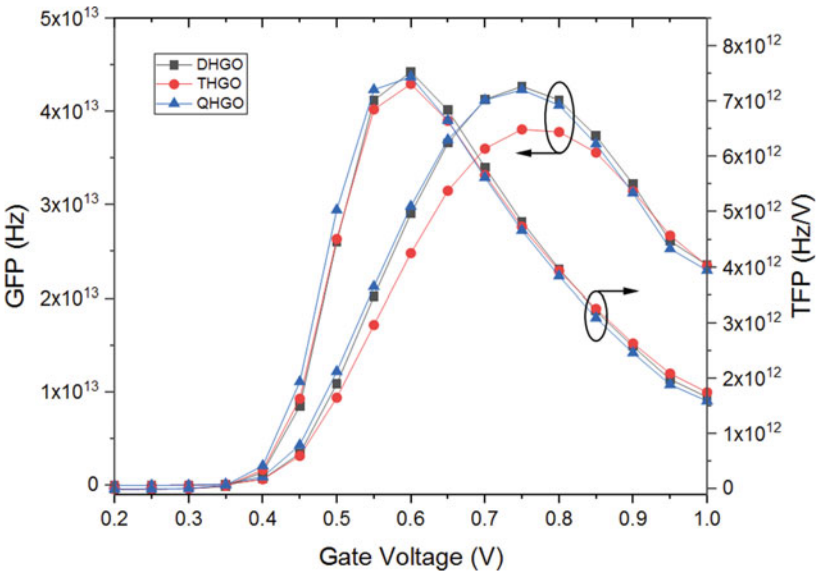


Fig. 6 Variation of GFP and TFP against V_{GS} at $V_{DS} = 1.0$ V

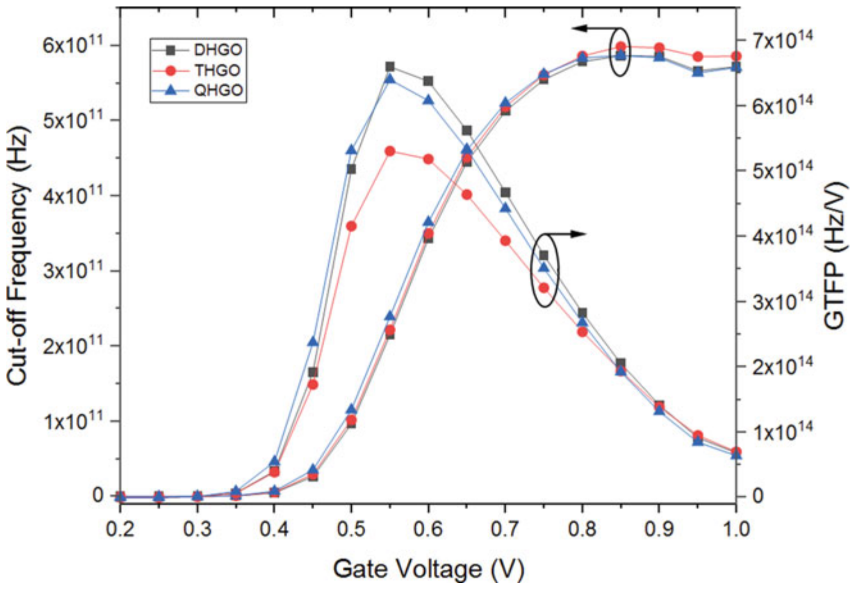


Fig. 7 Variation of cut-off frequency and V_{GS} at $V_{DS} = 1.0$ V

4 Conclusion

This paper has presented the three different Hetero- High-K gate stack engineering technique-based structures for simulation, namely DHGO, THGO, and QHGO. The simulation results expressed that the on/off current ratio is higher ~40% in the DHGO device than the QHGO device. In terms of DIBL, QHGO performed ~14% better than DHGO. However, DHGO exhibited ~30% lower leakage current than the QHGO device. The DHGO has also shown excellent analog/RF performance compared to other devices. Therefore, the DHGO device is the best candidate among the three devices for ultra-low power and analog/RF applications.

References

1. Lee, C.W., Afzalian, A., Akhavan, N.D., Yan, R., Ferain, I., Colinge, J.P.: Junctionless multigate field-effect transistor. Appl. Phys. Lett. **94**(5), 053511 (2009)
2. Lee, C.W., et al.: Performance estimation of junctionless multigate transistor. Solid State Electron. **54**(2), 97–103 (2010)
3. Lu, D.D., Dunga, M.V., Lin, C.H., Niknejad, A.M., Hu, C.: A multi-gate MOSFET compact model featuring independent-gate operation. IED meeting. https://doi.org/10.1109/IEDM.2007.4419001
4. Knoblingeer, G.: Multi-gate MOSFET design. ISQED. https://doi.org/10.1109/ISQED.2007.106

5. Singh, K.K., Yadav, N.: Study and analysis on multi-gate MOSFET design. J. Semi Device Circ. **7**(1), 1–7 (2020)
6. Colinge, J.P.: The SOI MOSFET: From Single Gate to Multigate. Springer, Cham (2008). https://doi.org/10.1007/978-0-387-71752-4_1
7. Shashank, N., Basak, R., Nahar, K.: Design and simulation of nano scale high-k MOSFETs with polysilicon and metal gate electrodes. IJOAT **1**(2), 252–262 (2010)
8. Bera, M.K., Maithi, C.K.: Electrical properties of SiO_2/TiO_2 High-K gate dielectric stack. Mater. Sci. Semicond. Process. **9**(6), 909–917 (2006)
9. Atan, N.B., Ahmad, I.B., Maijli, B.B.J.: Effects of high-k dielectrics with metal gate for electrical characteristics of 18 nm NMOS device. In: IEEE International Conference on Semiconductor Electronics (ICSE2014), pp. 56–59 (2014)
10. Das, R., Goswami, R., Baishya, S.: Tri-gate heterojunction SOI Ge-FinFETs. Superlattice Microstruct. **91**, 51–61 (2016)
11. Bousari, N.B., Mohammad, K., Haji-Nasiri, S.: Benefitting from high-k spacer engineering in ballistic triple gate junctionless FinFET – a full quantum study. Silicon **12**(9), 2221–2228 (2019). https://doi.org/10.1007/s12633-019-00318-y
12. Tayal, S., Nandi, A.: Comparative analysis of high-k gate stack based conventional and junctionless FinFET. In: INDICON (2017). https://doi.org/10.1109/INDICON.2017.8487675
13. Bha, J.K.K., Priya, P.A., Josheph, H.P., Trivadugal, D.J.: 10 nm TriGate High K underlap FinFETs: scaling effects and analog performance. Silicon **12**(9), 2111–2119 (2019). https://doi.org/10.1007/s12633-019-00299-y
14. Gupta, A., Rai, S., Kumar, N., et al.: A novel approch to investigate the impact of Hetero-high-k gate stack on SiGe junctionless gate-all-around (JL-GAA) MOSFET. Silicon **14**(2), 1005–1012 (2021). https://doi.org/10.1007/s12633-020-00860-0
15. Kumar, N., Purwar, V., Awasthi, H., Gupta, R., Singh, K., Dubey, S.: Modeling the threshold voltage of core-and-outer gates of ultra-thin nanotube junctionless-double gate-all-around (NJL-DGAA) MOSFETs. Electron. J **113**, 105104 (2021). https://doi.org/10.1016/j.mejo.2021.105104
16. Roy, N., Gupta, A., Rai, S.: Analytical surface potential modeling and simulation of junctionless double-gate MOSFET for ultra-low-power analog/RF circuits. Microelectron. J. **46**, 916–922 (2015)
17. Abhinav, Rai, S.: Reliability analysis of junctionless double-gate MOSFET for analog/RF circuits for high linearity applications. Microelectron. J. **64**, 60–68 (2017)

Bloch Surface Wave (BSW) Based Biosensor-An Alternative of SPR Technique

Akash Srivastava, Alka Verma, and Y. K. Prajapati

Abstract In plasmonic-based detection platforms the major disadvantage is involvement of materials which have ohmic losses. This is major cause of broad resonance angle and limited Figure of merit (FoM) value so it's required to develop an alternative technique of plasmonic sensor a photonic platforms such as use of one-dimensional photonic crystals (1DPhCs) may be prove a good alternative for it. In this paper, the truncated one dimensional photonic crystal (1DPC) based structure includes monolayer $MoSe_2$ as a BRE layer has been investigated. The shift of the resonance angle as a result of the change in the refractive index of the bio-sample is explored by using the transfer matrix method (TMM) in the near-infrared range. The impact of the change in the spacer layer thickness was discussed in detail and major performance parameters like sensitivity, FoM and detection accuracy (DA) has been obtained. It's shown that Bloch Surface Wave (BSW) based biosensor provides very high value of FoM and appreciable sensitivity in compare to metal layer based SPR configuration.

Keywords Bloch surface wave (BSW) · Surface plasmon wave (SPW) · 1D-PC · Sensitivity · Figure of Merit (FoM)

1 Introduction

Due to label free sensing optical biosensor are widely used, from long span of time surface plasmon resonance based sensor are used frequently. Surface plasmon polariton (SPP) [1, 2], propagates along the metal dielectric interface and significantly used to sense the change in refractive index of sample under consideration at

A. Srivastava (✉)
Department of Physics and Electronics KNIPSS Sultanpur, Sultanpur, Uttar Pradesh, India
e-mail: akash2011srivastav@gmail.com

A. Verma
Institute of Engineering and Rural Technology, Allahabad, Uttar Pradesh, India

Y. K. Prajapati
MNNIT Allahabad, Prayagraj, Uttar Pradesh, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_48

the sensor surface. No doubt it's a real-time and label free technique for biosample sensing purpose [3], even though some of the limitation with plasmonic sensors motivates the researcher for new development. Surface plasmon resonance takes place at limited number of suitable metals also called as noble metal like (Au, Ag) and user don't have the freedom to arbitrarily tuned the selection. Moreover, due to specific optical properties of metal layer (complex part in refractive index) a large absorption loss observed at optical frequencies which causes damping of SPPs on metals and hence wide and shallow resonance obtained. There are some specific technique are in existence to get a sharp resonance angle, increase sensitivity, get larger propagation length like use IR light and Long range surface plasmons (LRSPR) [4]. However the detector and light source are very expensive in IR, and in LRSPR for ideal coupling of modes (sometimes called as symmetric modes) refractive index of thin metal surrounding medium (generally some buffer solution) must be symmetric in terms of refractive index. This kind of requirement limit the user to elect category of usable buffer solutions [5] and hence for analytic detection. Bloch surface wave (BSW) [6, 7] offer several possible advantages compared to SPP like resonance angle gained at nearly any wavelength by well electing the refractive index and thickness of the layers creating the one dimensional photonic crystal (1DPhC) [8, 9]. Since refractive index of dielectrics has very lesser value of extinction coefficients (almost negligible) than metals in their transmission regime, this is the key reason that BSW resonances appear much narrower than those observed for SPP [10]. One more major advantage with BSW is that it support extreme EM field intensity caused by truncated photonic crystals which have photonic band gap. BSW were first ideally conveyed and tailed by experimental demo in 1978 [11]. Like SPR, the EM-waves or generally Bloch surface wave (BSW) travels at dielectric multilayers interface and the field decays exponentially inside the layered medium and in the liquid bio-sample under consideration [12]. Because of the periodic distribution of different refractive index materials along the stacking direction, multilayer interference within the films finally results in the elegant optical property of 1DPC called photonic band gap (PBG), which is defined as a waveband prohibiting light propagation. Bright and high-saturation color appears when PBG is located in the visible region, which makes 1DPCs a good candidate for applications such as a high-efficient reflector, chemical/physical sensor, dynamic color display, and so on. Due to the presence of the photonics band gap modes propagating on the multilayer interface decays exponentially inside the multilayers [13]. Due to total internal reflection (TIR) the internal optical modes bound close to the multilayer surface and not scattered outside. Like evanescent wave in plasmonic sensor the amplitude of bloch surface decays along the propagation direction and evanescent tail falls into the sensing medium like in SPR. Bloch wave shows resonance because of various reasons like absorption into the material, surface scattering, and leakage into the multilayer due to prism coupling. Detailed analysis about schematic diagram of periodic multilayers based BSW sensor given in the 'design consideration and mathematical analysis' section. Some of major advantage with BSW is that user can tune the local field confinement [14] more generally maximum field amplitude of the BSW mode can be altered at the surface of the dielectric multilayer (it is a favorable condition for sensing purpose [15]. In case of

BSW, surface wave excitation takes place both of the p and s polarized light and while SPP can excite only with p-polarized or TM polarized light [16, 17].

2 Design Consideration and Mathematical Analysis

As it's already described in previous sections that BSW stimulates on the superficial of an extremely reflective one-dimensional photonic band-gap intended such that the thickness of each layer in founded 1DPhC is approximately one quarter of the incident wavelength, λ. In proposed diagram multilayered 1DPhC gained by putting five duos of irregular layers of titanium dioxide (TiO_2) [18] and potassium fluoride (KF) [19] respectively on BK7 [20] substrate. By breaking the periodicity, the transmission mode inside the bandgap can be introduced and hence can travel on 1DPhC surface. In other way this will when dielectric crystals periodicity gets break. The periodic multilayer 1DPhC comprises of five duos of irregular layers of high (n_H) and low index (n_L) materials respectively. As a low index and high index material layers are potassium fluoride (KF) and titanium di oxide (TiO_2) are used in proposed diagram having the refractive indexes of 1.35 and 2.47 at the wavelength $\lambda = 1100$ nm. The thicknesses of the KF and TiO_2 layers are 240 nm each. To eliminate the periodicity of the multilayers, a 20 nm thick TiO_2 layer placed on top of schematic diagram. This upper layer is called the defect layer or buffer layer. The complete stack (periodic multilayer or 1DPhC + default layer) is called a bare multilayer platform (BML). The BML layer functionalized with $MoSe_2$ monolayer to enhance the performance for Biosample detection. $MoSe_2$ belongs to transition metal dichalcogenides (TMD) family and have a thickness dependent band gaps (<2 eV) and high electron mobility [21], greater absorption rate (higher than graphene and other TMD and heterostructures) with a wide incident angle. Figure 1 gives the schematic diagram of proposed BSW based sensing device. As a sensing medium a biological solution is taken, its refractive index can be given as $n_s = 1.330 + \Delta n$, where Δn signifies the variation of the refractive index of the biosample due to change in its physical conditions. N-layer transfer matrix method (TMM) [4] is used to calculate the reflectivity. For angular-interrogation significant performance parameters can calculate using mathematical formulas like sensitivity S (deg/RIU) can be defined [22], $S = \frac{\Delta\theta}{\Delta n}$ where $\Delta\theta$ is shift in SPR angle, and the FoM can be defined as $oM = \frac{S}{FWHM}$, where FWHM is the full width half maxima roots to sharpness of resonance angle. For ideal sensing sensitivity and FoM must be significantly high and value of FWHM must be as least as possible.

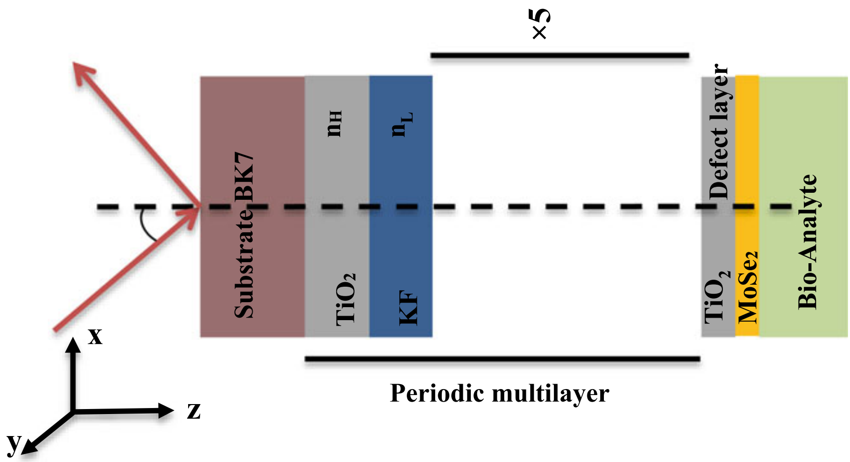


Fig. 1 Schematic diagram of proposed multilayered 1DPhC based BSW biosensor

3 Results and Discussion

Kretschmann configuration is used in this study for the coupling of incident light with the surface waves [23]. For the excitation of surface wave the wavevector of incident light must matched to wavevector of surface wave. The schematic diagram is shown in Fig. 1, glass prism play a significant role for phase-matching condition and to excite the BSWs. As an incident light source a fixed wavelength of 1100 nm is at an incident angle ‘θ’, the incident light angle must higher than the critical angle for coupling. Thickness of defect layer must be optimized in order to get reliable result hence author have varied the thickness with step size of 2 nm and calculated the value of minimum reflectance (R_{min}) and sensitivity for each case, the estimated values for each case are given in Table.1 and related graph is given in Fig. 2(a) and (b). From obtained values it’s observed that optimized thickness of spacer is 20 nm because at this value maximum value of sensitivity obtained. Although least value of reflectance obtained at 24 nm but sensitivity is not maximum.

Table 1 Optimization of spacer layer thickness by estimating the value of sensitivity and R_{min}

Spacer thickness (d_S)	Sensitivity (deg/RIU)	R_{min}
16	68	0.03269
18	62	0.0233
20	70	0.01137
22	58	0.0032
24	56	0.002
26	54	0.004
28	52	0.013
30	54	0.035

One kind of index sensor is examined with angular-interrogation method. Figure 3 it's concluded that the resonance curve obtained (maximum intensity with minimum reflectance) when incident light focused at an angle $\theta_{spr} = 62.85$ deg. The current position of angle gets shifted to 63.15 deg when sample refractive index changes from 1.33 to 1.335. Shift in resonance angle by changing in sample refractive index given in Fig. 3. From reflectance spectra its calculated that the value of Sensitivity is 70 deg/RIU while FWHM is 0.12 deg hence estimated values of FoM is 583.3 deg which is high than metallic plasmonic sensors. Detection accuracy of sensor increased with least value of FWHM, in this case DA value obtained 8.3 deg^{-1}.

As it's already discussed that the major drawback with plasmonic sensors is involvement of metal layer, Ohmic loss in metal layer degrades the performance of any application because propagation distances of SP polaritons are significantly reduced because of absorption losses. This is the major cause of increase in FWHM and reduces in FoM. In Fig. 4 author have shown a comparative study between proposed BSW biosensor and gold metal layer based plasmonic sensor having $MoSe_2$ as a BRE layer.

It's observed from Fig. 4 that BSW biosensor shows very sharp resonance curve and hence maximum value of FoM while plasmonic sensor is wide and shallow comparatively. The proposed scheme provides enhanced value of sensitivity and FoM than previously published work in this area. The proposed work shows better result than most previous published work which is most relevant to present scheme (Table 2).

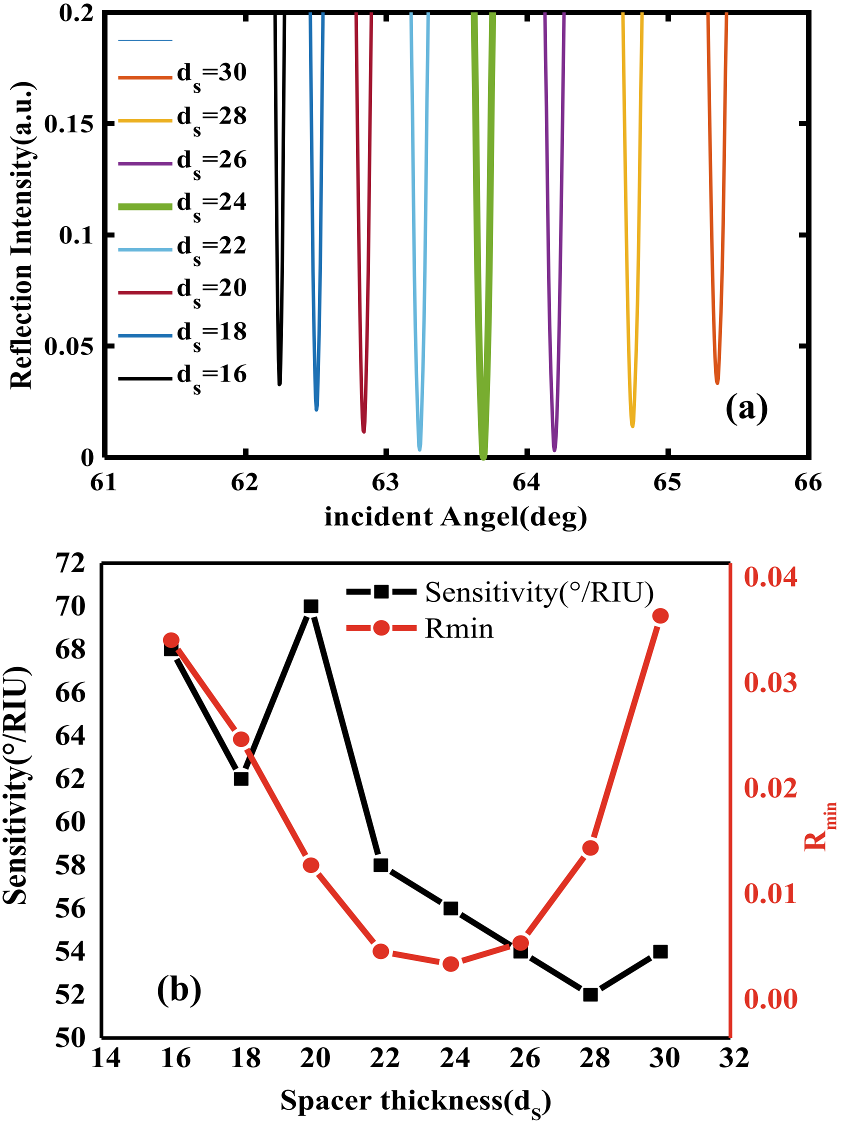


Fig. 2 **a** Minimum reflectance analysis in order to get optimized thickness of spacer (d_S) **b** Sensitivity and minimum reflectance curve analysis at different spacer thickness

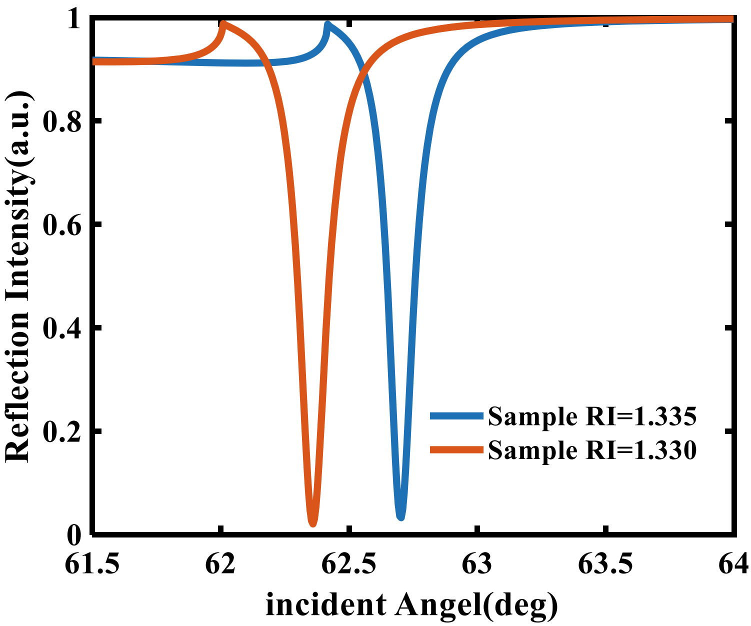


Fig. 3 Shift in reflectance curve by changing the refractive index of sample under consideration from 1.330 to 1.335

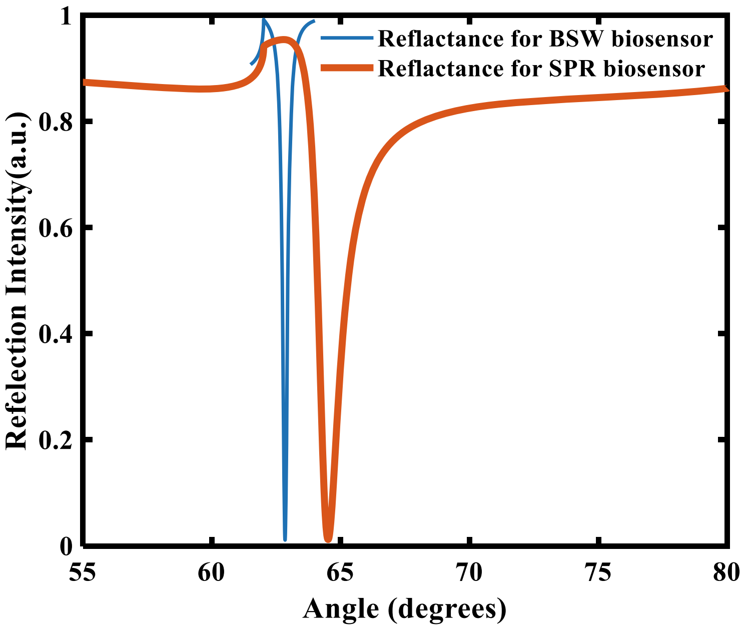


Fig. 4 Comparison study between reflectance curves obtained from BSW and SPR sensor at sample refractive index of 1.33

Table 2 Comparative results from previous works

BSW based work	Sensitivity (deg/RIU)	FoM (1/RIU)
S. Gan, H. Wang et al. 2019 [6]	65.829	467.57
Proposed scheme	70	583.3

4 Conclusion

In this paper we have examined the effect of 1DPC for generation of Bloch Surface Wave. Its shows by comparative study from plasmonic sensor that present study exhibit a sharp and smooth resonance angle which causes a high value of FoM and least value of FWHM**.** The proposed sensor is applicable to detect very minute change in biosample refractive index by relative shift in resonance angle. As a defect layer or spacer TiO_2 and as a BRE layer monolayer $MoSe_2$ applied in proposed sensor. Thickness of spacer is optimized by observing the value of sensitivity as well as R_{min}. It's expected that proposed sensing device may be alternative of plasmonic sensor which have suffer metallic ohmic loss.

References

1. Chatelier, R.C., Gengenbach, T.R., Griesser, H.J., Brighamburke, M., Oshannessy, D.J.: A general method to recondition and reuse biacore sensor chips fouled with covalently immobilized protein/peptide. Anal. Biochem. **229**(1), 112–118 (1995)
2. Homola, J.: Present and future of surface plasmon resonance biosensors. Anal. Bioanal. Chem. **377**(3), 528–539 (2003)
3. Liedberg, B., Nylander, C., Lunström, I.: Surface plasmon resonance for gas detection and biosensing. Sens. Actuators **4**, 299–304 (1983)
4. Srivastava, A., Prajapati, Y.K.: Effect of sulfosalt and polymers on performance parameter of SPR biosensor. Opt. Quantum Electron. **52**(10), 02545 (2020)
5. Ruan, B., et al.: Improving the performance of an SPR biosensor using long-range surface plasmon of Ga-doped zinc oxide. Sensors **18**(7), 2098 (2018)
6. Gan, S., Wang, H., Liang, J., Dai, X., Xiang, Y.: Ultra-sensitive refractive index sensors based on Bloch surface waves with transition metal dichalcogenides. IEEE Sens. J. **19**(19), 8675–8680 (2019)
7. Sinibaldi, A., Danz, N., et al.: Direct comparison of the performance of Bloch surface wave and surface plasmon polariton sensors. Sens. Actuators B Chem. **174**, 292–298 (2012)
8. Soboleva, I.V., Descrovi, E., Summonte, C.: Fluorescence emission enhanced by surface electromagnetic waves on one-dimensional photonic crystals. Appl. Phys. Lett. **94**(23), 231122 (2009)
9. Fenzl, C., Hirsch, T., Wolfbeis, O.S.: Photonic crystals for chemical sensing and biosensing. Angew. Chem. Int. Ed. **53**(13), 3318–3335 (2014)
10. Gryga, M., Ciprian, D., Hlubina, P.: Bloch surface wave resonance based sensors as an alternative to surface plasmon resonance sensors. Sensors **20**(18), 5119 (2020)
11. Yeh, P., Yariv, A., Cho, A.Y.: Optical surface waves in periodic layered media. Appl. Phys. Lett. **32**(2), 104–105 (1978)
12. Dubey, R., Marchena, M., Lahijani, B.V., et al.: Bloch surface waves using graphene layers: an approach toward in-plane photodetectors. Appl. Sci. **8**(3), 390 (2018)

13. Danz, N., Sinibaldi, A., Michelotti, F., Wachter, C.: Bloch surface wave based biosensing. In: 2018 Photonics North (PN) (2018)
14. Khan, M.U., Corbett, B.: Bloch surface wave structures for high sensitivity detection and compact waveguiding. Sci. Technol. Adv. Mater. **17**, 398–409 (2016)
15. Konopsky, V.N., Alieva, E.V.: Photonic crystal surface waves for optical biosensors. Anal. Chem. **79**, 4729–4735 (2007)
16. Dubey, R., Barakat, E., Herzig, H.P.: Bloch surface waves based platform for integrated optics. In: Proceedings of the IEEE Photonics Conference, Reston, 4–8 October 2015
17. Herzig, H.P., Barakat, E., Yu, L., Dubey, R.: Bloch surface waves, a 2D platform for planar optical integration. In: Proceedings of the 13th Workshop on Information Optics (WIO), Neuchatel, Switzerland, pp. 7–11 (2014)
18. Optical constants of TiO_2. https://refractiveindex.info/?shelf=main&book=TiO2&page=Devore-o
19. Li, H.H.: Refractive index of alkali halides and its wavelength and temperature derivatives. J. Phys. Chem. Ref. Data **5**, 329–528 (1976)
20. Optical constants of BK7. https://refractiveindex.info/?shelf=glass&book=BK7&page=SCHOTT
21. Wang, Q.H., Kalantar-Zadeh, K., Kis, A., Coleman, J.N., Strano, M.S.: Electronics and optoelectronics of two-dimensional transition metal dichalcogenides. Nat. Nanotechnol. **7**, 699–712 (2012)
22. Srivastava, A., Das, R., Prajapati, Y.K.: Effect of Perovskite material on performance of surface plasmon resonance biosensor. IET Optoelectron. **14**(5), 256–265 (2020)
23. Kretschmann, E.: The determination of the optical constants of metals by excitation of surface plasmons. J. Phys. **241**, 313–324 (1971)

Power Optimized 7 T SRAM Cell with Single Bit Line Using FinFET Technology

Vipin Kumar Verma, Manish Kumar Rai, and Sanjeev Rai

Abstract In the designing of an SRAM cell, the bit line plays an important role in the power optimization of the complete circuit. The critical parameters of a memory cell such as power dissipation in different modes of operation read, write, hold, and associated delays with them. Power Delay Product (PDP) and Static Noise Margin (SNM) both parameters play crucial role in the performance of SRAM memory cells. This paper explores the comparative study of Robust 8 T SRAM cell for low-voltage operation, Low power 7 T SRAM cell, and a Proposed SRAM cell with better results in delay and power usage by the memory cell with FinFET technology. Various critical parameters of SRAMs are analyzed in 18 nm finFET technology for the above-mentioned circuits and tabulated in this article. All the circuits have been executed with the power supply of 0.5 V. The digital parameters such as power, read/write/hold delay, Static Noise Margin and Power Delay Product are simulated calculated using cadence virtuoso tool and stability analysis done with the SNM. In the proposed SRAM cell power reduced by 20%, PDP and read delay reduced by one by fifth as compared to the 7 T model. To use the FinFET 18 nm technology, 18 nm PDK library (spice models) imported to the models folder of the cadence directory in root system.

Keywords Single-ended SRAM · FinFET · Memory cell design · Low power design

V. K. Verma (✉) · M. K. Rai · S. Rai
Motilal Nehru National Institute of Technology, Allahabad, Prayagraj, India
e-mail: vermavipin75@gmail.com

M. K. Rai
e-mail: mrai@mnnit.ac.in

S. Rai
e-mail: srai@mnnit.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_49

1 Introduction

According to Moore's Law, the number of transistors on an Integrated Circuit would be doubled every two years, as the complexity of the chip increases, improvement in speed and chip performance will go hand in hand. The present semiconductor industry is showing the trend itself, Today's single transistor came into the 5 nm regime which wasn't expected a couple of decades ago. The extensive scaling in Integrated Circuits pops up the problem of Short Channel Effects [1], which is a problem for the complete SOC. With technological advancements, the demand for better functionality of the device is increasing. To get enhanced and automatic performance, SOC needs more programmable memory space. Here comes the necessity to reduce the Short Channel Effects in lower node technologies, which is provided by the Multiple-gate field-effect transistors (MGFETs) [2]. Among different MGFETs, FinFETs [3] with the double gate and triple gate are the better alternatives to the conventional planner MOSFETs due to the similarity in fabrication process [4]. FinFET based designs over the past decades are providing better results as compared to the Planar MOSFET, with a significant reduction in short channel effects. As discussed above in this era of state-of-art, SRAM is a major component of on-chip caches. Today's high-performance processors contain more than half a chip dedicated to SRAMs. To improve the memory density, memory cells are built using the smallest possible technology node. The trade-off between the area, speed, and power should be done wisely.

Standard 6 transistor SRAM cell [5], as shown in Fig. 1, is less reliable at low voltage operation, due to the internal disturbance and voltage division between a pull-down and access transistors. Furthermore, there is a constraint on both transistors for stability in the write operation. To improve the read and write stability different assist methods are provided by different authors [6–8]. A few of them are virtual cell-ground, negative bit line voltage, adaptive cell-V_{DD}, etc. However, low voltage operation stability still a problem. As the chip density increased with scaling, short channel effects dominate the bit cell, selecting a suitable SRAM cell is an important step for the complete chip. The selection of SRAM cells is dependent upon three important parameters of any memory cell design, power, read/write delay, and Static Noise Margin (SNM) for the read, write, and hold operation.

2 Performance Parameters

- *Power Consumption*: There are many proposed methods to reduce power [9, 10]. One is to reduce the supply voltage. Reduction in supply voltage leads to a reduction in energy usage at various nodes. Due to this low power consumption circuits [11, 12], operating near threshold are more popular now. Nevertheless, a low voltage supply can lead to system failure. On the other hand, single-ended SRAM (single bit line) is an alternate proposed by different authors. Bit lines are the major

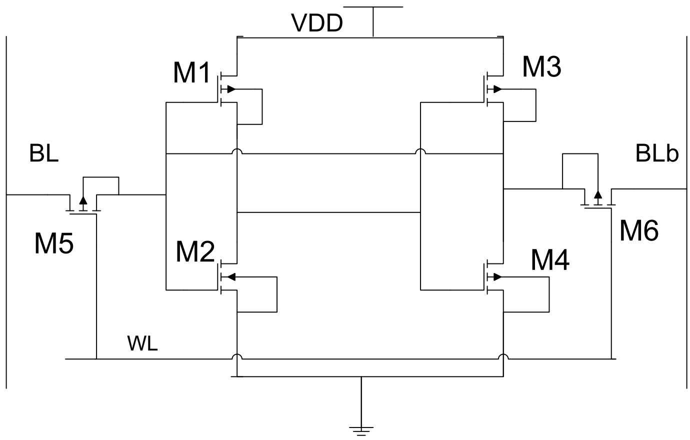


Fig. 1 Standard 6 T SRAM cell

reason behind the parasitic capacitances which can be improved with single-ended cells.

- *Delay:* This parameter decides how fast the circuit is performing the desired operations. Delay in SRAM cell constitutes read delay and write delay. Read delay is the time taken by the circuit to reproduce the node Q (the node where data bit will be stored) value at the bit line provided and Write delay is the time taken to store the value at node Q from the bit line. Improvement in the delay is directly a deciding factor for the speed, which can be reduced by the simplification of the read/write path in the bit cell.
- *Static Noise Margin (SNM):* Minimum noise voltage at each node of storage which can flip the actual value. This decides the stability of the memory cell which can be simulated by drawing inverter characteristics of the latch present in SRAM and mirroring with itself. SNM dominantly depends on the ratio of cell, voltage supply provided. After getting the mirrored graph, the size of the maximum size square that can be inserted will give the SNM value shown in Fig. 1 (Fig. 2).

The minimum value of the size of both squares in the butterfly curve will be the worst-case SNM. This can be defined for all three modes of operation read, write and hold. In single-ended models during the write operation, SNM degrades. In the 7 T model [13] feedback path is disconnected while writing operation which enhances the WSNM (write static noise margin). Similar behavior was observed in the 8 T model [14]. But in reading mode feedback is present, so SNM would not be affected. Single-ended SRAMs provide less power consumption in dynamic as well as static power. Both read and write operations in single-ended SRAM performed by a single bit line. Bit lines have capacitances of larger value, which contributes to the major portion of dynamic power consumption in SRAM. So, by using the only one-bit line a major portion of power consumption can be removed. However, Single bit line SRAM cells suffer from lower WSNM. Several approaches were proposed to

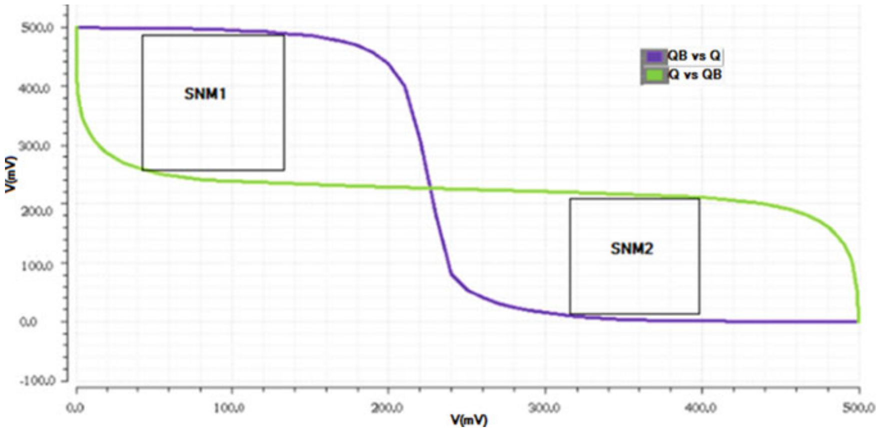


Fig. 2 SRAM butterfly curve

suppress this lowering in Single-ended models. Write assist circuit is one of the methods and using transistors of the different threshold voltage is the other, but due to width quantization in FinFET transistors of different width is not a wise option.

This paper presents a proposed circuit with a 7 T single-ended SRAM cell, which operates near-threshold region, providing reduced average power consumption and the Read timing, which provides lower PDP for the circuit. An under-shoot in bit line voltage during read-operation provides better RSNM results. The rest of paper structure as follows: The third section contains the previous single-ended SRAM cells with advantages and disadvantages. It also contains the proposed circuit explanation. In the fourth section simulation results and tabulated comparison with previous circuits are mentioned, and finally, the conclusion of all circuits discussed in the paper is provided in the fifth section.

3 Different Single-Ended Sram Cells

3.1 Low-Voltage Robust 8 T SRAM Cell

In this robust 8 T SRAM model, shown in Fig. 3 design is based on MOSFET technology. In this paper, the circuit is designed with 18 nm FinFET technology. This robust model is using 8 transistors for a single SRAM cell providing the disconnected feedback path by M5 during the write operation and feedback is enabled by the M5 during read-operation.

In the above 8 T SRAM cell, data saved by nonsymmetrical cross-coupled pair of inverter. Different modes of operation are controlled by the signals mentioned in Table 1, WWL/WBL (write control/bit line), RWL/RBL (read control/bit line),

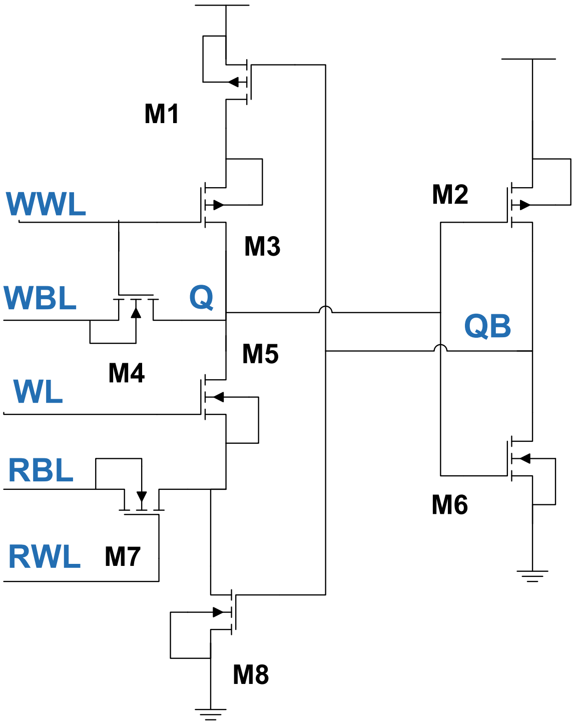


Fig. 3 Schematic of 8 T SRAM cell [14]

WL (feedback control). For the write-1, WWL is activated and WL is connected with logic low. Further, M5 and M3 are off and the value associated with WBL is transferred to node Q through M4 (Fig. 3). In write mode, the disturbance is very low due to the feedback cut of the inverter loop. But the problem here is that node Q is not able to achieve V_{DD} because of the weak pass of "1" in NMOS. After WWL deactivated, the cell enters into a hold state and the node Q charged to V_{DD} by M1 and M3. Similarly, the write "0" operation can be performed by giving a low value to WWL. In Read-operation RBL needs to be charged with V_{DD}. Turn RWL signal high and WL signal low, data will be transferred to bit line by M7 and M8. Here the important part is read path gets disconnected from the storage node, which will provide a stable read operation. But the cell is prone to noise because Q is floating when read "0" is performed. However, this circuit suppresses the DIBL effect by the use of cascaded M1 and M3 (Fig. 4).

3.2 Low Power Single-Ended 7 T SRAM Cell

Table 1 Operation modes of 8 T SRAM cell

Controls	Read	Write-0	Write-1	Hold
RWL/RBL	1/1	0/1	0/1	0/1
WWL/WBL	0/0	1/0	1/1	0/0
WL	0	0	0	1

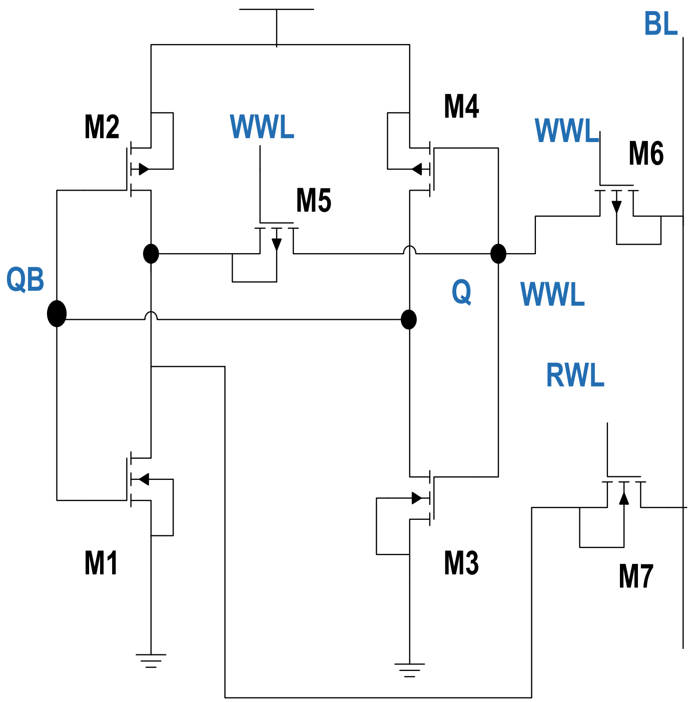


Fig. 4 Low power single-ended 7 T SRAM cell [13]

This model come up with 7 transistor SRAM model which provides much better results as compared to 8 T SRAM model. Firstly it is using less chip area because of the saving one transistor area in each cell as compared to the 8 transistors. It is also a design that uses one bit-line for both the operations read and write. This cell contains two cross-couple inverters. Moreover, a p-type transistor (M5) is of low V_{th}. Input to the inverter (M3, M5) controls the feedback path. For the write operation WWL signal is provided with V_{dd}, Read operation is controlled with the RWL signal, for the hold operation both WWL and RWL are grounded (Table 2).

Table 2 Operation modes of 7 T SRAM cell

Control	Read	Write-0	Write-1	Hold
WWL	0	1	1	0
RWL	1	0	0	0
RBL/WBL	–	0	1	–

Read Operation: To perform read operation this cell does not needs pre-charging of bit line. Simply RWL needs to be asserted, data at the node Q will be drawn to the bit line directly through M7. However, for the passing of 1, there is a limitation because the NMOS cannot pass strong '1'. When BL = '1' and the input node of the left inverter is '0', there is a small increase in that node voltage due to the voltage division between M7 and M1. In addition for the case, BL = '0' and the input node of the left inverter is '1', M2 will be in a race with M7.

Write operation: For the write, the operation provides data that needs to be store at node Q, and WWL activated, feedback path will cut by M5 and both inverters will be in series. Now only V_{dd} is connected to node Q, however, Q will get V_{dd}-V_{th} in this cycle when WWL is '1', In the next cycle when WWL is '0' feedback is present and the voltage of node Q gets V_{dd}. To write '0', M6 can easily pass strong '0'. Node Q will store '0' but as WWL asserted '0' in the next cycle, there will be an under shoot in node Q voltage, which is better for noise tolerance.

3.3 Proposed 7 T SRAM Cell with Better Power/delay

This cell provides optimized power usage using the single ended technique [15], improved read delay, and a stable bit line with reduced PDP. The majority of the power usage is dependent on the number of transistors ON in a particular operation. The number of transistors ON/OFF for a particular operation is similar to the 7 T [13] model discussed earlier but the power loss at switching nodes has been optimized in this circuit. During the read operation signal is transferred from node Q to the RBL through M3 and M6 which reduces the race around condition in the 7 T model. As compared to the 8 T model one of the important advantages is the number of transistors used is one less which is insignificant for a single cell, but it will make a difference when complete memory is designed with this SRAM cell. This model provides a reduction in reading delay, which leads to lower PDP. For the signal stability, during write-operation, two different signals are involved in the 8 T model, and loop cutting is involved in the 7 T model. This proposed model is also using the loop cutting technique but here the advantage is loop cutting does not affect the node Q connection from V_{dd}. During the write, '1'operation M1 is also ON and providing the path for the node Q to attain strong '1', which was a problem in previously discussed 7 and 8 T models. During writing '0' M5 is connected with bit line will provide the path for the strong '0'. So this will improve the writing operation for the proposed circuit with very little disturbance at the node Q.

Table 3 Operation modes for 7 T proposed circuit

	Read	Write-0	Write-1	Hold
WWL	0	1	1	0
RWL	1	0	0	0
WL	1	0	0	1
RBL/WBL	–	0	1	–

The above model is working according to Table 3. For the write operation, RWL should be grounded, bit line is provided with the value that needs to be stored at node Q. After asserting the WWL signal M5 will pass the value directly to the node Q, at the same time node Qb will get the opposite value of Q through M4. There is no wait for the node Q/Qb for the next cycle to stabilize the signal value. While read operation WWL signal should be grounded. RWL and WL are asserted with V_{dd}, feedback path again comes in the picture now the value stored at node Q will be read out to the bitline through M3 and M6. During read-operation in the 7 T [10] circuit, there was a race around condition which is resolved in this circuit (Figs. 6 and 7).

Write-0: Signal WL is connected with 0 V, which will cut off the feedback of both inverters because M3 is OFF in the circuit Fig. 5. RWL signal is also asserted with 0 V. Bit line RBL/WBL is provided with 0 V and WWL signal is asserted with V_{dd}. Now M5 will pass the 0 V to the input of M2 and Qb will be set high. Since the Qb is high M1 will be Off hence M5 will directly pass the 0 V to node Q, As M3 and M1 are off node Q is isolated hence the chances of fluctuations are reduced.

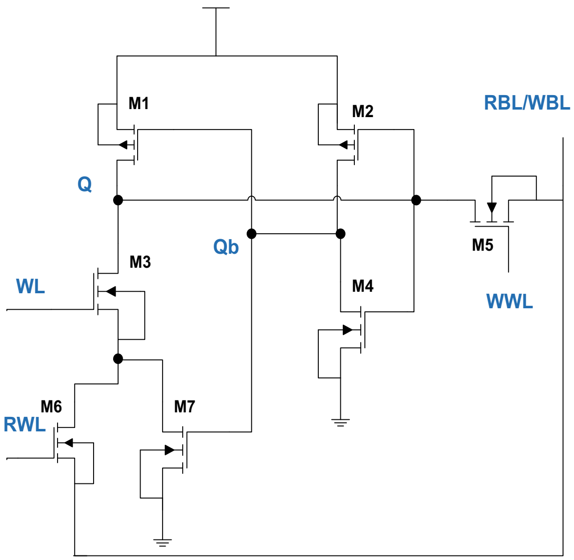


Fig. 5 Proposed 7 T SRAM with better power/delay

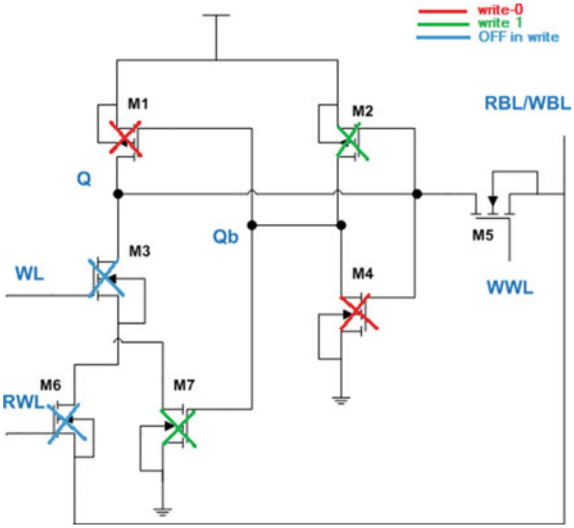


Fig. 6 Proposed 7 T SRAM with write-0/1, 3 transistors ON

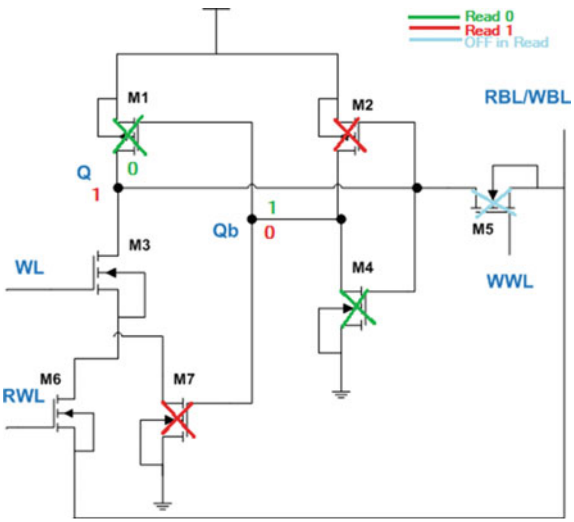


Fig. 7 Proposed 7 T SRAM with read-operation '0/1', 4transistors ON

Write-1: Similar to write-0, WL and RWL are connected to 0 V, and data 1(V_{dd}) is provided at RBL/WBL. Here M4 is ON and Qb will be connected to the ground through M4. Since Qb is 0 V M1 is ON and node Q is directly connected with V_{dd} through M1. Node Q is connected to V_{dd} through M5 and M1 so the signal stability at node Q is high.

Table 4 FinFET model parameters of used technology

Parameters	n-type	p-type
Model name	Cds_ff_mpt_n1lvt	Cds_ff_mpt_p1lvt
Fin pitch	48 nm	48 nm
Gate length	18 nm	18 nm
Gate connections	2	2
Voltage	0,0.5 V	0,0.5 V

Table 5 Simulation result comparison of all models

SRAM cells	Write delay (ps)	Read delay (ps)	Write power (μW)	Read power (μW)	Average power (μW)	PDP (fJ)
8 T[2]	33.54	53.36	0.526	0.941	0.671	0.035
7 T[1]	33.70	104.40	0.353	0.732	0.543	0.056
Proposed cell	34.56	22.42	0.328	0.540	0.434	0.015

Read Operation: Firstly WWL should be connected to 0 V, WL and RWL should be asserted with V_{dd} this will ON the M3 and M6 respectively. Since feedback path is present, if Q node is storing the value '1' (high) then Qb will be storing '0', 4 transistors are ON M1, M3, M4, and M6, which is similar to the 7 T model [13], due to less leakage power dissipation reduced in this circuit.

4 Simulation and Result Analysis

In this section above explained critical parameters of SRAM circuits are simulated, evaluated, and compared using the Cadence Virtuoso tool using 18 nm FinFET technology. Parameters of the FinFET technology file used in the above circuits are listed in Table 4. The results include the performance analysis of the read, write and hold operations, worst-case delays in the Read and Write mode. In power analysis, the average power and power delay product given in Table 5. Stability analysis is done for the three modes of operation with temperature (–50–50 °C).

In Fig. 8 undershoot on bit line near the threshold value is contributing to the better stability and noise margin in read-operation. Figure 9 shows the signal strength at various supply voltage varying from 0–2 V, it was observed voltage range 0.45–0.65 V providing the strong signal strength. In above Table 4, it can be observed that read delay, average power and, the PDP improved significantly. Figure 10 shows the HSNM at room temperature, Fig. 11 and Fig. 13 shows that WSNM and RSNM respectively. Figure 12 and Fig. 14 show the variation in WSNM and RSNM respectively concerning temperature variation from –50 to 50 °C. The below table represents the SNM values of the various operation it is observed that the RSNM and HSNM showing an improvement of around 10 mV, However WSNM is around 13 mV less than the 7 T [13] model. Overall results have improved considering the trade-off (Table 6).

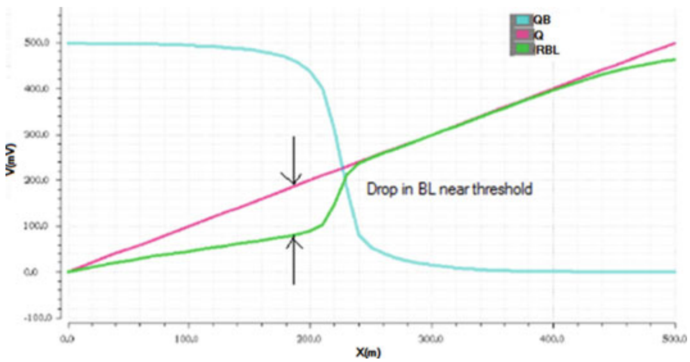


Fig. 8 Read operation DC response of the proposed circuit

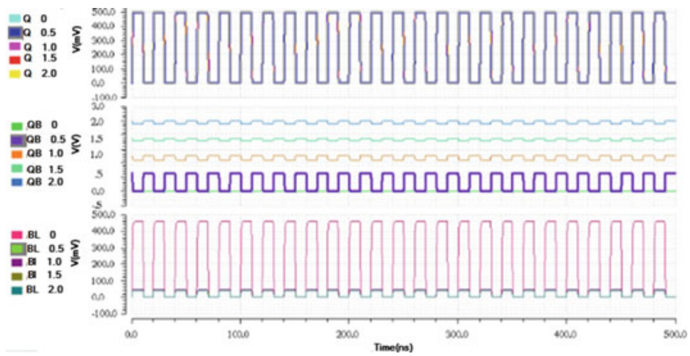


Fig. 9 Voltage variation in read-operation of the proposed circuit

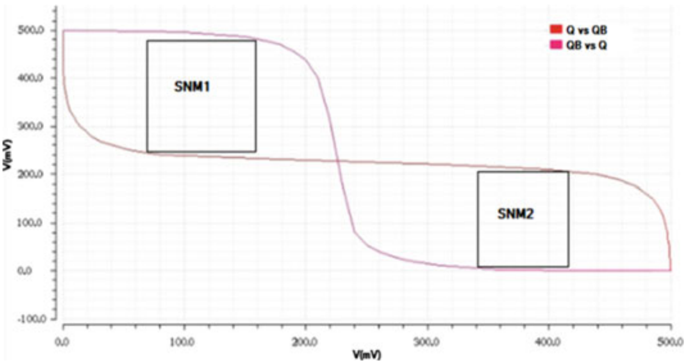


Fig. 10 Hold butterfly curve of the proposed circuit

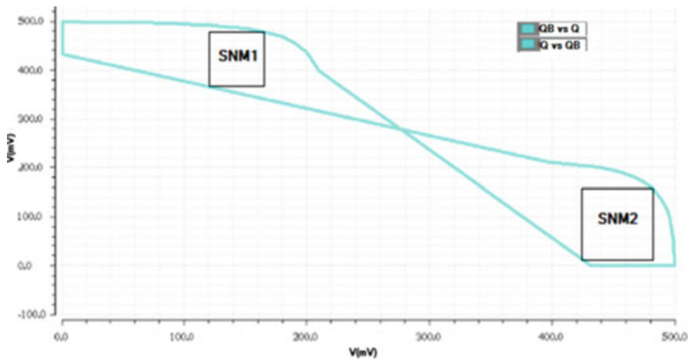


Fig. 11 Write butterfly curve of the proposed circuit

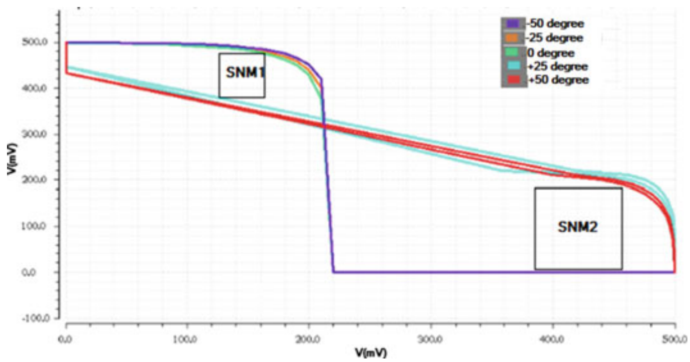


Fig. 12 Write butterfly curve with temperature variation

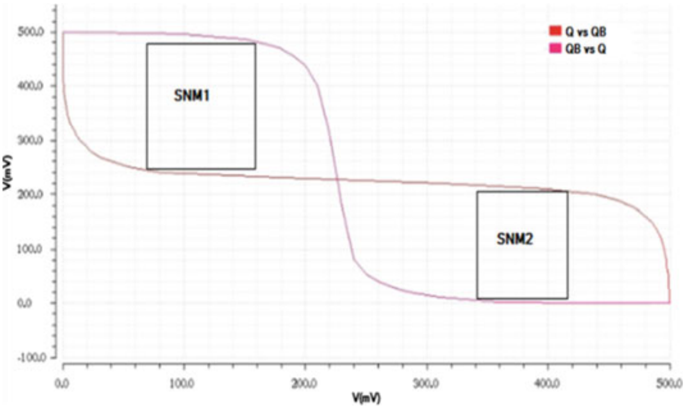


Fig. 13 Read butterfly curve of the proposed circuit

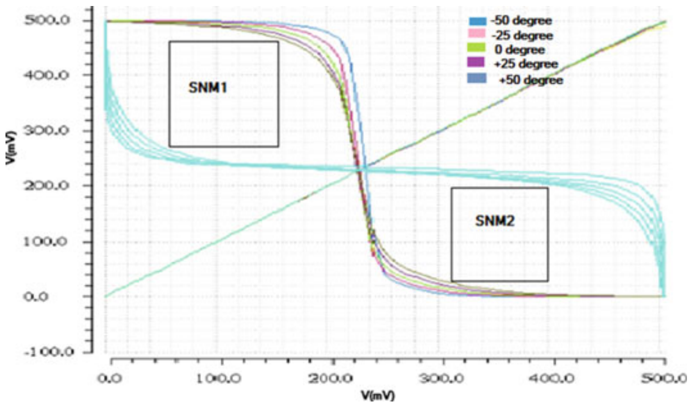


Fig. 14 Read butterfly curve with temperature variation

Table 6 SNM simulation result comparison of all models

SRAM cell	HSNM	RSNM	WSNM
8 T[14]	127	134	52
7 T[13]	210	210	123
Proposed cell	222	220	110

5 Conclusion

Various SRAM cells have been discussed. All the read, write, and hold operations have been analyzed and tabulated. Simulation results are shown for the proposed circuit. Single-ended, robust 8 T SRAM cell consists of 8 transistors, Low power single-ended SRAM cell has only 7 transistors, as compared to 7 T model proposed cell is providing read delay reduced by 5x, average power reduced by half times and PDP reduced by 4x. The proposed cell also provides the solution to the weak '1' signal in a write operation by the dual connection of PMOS and NMOS. The simulation result shows that the proposed circuit is providing better results for timing analysis, the only problem observed in this circuit is reduced WSNM. It can be improved by the use of the write assist circuit technique. However, with the use of the assist technique, there is an increase in the number of transistors.

References

1. Gupta, S.K., Roy, K.: Low power robust FinFET-based SRAM design in scaled technologies. In: Reis, R., Cao, Y., Wirth, G. (eds.) Circuit Design for Reliability, pp. 223–253. Springer, New York, NY (2015). https://doi.org/10.1007/978-1-4614-4078-9_11
2. Bhattacharya, D., Jha, N.K.: FinFETsL: from devices to architectures. Adv. Electron. **2014**, 21, Article ID 365689
3. Salahuddin, S.M., Chan, M.: Eight-FinFET fully differential SRAM cell with enhanced read and write voltage margins. IEEE Trans. Electron. Devices **62.6**, 2014–2021 (2015)
4. Chang, J.B., Guilorn, M., Solomon, P.M., et al.: Scaling of SOI FinFETs down to fin width of 4 nm for the 10 nm technology node. In: Proceedings of the Symposium on VLSI Technology, Systems and Applications (VLSIT 2011), pp. 12–13, June 2011
5. Mann, R,W., et al.: Impact of circuit assist methods on margin and performance in 6T SRAM. Solid-State Electron. **54**(11) (2010)
6. Tu, M., Lin, J., Tsai, M., Jou, S., Chuang, C.: Single-ended subthreshold SRAM with asymmetrical write/read-assist. IEEE Trans. Circuits Syst. I Regul. Pap. **57**(12), 3039–3047 (2010). https://doi.org/10.1109/TCSI.2010.2071690
7. Nayak, D., Acharya, D.P., Rout, P.K., Nanda, U.: A novel charge recycle read-write assist technique for energy-efficient and fast 20 nm 8T-SRAM array. Solid-State Electron. **148** (2018)
8. Tu, M.-H., et al.: Single-ended subthreshold SRAM with symmetrical write/read-assist. IEEE Trans. Circuits Syst. I Regul. Pap. **57**(12), 3039–3047 (2010)
9. Hajare, R., Lakhminarayana, C.: Design and evaluation of FinFET based SRAM cells at 22 and 14 nm node technologies. Int. J. Comput. Appl. **171**, 26–31 (2017). https://doi.org/10.5120/ijca2017915005
10. Moaiyeri, M.H., Akbari, H., Moghaddam, M.: An ultra-low-power and robust ternary static random access memory cell based on carbon nanotube FETs. J. Nanoelectron. Optoelectron. **13**(4), 617–627 (2018)
11. Moghaddam, M., et al.: An ultra-low-power 9T SRAM cell based on threshold voltage techniques. Circuits Syst. Signal Process. **35**(5), 1437–1455 (2016)
12. Singh, A., Khosla, M., Raj, B.: Design and analysis of electrostatic doped Schottky barrier CNTFET based low power SRAM. AEU Int. J. Electron. Commun. **80**, 67–72 (2017)
13. Ensan, S.S., Moaiyeri, M.H., Moghaddam, M., Hessabi, S.: A low-power single-ended SRAM in FinFET technology. AEU Int. J. Electron. Commun. **99** (2019)
14. Wen, L., Li, Z., Li, Y.: Single-ended, robust 8T SRAM cell for low-voltage operation. Microelectron. J. **44**(8), 718–728 (2013)
15. Ojha, S.K., Singh, O.P., Mishra, G.R., Vaya, P.R.: Analysis and design of single ended SRAM cell for low-power operation. In: 2016 11th International Conference on Industrial and Information Systems (ICIIS), pp. 178–181 (2016). https://doi.org/10.1109/ICIINFS.2016.8262929

Comparative Analysis of Negative Capacitance Double Gate Junctionless Transistor and Tri-Gate FinFET for Analog Performance

Deepak Bairwa, Manish Kumar Rai, and Sanjeev Rai

Abstract In this paper, the comparative analysis of two device structures has been proposed for analog and RF performance i.e. double gate junctionless transistor DGJLT and TG-FinFET with the presence of the negative capacitance phenomena as all the results are getting at the constant thickness of 1 nm of HfO_2 material, Silvaco atlas tool has been used. In this paper, the simulated parameters are the gate to source capacitance C_{gs}, the gate to drain capacitance C_{gd}, trans-conductance g_m, output conductance g_{ds}, Early voltage V_{EA}, cut off frequency f_T, intrinsic gain A_V, trans-conductance generation factor TGF, trans-conductance frequency product TFP, gain frequency product GFP. Simulated analog and RF performance parameters are showing better results for both structures when having negative capacitance material.

Keywords Double gate junction less transistor · Triple gate FinFET · Negative capacitance · Analog · RF FOMs

1 Introduction

For making the device smaller scaling is the major phenomenon but when scaling down the device under nanometer region power supply is not scale down w.r.t. scaling and also at nanometer region serval short channel effect come into the picture. Such as poor sub-threshold swing, leakage current, increased DIBL, and maintain V_{th} as low are major challenges [1–3]. In conventional FETs, the fundamental limit of sub-threshold swing is 60 mv/dec. This limit allows minimum operating voltage and switching energy for the operation of the bulk MOSFET. So we need new technology to reduce the operating supply voltage for low-power devices. For that, negative capacitance-based FET (NCFET) is being used, having the capability to work on a very low operating point with low power dissipation. NCFET is designed using the ferroelectric material in the gate stack and gives a voltage amplification due to the negative capacitance effect [4, 5]. In this paper, two structures have been designed. In

D. Bairwa (✉) · M. K. Rai · S. Rai
Department of Electronics and Communication Engineering, MNNIT Allahabad, Prayagraj, Uttar Pradesh, India
e-mail: bairwadeepak6207@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_50

both device structures, a very thin layer of FE Material (HfO_2) is used to generate a negative capacitance effect in both structures. A lot of research has been working on different devices and used to come up with different device structures. An alternative way to describe more control on the channel region has been proposed by many researchers [6] such as double-gate junction-less transistors, FinFET, gate all around FET. As the channel is normal controlled by the gate. Multigate device structure having more control on the channel region, as they control the channel from multiple sides. DG and FinFET provide more control on the channel region and junction-less transistors improve the packing density, more control on the channel region to operate the device in the negative capacitance region. In comparison to conventional MOSFET; an achievable sub-threshold swing is 42 mv/dec in DGJLT and 37 mv/dec in FinFET [7, 8]. Junction-less transistor or vertical FET with no source and drain junction is a good performance device for making high packing density devices [9]. The regular device having a source and drain junction and to turn off the device reversed bias the junction mechanism is used. But DGJLT turns off by full depletion of the channel region mechanism and this depletion mechanism is happen by work function difference between gate material and doped silicon. Bulk current conduction mechanism has into the DGJLT, the transistor is in off condition under the $V_{gs} < V_{th}$ the channel has fully depleted known sub-threshold state, as gate voltage starts to increases form $V_{gs} < V_{th}$ to $V_{gs} > V_{th}$ device channel region start to partially depleted and current start to flows through the center of the channel as further V_{gs} increases and when $V_{gs} > V_{FB}$, DGJLT is now in on condition, considerable current flows through the entire channel region [10]. Doping concentration into DGJLT for source, drain and channel are the same. Due to the same doping concentration, source-gate and drain gate junctions have never been formed [11–13]. The leakage current into DGJLT is very less due to this bulk current conduction mechanism. Due to negative capacitance total capacitance is increased, the resultant charge is increased into the channel region, I_{on} current is increased into device due to that trans-conductance increases, overall analog and RF performance of both devices getting better to compare to regular FETs. At present time FinFET technology plays a very important role to be significantly used in the electronic domain but for analog perspective, very little work had been carried out so in the future if we need a SOC system where analog and digital circuit fabrication on the same wafer that time it needed to FinFET analog performance-enhancing circuit [14]. A lot of research has been done on FinFET. FinFET technology is needed to address the short channel effect and leakage issue into the device at the nanometer region [16]. Fabrication of the FinFET is comparably the same as the CMOS fabrication process also FinFET is having ultra-thin body device structures beneficial for high packing density, as well as reducing leakage because due to of the high electric field. A negative capacitance material layer constant thickness of 1 nm has been used during designing of DGJLT and FinFET due to this negative capacitance phenomenon, both the device having a steeper sub-threshold slope and very low leakage current also reducing the short channel effect and better result for analog and RF application.

2 Device Structure and Simulation Setup

To fabricate a large number of transistors on a single chip (high packing density) required and to reduce the fabrication complexity junction-less transistor concept comes into the picture. In junction-less transistors, the absence of source and drain junction provides better scaling and reduced cost. In bulk FET, The leakage current is very large when $V_{GS} < V_{Th}$ (the device is off). As we reduced the dimension of the device, up to a channel length of 20 nm undesirable short channel effect comes into the picture, and steep doping is necessary for the source and drain junction in conventional FET. To make a low power-consuming structure, ferroelectric material is used in gate stacks in FETs. According to Landau's theory, ferroelectric material shows the negative capacitance behavior, in charge voltage characteristics of ferroelectric materials. Multigate device structure that replaces the bulk FET provides lower short channel effect and very low leakage current compared to bulk MOSFET. The rise in the temperature increases the power dissipation in the device. Negative capacitance devices are provided lower sub-threshold swing, low power dissipation, better performance, and enhance electrical characteristics of devices.

The device structure shown in Figs. 1 and 2 is DGJLT and FinFET. The junction-less transistor has an n layer in the channel and FinFET has an n^+ layer in the channel. DGJLT has the channel doping of 10^{19} atoms per cm^3, the work function is 4.2 and 4.1 for DGJLT and FinFET, respectively. Here we are using Si-doped HfO_2 ferroelectric material to generate negative capacitance with a thickness of 1 nm. All the dimension and structure-related parameters are represented in Table 1. SRH (Shockley- Read-Hall) model is used to calculate conduction band and valance band energy, and interface charge density is 3×10^{10} C/cm^2. The newton trapped method and Gummel method is used to extract the parameters from the silvaco tool.

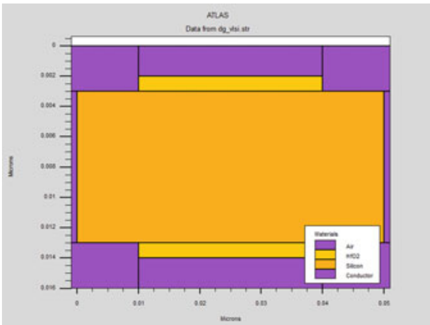


Fig. 1 Geometric view of DGJLT structure design in silvaco atlas tool

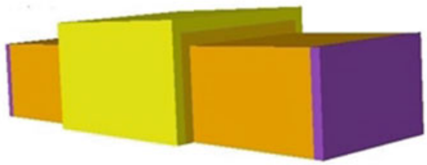


Fig. 2 Geometric view of TG-FinFet structure design in silvaco atlas tool

Table 1 Device dimensions of DGJLT AND FinFET

Parameter	DGJLT	FINFET
Doping Concentration (atoms per cm^3)	10^{19}	10^{19}
WF for the gate material	4.2	4.1
Channel Length (nm)	20	20 (Fin Width = 6 nm Fin Height = 16 nm)
Surface charge density (coulomb per cm^2)	3×10^{10}	3×10^{10}

2.1 Newton Methods

The newton methods solve the total system of unknowns together. It is useful when the system of equations is powerfully joined and has quadratic conjunction. It may devote extra time resolving for quantizes that are essential content or weakly joined. This is by default method and used almost in all calculations, this method is preferable. The following cases need the Newton methods to be set for isothermal drift-diffusion simulations.

- Currents boundary circumstances
- Exterior essentials
- Ac Analysis

2.2 Gummel and Block Methods

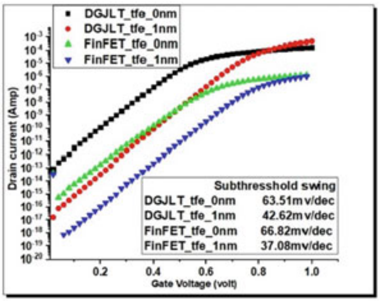
The Gummel method resolves for each pending in turn, possession the other variables constant and retelling the process until a stable solution is conquered. A good result is not obtained by this method. Normally, it is useful where the system of calculation is weakly engaged but has only linear conjunction. It cannot be used with lumped elements or currents boundaries conditions. The block method resolves some equations that are based on the previous calculation. It is useful when lattice heating and energy well-adjusted equation are encompassed.

3 Simulation and Discussion

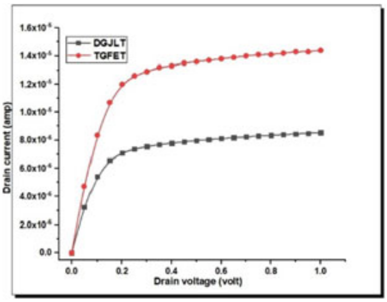
Figure 3a shows the I_d-V_{gs} characteristic with the value of sub-threshold voltage swing for two different structures of DGJLT and FinFET with having a thickness of ferro layer is 0 to 1 nm for both structures. V_{gs} Voltage variation of 0 to 1 V at constant V_{ds} = 1 V. Figure 3b shows the I_d-V_{ds} characteristic for two different structures of DGJLT and FinFET with having a constant thickness of ferro layer is 1 nm for both structures. V_{ds} voltage variation of 0 to 1 V at constant Vgs = 1 V for both structures (Table 2).

3.1 Analog Performance

The Integrated circuits that are used in analog and RF design systems are considered analog integrated circuits which are the most difficult circuit to stabilize. In this paper, analog performance parameters i.e. trans-conductance g_m, Output conductance g_{ds}, early voltage V_{EA}, intrinsic gain A_V, trans-conductance generation factor TGF for both the devices i.e. DGJLT and FinFET are plotted.



a variation of I_{ds} vs V_{gs} as a function of gate voltage from V_{gs}= 0 to 1 volt with FE thickness at V_{ds} = 1 volt.



b variation of I_{ds} vs V_{ds} as function of drain voltage from V_{ds}= 0 to 1 volt at V_{gs} = 1 volt

Fig. 3 **A** variation of I_{ds} vs V_{gs} as a function of gate voltage from $V_{gs} = 0$ to 1 V with FE thickness at $V_{ds} = 1$ V **b** Variation of I_{ds} vs V_{ds} as function of drain voltage from $V_{ds} = 0$ to 1 V at $V_{gs} = 1$ V

Table 2 Extracted parameters of DGJLT and FinFET

Parameter	DGJLT		FinFET	
Sub-threshold Swing (mv/dec)	$t_{fe} = 0$ nm	$t_{fe} = 1$ nm	$t_{fe} = 0$ nm	$t_{fe} = 1$ nm
	63.51	42.62	66.82	37.08
I_{off} current (amp)	10^{-13}	10^{-17}	10^{-15}	10^{-18}

Fig. 4 Variation of trans-conductance as a function of gate voltage from $V_{gs} = 0$ to 1 V at $V_{ds} = 1$ V

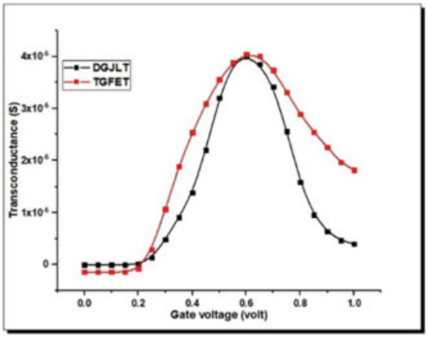


Figure 4 variations of the trans-conductance g_m as a function of the gate voltage in respective of constant $V_{ds} = 1$ V, constant threshold voltage, and doping concentration for both devices DGJLT and FinFET. From the figure, it can be observed that DGJLT having the low trans-conductance g_m value compare to the FinFET device. From the figure, we can also conclude the trans-conductance of both the devices increases as gate voltages increases and after getting the certain value of g_m, it starts to decreases. Trans-conductance g_m of the FET is used to calculate the gain of the Amplifier for that always try to increase the g_m and for improving the g_m, increment the no. of the fin into FinFET structure is used.

$$g_m = \frac{\partial I_{ds}}{\partial V_{gs}} \tag{1}$$

Figure 5 variations of the output conductance g_d as a function of the drain voltage in respective of constant $V_{gs} = 1$ V, constant threshold voltage, and doping concentration for both devices DGJLT and FinFET. From the figure, we can also conclude the output-conductance of both the devices is getting high value for low drain voltages and as the drain voltage increases output-conductance of both the devices starts to decreases because current starts to flow into the channel region. For increases, the gain we need the lower output conductance g_d, so that output resistance is high. Result achieve the high gain device but for higher current handling capability required g_d high respectively, r_o has lower value result I_{ds} increases.

$$g_{ds} = \frac{\partial I_{ds}}{\partial V_{ds}} \tag{2}$$

The variation of the early voltage V_{EA} as the function of gate voltage is shown in Fig. 6 for both the device at constant $V_{ds} = 1$ V. The early voltage of the device can be calculated by using Eq. 3 as a result of the lower gate V_{gs}, the early voltage is a very less nearly constant value, as the gate voltage increases early voltage increases as linear configuration and at on larger gate voltage it will reach an almost constant value.

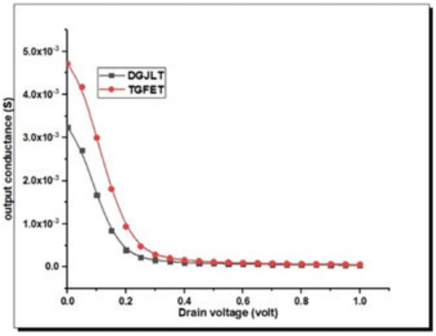


Fig. 5 Variation of output-conductance as a function of gate voltage from $V_{gs} = 0$ to 1 V at $V_{ds} = 1$ V

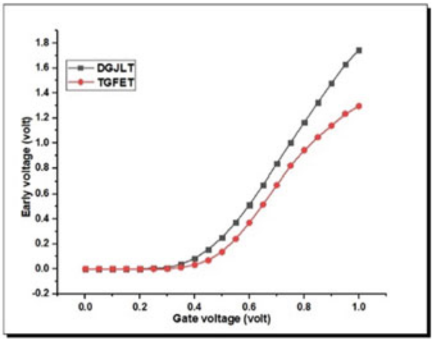


Fig. 6 Variation of early voltage V_{EA} as a function of gate voltage from $V_{gs} = 0$ to 1 V at $V_{ds} = 1$ V

$$V_{EA} = \frac{I_{ds}}{g_{ds}} \tag{3}$$

Intrinsic gain A_V is an important parameter of analog and RF system applications. The variation of the gain w.r.t gate to source voltage as is obtained from Fig. 7 for both device structures at constant $V_{ds} = 1$ V. As the gate voltage increase, the gain of both structures start to increase, and at a certain value of V_{gs} the gain will achieve a high value and then it will again start to decrease w.r.t. V_{gs}, the gain can be calculated by Eq. 4.

$$\text{Gain} = \frac{g_m}{g_{ds}} = \frac{g_m}{I_{ds}} * V_{EA} \tag{4}$$

Figure 8 shows the TGF with the variation of the V_{gs} at constant drain voltage Vds = 1 V for both structures i.e. DGJLT and FINFET. TGF is defined as the trans-conductance per unit current and it is given as Eq. 5. TGF is described how efficiently it translates the current into g_m less value of TGF is favorable for the microwave application system.

$$\text{TGF} = \frac{g_m}{I_{ds}} \tag{5}$$

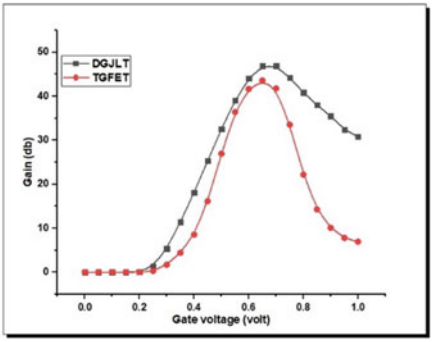


Fig. 7 Variation of intrinsic gain A_V as a function of gate voltage from $V_{gs} = 0$ to 1 V at $V_{ds} = 1$ V

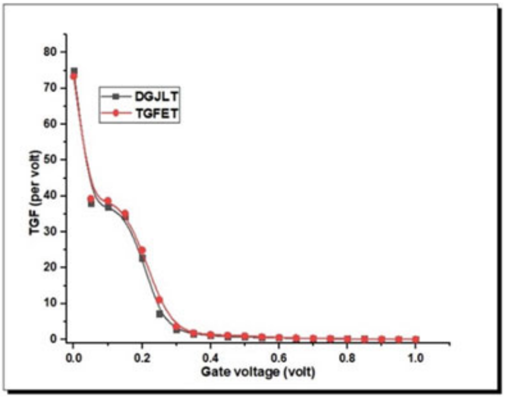


Fig. 8 Variation of trans-conductance generation factor TGF as a function of gate voltage from $V_{gs} = 0$ to 1 V at $V_{ds} = 1$ V

3.2 RF Performance

The capacitances i.e. gate to source capacitance C_{gs}, the gate to drain capacitance C_{gd}, cut off frequency f_T, trans-conductance frequency product TFP, gain frequency product GFP, values are extracted at a constant value of frequency of 1 MHz, as the function of V_{gs} from 0 to 1 V with the steps size of 0.05 V at constant $V_{ds} = 1$ V and constant thickness of ferro layer at 1 nm, for both device structures i.e. DGJLT and FinFET.

The outcomes are observed from Fig. 9 that are represent the gate to source capacitance C_{gs} as the function of gate voltage V_{gs} for both device structure DGJLT and FinFET. Outcomes observed at the $V_{ds} = 1$ V, constant threshold voltage, and constant doping concentration. From starting the value C_{gs} is nearly constant for less gate voltage V_{gs} and as the gate voltage increases the capacitance value increases linearly for both devices after at a certain value of V_{gs} the value of C_{gs} for both, the device is being started to saturate, after that point if we increase the V_{gs} no changes occur into the gate to source capacitance C_{gs}.

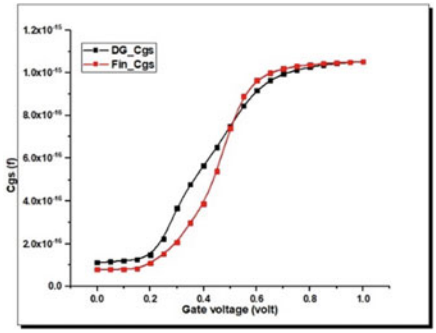


Fig. 9 Variation of gate to source capacitance C_{gs} as a function of gate voltage from $V_{gs} = 0$ to 1 V at $V_{ds} = 1$ V

The results are detected from Fig. 10 that is symbolizes the gate-to-drain capacitance C_{gd} as the function of gate voltage for both device structure DGJLT and FinFET. Results witnessed at the $V_{ds} = 1$ V, and constant doping concentration. Gate to drain capacitance increases when parasitic capacitance increases between drain to gate. Gate to drain capacitance is responsible for found the output load capacitance that is used to find the frequency of the device. Gate to drain capacitance C_{gd} up to $V_{gs} = 0.25$ V, nearly constant and as the value of V_{gs} increases the C_{gd} for both the device DGJLT and FinFET increase linearly up to a certain value of V_{gs}. When gate voltage reaches at $V_{gs} = 0.625$ V then after increases the value of V_{gs}, C_{gd} of the FinFET device start to decreases and the C_{gd} of DGFET is being started to move into saturation.

The result obtained from Fig. 11 shows the cut-off frequency as a function of gate voltage at constant doping concentration and content $V_{ds} = 1$ V. Cut off frequency of FinFET having low value compare to DGJLT this happens because of trans-conductance having the nearly same values for both structures i.e. g_m directly proportional to f_T but the summation of the capacitance of C_{gs} and C_{gd} in the case of FinFET getting the high value compare to DGJLT i.e. inversely proportional to Cut off frequency values varies from 2×10^{11} to 3×10^{11} Hz for both structures as DGJLT and FinFET.

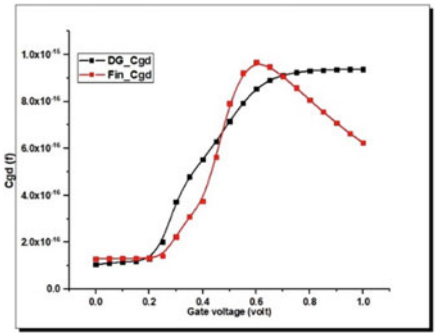


Fig. 10 Variation of gate to drain capacitance C_{gs} as a function of gate voltage from $V_{gs} = 0$ to 1 V at $V_{ds} = 1$ V

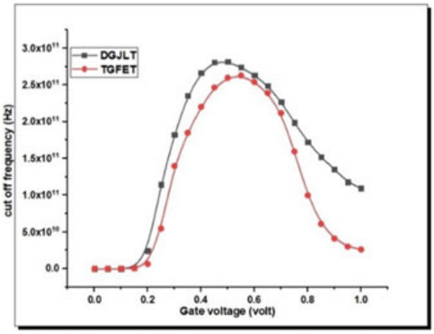


Fig. 11 Variation of Cut off frequency as a function of gate voltage from $V_{gs} = 0$ to 1 V at $V_{ds} = 1$ V

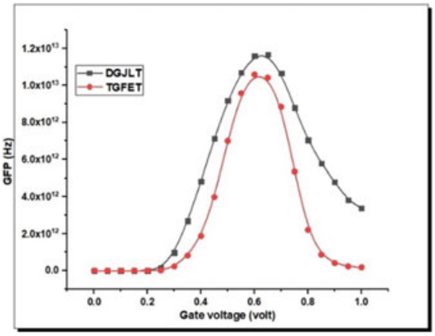


Fig. 12 Variation of gain frequency product GFP as a function of gate voltage from $V_{gs} = 0$ to 1 V at $V_{ds} = 1$ V

$$\mathrm{GFP} = \frac{\mathrm{g_m}}{\mathrm{g_d}} * \mathrm{f_T} \tag{6}$$

$$\mathrm{TFP} = \frac{\mathrm{g_m}}{\mathrm{I_{ds}}} * \mathrm{f_T} \tag{7}$$

The product of TGF and cut-off frequency describes TFP. After calculation, the variation of TFP w.r.t function of V_{gs} from 0 to 1 V, and GFP has described by, a product of gain and cut off frequency. Figure 12 shows the variation of GFP w.r.t function of V_{gs} from 0 to 1 V TFP and GFP are two important parameters for analog circuit design and RF application systems. TFP and GFP contribute a major role when the op-amp circuit is designed for a high-frequency application.

4 Conclusion

In this paper, we designed two device structures i.e. DGJLT and TG-FinFET and done a comparison of both structures w.r.t constant thickness of ferroelectric material t_{fe} = 1 nm and constant V_{ds} = 1 V. Simulation of both device structures has been done into the silvaco atlas tool. The simulated analog FOMs are trans-conductance

Table 3 Extracted analog FOMs of DGJLT and FINFET

Parameter	DGJLT	FINFET
Trans-conductance (S)	3.8×10^{-5}	3.9×10^{-5}
Output-conductance (S)	3.4×10^{-3}	4.6×10^{-3}
Early voltage (volt)	1.6	1.2
Intrinsic gain (dB)	75.36	74.20
TGF (V^{-1})	45.62	44.37

Table 4 Extracted RF FOMs of DGJLT and FINFET

Parameter	DGJLT	FINFET
Cut off frequency (Hz)	2.8×10^{11}	3.1×10^{11}
C_{gs}(F)	8×10^{-16}	1.0×10^{-15}
C_{gd}(F)	1×10^{-15}	1.1×10^{-15}
GFP (Hz)	1.2×10^{13}	1×10^{13}
TFP (Hz)	8×10^{12}	6×10^{12}

g_m (~+35%) [17], output conductance g_{ds}, early voltage V_{EA}(~–33%) [17], intrinsic gain A_V(~+45%) [17], trans-conductance generation factor TGF(~–38%) [17] and peak value of these parameters are showing in Table 3, simulated RF FOMs is the gate to source capacitance C_{gs}, the gate to drain capacitance C_{gd}, cut off frequency f_T(~–80%) [17], trans-conductance frequency product TFP, gain frequency product GFP(~+86%) [17] and peak value of these parameters are showing in Table 4. For recent work suggestions, paper number [17] is used for comparative analysis. After analysis of results, we can conclude that NC-FinFET provides a better result as compare to DGJLT.

References

1. Lee, C.-W., et al.: Performance estimation of junction less multigate transistors. Solid State Electron. **54**, 97–103 (2010)
2. Jazaeri, F., Barbut, L., Sallese, J.-M.: Modeling and design space of junctionless symmetric DG MOSFETs with long channel. IEEE Trans. Electron. Devices **60**(7), 2120–2127 (2013)
3. Koukab, A., Jazaeri, F., Sallese, J.-M.: On performance scaling and speed of junctionless transistors. Solid State Electron. **79**, 18–21 (2013)
4. Mehtaand, H., Kaur, H.: Modeling and simulation study of novel double gate ferroelectric junctionless (DGFJL) transistor. Superlattices Microstruct. **97**, 536547 (2016)
5. Zhu, Y., et al.: Negative capacitance characteristics in a steady state ferroelectric capacitor made of parallel domains. IEEE Electron. Device Lett. **38**(8), 1176–1179
6. Yan, R., et al.: Nanowire transistor without junctions. Nat. Nano Technol. **5**, 225–229 (2010)
7. Jo, J., Shin, C.: Negative capacitance field effect transistor with hysteresis-free sub-60 mV/decade switching. IEEE Electron. Device Lett. **37**(3), 245–248 (2016). https://doi.org/10.1109/LED.2016.2523681
8. Lee, C.W., Afzalian, A., Akhavan, N.D., Yan, R., Ferain, I., Colinge, J.P.: Junctionless multigate field effect transistor. Appl. Phys. Lett. **94**(5), 053511-1–053511-2 (2009)

9. Holtij, T., Schwarz, M., Kloes, A., Iniguez, B.: Threshold voltage and 2D potential modelling within short channel junctionless DG MOSFETs in subthreshold region. Solid State Electron. **90**, 107–115 (2013)
10. Duarte, J.P., Choi, S.J., Choi, Y.K.: A full range drain current model for double gate junctionless transistor. IEEE Trans. Electron. Devices **58**(12), 4219–4225 (2011)
11. Kranti, A., et al.: Junctionless nano wire transistor (JNT): properties and design guidelines. In: Proceedings of the ESSDERC, pp. 357–360, 14–16 September 2010
12. J.P., et al.: Junctionless nano wire transistor (JNT): properties and design guidelines. Solid State Electron. **65–66**, 33–37 (2011)
13. Subramanian, V.: Multiple gate field-effect transistors for future CMOS technologies. IETE Tech. Rev. **27**(6), 446–454 (2013)
14. Huang, X., et al.: Sub 50-nm FinFET: PMOS. In: Technical Digest of IEDM, vol. 67 (1999)
15. Hisamoto, D., et al.: FinFET-a self-aligned double-gate MOSFET scalable beyond 20 nm. IEEE Trans. Electron. Devices **47**, 2320–2325 (2000)
16. Kumar, S., Goyal, S.: Design, simulation and parameter extraction of a TG-FinFET Device. Int. J. Sci. Eng. Technol. Res. (IJSETR) **4**(8) (2015)
17. Roy, N.Ch., Gupta, A.: Analytical surface potential modeling and simulation of junction-less double gate (JLDG) MOSFET for ultra low-power analog/RF circuits. Microelectron. J. (2015). www.elsevier.com/locate/mejo, Accepted 28 July 2015, Available online 24 August 2015
18. Yu, B., et al.: FinFET scaling to 10 nm gate length. In: Technical Digest of IEDM, vol. 251 (2002)

Comparative Analysis of Different Low Band Gap Source Materials in Heterojunction Hetero-Dielectric GAA TFET

Chandni Tiwari, Varun Mishra, Afreen Anamul Haque, Yogesh Kumar Verma, and Santosh Kumar Gupta

Abstract In this paper, we have done a comparative analysis of three source materials which are Si, SiGe and InAs. The proposed structure of the device is hetero-dielectric, heterojunction GAA-TFET. The energy band and I_D-V_{Gs} analysis substantiates that InAs source material results in lower tunnelling path which leads to higher Ion. Additionally, application of hetero dielectrics HfO_2 and SiO_2 has significant impact on increasing the I_{on} of the device. The surface potential and electrostatic analysis of the InAs source-based device verifies the results obtained from I_D-V_{GS} curve.

Keywords Gate all around · Tunnel FET · Low band gap · Hetero-dielectric · Hetero-junction

1 Introduction

The increase in transistor density on a single chip and scaling of dimension have led to enhanced power density. Scaling of supply voltage V_{DD} results in reduction of power consumption which is highly desirable in low power applications [1]. The threshold voltage V_{TH} needs to be scaled properly while limiting the V_{DD} of the transistor to have same on current I_{on}. This results in increased leakage current

C. Tiwari · V. Mishra (✉) · A. A. Haque
Graphic Era (Deemed to be University), Dehradun, Uttarakhand, India
e-mail: rel1605@mnnit.ac.in

C. Tiwari
e-mail: Chandnitiwari.ece@geu.ac.in

Y. K. Verma
Lovely Professional University, Phagwara, Punjab, India
e-mail: yogesh.25263@lpu.co.in

S. K. Gupta
Motilal Nehru National Institute of Technology, Allahabad, Uttar Pradesh, India
e-mail: skg@mnnit.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_51

I_{off} and static leakage power exhibiting exponential relation. The increased I_{off} in MOSFET is attributed to sub-threshold slope which is equal to 60 mV/decade. The scaling of V_{DD} is limited by a contrast relationship between decreasing V_{TH} and small leakage power [2]. This puts a restriction on reduction of power for low power and high-performance requiring applications.

The research has been conducted over the years to design devices with steep sub-threshold swing (SS). Tunnel field effect transistor (TFET) is on such device which have band to band tunnelling (BTBT) conduction mechanism and results in lower supply voltage and I_{off}, in contrast to MOSFETs [3].

TFET exhibits BTBT conduction mechanism. It means that in n-channel TFET, the drain current arises due to charge carriers tunnelling from valance band of the source to conduction band of channel [4]. Also, TFET can achieve 60 mV/decade SS and has low I_{off}. Numerous researches have been conducted on reducing the SS value of TFET which is quite a challenging task. The reported literature has shown minimum point subthreshold (SS_{min}) as 5 and 11 mV/decade [5, 6]. The average sub threshold swing (SS_{avg}) has far more importance than SS_{min}. The value of SS_{avg} is considerably large with respect to SS_{min} in conventional silicon TFET. Whereas, MOSFET results in identical values of SS_{avg} and SS_{min}. In Si TFET, increasing V_{GS} results in a dramatic increase in SS $_{avg}$. This results in a larger difference between SS_{avg} and SS_{min} which leads to unstable value of SS_{avg} [7]. Another disadvantage associated with Si TFET is lower I_{on} which is attributed to larger band gap and mass of charge carriers of silicon materials. To avoid this problem, several techniques have been adopted by the researchers. Using lower band gap material in the place of silicon, using different gate material for work function engineering, use of lower band gap material for source and drain regions are a few focused research areas to improve I_{on} [8].

Apart from that, various device structures have been suggested to overcome the problems faced by conventional silicon as source material. It has been pointed that heterojunction TFET using Germanium as source material have shown enhanced I_{on}. Also, it is reported that a high K dielectric should be introduce at gate which leads to enhanced I_{on} and SS and decreased leakage current [9–12].

Numerous device structures such as Pi gate, omega gate, dual gate, triple gate, stack gate and surrounding gate have been investigated by the researchers for better performance and to avoid scaling limitations of the devices [13]. Gate all around TFET is one such structure which has high gate controllability and small screening length resulting in considerably higher I_{on}. It also offers high packaging density and steep SS characteristics.

Although the heterojunction TFET shows higher I_{on}, the values of SS_{min} and SS_{avg} needs to be improved [10, 14]. Various results are published on using different source material in order to overcome the problems faced by heterojunction TFET but there is lack of consensus on the outcomes. Hence, an extensive research needs to be done in heterojunction TFET to acquire a stable value of SS_{avg} with enhanced I_{on}.

In this paper, we report a comparative study to understand the different electronic properties of GAA TFET by varying the source materials. We have used Si, SiGe and InAs as source materials and have done analysis to understand its impact on various device characteristics and performances. Hetro-dielectric structure of the device is implemented in the proposed work by using HfO_2 and SiO_2 as high k dielectric materials at source and drain, respectively.

2 Device Dimensions and Model

The 2D schematic of the device under investigation is shown in Fig. 1. The pristine regions of source, drain and channel, of the proposed device are fabricated on intrinsic silicon film. The intrinsic carrier concentration of the used Si is $n_i = 10^{15}$ cm^{-3}. The thickness of the pristine silicon under consideration is 10 nm.

The various device parameters are listed in Table 1. HfO_2 and SiO_2 are used as high K dielectric in source and drain side respectively to implement hetero-dielectric structure.

To have a comparative analysis of effect of different source materials, same drain material has been used for all the devices. This step also helps in felicitation of low I_{off}. The TCAD tool Silvaco ATLAS is used throughout this study to perform

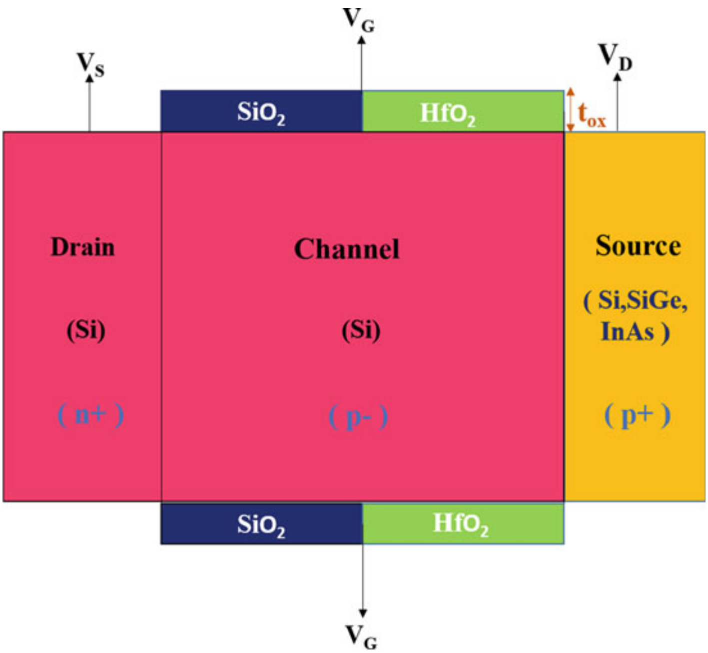


Fig. 1 Schematic of the investigated device with different source materials

Table 1 Parameters of devices under consideration

Parameters	Numerical values
Gate length	22 nm
Drain Length	20 nm
Source Length	20 nm
Oxide thickness	2 nm
Doping Concentration in Source (N_A)	10^{20} cm^{-3}
Doping Concentration in Drain (N_D)	10^{18} cm^{-3}
Channel doping concentration	10^{17} cm^{-3}
Work-function of Gate	4.7 eV

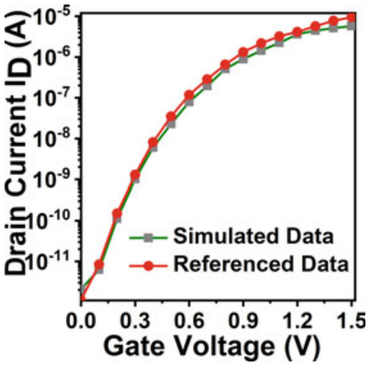


Fig. 2 Comparative analysis of simulated and reference data

all the numerical analysis of the devices. The analytical models used in various simulations are Auger recombination model, Shockley-Read Hall recombination model, quantum model, Lombardi mobility model, non-local BTBT model, non-local derivative model and phonon-assisted tunneling model. Spatial energy distribution associated with energy levels are analysed using Non local model, resulting in better accuracy as compared to the local models.

3 Results and Discussions

This research works analyses the performance of GAA TFET with Si, SiGe and InAs as source materials in terms of electrical and analog/RF parameters. The numerical model used in this work is calibrated again the previously published work, which is shown in Fig. 2.

The change in energy band with respect to channel length is shown in Fig. 3(a) and (b) for off state and on state, respectively, for all the devices. The gate voltage

V_{GS} is 0 and 0.5 V for off and on state of the devices, respectively. The drain voltage V_{DS} is fixed at 0.5 V, for on and off states. Silicon, SiGe and InAs has band gaps 1.1, 0.94 and 0.4 eV respectively. The inclusion of low band gap material at the source side has decreased the tunnelling path which is evident from Fig. 3(b). In comparison to Si, SiGe has resulted in lower tunnelling path which further reduces for InAs. This phenomenon enhances the tunnelling of charge carriers by reducing the offered tunnelling resistance. Another reason attributed to this phenomenon is usage of hetero-dielectric configuration with high k-material which enhances I_{on} and reduces I_{off}. Table 2 presents the comparison of performance parameters of proposed device with the formerly reported literatures.

The narrow band gap of InAs results in higher surface potential which is evident from Fig. 4(a). As the energy band gap of the source materials are increasing, the surface potential declines, which can be inferred from Fig. 4(a). The direct relation between surface potential and electric field leads to larger electrostatic field of at the source channel junction of InAs, in contrast to other source materials. The results can be accessed from Fig. 4(b).

The transfer characteristics of the proposed devices are shown in Fig. 5(a). The InAs source material results in higher I_{on} in comparison to the source materials. The reason is credited smaller tunnelling width which leads to lesser tunnelling resistance. Transconductance characteristics curve for all the proposed devices are given in

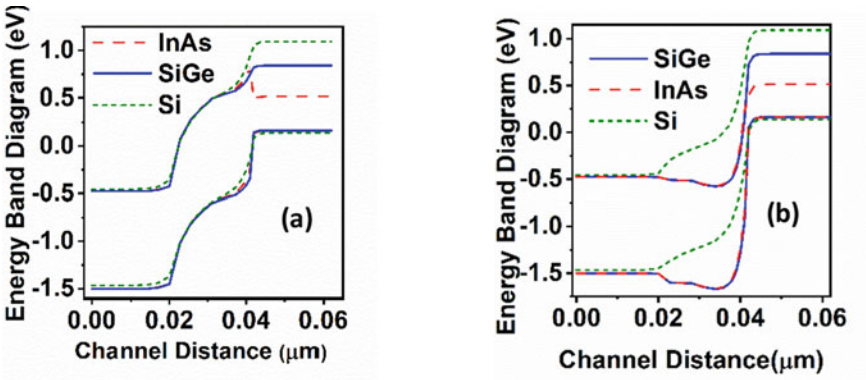


Fig. 3 Energy band diagram of the proposed devices in **a** off-state and **b** on-state

Table 2 Comparison of performance metrics

Device	I_{ON} (A)	I_{ON}/I_{OFF}	SS (mV/dec)
Proposed Si source GAA-TFET	2.10×10^{-8}	2.5×10^{5}	14.98
Reference [15]	7.65×10^{-5}	2.4×10^{8}	62.57
Reference [16]	4.23×10^{-5}	6.09×10^{12}	62.92
Reference [17]	$\sim 1 \times 10^{-7}$	$\sim 10^{11}$	51
Proposed InAs source GAA-TFET	1.78×10^{-4}	3.58×10^{13}	21.4

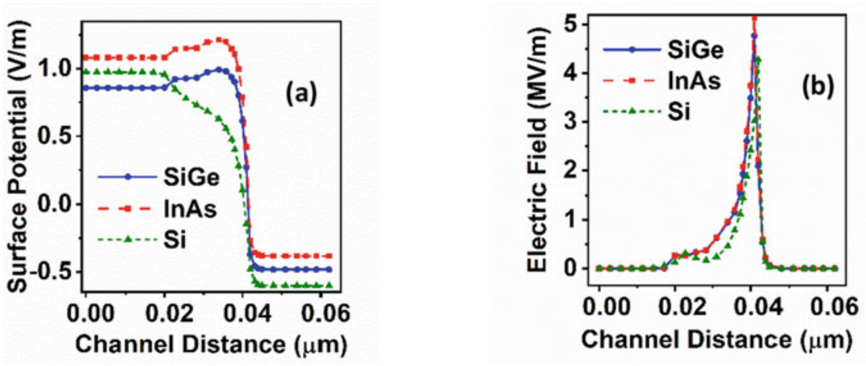


Fig. 4 Comparison of **a** Surface potential, and **b** Electric field of the proposed devices

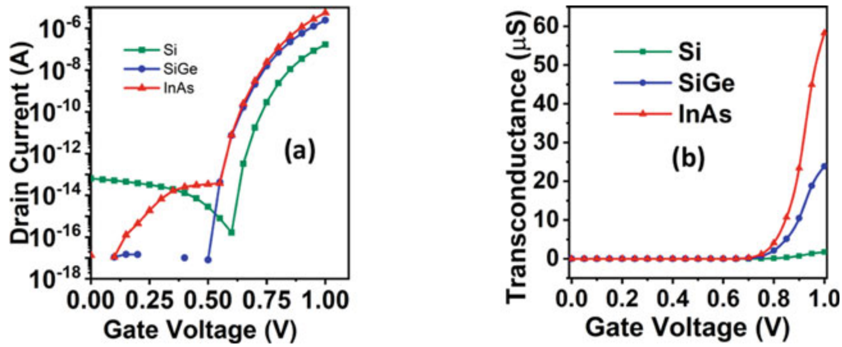


Fig. 5 **a** I_D-V_{GS} curve, and **b** Transconductance characteristics of the proposed devices

Fig. 5(b). Transconductance is an important device parameter which estimates the current driving capability of the investigated device. The value of transconductance is evaluated by $\partial I_D/\partial V_{GS}$, where V_{DS} is kept constant. The direct dependence of g_m on I_D leads to a higher value of transconductance for InAS source material-based device, as I_D of this device is larger. That is evident from Fig. 5(a).

The gate capacitance Vs. VGS of all the three devices is presented in Fig. 6(a). Summation of gate to drain capacitance and grate tp source capacitance leads to total capacitance of TFET. The Cut-off frequency of the device is defined as the value of frequency the gain attains reaches unity. This is a substantial parameter to understand its usage in RF applications.

$$f_T = \frac{g_m}{2\pi\ C_{gg}} \tag{1}$$

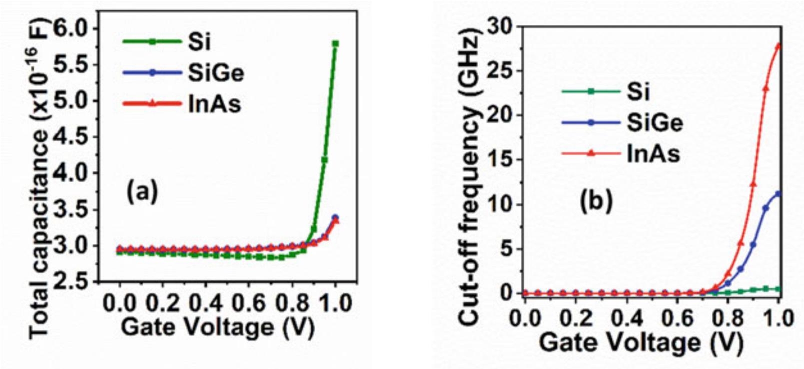


Fig. 6 **a** Capacitance **b** Cut off frequency of the proposed devices

The parasitic capacitance exhibits an inverse relationship with cut off frequency which can be deducted from above Eq. (1). The graph between cut-off frequency and gate voltage is shown in Fig. 6(b).

4 Conclusion

In summary, we have demonstrated the effect of narrow band gap materials as source in heterojunction heterodielectric GAA-TFET. The source materials under investigation are Si, SiGe and InAs. The energy band analysis showed that the tunnelling path is becoming narrower for InAs, which further resulted in decreased tunnelling resistance and increased Ion. Further, I_D-V_{GS} and transconductance curve substantiates higher drain current and transconductance for InAs in comparison to other materials under investigation.

References

1. Dreslinski, R.G., Wieckowski, M., Blaauw, D., Sylvester, D., Mudge, T.: Near-threshold computing: reclaiming Moore's law through energy efficient integrated circuits. Proc. IEEE **98**(2), 253–266 (2010). https://doi.org/10.1109/JPROC.2009.2034764
2. Datta, S., Liu, H., Narayanan, V.: Tunnel FET technology: a reliability perspective. Microelectron. Reliab. **54**(5), 861–874 (2014). https://doi.org/10.1016/j.microrel.2014.02.002
3. Seabaugh, A.C., Zhang, Q.: Low-voltage tunnel transistors for beyond CMOS logic. Proc. IEEE **98**(12), 2095–2110 (2010). https://doi.org/10.1109/JPROC.2010.2070470
4. Beohar, A., Yadav, N., Vishvakarma, S.K.: Analysis of trap-assisted tunnelling in asymmetrical underlap 3D-cylindrical GAA-TFET based on hetero-spacer engineering for improved device reliability. Micro Nano Lett. **12**(12), 982–986 (2017). https://doi.org/10.1049/mnl.2017.0311

5. Musalgaonkar, G., Sahay, S., Saxena, R.S., Kumar, M.J.: A line tunneling field-effect transistor based on misaligned core-shell gate architecture in emerging nanotube FETs. IEEE Trans. Electron. Devices **66**(6), 2809–2816 (2019). https://doi.org/10.1109/TED.2019.2910156
6. Shirazi, S.G., Karimi, G.R., Mirzakuchaki, S.: GAA CNT TFETs structural engineering: a higher ON current, lower ambipolarity. IEEE Trans. Electron. Devices **66**(6), 2822–2830 (2019). https://doi.org/10.1109/TED.2019.2912950
7. Lu, H., Seabaugh, A.: Tunnel field-effect transistors: state-of-the-art. IEEE J. Electron. Devices Soc. **2**(4), 44–49 (2014). https://doi.org/10.1109/JEDS.2014.2326622
8. Vanlalawpuia, K., Bhowmick, B.: Investigation of a Ge-source vertical TFET with delta-doped layer. IEEE Trans. Electron. Devices **66**(10), 4439–4445 (2019). https://doi.org/10.1109/TED.2019.2933313
9. Choi, W.Y., Lee, W.: Hetero-gate-dielectric tunneling field-effect transistors. IEEE Trans. Electron. Devices **57**(9), 2317–2319 (2010). https://doi.org/10.1109/TED.2010.2052167
10. Gupta, A., et al.: A novel approach to investigate the impact of hetero-high-K gate stack on SiGe junctionless gate-all-around (JL-GAA) MOSFET. Silicon **14**(3), 1–8 (2021)
11. Rai, S.: Reliability analysis of junction-less double gate (JLDG) MOSFET for analog/RF circuits for high linearity applications. Microelectronics J. **64**(April), 60–68 (2017). https://doi.org/10.1016/j.mejo.2017.04.009
12. Gupta, A., Singh, A., Gupta, S.K., Rai, S.: Potential modeling of oxide engineered doping-less dual-material-double-gate Si–Ge MOSFET and its application. J. Nanoelectron. Optoelectron. **13**(8), 1115–1122 (2018)
13. Balamurugan, N.B., Priya, G.L., Manikandan, S., Srimathi, G.: Analytical modeling of dual material gate all around stack architecture of tunnel FET. In: Proceedings IEEE International Conference VLSI Design, vol. 2016, pp. 294–299, March 2016. https://doi.org/10.1109/VLSID.2016.74
14. Han, K., Long, S., Deng, Z., Zhang, Y., Li, J.: A novel germanium-around-source gate-all-around tunnelling field-effect transistor for low-power applications. Micromachines **11**(2), 164 (2020). https://doi.org/10.3390/mi11020164
15. Tripathi, S.L., Patel, G.S.: Design of low power Si0.7Ge0.3 pocket junction-less tunnel FET using below 5 nm technology. Wirel. Pers. Commun. **111**(4), 2167–2176 (2020). https://doi.org/10.1007/s11277-019-06978-8
16. Anand, S., Amin, S.I., Sarin, R.K.: Analog performance investigation of dual electrode based doping-less tunnel FET. J. Comput. Electron. **15**(1), 94–103 (2016). https://doi.org/10.1007/s10825-015-0771-4
17. Sahay, S., Kumar, M.J.: Nanotube junctionless FET: proposal, design, and investigation. IEEE Trans. Electron. Devices **64**(4), 1851–1856 (2017). https://doi.org/10.1109/TED.2017.2672203

Eigen Space and ANN Based Approach to Synthesize 12-Lead ECG

Deepak Sigroha, Kuldip Singh, and Ram Ishwar Vais

Abstract In this work, we propose and evaluate a novel method to synthesize 12 lead ECG from 3 lead ECG by combining principal component analysis (PCA) and artificial neural network (ANN). PCA is applied on ECG signals to obtain the principal components which are then used to train an ANN. Performance is evaluated using cross correlation coefficient, root mean square error, and wavelet energy based diagnostic distortion measure between original and synthesized leads. Performance evaluation show that the proposed method is an efficient way to synthesize 12-lead ECG from 3-lead ECG without loss of significant clinical information.

Keywords Electrocardiogram · ECG synthesis · Derived 12-lead ECG · Principal component analysis · Artificial neural network

1 Introduction

Analyzing the condition of heart from 12-lead Electrocardiogram (ECG) is very common for the diagnosis of various cardiovascular diseases (CVD). The gold standard 12-leads, lead I, II, III, aVR, aVL, aVF and V1 to V6, require 10 electrodes to be placed precisely over predefined body positions for recording [1]. This may disturb the comfort level of patients and increases the cost and complexity of the device. Experimental results [2] show that the 12-leads are having redundant information about heart and has high inter-lead correlation [3]. This factor can be utilized to minimize the number of electrodes for recording and synthesize the 12 lead system from those fewer leads.

The system that synthesizes 12-leads from reduced number of leads are often mentioned as derived 12-lead ECG system and the set of leads used to derive the 12-leads are referred as reduced lead set [4]. Reduced lead set normally consist of

D. Sigroha (✉) · R. I. Vais
Rajkiya Engineering College Sonbhadra, Sonbhadra 231206, Uttar Pradesh, India
e-mail: sigroha.deepak@gmail.com

K. Singh
Sant Longowal Institute of Engineering and Technology, Longowal 148106, Punjab, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_52

a subset of the 12-lead ECG. The system that uses special positioning of electrodes which reduces the leads are also available. The development of methods and system with reduced number of leads has been going on for the last seven decades. The first major derived 12-lead ECG system is introduced by Dower in 1968 [5]. He synthesized the 12-lead ECG from the orthogonal lead system introduced by Frank [5]. Synthesis of 12 lead ECG from a 3-lead semi-orthogonal subset using linear transformation arrays has been published by J. A. Scherer and J. M. Nicklas [3]. This work is based on the study that standard 12-lead ECG is redundant and smaller subsets of electrodes can be used for constructing 12-leads. Reconstruction of 12-lead ECG from reduced lead sets is published by S. P. Nelwan et al. [2]. They used linear regression for obtaining the derived leads from a subset of the set of leads (I, II, V1–V6) with lead I, II and at least one precordial leads always selected [2]. The idea of deriving standard 12-lead ECG from four standard leads using information redundancy in 12-lead system is proposed by Wei [6] in 2002. This model is based on Frank torso model. S. C. Man et al. published the reconstruction of standard 12-lead ECG from Mason-Likar electrode configuration in 2008 [7]. An artificial neural network model for deriving standard 12-lead ECG from three lead ECG is proposed by Atoui et al. in 2010 [8]. The 12-lead ECG reconstruction from sparse electrodes using Independent component Analysis is published by Tsouri et al. in 2013 [9].

Synthesis of ECG leads by transforming it into another domain and trying to learn the relationship may yield better results. In this experiment, we are trying to learn the relationship between standard 12-lead ECG and a subset of it by transforming it into PCA domain. Since most of the information is accumulated over first few principal components, learning between the components of 12-leads and its subset will produce better results. In Sect. 2, we discuss in detail about the proposed synthesis method and the database used. Results and discussion are given in Sect. 3 and Sect. 4 respectively. Section 5 concludes the work.

2 Method

The most popular and commonly available PTB database [10, 11] is used for this experiment. The main focus is to synthesize the eight independent leads of pathological ECG signal. Thus various pathological data recording such as anterior myocardial infraction, infero-lateral myocardial infraction along with healthy control are used for experimenting. Data of two minute duration is used for training and testing the model due to unavailability of longer period ECG.

The proposed method can be divided into three stages as shown in Fig. 1. Each ECG data is pre-processed for removing baseline wander and power line noise. The principal components (PC) from the pre-processed data is calculated and selected principal components are then used to train a feed forward ANN. The trained model is then used to synthesize the 12-lead ECG. Detailed explanation of each stages in Fig. 1 provided in the following subsections.

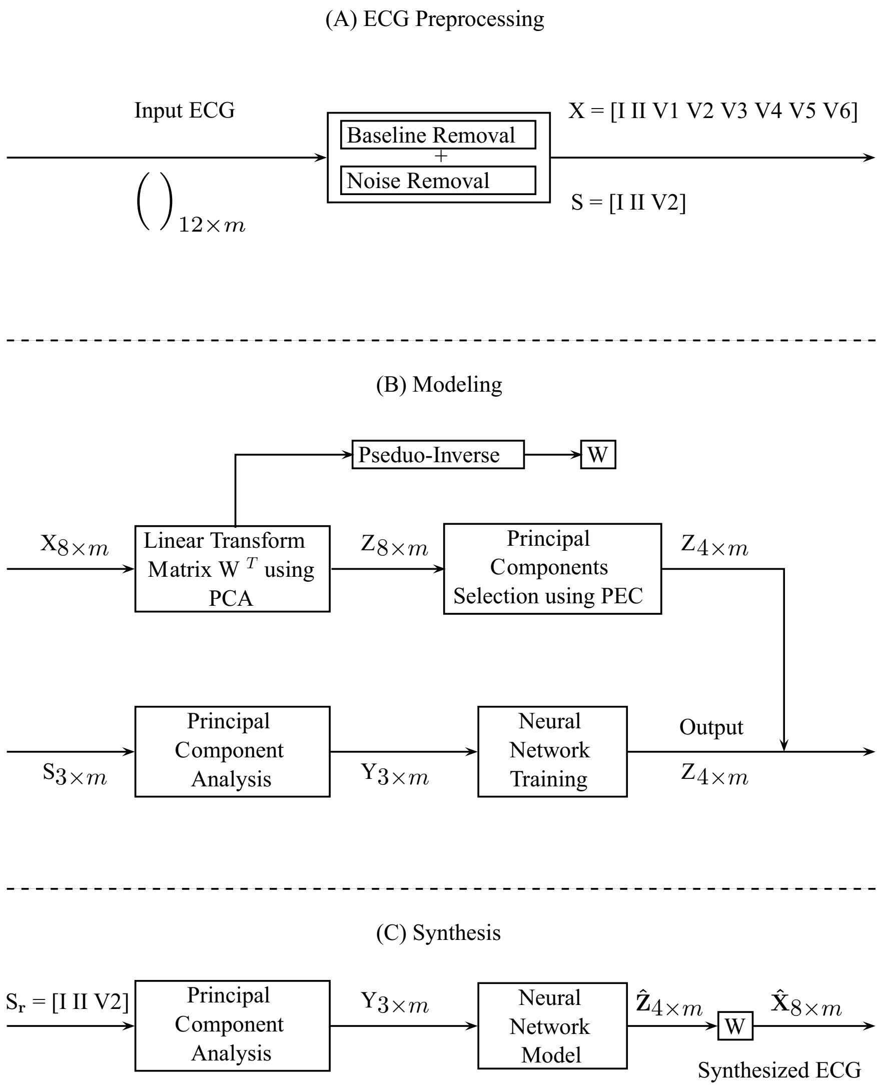


Fig. 1 Detailed representation of the proposed model. **A** Preprocessing, **B** Training the model and **C** Synthesis of eight Independent leads

2.1 Preprocessing

In the preprocessing stage, the baseline wandering effect is removed from each ECG data by filtering it with two median filters [12]. The QRS complex and P-wave is removed by using a 200 ms window and T-wave by a 600 ms window [12]. The resulting signal is the baseline and is subtracted from original signal to remove the baseline wander [12]. An 8^{th} order Butterworth low pass filter with a cutoff frequency 100 Hz is designed to remove the power line interference. The pre-processed ECG signal is divided into a learning set of duration $n_t \in \{5, 10, ...60\}$ seconds and the remaining signal is used to test the model.

2.2 PCA and ANN Based Model

ECG signal exhibits inter-beat and inter-lead redundancy and hence it is possible to remove the redundant information present in the 12-lead ECG using principal component analysis [13]. Let **X** be a matrix consisting of the eight independent leads given as $\mathbf{X} = [I\ II\ V1\ V2\ V3\ V4\ V5\ V6]$ and **S** be the matrix containing the reduced set of leads which is a subset of the eight independent leads given as $\mathbf{S} = [I\ II\ V2]$. Principal components are obtained by performing PCA over **X** and **S** separately as given in Eqs. 1 and 2.

$$\mathbf{Z} = \mathbf{W^T X}. \tag{1}$$

$$\mathbf{Y} = \mathbf{A^T S}. \tag{2}$$

Linear orthonormal transformation matrices **W** and **A** transforms **X** and **S** respectively to its principal components. Here, **Z** represents the eight principal components corresponding to the eight leads and **Y** represents three principal components corresponding to reduced leads.

After computing the eight principal components of **X**, the most representative principal components are to be selected. The selection of Principal components are based on the percentage energy contribution by principal components which is given by Eq. 3 as

$$PEC_i = \frac{\lambda_i}{\sum_{i=1}^{8} \lambda_i}. \tag{3}$$

where λ_i is the energy of the i^{th} principal component. Figure 2 shows the percentage energy contribution by principal components versus number of principal components. From Fig. 2, it is clear that most of the energy is contributed by the first four principal components. In this experiment, the first four principal components are used to train the model after discarding the last four principal components. After selecting the most representative principal components, the orthonormal matrix **W** becomes an '8×4' matrix which is then stored for applying in synthesis. In this work, an assumption is taken into account that the linear orthonormal transform matrix **W** is not changing over time for a particular patient.

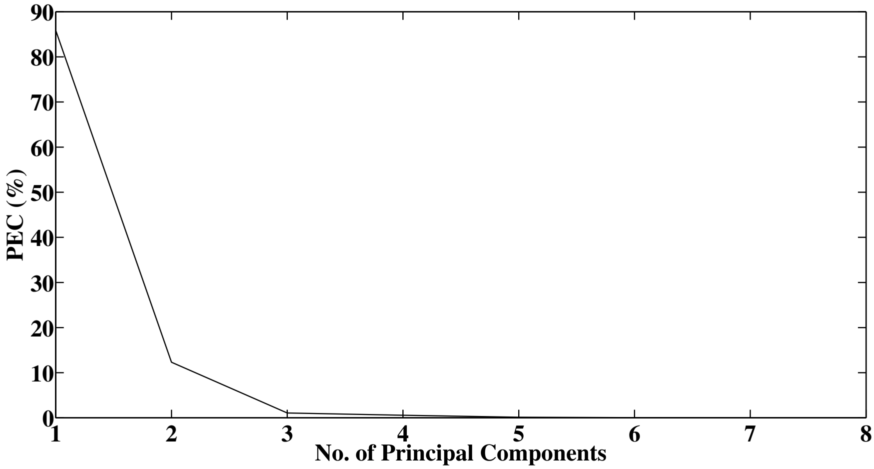


Fig. 2 Percentage energy contribution (PEC) against principal components of **X**

The principal components **Y** of the reduced lead set is calculated as per Eq. 2 and is provided as the input to a feed-forward ANN, keeping the four selected principal components from **Z** as the target output. The ANN is trained by the well known supervised back-propagation algorithm. Experiments are conducted with various learning set with time duration $n_t \in \{5, 10, ...60\}$ seconds and with varying number of hidden layer neurons $N_H = 2, 4, ..40$ to study the behavior and accuracy of proposed work.

2.3 *Synthesis of Independent Leads*

Synthesis is performed over the remaining duration of data after training the model. Principal components $\hat{\mathbf{Y}}$ of the remaining part of the reduced set $\mathbf{S_r}$ is calculated in a similar way as per Eq. 2. The principal components of reduced set, $\hat{\mathbf{Y}}$ is then applied as the input to the trained ANN model to obtain $\hat{\mathbf{Z}}$. This is an estimate of the principal components corresponding to the eight independent leads. Since the variation of orthonormal transformation matrix **W** is nominal over time period for a specific individual and the trained model is patient specific, it can be assumed that **W** is time invariant.

$$\hat{\mathbf{X}} = \mathbf{W}\hat{\mathbf{Z}}. \tag{4}$$

Thus the eight independent leads can be synthesized from the estimated principal components $\hat{\mathbf{Z}}$ using the linear orthonormal transformation matrix calculated during the training phase as given by Eq. 4.

2.4 Performance Evaluation

Performance evaluation is calculated by comparing the original leads with the reconstructed leads. Standard evaluation techniques such as correlation coefficient, root mean square error and wavelet energy based diagnostic distortion measure [14, 15] are used to evaluate the performance of synthesized ECG signals. Correlation coefficient (r_x) given by Eq. 5 is used to evaluate the similarity between the original and reconstructed leads.

$$r_x = \frac{\sum_{k=1}^{K}[x(k) - \mu_o][\hat{x}(k) - \mu_r]}{\sqrt{\sum_{k=1}^{K} x(k)^2}\sqrt{\sum_{k=1}^{K} \hat{x}(k)^2}}. \tag{5}$$

Amplitude difference between original and reconstructed signal is also evaluated by calculating root mean square error (RMSE) given by Eq. 6.

$$RMSE = \sqrt{\frac{1}{K}\sum_{k=1}^{K}[x(k) - \hat{x}(k)]^2}. \tag{6}$$

Here, $x(k)$ represents the original signal, $\hat{x}(k)$ represents the reconstructed signal, k denotes the k^{th} sample and K denotes the total number of samples in the signal. Mean value of original signal is denoted as μ_o and mean value of synthesized signal as μ_r. In this experiment, performance evaluation is done over mean free data and hence the values for μ_o and μ_r are zero.

ECG signal contains important diagnostic information and hence evaluating the synthesized signal for diagnostic quality is necessary. Here, wavelet energy based diagnostic distortion (WEDD) [15] given in Eq. 7 is used to evaluate the original and reconstructed signals.

$$WEDD = \sum_{l=1}^{L+1} w_l WPRD_l. \tag{7}$$

where w_l is weight for l^{th} sub-band given by Eq. 8 and WPRD is the wavelet percentage root mean square difference given by Eq. 9.

$$w_l = \frac{\sum_{k=1}^{K_l} d_l^2(k)}{\sum_{m=1}^{L+1}\sum_{k=1}^{K_l} d_m^2(k)}, l = 1, 2, ..(L + 1). \tag{8}$$

$$WPRD_l = \sqrt{\frac{\sum_{k=1}^{K_l}[d_l(k) - \hat{d}_l(k)]^2}{\sum_{k=1}^{K_l}[d_l(k)]^2}}. \tag{9}$$

Here, K_l is the number of wavelet coefficient in l^{th} sub-band and $d_l(k)$ is the k^{th} wavelet coefficient of signal in l^{th} sub-band. In this experiment, mean free original and synthesized ECG is decomposed to the 7^{th} level using Daubechies biorthogonal 4/4 wavelet filters.

3 Results

ECG signals taken from PTB diagnostic database is used in this study. Experimental results shows that different limb lead combination generates identical principal components since limb lead I, II and III are interchangeable. Here limb leads I and II are selected. Precordial lead V2 is selected since it is the most orthogonal lead and has been used successfully in the work by S. P. Nelwan et al. [2]. Lead set I, II, V2 contain as much diagnostic information as 12-lead ECG for diagnosing myocardial infraction and myocardial ischemia [16]. Limb lead III, augmented leads aVR, aVF and aVL are left out because these leads can be synthesized by linearly combining limb leads I and II. In this work, it is assumed that these three leads have sufficient information for synthesizing 12-lead ECG and this study is mainly focused on synthesizing the five precordial leads.

3.1 Artificial Neural Network Parameter Optimization

Performance of ANN based model depends upon the tuning of ANN parameters, i.e. the number of hidden layer neurons N_H and training set duration n_t. A parameter optimization analysis is performed on datasets in order to obtain the optimum parameter values. For this purpose, the mean square error (MSE) is plotted against the number of hidden layer neurons $N_H \in \{2, 4, ..., 40\}$ with a fixed training set duration of $n_t = 10$ s. Figure 3 shows that MSE decreases as the number of hidden layer neurons N_H increases. It can be observed that the mean square error is almost constant above 20 neurons and hence any number of hidden layer neurons above 20 can be used. A higher number of hidden layer neurons will increase computational complexity with slight improvement in the model. Thus, in this experiment, the number of hidden layer neurons N_H is tuned at 20. In order to find the training set duration, performance evaluation criteria as discussed in the previous section is calculated for various training set of duration $n_t \in \{5, 10, ...60\}$ seconds with the optimized number of hidden layer neuron N_H. Average value of the performance evaluation parameters for various training set duration n_t with optimal hidden layer neurons $N_H = 20$ is shown in Table 1. From Table 1, it can be observed that, the

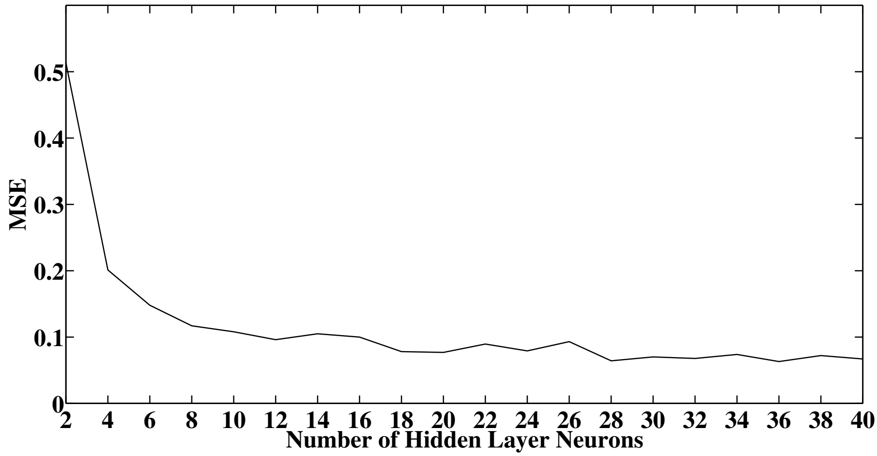


Fig. 3 Effect of the number of hidden layer neurons: MSE against the number of hidden layer neurons

Table 1 Averaged correlation coefficient (r_x), RMSE (μV) and WEDD (%) as function of training data set length n_t in sec for missing precordial leads

$n_{t(sec)}$	r_x	RMSE (μV)	WEDD (%)
05	0.9934	17.1451	9.6481
10	0.9940	15.9272	8.8393
15	0.9941	15.6772	8.8099
20	0.9944	15.3625	8.5303
25	0.9940	15.8109	8.7293
30	0.9943	15.4000	8.5935
35	0.9946	15.0936	8.2897
40	0.9946	14.9797	8.2942
45	0.9946	15.0512	8.2414
50	0.9943	15.4167	8.5165
55	0.9943	15.4133	8.5232
60	0.9942	14.9421	8.7188

optimum values for r_x, $RMSE$ and $WEDD$ are obtained at $n_t = 40$ s. Thus the number of hidden layer neurons N_H is tuned at 20 and training set duration n_t is tuned at 40 s.

3.2 Selection of 12-Lead ECG Subset

As per the work by S. P. Nelwan et al. [2], lead subset I, II and V2 are the most successful leads to synthesize the missing precordial leads. The performance evaluation of the synthesis process with lead I, II and any one precordial leads always included is performed in order to find the best lead subset. Experiment is performed on datasets patient05/s0025lrem with acute myocardial infraction (MI). Lead-wise performance evaluation for synthesizing missing precordial leads for different precordial leads is shown in Table 2. High r_x, low $RMSE$ and low $WEDD$ for lead subset I, II, V2 can be observed from Table 2 and hence lead subset I, II and V2 is the best lead configuration to synthesize the missing precordial leads for a patient. It can also be observed that lead configurations (I, II, V5 and I, II, V6) provides poor performance since V5 and V6 are much away from other precordial leads. Thus, based on this experiment, lead subset (I, II and V2) is selected for synthesizing the missing leads.

3.3 Performance Evaluation of Synthesized ECG

The standard 12-lead ECG signals used for this experiment is obtained from the PTB diagnostic ECG database. The ECG data which are used for experiments are acute myocardial infraction (MI) and healthy control (HC). The datasets patient05 /s0025lrem, patient017/s0053lrem and patient116/s0302lrem contains the multilead ECG signals with pathologies like anterior MI, infero-lateral MI and healthy ECG receptively. They are tested according to the proposed work and performance evaluation is executed over the original and synthesized ECG.

Tables 3, 4 and 5 show the correlation coefficient, root mean square error and wavelet energy based diagnostic distortion measure between original precordial leads and synthesized precordial leads for patient005/s0025lrem, patient017/s0053lrem and patient116/s0302lrem respectively. Proximity effect explained in [17] can be observed from these three tables. Lead in close proximity from basis lead have better reconstruction than away from it. Lead V1 and V3 have higher correlation coefficient, lower RMSE and lower WEDD than V5–V6. The lower synthesis correlation of leads V5–V6 suggest that it will be good to include one more precordial lead based on proximity effect for better synthesis accuracy with higher correlation. Tables 3, 4 and 5 shows an average $r_x = 0.9945$, $RMSE = 15.1989\,\mu$V and $W = 8.4140\%$ for patient005/s0025lrem, an average $r_x = 0.9963$, $RMSE = 18.0157\,\mu$V and $W = 7.9367\%$ for patient017/s0053lrem and an average $r_x = 0.9942$, $R = 14.0932\,\mu$V and $W = 8.5644\%$ for patient116/s0302lrem respectively. In this work, lead V3 showed strongest similarity between original and synthesized ECG with highest r_x and lowest $WEDD$ for almost all patients as shown in Table 3.

An additional precordial lead, i.e. lead V5, is added to existing lead subset I, II and V2 to synthesize the missing precordial leads for investigating the proximity effect and a phenomenal improvement in accuracy is observed. Lead-wise

Table 2 Lead-wise correlation coefficient r_x, RMSE and WEDD for synthesized missing precordial leads as function of different lead configuration (Lead I, II and one precordial lead is selected)

Lead subset	Synthesized lead	r_x	RMSE (μV)	WEDD (%)
I, II, V1	V2	0.9962	25.1786	6.2496
	V3	0.9962	25.8862	6.3934
	V4	0.9943	24.5051	8.5582
	V5	0.9872	21.4764	14.640
	V6	0.9873	17.3902	13.7937
I, II, V2	V1	0.9953	16.1163	8.0462
	V3	0.9994	10.4748	3.3136
	V4	0.9976	16.2522	6.1770
	V5	0.9916	17.7685	12.4440
	V6	0.9889	16.6138	12.3237
I, II, V3	V1	0.9950	16.4490	8.7758
	V2	0.9991	12.3016	3.6516
	V4	0.9987	12.0354	4.3273
	V5	0.9929	16.0700	10.7852
	V6	0.9885	16.5616	13.0034
I, II, V4	V1	0.9916	46.1016	29.2947
	V2	0.9969	83.8550	29.7715
	V3	0.9958	89.7500	29.8102
	V5	0.9641	50.5601	35.2846
	V6	0.9823	24.4001	19.9916
I, II, V5	V1	0.9401	149.4027	93.0436
	V2	0.9264	274.6196	87.3561
	V3	0.9410	266.0655	84.6327
	V4	0.9277	124.3353	91.9662
	V6	0.8915	87.7412	62.2828
I, II, V6	V1	−0.0342	412.3541	190.4619
	V2	−0.1280	794.8023	205.4803
	V3	−0.0733	815.1523	204.1508
	V4	0.1392	369.8659	219.0916
	V5	−0.3443	276.0097	223.5341

correlation coefficient, $RMSE$ in μV and $WEDD$ in % of missing precordial leads for patient005/s0025lrem with lead subset I, II, V2 and V5 is shown in Table 6. It is observed that, by including lead V5 we can get an average $r_x = 0.9968$ and $WEDD$ low as 2.5021% for lead V3, with an average $WEDD$ as 6.0232% for patient005/s0025lrem. Nonetheless, synthesis accuracy by lead subset I, II and V2 is very high and observed better performance as compared to the results reported in the literature.

The original and synthesized precordial leads with proposed method using reduced leads set I, II and V2 are specified in Figs. 4, 5 and 6 for patient005/s0025lrem,

Table 3 Lead-wise correlation coefficient r_x, RMSE and WEDD for patient005/s0025lrem with anterior MI

Synthesized lead	r_x	RMSE (μV)	WEDD (%)
V1	0.9953	15.9042	8.1906
V3	0.9994	10.0812	3.1411
V4	0.9976	15.8948	6.0139
V5	0.9916	17.5185	12.2192
V6	0.9887	16.5959	12.5050
Average	0.9945	15.1984	8.4140
Minimum	0.9887	10.0812	3.1411

Table 4 Lead-wise correlation coefficient r_x, RMSE and WEDD for patient017/s0053lrem with infero-lateral MI

Synthesized lead	r_x	RMSE (μV)	WEDD (%)
V1	0.9978	13.6852	6.1971
V3	0.9970	18.6719	7.6258
V4	0.9958	24.9326	8.3213
V5	0.9963	18.3138	7.7466
V6	0.9845	14.4751	9.7929
Average	0.9963	18.0157	7.9367
Minimum	0.9845	13.6852	6.1971

Table 5 Lead-wise correlation coefficient r_x, RMSE and WEDD for patient116/s0302lrem (HC)

Synthesized lead	r_x	RMSE (μV)	WEDD (%)
V1	0.9866	12.4739	13.1879
V3	0.9981	9.2856	4.5976
V4	0.9959	14.3710	8.1658
V5	0.9952	16.9995	8.2221
V6	0.9844	17.3362	8.6486
Average	0.9942	18.0157	8.5644
Minimum	0.9844	9.2856	4.5976

patient017/s0053lrem and patient116 /s0302lrem respectively. In Fig. 4, the original and synthesized precordial leads from patient05/s0025lrem with anterior myocardial infarction, are plotted and the diagnostic segments are highlighted. The original and synthesized leads V1, V3, V4, V5 in Fig. 4 reflects the electrocardiogram changes (Fig. 4(a)–(h), marked with ellipse) associated with anterior infraction. ST-elevation in original and synthesized V1, V3, V4, V5 can be observed with great symmetry in shape. Abnormal P-wave can be observed in original and synthesized V3 and V4 marked by ellipse. In Fig. 5, original and synthesized precordial leads

Table 6 Lead-wise correlation coefficient r_x, RMSE and WEDD of missing precordial leads for patient005/s0025lrem with lead subset I, II, V2 and V5

Synthesized lead	r_x	RMSE (μV)	WEDD (%)
V1	0.9963	14.1075	7.2588
V3	0.9997	7.8548	2.5021
V4	0.9988	11.2619	4.1860
V6	0.9925	13.4918	10.1457
Average	0.9968	11.6790	6.0232
Minimum	0.9925	7.8548	2.5021

from patient017/s0053lrem with infero-lateral myocardial infarction, is plotted with marked diagnostic segments. Inverted P-wave can be observed from Fig. 5(a), (b) and ST-elevation in original and synthesized V1 and V3 can be observed in Fig. 5 marked with ellipse. The original and synthesized ECG of a healthy patient116/s0302lrem is plotted in Fig. 6 with diagnostic segments marked as inverted P-wave, P-wave, QRS complex and T-wave. As shown in Figs. 4, 5 and 6 all diagnostic information such as P-wave, PR interval, QRS complex, ST segment change (elevation or depression), T-wave and QRS interval are preserved and shows high similarity between original and synthesized leads for all three patients. Proposed method is also tested for other patients in the PTB database. The results demonstrate a strong accuracy without loss of any significant clinical information.

4 Discussion

4.1 Comparison with Linear Regression Model

Most of the methods to synthesize missing precordial leads reported in literature are based on linear transform or linear regression. A patient specific transformation matrix is computed that synthesizes the missing precordial leads V1, V3 to V6 of 12-lead ECG, using a multiple linear regression algorithm as per Eq. 10

$$V_i = a_{i0} + a_{i1}I + a_{i2}II + a_{i3}V2 \tag{10}$$

The performance evaluation of proposed method and multiple linear regression for synthesizing missing precordial leads V1, V3–V6 for patient05/s0025lrem, patient017/s0053lrem and patient116/s0302lrem is shown in Table 7. High correlation coefficient r_x, very low $RMSE$ and low $WEDD$ can be observed for the proposed work as compared to multiple linear regression. It can also be observed from Table 7 that the accuracy of proposed method for lead V4–V6 is much better. Since the performance evaluation parameters for proposed method are superior than

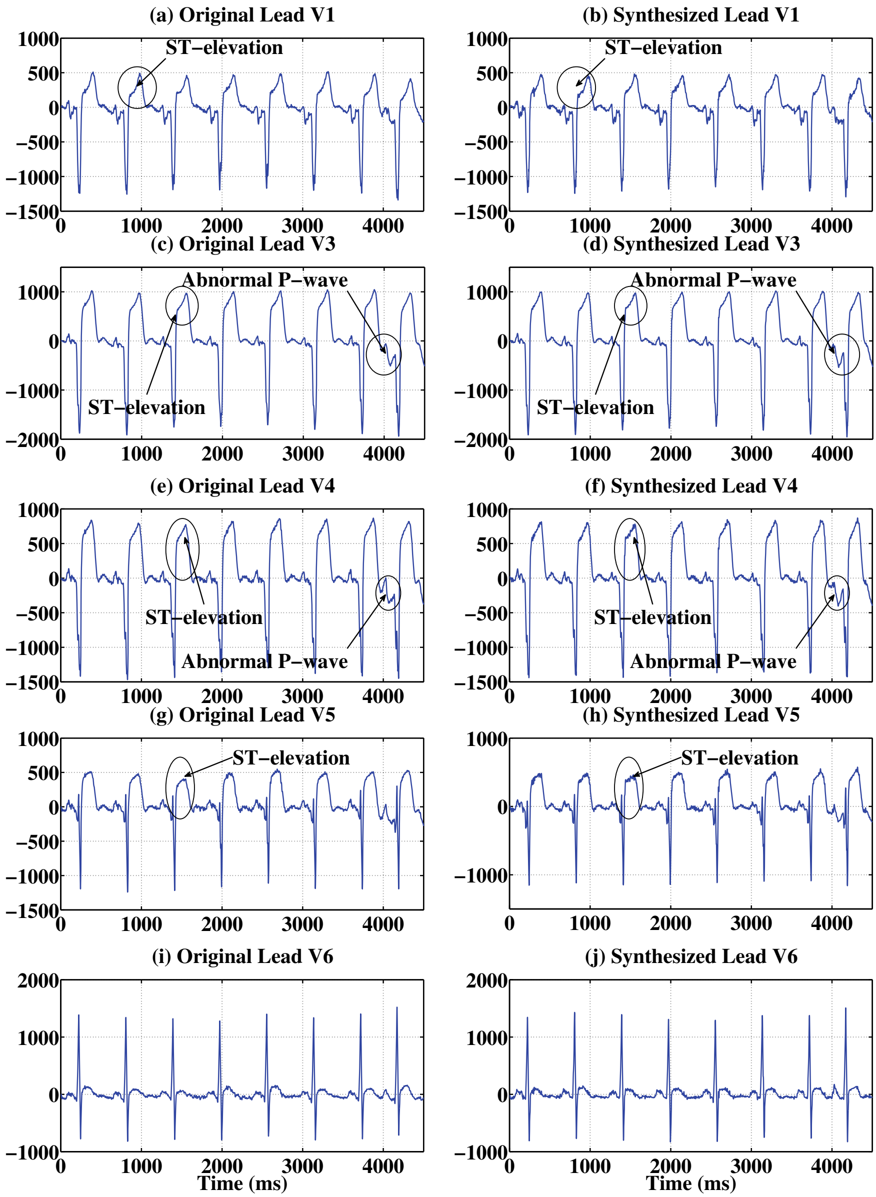


Fig. 4 Original and synthesized ECG leads V1, V3, V4, V5, and V6 for patient005/s0025lrem with lead subset I, II, and V2

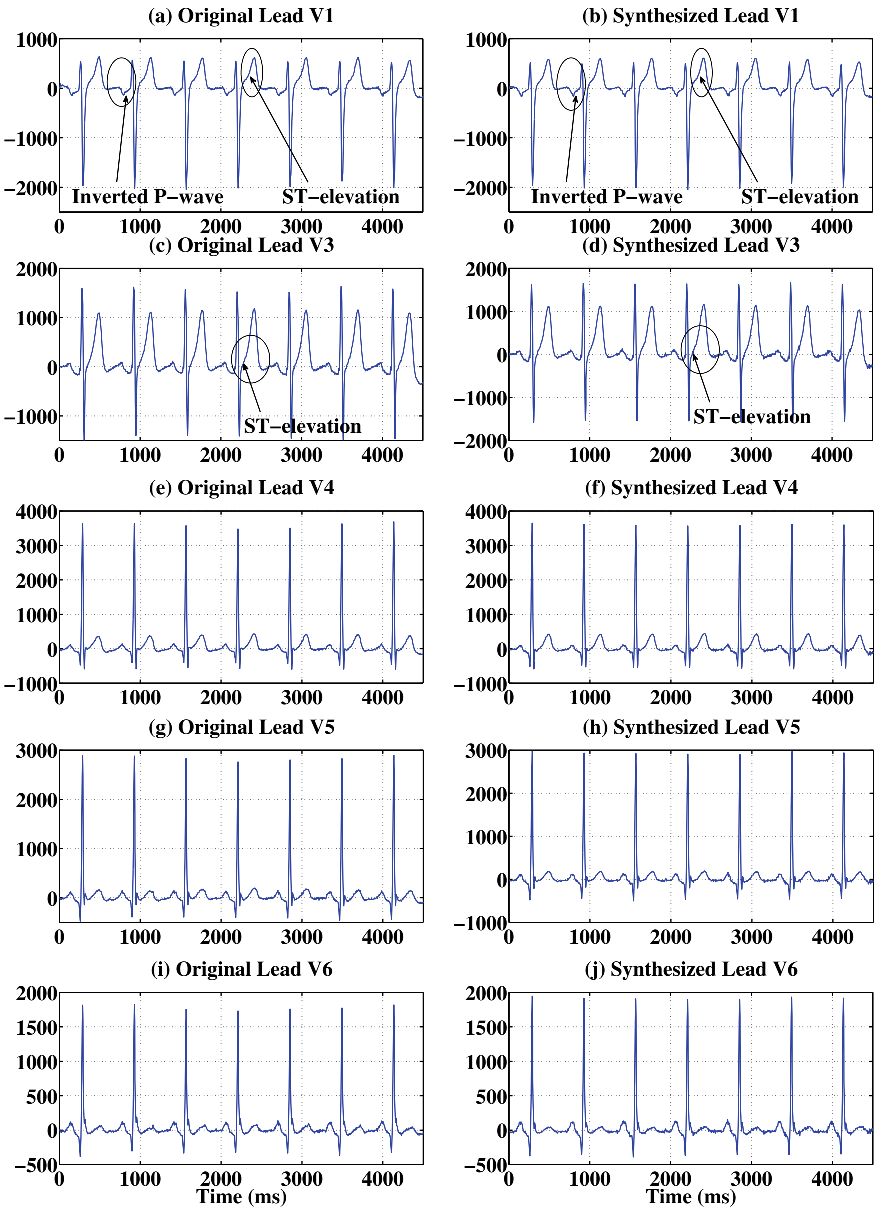


Fig. 5 Original and synthesized ECG leads V1, V3, V4, V5, and V6 for patient017/s0053lrem with lead subset I, II, and V2

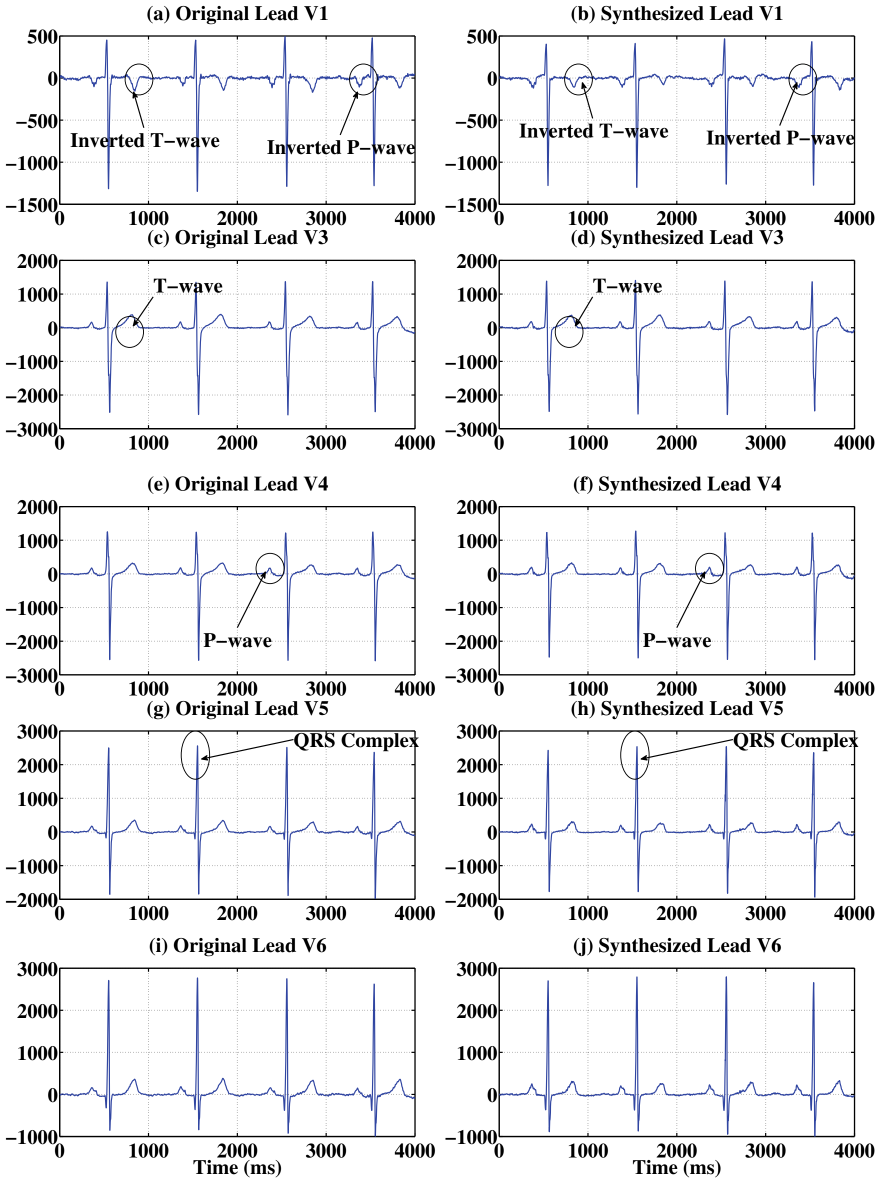


Fig. 6 Original and synthesized ECG leads V1, V3, V4, V5, and V6 for patient116/s0302lrem with lead subset I, II, and V2

Table 7 Lead-wise comparison of correlation coefficient r_x, RMSE and WEDD for three patients using proposed work and multiple-linear regression

Proposed method					Linear regression		
Patient	Synthesized leads	r_x	RMSE (μV)	WEDD (%)	r_x	RMSE (μV)	WEDD (%)
Patient05/ s0025lrem	V1	0.9953	15.9042	8.1906	0.9844	39.4965	19.8894
	V3	0.9994	10.0812	3.1411	0.9980	21.4827	5.7428
	V4	0.9976	15.8948	6.0139	0.9915	34.7511	12.2538
	V5	0.9916	17.5185	12.2192	0.9692	35.4554	23.4984
	V6	0.9887	16.5959	12.5050	0.9714	27.6278	18.4566
	Average	0.9945	15.1984	8.4140	0.9829	31.7627	15.9682
Patient017/ s0053lrem	V1	0.9978	13.6852	6.1971	0.9973	17.8618	6.5827
	V3	0.9970	18.6719	7.6258	0.9838	43.3142	15.9871
	V4	0.9958	24.9326	8.3213	0.9868	48.1503	15.2798
	V5	0.9963	18.3138	7.7466	0.9870	39.2150	16.4531
	V6	0.9845	14.4751	9.7929	0.9915	22.1686	15.7976
	Average	0.9963	18.0157	7.9367	0.9892	34.1420	14.0201
Patient116/ s0302lrem	V1	0.9866	12.4739	13.1879	0.9855	13.1224	13.6671
	V3	0.9981	9.2856	4.5976	0.9933	22.5160	10.3564
	V4	0.9959	14.3710	8.1658	0.9462	52.1667	28.1794
	V5	0.9952	16.9995	8.2221	0.9247	65.1181	32.4243
	V6	0.9844	17.3362	8.6486	0.9561	47.7656	24.4488
	Average	0.9942	14.0932	8.5644	0.9611	40.1377	21.8150

multiple linear regression, it can be concluded that, proposed model is a better model than the existing models.

5 Conclusion

A patient-specific method for synthesizing 12-lead ECG from reduced lead set I, II and V2 is proposed by combining principal component analysis and artificial neural network. Performance evaluation is carried out over publicly available PTB database. Proposed method utilize patient-specific transform which is more accurate than generic transform. Once calculated, the transformation parameter can be used over time assuming that it is time invariant. The proposed method show great similarity between original and synthesized ECG without loss of significant clinical information. Also, the performance evaluation criteria shows that the proposed method is superior to the existing models. The small variation in **W** with time is due to many factors like body posture, physical work, day to day ECG variation, respiration etc. A method to develop an adaptable linear orthonormal transform matrix can be considered which may give even better results.

References

1. Malmivuo, J., Plonsey, R.: Bioelectromagnetism: Principles and Applications of Bioelectric and Biomagnetic Fields, 1st edn. Oxford University Press, New York (1995)
2. Nelwan, S.P., Kors, J.A., Meij, S.H., van Bemmel, J.H., Simoons, M.L.: Reconstruction of the 12-lead electrocardiogram from reduced lead sets. J. Electrocardiol. **37**(1), 11–18 (2004)
3. Scherer, J., Nicklas, J.: Synthesis of the 12 lead electrocardiogram from a 3 lead semi-orthogonal subset using patient-specific linear transformation arrays. In: Computers in Cardiology. Proceedings, pp. 449–451, September 1988
4. Tomasic, I., Trobec, R.: Electrocardiographic systems with reduced numbers of leads - synthesis of the 12-lead ECG. IEEE Rev. Biomed. Eng. **7**, 126–142 (2014)
5. Dower, G.E.: A lead synthesizer for the frank system to simulate the standard 12-lead electrocardiogram. J. Electrocardiol. **1**(1), 101–116 (1968)
6. Wei, D.: Deriving the 12-lead electrocardiogram from four standard leads using information redundancy in the 12-lead system. Int. J. Bioelectromagn. **4**, 127–128 (2002)
7. Man, S.C., Maan, A.C., Kim, E., Draisma, H.H., Schalij, M.J., van der Wall, E.E., Swenne, C.A.: Reconstruction of standard 12-lead electrocardiograms from 12-lead electrocardiograms recorded with the Mason-Likar electrode configuration. J. Electrocardiol. **41**(3), 211–219 (2008)
8. Atoui, H., Fayn, J., Rubel, P.: A novel neural-network model for deriving standard 12-lead ECGs from serial three-lead ECGs: application to self-care. IEEE Trans. Inf. Technol. Biomed. **14**(3), 883–890 (2010)
9. Tsouri, G., Ostertag, M.: Patient-specific 12-lead ECG reconstruction from sparse electrodes using independent component analysis. IEEE J. Biomed. Health Inform. **18**(2), 476–482 (2014)
10. Bousseljot, R., Kreiseler, D., Schnabel, A.: Nutzung der ekg-signaldatenbank cardiodat der ptb über das internet. Biomed. Tech. **40**(1), 317–318 (1995)
11. Goldberger, A.L., Amaral, L.A.N., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E.: PhysioBank, physioToolkit, and physioNet: components of a new research resource for complex physiologic signals. Circulation **101**(23), e215–e220 (2000)
12. De Chazal, P., O'Dwyer, M., Reilly, R.: Automatic classification of heartbeats using ECG morphology and heartbeat interval features. IEEE Trans. Biomed. Eng. **51**(7), 1196–1206 (2004)
13. Castells, F., Laguna, P., Sornmo, L., Bollmann, A., Roig, J.M.: Principal component analysis in ECG signal processing. EURASIP J. Adv. Signal Process. **2007**(1), 074580 (2007)
14. Levkov, C.: Orthogonal electrocardiogram derived from the limb and chest electrodes of the conventional 12-lead system. Med. Biol. Eng. Comput. **25**(2), 155–164 (1987)
15. Manikandan, M.S., Dandapat, S.: Wavelet energy based diagnostic distortion measure for ECG. Biomed. Signal Process. Control **2**(2), 80–96 (2007)
16. Atoui, H., Fayn, J., Rubel, P.: A neural network approach for patient-specific 12-lead ECG synthesis in patient monitoring environments. In: Computers in Cardiology, pp. 161–164, September 2004
17. Feild, D.Q., Zhou, S.H., Helfenbein, E.D., Gregg, R.E., Lindauer, J.M.: Technical challenges and future directions in lead reconstruction for reduced-lead systems. J. Electrocardiol. **41**(6), 466–473 (2008)

Pentacene-Based Hybrid Photodetectors: A Review

Saumya Tripathi, Ashish Raman, and Ramesh Kumar Sunkaria

Abstract This paper presents a review which focuses on the combination of organic and inorganic materials in Photodetectors for opto-electronic applications. Photodetectors are the primary component of most optoelectronic devices, since they are the interface between the optical and electrical energy of any optoelectronic system. Both organic and inorganic materials have been explored for the fabrication of hybrid photodetectors. Such hybrid photodetectors have shown excellent performance by merging the advantages of both classes of materials such as high carrier mobility, flexibility, ease of fabrication, greater spectral response etc. Further, the effect of Pentacene as an organic material in the photodetector structure is reviewed. Finally, the shortcomings of Pentacene and inorganic semiconductor combinations in terms of performance and practical application are discussed, as well as the future prospects.

Keywords Photodetectors · Pentacene · Responsivity · Hybrid photodetectors

1 Introduction

Photodetector is widely used and one of the key components in optoelectronic integrated circuits, optical communication and sensing field due to high responsivity, fast response, etc. [1]. In the recent decade, research in the field of photodetectors gained impetus with John N. Shive's work on phototransistors and photodiodes in Bell Telephone Laboratories [2]. From there onwards, intensive research work began on devices that detect and convert light signals into electrical signals. Several configurations/structures such as nanowires, nanoparticles, p-i-n diode, p–n junction, homojunction, heterojunction, CMOS inverter, etc. using different kinds of materials including organic and inorganic materials have evolved for the fabrication of photodetectors [4–8]. However, these photodetectors have the limitation in low-light detection with better pulse height resolution and poor temperature stability [5].

S. Tripathi (✉) · A. Raman · R. K. Sunkaria
Department of Electronics and Communication, Dr. B. R. Ambedkar National Institute of Technology, Jalandhar, India
e-mail: saumyat.vl.19@nitj.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_53

Recently, researchers used the hybrid photodetectors to improve the detection of low light and temperature stability [6–8]. In case of hybrid photodetectors, two different materials come together to form a junction at which the photodetection activity takes place. Heterojunction structures are commonly used for hybrid photodetectors due to high-mobility and high speed [6]. Based on their advantage and clear inference drawn from literature on the photodetectors, they can be applied in applications like environmental monitoring, image-based processing operations, health monitoring, space applications, defense applications etc. [6–13]. Photo detecting devices possess the capability to convert light energy into electrical energy, and hence they can also be instrumental as a clean energy source [14, 15].

Since the discovery of photo detecting devices, a multitude of materials, both organic and inorganic, have been studied and used in their fabrication. Inorganic semiconductors have properties like high carrier mobility, lower exciton binding energies, are fairly stable silicon is one of the most widely used inorganic materials [4, 16–19]. In addition to that, ZnO [2, 3], Ge [4, 17, 19], CuO [20, 21], WSe_2 [5, 22, 23], MoS_2 [24, 25] etc. have also been explored. In the category of organic materials, Pentacene [6, 7], fullerene [26], P3HT [8, 27] etc. have been used. Organic semiconductors have a very wide range of absorption, and hence high absorption coefficients. Due to weak Van der Waals forces, the processing and fabrication becomes easier. Their versatile nature allows for their functionality to be tuned according to their application and purpose [15]. The flexible nature of organic semiconductors have made them the primary choice for use in flexible or wearable electronic devices these days. However, they also have certain disadvantages, like mechanical instability, low carrier mobilities and an improper alignment of organic molecules which results in lesser response speed and reduction in generation of carriers [8, 11, 16]. On the other hand, inorganic semiconductors are made of strong covalent bonds, which gives them mechanical stability. They also have higher carrier mobilities and rate of carrier generation, along with low exciton binding energy. In spite of all these, they come with their own disadvantages like complex fabrication process, material inflexibility and brittleness [28, 29].

This inconsistency in the electrical and behavioural properties of both the materials has led researchers to come up with the idea of combining both the materials in devices. Lately, there has been a steadily rising trend of hybrid devices having both organic and inorganic materials [30–34]. These hybrid devices can combine the advantages offered by both classes of materials and help make up for the disadvantages. This amalgamation of organic and inorganic material can result in devices that have high carrier mobility, flexibility, lighter weight and simpler processing methods [32]. The combination also provides opportunity for bandwidth tailoring of the photo detecting devices [30]. For a wide readership, the photodetector with hybrid structure in recent years is described in depth in this paper.

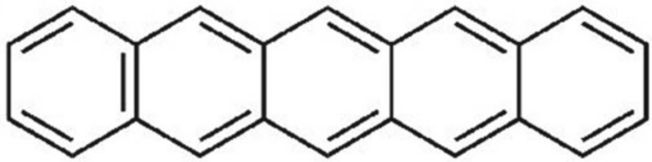


Fig. 1 Pentacene structure

1.1 Advantages of Pentacene

In this section, the focus is on organic material, i.e., Pentacene-based hybrid photodetectors. Pentacene is a very widely used p-type semiconductor, with a triclinic structure. It constitutes five linearly connected benzene rings which form a planar organic molecule [14], as shown in Fig. 1.

Pentacene can easily form films or discontinuous patches or islands and be integrated on to the photodetector devices [11], which means that it has an easy coating behavior. The hole mobility of Pentacene is considerably high (>5 V − 1 s − 1 cm^2), which makes it an excellent choice for the hole transport layer in electronic devices. It has a good photoresponsivity and can give external quantum efficiency much higher than 100% in photo detecting devices, such as the ones in [11, 13] and [10]. Pentacene with lower thickness offers high photo transparency [7]. Another very favorable feature of this material is its long exciton diffusion length, which enhances its spectral range and improves the efficiency of the device [11, 13, 14, 28]. So when Pentacene is fused with inorganic materials to form the hybrid photodetectors, it enhances the overall range of absorption of the device due to its intrinsic high absorption coefficient. The fabrication of devices when using Pentacene becomes quite convenient due to its easy solubility in organic solvents, such as chloroform and di-chloro benzene [14]. This ease in fabrication helps in making the manufacturing process more cost effective, which in turn makes it viable for commercial use.

1.2 Structures

The most commonly used structure out of all the papers referred to here is the heterojunction. It is a very basic structure that consists of two dissimilar semiconductors having different band gaps, but firmly matching lattice parameters, fused together. Most of the heterojunctions are fabricated using the spin-coating method to deposit the films on the substrate [7, 9, 10, 12, 13, 30, 35]. The ease of fabrication is a major plus point with this structure.

Thin-film transistor (TFT) structures have also been explored for fabrication of photodetectors [11, 33]. A TFT is a three-terminal device that is formed by depositing a layer of a semiconductor material, metallic contacts and a dielectric layer on top of an insulating substrate. They can be fabricated using the thermal evaporation and spin coating methods. TFTs have emerged as a favorable structure because of its capacity to self-amplify the photocurrent of the device through the controlled gate voltage [36]. They are light-weight devices that are easy to integrate in other circuits. The TFT structures proposed by Healin Im et al. [11] using WSe_2 and Pentacene, by Xi Xiao et al. [36] using Ga_2O_3 and CdO in the years 2021 and 2020 respectively have both reported very high efficiencies (>1000%) in the visible region.

Complementary metal–oxide–semiconductor (CMOS) technology-based photo detecting structures have also been researched upon in the last few years [6] [37–39]. They are mostly used in areas where there is a low power requirement. Their low costs and compatibility due to easy system integration has made them a suitable choice. The CMOS converts the optical signals received as input into amplified voltage signals, which in turn enhances the detectivity of the device [6]. The CMOS photodetectors have found abundant use in image sensing applications in the last few years, replacing charge coupling devices (CCDs) [38].

Vertical field effect transistor (VFET) structures have also been explored by Dan Yang et al. [34] and Muhammad Sulaman et al. [40] in their reported works. Lateral field-effect transistors can also be used to make photodetectors, but they have certain shortcomings such as a low sensitivity to optical signals, less responsivity, degraded operating speed and inferior mobility. VFETs have been reported to overcome these defects due to their structural supremacy. VFETs, due to their short channel lengths, have the ability to produce high output current at very low input voltages [41]. A comparison of planar FET and VFET was done by Dan Yang et. al. [34], and it was reported that the VFET structure had much better sensitivity and response to optical energy, with a better light selectivity in the visible region at low voltages.

1.3 Parameters

The output parameter values of an optoelectronic device help to analyze its behavior and also determine the applications of the device. For photodetectors, the primary performance parameters are external quantum efficiency (EQE), responsivity, spectral range and speed of response.

The spectral response range specifies the range of the wavelength that the device can detect and respond to. Photodetectors show different electrical responses to different wavelengths. That response determines the applications that they are to be used in. Almost all the photodetectors reviewed in the table below cover the spectral range from ultraviolet (UV) to near infrared (NIR). For instance, UV photodetectors are used in flame-sensing applications, biological and chemical sensing, ozone detection etc. [41, 42]. The visible range photodetectors can be used in image sensing applications, photography, obstacle sensing etc. [6, 11, 13]. NIR photodetectors have been used in military applications, optical fiber communications, satellite communication etc. [10, 43].

Responsivity gives us a measure of the efficiency with which a photodetector can convert the optical input received by it to electrical output. The formula for responsivity [2] is:

$$R_s = \frac{I_{light} - I_{dark}}{P_{input}} \tag{1}$$

where I_{light} and I_{dark} are the light and dark currents and P_{input} is the input power value. Higher responsivity at a given wavelength means that the photodetector operates best in that wavelength range.

Dark current is the current that is generated in the device when no light is falling on it. For an ideal device, the dark current should be minimum. The devices reported in [10–12, 33] give dark current of approximately 10^{-8} A, whereas the photodetector in [13] reports a dark current value of 10^{-9} A.

EQE is also a very important figure of merit for a photodetector. EQE of a device gives the number of carriers generated per each incident photon falling on it. It is the measure of the conversion efficiency from photons to charges. The formula for EQE [2] is:

$$EQE = \left(\frac{hc}{\lambda q}\right) R_S \tag{2}$$

With the emerging new technologies and fabrication methods, the EQE of devices has been reported to be greater than 1000%, such as those reported in [10, 11, 13, 31].

Response time is another parameter of photodetectors, which is used to determine how rapidly the device responds to a change in the input signal. It measures the response of the device in the time domain. It comprises of both rise time and fall time, where rise time is the time required for the output signal to change from 10 to 90% of its maximum fixed value and fall time is the time it takes to go from 90 to 10% [29]. Finally, performance parameters and design composition schemes of photodetectors structure including Pentacene are shown in Table 1.

1.4 Existing Challenges and Prospects

Despite the fact that Pentacene-based photodetectors have been developed and perfected through many years of research and work, there are still significant flaws that limit their application. The low kinetic stability of Pentacene due to its easy oxidation can pose a problem for Pentacene based devices. It also has low solubility in most of the organic solvents. The number of organic solvents in which Pentacene can completely dissolve is limited. These limitations can make the processing and fabrication complicated. However, a number of techniques have been developed to overcome these limitations of Pentacene [44, 45].

Table 1 Summary of the different hybrid photodetectors using pentacene based on composition and structure and their performance parameters

Composition and structure	Parameters						
	Year	Ref	Spectral response range (nm)	Responsivity (A/W)	Dark current	External quantum efficiency	Rise time/Decay time
WSe_2/Pentacene (TFT)	2021	[11]	UV–Vis	19.313 (at 638 nm) 1.16 (at 852 nm)	~4.28 × 10^{-7}A (at 638 nm) ~3.23 × 10^{-8}A (852 nm)	3753% (at 638 nm) 168.82% (at 852 nm)	0.34 s/0.30 s
ZnO/Pentacene/PEDOT:PSS (Heterojunction)	2021	[13]	UV–Vis	25.51 (at 311 nm) 0.36 (at 565 nm)	~10^{-9}A	10,171% (at 311 nm) 79.65% (at 565 nm)	
Al/Pentacene/CuO/PEDOT: PSS/ITO (Heterojunction)	2021	[10]	UV–Vis-NIR	35.25 (at 380 nm) 8.74 (at 600 nm) 2.75 (at 980 nm)	~5.67 × 10^{-8}A	11,700% (at 380 nm) 1900% (at 600 nm) 375% (at 980 nm)	12.16 µs/26.71 µs
Pentacene/MoS_2 (Hybrid CMOS inverter)	2020	[6]	Vis	0.14		27.42% (at 633 nm)	1.1 s/7.7 s
Au/Pentacene/n-Si Au/Au-NPs/Pentacene /n-Si Au/Ag-NPs/Pentacene/n-Si (Heterojunction)	2020	[7]	UV–Vis	0.0412 0.18871 0.47982		7.86% (at 650 nm) 36% 91.53%	766.7 ms/79.12 ms 84.7 ms/81.35 ms 219.9 ms/80.89 ms

(continued)

Table 1 (continued)

Composition and structure	Parameters						
	Year	Ref	Spectral response range (nm)	Responsivity (A/W)	Dark current	External quantum efficiency	Rise time/Decay time
3D TI Bi_2Te_3/Pentacene (Thin film heterojunction)	2019	[31]	450–3500	14.89		2840%	2.98 ms/3.57 ms (at 450 nm) 3.04 ms/4.41 ms, (at 780 nm)
Cd_3As_2 Thin Film/Pentacene (Heterojunction)	2018	[9]	Vis–NIR	0.03615 (at 650 nm) 1.55×10^{-3} (at 10600 nm)		7.29% (at 650 nm)	
ITO/Pentacene/Al/Pentacene/Au (VFET)	2015	[34]	Vis	0.188 (at 350 nm)		66.60% (at 350 nm)	
ZnO Nanoparticle/Pentacene (Heterojunction)	2014	[12]	380–750	17.3	~1.02×10^{-8} A	3.3% (at 650 nm)	
Single-crystal Si nanomembrane/ Pentacene (Heterojunction)	2013	[30]	400–800 (Vis)	0.7 (at 633 nm)	~3.64×10^{-7} A	21.9% (at 633 nm)	
PbSe QDs into poly-N-vinyl carbazole/ Pentacene (thin-film transistor)	2006	[33]	NIR		~10^{-8} A	8%	
Au/Pentacene/Si (Heterojunction)	2003	[35]	Vis	2.6 (at 650 nm)		496%	35 ns/40 ns

2 Conclusion

An overview of the latest trends and performances shown by Pentacene based organic–inorganic photodetectors has been done. It has been realized that by tuning the materials used in combination with Pentacene, an enhancement in parameters such as EQE, responsivity etc. can be achieved. The tunability of Pentacene can also enhance the spectral range and give way for broadband photodetectors, which can detect a wider range of wavelength. If in the coming future, focus will be placed on improving the bandwidth of the photodetector, then we can possibly save on manufacturing costs by using one device for multiple functions. More extensive research needs to be done on the different combinations of Pentacene and inorganic semiconductors that can be fabricated. Finally, the objective of this paper is to let the readers better understand the development of organic–inorganic photodetectors during the last several years and to predict future trends.

References

1. Oehme, M., Werner, J., Kasper, E., Jutzi, M., Berroth, M.: High bandwidth Ge p-i-n photodetector integrated on Si. Appl. Phys. Lett. **89**(7), 071117 (2006)
2. John: Invention history of phototransistors and photodiodes. Circ. Today **30** (2013)
3. Soci, C., Zhang, A., Xiang, B., Dayeh, S.A., Aplin, D.P., Park, J., Bao, X.Y., Lo, Y.H., Wang, D.: ZnO nanowire UV photodetectors with high internal gain. Nano Lett. **7**(4), 1003–1009 (2007). https://doi.org/10.1021/nl070111x
4. Liu, K., Sakurai, M., Liao, M., Aono, M.: Giant improvement of the performance of ZnO nanowire photodetectors by Au nanoparticles. J. Phys. Chem. C **114**(46), 19835–19839 (2010)
5. Tan, C., Wang, H., Zhu, X., Gao, W., Li, H., Chen, J., Li, G., Chen, L., Xu, J., Hu, X., Li, L., Zhai, T.: A Self-powered photovoltaic photodetector based on a lateral WSe2-WSe2 homojunction. ACS Appl. Mater. Interfaces. **12**(40), 44934–44942 (2020)
6. Liu, F., Zhang, Y., Wang, J., Chen, Y., Wang, L., Wang, G., Dong, J., Jiang, C.: MoS_2/Pentacene hybrid complementary inverter based photodetector with amplified voltage-output. Nanotechnology **32**(1), 015203 (2020). https://doi.org/10.1088/1361-6528/abb9da
7. El-Mahalawy, A.M., Wassel, A.R.: Enhancement of organic/inorganic hybrid photodetector based on Pentacene/n-Si by surface plasmonic effect of gold and silver nanoparticles: a comparative study. Opt. Laser Technol. **131**, 106395 (2020)
8. Wang, X., Song, W., Liu, B., Chen, G., Chen, D., Zhou, C., Shen, G.: High-performance organic-inorganic hybrid photodetectors based on P3HT:CdSe nanowire heterojunctions on rigid and flexible substrates. Adv. Funct. Mater. **23**, 1202–1209 (2013). https://doi.org/10.1002/adfm.201201786
9. Yang, M., Wang, J., Han, J., Ling, J., Ji, C., Kong, X., Liu, X., Huang, Z., Gou, J., Liu, Z., Xiu, F., Jiang, Y.: Enhanced performance of wideband room temperature photodetector based on Cd3As2 thin film/pentacene heterojunction. ACS Photonics **5**(8), 3438–3445 (2018)
10. Srivastava, A., Singh, R., Jit, S., Tripathi, S.: Pentacene and CuO nanocomposite based self-powered broadband photodetector. IEEE Electron Device Lett. **42**(6), 875–878 (2021)
11. Im, H., Kim, S.: Effectively enhanced broadband phototransistors based on multilayer WSe2/Pentacene. Adv. Electron. Mater. **7**(5), 2100003 (2021)
12. Yuan, Z., Fu, M., Ren, Y.: Optoelectronic properties of ZnO nanoparticle/pentacene heterojunction photodiode. J. Electron. Mater. **43**, 3270–3275 (2014)

13. Srivastava, A., Jit, S., Tripathi, S.: High-performance Pentacene/ZnO UV-visible photodetector using solution method. IEEE Trans. Electron. Dev. **68**(7), 3439–3445 (2021)
14. Kacus, H., Yilmaz, M., Kocyigit, A., Incekara, U., Aydogan, S.: Optoelectronic properties of Co/Pentacene/Si MIS heterojunction photodiode. Physica B **597**, 412408 (2020)
15. Ocaya, R.O., Al-Ghamdi, A., Kwadwo Mensah-Darkwa, R.K., Gupta, W.F., Yakuphanoglu, F.: Organic photodetector with coumarin-adjustable photocurrent. Synth. Met. **213**, 65–72 (2016)
16. Ortega, M., Santana, G., Morales-Acevedo, A.: Optoelectronic properties of CdO/Si photodetectors. Solid-State Electron. **44**(10), 1765–1769 (2000)
17. Byrd, M.J., Timurdogan, E., Su, Z., Poulton, C.V., Fahrenkopf, N.M., Leake, G., Watts, M.R.: Mode-evolution-based coupler for high saturation power Ge-on-Si photodetectors. Opt. Lett. **42**(4), 851–854 (2017)
18. Feng, B., Zhu, J., Xu, C., Wan, J., Gan, Z., Lu, B., Chen, Y.: All-Si photodetectors with a resonant cavity for near-infrared polarimetric detection. Nanoscale Res. Lett. **14**(1), 1–9 (2019)
19. Bandaru, P.R., Sahni, S., Yablonovitch, E., Liu, J., Kim, H.J., Xie, Y.H.: Fabrication and characterization of low temperature (<450 C) grown p-Ge/n-Si photodetectors for silicon based photonics. Mater. Sci. Eng., B **113**(1), 79–84 (2004)
20. Wang, S.B., Hsiao, C.H., Chang, S.J., Lam, K.T., Wen, K.H., Hung, S.C., Huang, B.R.: A CuO nanowire infrared photodetector. Sens. Actuators, A **171**(2), 207–211 (2011)
21. Hong, Q., Cao, Y., Xu, J., Lu, H., He, J., Sun, J.L.: Self-powered ultrafast broadband photodetector based on p–n heterojunctions of CuO/Si nanowire array. ACS Appl. Mater. Interfaces. **6**(23), 20887–20894 (2014)
22. Zheng, Z., Zhang, T., Yao, J., Zhang, Y., Xu, J., Yang, G.: Flexible, transparent and ultra-broadband photodetector based on large-area WSe2 film for wearable devices. Nanotechnology **27**(22), 225501 (2016)
23. Chen, J., Wang, Q., Sheng, Y., Cao, G., Yang, P., Shan, Y., Qiu, Z.J.: High-performance WSe2 photodetector based on a laser-induced p–n junction. ACS Appl. Mater. Interfaces. **11**(46), 43330–43336 (2019)
24. Xie, Y., Zhang, B., Wang, S., Wang, D., Wang, A., Wang, Z., Wang, J.: Ultrabroadband MoS2 photodetector with spectral response from 445 to 2717 nm. Adv. Mater. **29**(17), 1605972 (2017)
25. Wang, X., Wang, P., Wang, J., Hu, W., Zhou, X., Guo, N., Chu, J.: Ultrasensitive and broadband MoS2 photodetector driven by ferroelectrics. Adv. Mater. **27**(42), 6575–6581 (2015)
26. Tang, Q., Zhang, G., Jiang, B., Ji, D., Kong, H., Riehemann, K., Fuchs, H.: Self-assembled fullerene (C60)-Pentacene superstructures for photodetectors. SmartMat **2**(1), 109–118 (2021)
27. Wu, Y.L., Fukuda, K., Yokota, T., Someya, T.: A highly responsive organic image sensor based on a two-terminal organic photodetector with photomultiplication. Adv. Mater. **31**(43), 1903687 (2019)
28. Sajjad, M., Ruseckas, A., Samuel, I.: Enhancing exciton diffusion length provides new opportunities for organic photovoltaics. Matter **3**, 341–354 (2020)
29. Ren, H., Chen, J.D., Li, Y.Q., Tang, J.X.: Recent progress in organic photodetectors and their applications. Adv. Sci. **8**(1), 2002418 (2021)
30. Seo, J.H., Oh, T.Y., Park, J., Zhou, W., Ju, B.K., Ma, Z.: A multifunction heterojunction formed between pentacene and a single-crystal silicon nanomembrane. Adv. Func. Mater. **23**(27), 3398–3403 (2013)
31. Yang, M., Wang, J., Zhao, Y., He, L., Ji, C., Liu, X., Zhou, H., Wu, Z., Wang, X., Jiang, Y.: Three-dimensional topological insulator Bi2Te3/Organic thin film heterojunction photodetector with fast and wideband response from 450 to 3500 nanometers. ACS Nano **13**(1), 755–763 (2019)
32. Pei, K., Zhai, T.: Emerging 2D organic-inorganic heterojunctions. Cell Rep. Phys. Sci. **1**(8), 100166 (2020)
33. Roy, C.K., Kim, W., Sahoo, Y., Lee, K.S., Prasad, P.: Solution-processed Pentacene quantum-dot polymeric nanocomposite for infrared photodetection. Appl. Phys. Lett. **89**(5), 051109 (2006)
34. Yang, D., Zhang, L., Wang, H., Wang, Y., Li, Z., Song, T., Fu, C., Yang, S., Zou, B.: Pentacene-based photodetector in visible region with vertical field-effect transistor configuration. IEEE Photonics Technol. Lett. **27**(3), 233–236 (2015)

35. Kim, S., Choi, Y., Kim, K., Kim, J., Im, S.: Fabrication of p-Pentacene/n-Si organic photodiodes and characterization of their photoelectric properties. Appl. Phys. Lett. **82**(4), 639–641 (2003)
36. Xiao, X., Liang, L., Pei, Y., Yu, J., Duan, H., Chang, T.C., Cao, H.: Solution-processed amorphous Ga2O3:CdO TFT-type deep-UV photodetectors. Appl. Phys. Lett. **116**(19), 192102 (2020)
37. Titus, A.H., Cheung, M., Chodavarapu, V.: CMOS Photodetectors. Photodiodes - World Activities in 2011, Jeong-Woo Park, IntechOpen (2011)
38. Chen, Q., Chitnis, D., Walls, K., Drysdale, T., Collins, S., Cumming, D.: CMOS photodetectors integrated with plasmonic color filters. IEEE Photonics Technol. Lett. **24**(3), 197–199 (2012)
39. Feruglio, S., Lu, G.N., Garda, P., Vasilescu, G.: A review of the CMOS buried double junction (BDJ) photodetector and its applications. Sensors **8**(10), 6566–6594 (2008)
40. Sulaman, M., Song, Y., Yang, S., Li, M., Saleem, M.I., Chandraseakar, P.V., Jiang, Y., Tang, Y., Zou, B.: Ultra-sensitive solution-processed broadband photodetectors based on vertical field-effect transistor. Nanotechnology **31**(10), 105203 (2019)
41. Ben-Sasson, A., Tessler, N.: Unraveling the physics of vertical organic field effect transistors through nanoscale engineering of a self-assembled transparent electrode. Nano Lett. **12**(9), 4729–4733 (2012)
42. Omnès, F., Monroy, E., Muñoz, E., Reverchon, J.L.: Wide bandgap UV photodetectors: a short review of devices and applications. In: Gallium Nitride Materials and Devices II, vol. 6473, p. 64730E. International Society for Optics and Photonics (2007)
43. Maity, P., Singh, S.V., Biring, S., Pal, B.N., Ghosh, A.K.: Selective near-infrared (NIR) photodetectors fabricated with colloidal CdS: Co quantum dots. J. Mater. Chem. C **7**(25), 7725–7733 (2019)
44. Bhatia, R., Durgesh, W., Gargi, G., Gaur, J., Gupta, D.: Methodologies for the synthesis of Pentacene and its derivatives. J. Saudi Chem. Soc. **23**(7), 925–937 (2019)
45. Zhao, L., Kaiser, R.I., Lu, W., Ahmed, M., Evseev, M.M., Bashkirov, E.K., Azyazov, V.N., Tönshoff, C., Reicherter, F., Bettinger, H.F., Mebel, A.M.: A Free-radical prompted barrierless gas-phase synthesis of pentacene. Angew. Chem. **132**(28), 11430–11434 (2020)

Cochlear Acoustic Model that Improves the Speech Perception in Noise by Encoding TFS

Poluboina Venkateswarlu, Aparna Pulikala, and Arivudai Nambi Pitchaimuthu

Abstract People with cochlear implants accomplish good speech recognition scores in quiet. Temporal envelope (ENV) is encoded primarily in cochlear implant (CI), and it is sufficient for recognizing speech in quiet. However, temporal fine structures (TFS) are needed for better recognition of speech in noise. Some fine structure coding strategies tried to modulate temporal envelope with TFS. In such coding strategies, FS4 is one that tried to encode fine structures up 950 Hz. In this study, the performance of FS4 with speech recognition in noise was investigated by using acoustic simulation. The speech intelligibility of this study was conducted on five normal-hearing (NH) persons. This performance was compared with 16 channel sinewave vocoder and with the Full band TFS condition. The variance of these three conditions was analyzed using the SNR 50. These results indicate that the fine structure (FS4) coding (up to 1078 Hz Hz) has improved speech recognition in noise compared to the sinewave vocoder.

Keywords Cochlear implants · Temporal fine structures · FS4 · Vocoder simulation · Quick-SIN · SNR50

1 Introduction

The best option for the treatment of profound hearing impairment individuals is cochlear implants, in which the speech processor plays a significant role in the effectual extraction and delivery of the information from the input speech signal.

Most of the sound coding strategies of cochlear implant (CI), basically encode the envelope cue and discard the TFS. CIs have been successful in providing good speech recognition outcomes in quiet environments. Envelope cue is adequate for speech recognition in quiet. However, envelope cues alone are insufficient for speech

P. Venkateswarlu (✉) · A. Pulikala
National Institute of Technology Karnataka, Surathkal, Mangalore, India
e-mail: venki.187ec009@nitk.edu.in

A. N. Pitchaimuthu
Department of audiology and speech language pathology, KMC, Mangalore, Manipal Academy of Higher Education, Manipal, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_54

recognition in background noise [1], and also the TFS is crucial for better speech recognition in noise. This is because the temporal envelope does not effectively convey the essential cue for speech recognition in noise such as F0 [2] and interaural time difference [3].

Acoustic modeling studies have mentioned that providing an extra cue temporal fine structure (TFS) would enhance speech recognition in noise [4]. Hence, developing a sound coding strategy to code TFS along with an envelope would improve the speech recognition in noise ability of cochlear implantees. Some coding methods tried to encode both TFS and temporal envelope information by modulating the envelope component with the instantaneous frequency of the corresponding band signal in each processing channel [5]. Fine structure processing (FSP) is a commercially available sound coding strategy implemented by MED-EL in which especially tried to encode TFS. The pre-processed incoming speech signal is given to a 12 channel filter bank. A rectifier followed by a low pass filter extracts the temporal envelope for each channel. The TFS information extracted for the lower apical channels through positive zero crossings occurs in the analysis time frame [6]. Upper basal channels follow the continuous interleaved sampling algorithm. The sampling rate available to identify zero-crossings limit the correctness of TFS coding. In FSP, TFS is coded up to three lower apical channels. The TFS availability depends on the stimulation rate of individual channels, and the TFS frequency range covered up 495 Hz. The FSP speech perception in the noise was not appreciated statistically after twelve months of follow up [7]. The frequency listening threshold of CI users is around 1000 Hz Hz. The more recent TFS coding strategy is FS4, in which TFS frequency expands up 950 Hz. The individual sampling rate of channel for positive zero crossing detection is increased compared to FSP [8]. This study was conducted to observe the speech intelligibility of the FS4 (up to 1078 Hz Hz TFS) acoustic simulation model in a speech in noise. The performance of the FS4 (up to 1078 Hz Hz TFS) was compared with the sinewave vocoder and full band TFS by using acoustic simulation of CI.

The rest of the paper is organized as follows: In Sect. 2, we describe the details of the methodology. The experimental design and results are discussed in Sect. 3 and we draw some conclusions in Sect. 4.

2 Methodology

2.1 Signal Processing

The input speech was analyzed using 16 channel band-pass filters in between 80 and 7562 Hz Hz. The filter bank corner frequencies were decided by using the greenwood function [9].

$$D_{min} = log((F_l/165.4) + 1)/2.1 \tag{1}$$

Table 1 Frequency division of 16 band pass filters

Band number	Centre frequency (Hz)	Band number	Centre frequency (Hz)
1	109.52	9	1377.3
2	175.67	10	1748.5
3	257.73	11	2209
4	359.55	12	2780.3
5	485.85	13	3489
6	642.55	14	4368.4
7	836.95	15	5459.3
8	1078	16	6812

$$D_{max} = log((F_h/165.4) + 1)/2.1 \tag{2}$$

$$D_x = (D_{max} - D_{min})/bands \tag{3}$$

$$x = D_{min} : D_x : D_{max} \tag{4}$$

The band pass filter which passes frequencies in between [f(k)/(f_s/2)] and [f(k+1)/((f_s/2)], where k is the number of channels and f_s is sampling rate. The centre frequencies (f_c) of the 16 bands are derived from the following equation

$$f_c = f(k) + [(f(k+1) - f(k))/2] \tag{5}$$

The frequency division of 16 band pass filters is as shown in Table 1.

The speech signal within each frequency band is transformed into the corresponding analytical signal $X_a(t)$ using the Hilbert transform (HT), where $X_a(t)$ is

$$X_a(t) = x_r(t) + ix_i(t) \tag{6}$$

The speech envelope was computed from the Hilbert transform output by using an envelope detector, which consists of a half wave rectifier followed by a low-pass filter with a cut-off frequency 400 Hz. The extracted envelope from all sub-bands were modulated by the carrier signal. The carrier signal for sinewave vocoder is sinusoidal, which is extracted from each band center frequency (f_c).

$$C_{r1}(t) = cos(2\pi * f_c t) \tag{7}$$

where t is length of the band's.

In full band TFS and eight band TFS (FS4) conditions, the temporal fine structures were used as a carrier for modulating envelope. The temporal fine structures were extracted from the Hilbert bands,

$$C_{r2}(t) = \cos(\phi(t)) \tag{8}$$

where $\phi(t)$ is phase information of the Hilbert bands. The temporal envelope and the carrier signals were extracted from the input speech as shown in Fig. 2.

The modulated speech is the product of envelope (ENV) cue and carrier signal i.e.;

$$S_1(t) = ENV * C_{r1}(t) \tag{9}$$

where S_1(t) is modulated output 1.

$$S_2(t) = ENV * C_{r2}(t) \tag{10}$$

where S_2(t) is modulated output 2.

$$S_3(t) = ENV * C_{r1}(t) + ENV * C_{r2}(t) \tag{11}$$

where S_3(t) is modulated output in which TFS (C_{r2}(t)) was used as a carrier up to lower eight bands (1078 Hz) and for the rest of bands C_{r1}(t) was consider as carrier.

For the reconstruction of the original input speech, the modulated speech was processed through the synthesized filter bank. In sinewave vocoder, the modulated output 1 from all bands are added and the resultant synthesized output1 as shown in Fig. 1. Similarly, the synthesized output 2 represents Full band TFS, and synthesized output 3 represents fs4 (TFS up to 1078 Hz Hz).

The Quick speech in noise (Quick-SIN) test was conducted to evaluate the above acoustic simulation methods.

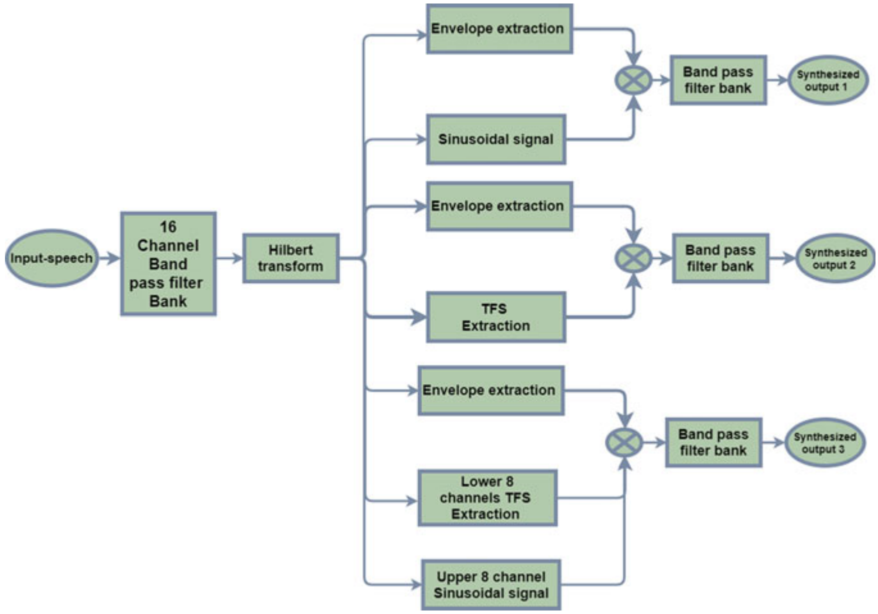


Fig. 1 The block diagram of the three signal processing methods

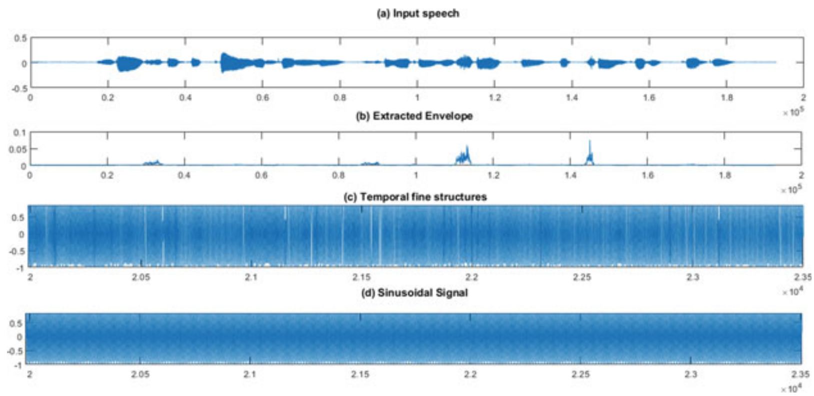


Fig. 2 Extracted envelope and TFS from input speech

2.2 *Quick Speech in Noise (QuickSIN)*

Speech-in-Noise tests are designed to mimic real-life circumstances. In each list, there are 7 sentences, and in each sentence, there are 5 keywords highlighted in bold. Thus based on the response of the subject, the score was given between 0 to 5. In this study, each sentence was presented to the listeners at different SNR levels. The variance of the three methods was measured using the SNR50.

2.3 Subjects

Five normal hearing (NH) persons (average age: 29.5 years; age ranges from 24 to 32 years) participated in the current study with their self-report, and screening audiometry for hearing of the listeners within 15 dBHL at octave frequencies range 250–8000 HZ.

2.4 Stimuli

The speech intelligibility test was performed to verify this simulation methods through QuickSIN standard Kannada sentences [10]. The QuickSIN test comprises two lists which consist of seven sentences in each list. For the perceptual test, two lists were used with six different combinations (2*3) of the three signal processing methods using MATLAB R2017a. To the target speech, a four-talker babble noise was added and the processed speech was given at 7 different SNR levels (+20 dB, +10 dB, +05 dB, 0, −05 dB, −10 dB).

3 Experimental Design and Results

Each sentence was chosen from two lists of the QuickSIN and processed under 6 test conditions. The participants were seated in the middle of a soundproof room. The speech stimuli were conferred to each listener at 65 dB SPL through the Sennheiser HD280pro headset. The sequence of presentation of two processed speech lists was varied for each subject. Participants were instructed to listen carefully to the target speech and repeat them in written form. If participants were unsure about the sentences, they were permitted to guess. Their responses were evaluated based on the number of correctly identified words in each sentence. A score of 1 was given to each correctly identified keyword. Finally, the total score was counted in three signal processing methods in each SNRs. The scores were converted to a form of proportion correct for fitting the psychometric function.

The proportion correct score of three signal processing methods was plotted on the Gaussian psychometric function as shown in Fig. 3. The SNR50 was measured with respect to the midpoint of the proportion correct, which represents the minimum SNR is required for 50% speech perception.

The yellow curve in the above Fig. 3. represents the Sinewave vocoder, the required SNR for 50% perception was 13.53 dB, when compared to TFS conditions, this was a high number. The Full band TFS represents the red curve, which required less SNR (−3.273 dB) because all bands envelopes were modulated with the original TFS. Finally, the required SNR for FS4 (TFS coded up to 1078 Hz) was 5.926 dB, which represents the blue curve in Fig. 3, which is less compared to the sinewave vocoder.

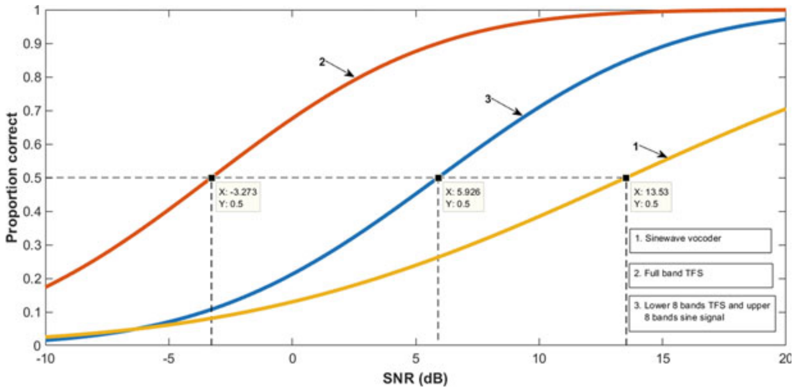


Fig. 3 Fitting the psychometric function for three methods

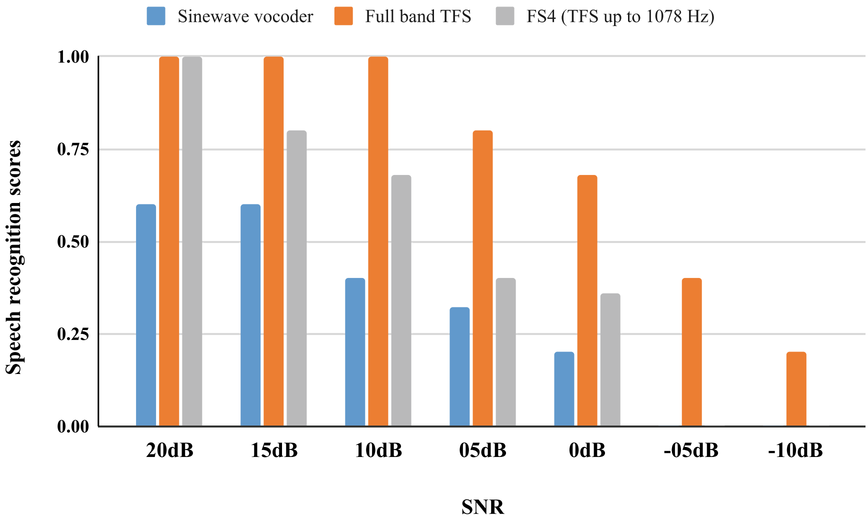


Fig. 4 Speech recognition score of the three methods

The listener's mean opinion score (MOS) of the three signal processing methods was represented in terms of speech recognition score vs seven different SNRs as shown in Fig. 4. In which full band TFS provides maximum score irrespective of 7 SNR levels compared to the rest of two methods. In sinewave vocoder, TFS was discarded and only the envelope was coded so the speech recognition score was poor in noise. Hence TFS is required for better speech recognition in noise. Finally, the TFS coding up to 1078 Hz was providing average speech recognition of sinewave vocoder and full band TFS at 15, 10, 05, 0 dB SNR levels.

4 Conclusion

This paper shows the speech perception in noise for three models namely, sinewave vocoder, full band TFS and FS4. In sinewave vocoder, the envelope is modulated with sinusoidal signal and in full band TFS model the envelope is modulated with TFS. In the FS4 (up to 1078 Hz) method, temporal envelopes of the lower eight bands were modulated with TFS and the upper band envelopes were modulated with a sinusoidal signal. The speech recognition score of the FS4 (TFS up to 1078 Hz) strategy gives a better response than the sinewave vocoder where as gives a significantly low performance than full band TFS in the case of speech recognition with noise. The full band TFS method was used here to observe the importance of the fine structure for speech recognition in noise, but it was not used in real-time, due to the frequency listening threshold of CI users. Hence FS4 coding up to 1078 Hz gives better perception in noise within the frequency listening threshold of CI users.

References

1. Stickney, G.S., Nie, K., Zeng, F.-G.: Contribution of frequency modulation to speech recognition in noise. J. Acoust. Soc. Am. **118**(4), 2412–20 (2005)
2. Manjunath, D., et al.: Contribution of amplitude modulations for concurrent stream segregation: a probe through object-related negativity. Speech Lang. Hear. **18**(1), 55–61 (2015)
3. Wouters, J., McDermott, H.J., Francart, T.: Sound coding in cochlear implants: from electric pulses to hearing. IEEE Signal Process. Mag. **32**(2), 67–80 (2015)
4. Nambi, P.M.A., Mahajan, Y., Francis, N., Bhat, J.S.: Temporal fine structure mediated recognition of speech in the presence of multitalker babble. J. Acoust. Soc. Am. **140**(4), EL296–EL301 (2016)
5. Nie, Kaibao, Stickney, Ginger, Zeng, Fan-Gang.: Encoding frequency modulation to improve cochlear implant performance in noise. IEEE Trans. Biomed. Eng. **52**(1), 64–73 (2004)
6. Zierhofer, C.: Electrical nerve stimulation based on channel specific sampling sequences, pp. US 6,594,525 B1 (2003)
7. Riss, D., et al.: A new fine structure speech coding strategy: speech perception at a reduced number of channels. Otol. Neurotol. **29**(6), 784–788 (2008)
8. Riss, D., et al.: FS4, FS4-p, and FSP: a 4-month crossover study of 3 fine structure sound-coding strategies. Ear Hear. **35**(6), e272–e281 (2014)
9. Greenwood, Donald D.: A cochlear frequency-position function for several species-29 years later. J. Acoust. Soc. Am. **87**(6), 2592–2605 (1990)
10. Geetha, C., et al.: Development and standardisation of the sentence identification test in the Kannada language. J. Hear. Sci. **4**(01), 18–26 (2014)

Simulation and Performance Analysis of Gate Source Overlapped Dual Material Double Gate Tunnel Field Effect Transistor with SiO_2/HfO_2 Stacked-Gate Oxide Structure

Kavindra Kumar Kavi, Saumya Tripathi, and R. A. Mishra

Abstract In this article, we proposed Gate-Source Overlapping Dual Material Double Gate Tunnel Field Effect Transistor (GSDMDGFET) with SiO_2/HfO_2 Stacked Gate Oxide structure. The proposed device is used to effectively reduce the ambipolar current, improve the device performance with higher device drain current, lower leakage current, lower subthreshold swing (SS) below 60 mV/dec and also improves RF performance. Starting from a conventional TFET structure with double-gate, the optimization of device described as a result of the research presented in this paper. Device performance has steadily increased. in terms of the RF parameters, higher I_{ON}, lower I_{OFF}, higher I_{ON}/I_{OFF} led optimized GSDMDGTFET with SiO_2/HfO_2 stacked gate oxide structure are analyzed, e.g. transconductance gm, gate to source capacitance C_{GS} and gate to drain capacitance C_{GD}. We have optimized the tunnel FET component with the industry standard SILVACO tool. We show that the suggested GSDMDGTFET with SiO_2/HfO_2 stacked gate oxide structure reduces the device's ambipolar behavior at negative voltages without compromising its performance.

Keywords Stacked gate oxide · Ambipolar current · RF parameters

1 Introduction

At nanoscale levels, sub-threshold leakage is especially detrimental to device functioning. As a result, researchers all around the world are interested in the TFET technology, which uses band-to-band tunneling to replace diffusion-based minority carrier injection in ordinary MOSFETs [1–3]. Band to band tunneling between the conduction band of source to the valence band channel regions in TFET causes the higher ON- current ION. When compared to traditional CMOS technology, this leads by significant improvement in subthreshold swing and leakage current. According

K. K. Kavi (✉) · S. Tripathi · R. A. Mishra
Motilal Nehru National Institute of Technology, Allahabad, India
e-mail: kkavi@mnnit.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_55

to a new study, TFETs might be a potential choice for low power and RF applications [4, 5]. However, a key disadvantage of traditional TFETs is that, in addition to having lower IOFF, they also have lower ION levels than normal MOSFETs. Several TFET architectures (both dual and tri gate) have been described previously for achieving greater ION and lower IOFF. Unlike traditional TFETs, new research has demonstrated that TFETs with greater ION can be made in a variety of ways.

Ambipolar current is second major limitation of TFET, as a result, it is less effective when employed as a MOSFET complement. The enhanced subthreshold swing (SS) of TFETs over MOSFETs is the major reason for researching them as an alternative to MOSFETs. While MOSFETs have a limit of 60 mV/dec on SS, TFETs can have an SS below 60 mV/dec because they use a different tunneling mechanism as compare to MOSFETs, which use thermionic injection [5]. So to overcome these two limitations two major changes have been done in the conventional TFET. In place of Single metal gate, a double metal gate has been used which completely overcome the ambipolar current at negative bias [6]. One more method is available to reduce the ambipolar current is gate drain overlapping, but we suggested this method because it is having less leakage current Apart from this, to achieve higher ION, in place of SiO2 high dielectric gate oxide layer HfO2 have been used [7, 8]. To further increase the on current, improves the vertical tunneling and get a better ION/IOFF ratio using gate overlapping on source side.

2 Proposed GSDMDGTFET Structure

Figure 1 shows the suggested GSDMDGTFET with SiO2/HfO2 stacked gate oxide structure. We optimized the device by looking at the effects of modifications in channel thickness, gate material selection, semiconductor material doping concentration, and overlapping of the gate and source regions. The electrical properties of the proposed device were simulated using the industry-standard SILVACO tool, with a channel length of 45 nm, a thickness of 12 nm, Si channel doping Nch = $10^{15}/\text{cm}^3$, Ge source doping NS = $10^{20}/\text{cm}^3$ and Si drain doping ND = $10^{19}/\text{cm}^3$. The thickness of high-κ oxide (HfO2) is 2 nm and thickness of low-κ oxide (SiO2) is 1 nm. As the dielectric oxide thickness drops, the drain current ID increases. However, because of the low dielectric thicknesses might result in direct gate tunneling, by keeping dielectric thickness below 1.5 nm to ensure reliability. Now we will look at the work function of a TFET with a double material gate. The tunnel gate is located closer to the source end, and its work-function mostly controls the ON-state current. The auxiliary gate is located closer to the drain, and its work-function primarily controls the OFF-state current. Higher ION, better SS and lower IOFF, are attained when the tunneling gate having a lower work function as compare to auxiliary gate [9, 10]. Electrons tunnel from valence band of the source to the intrinsic body's conduction band, then drift towards the drain end through a drift– diffusion process [11].

Tunneling takes place in the presence of a strong electric field (source–body interface), when local band bending decreases the energy barrier's breadth. Tunnel

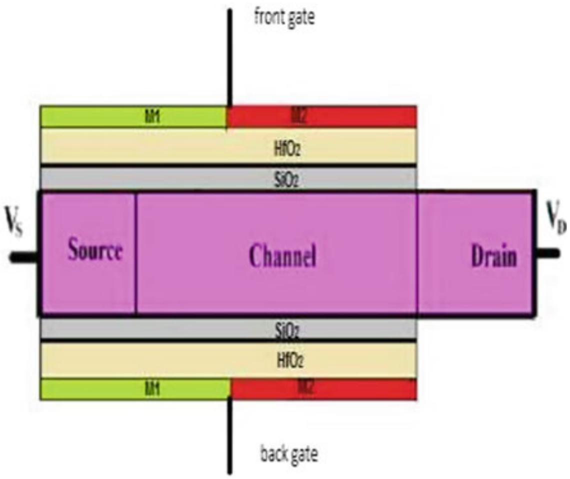


Fig. 1 Device structure of GSDMDGTFET with SiO2/HfO2 stacked gate oxide

gate work-function and auxiliary gate work- function are 4.5 eV (e.g. Cr) and 4.7 eV (e.g., Fe) respectively. These are the metal work functions utilized in this study.

3 Results and Discussions

This article contains a simulation which were performed using the SILVACO TCAD tool. The simulation deck includes the traditional Kane, SRH, and Auger recombination models [12].

3.1 Suppression of Ambipolar Nature

We have compared the proposed GSDMDGTFET device with GSSMDGTFET and observed that if we use dual metal gate at the place of single metal gate then at negative bias we can completely suppress the ambipolar current (Fig. 2).

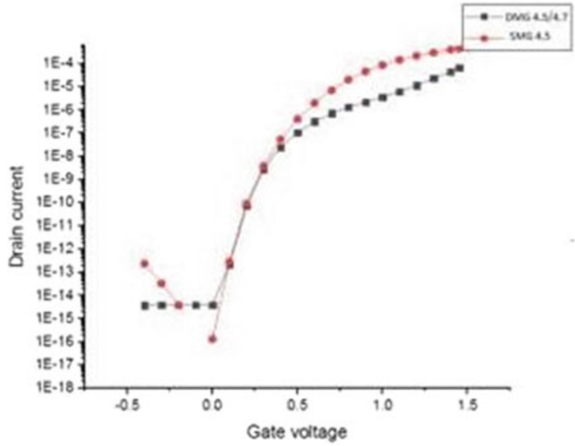


Fig. 2 Comparison of Single material gate with double material gate

3.2 Gate Overlapped Source vs Without Gate Over-Lapped Source

We did the analysis of the proposed device with gate overlapped source and without gate overlapped source and we found that with gate overlapped source drain current has been increased significantly with no change in leakage current (Fig. 3).

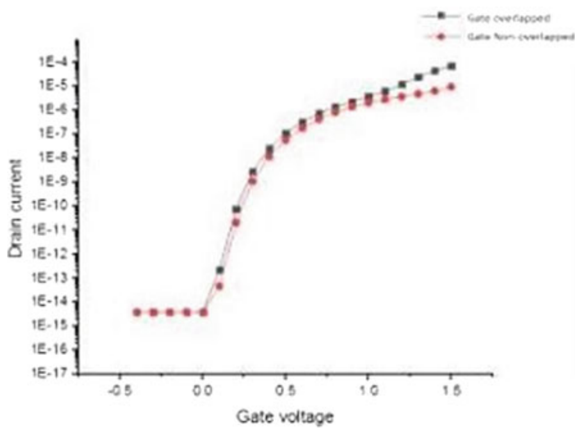


Fig. 3 Comparison of gate overlapped source with without gate overlapped source

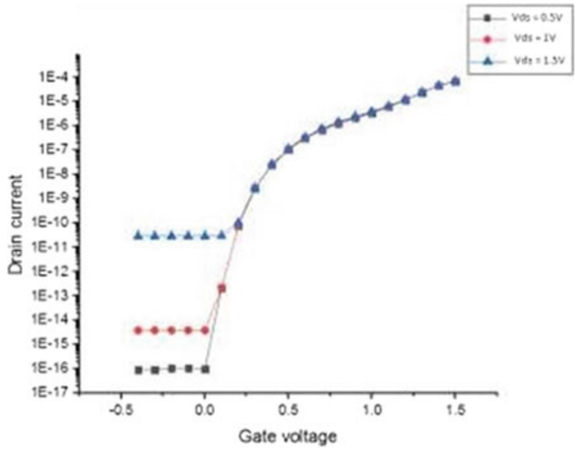


Fig. 4 Transfer characteristics with different values of V_{ds}

3.3 Variation of Transfer Characteristics by Changing the Vds

when we increase the Vds value, then leakage current is also increased. We can see the changes in transfer characteristics in Fig. 4. It occurs due to the effective charge inversions.

3.4 Change in Thickness

When we changed the thickness of the source, drain and channel region we observed that gate control has been increased by increasing the thickness, so we are getting more ON-current, more leakage current also. Although ON to OFF-current ratio remains almost similar (Fig. 5).

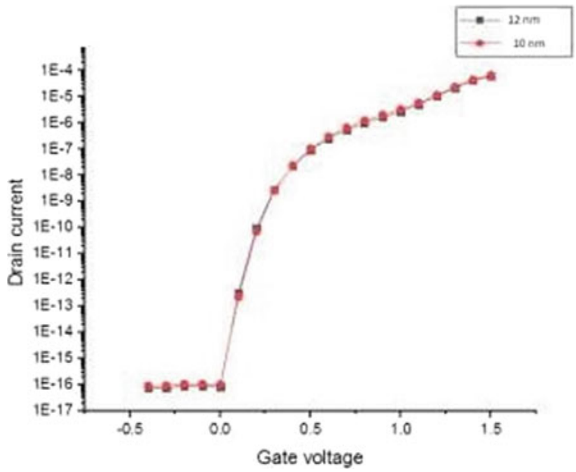


Fig. 5 Comparison of Si region thickness

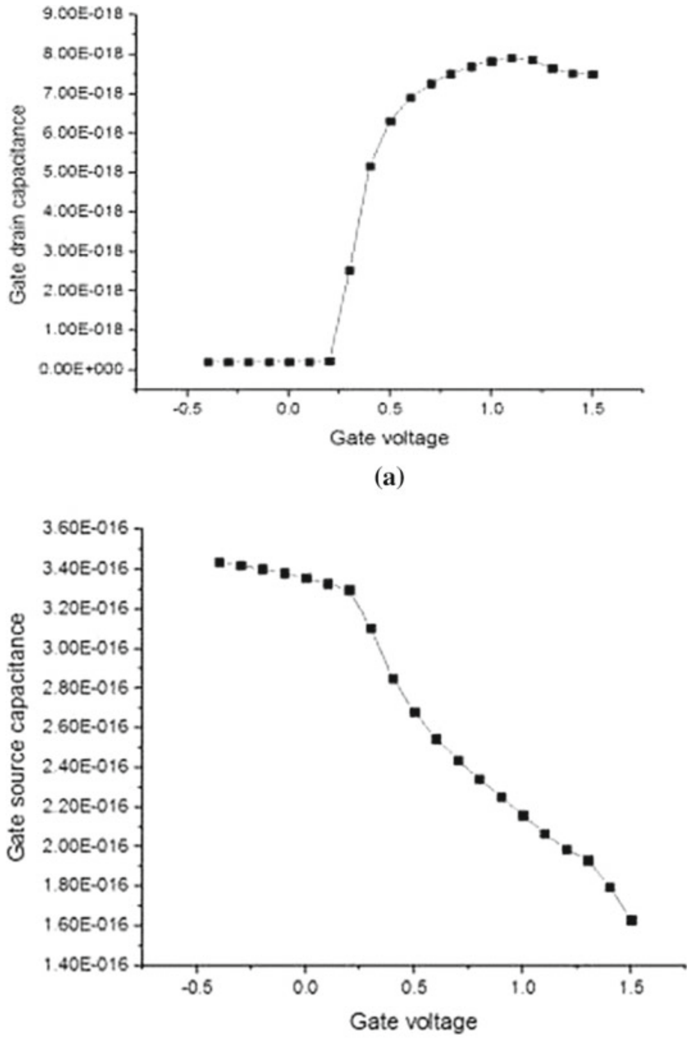
4 RF Analysis

The proposed GSDMDGTFET's RF analysis has been discussed in terms of capacitances and transconductance g_m. There are two type of gate parasitic capacitance that is gate to drain capacitance C_{GD} and gate to source capacitance C_{GS} discussed in this section. The variation in different capacitance and transconductance are shown in graph that is varying with respect to the gate voltage.

The C_{GD} and C_{GS} are shown in Figs. 6(a) and (b). Up to V_{GS} 0 V, the C_{GD} of the proposed GSDMDGTFET is nearly constant. C_{GD} grows fast with V_{GS} for $V_{GS} >$ 0 V, less gate to drain potential barrier, hence stronger capacitive coupling between gate and drain, while C_{GS} drops with V_{GS}, due to higher gate to source barrier.

As we can see from the below Fig. 7 as we increase the V_{GS} simultaneously g_m gets increased, but after a certain value of gate voltage it will decrease due to mobility degradation.

Here we have compared the reference device [9] with the proposed device. As we can see that there is a tradeoff between ON-current and subthreshold swing. As we decreased the V_{ds}, ON-current increased and also getting less subthreshold swing compared to reference device [9] (Table 1).



(a)

(b)

Fig. 6 **a** C_{GD} variation with respect to V_{GS} **b** C_{GS} variation with respect to V_{GS}

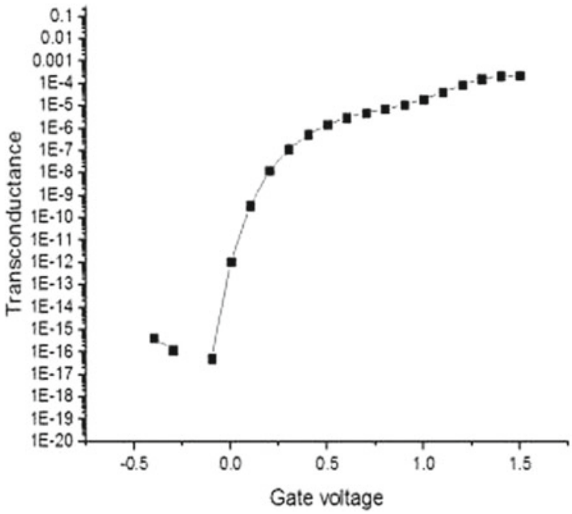


Fig. 7 Transconductance with respect to V_{GS}

Table 1 Comparison of the proposed device with reference device

Devices	L_G (nm)	V_{DS} (V)	I_{ON} (mA)	I_{OFF} (fA)	Subthreshold swing (mV/dec)
Reference device [1]	45	1	1.04	270	32
Proposed device without source gate overlap	45	1	4.45	1650	39
Proposed device with source gate overlap	45	0.8	5.2	120	30

5 Conclusion

The ION/IOFF ratio, ambipolar effect, and SS (subthreshold swing) of the proposed structure have been improved by optimizing the work functions of the auxiliary and tunneling gates. The proposed device has tunnel gate work function 4.5 eV and auxiliary gate work function 4.7 eV. The device's RF analysis- Because the trans conductance is greater, the gain will be higher. On-current $I_{ON} = 10$ mA, OFF-current $I_{OFF} = 84.18$ fA, and subthreshold slope SS = 28 mV/dec were reported in the proposed device.

References

1. Luo, Z., Wang, H., An, N., Zhu, Z.: A tunnel dielectric-based tunnel FET. IEEE Electron Device Lett. **36**, 966–968 (2015)
2. Colinge, J.-P., Lee, C.-W., Afzalian, A.: Nanowire transistors without junctions. Nat. Nanotechnol. **5**, 225229 (2010)
3. Saurabh, S., Kumar, M.J.: Fundamentals of Tunnel Field-Effect Transistors. CRC Press, Boca Raton (2016)
4. Kumar, S., Goel, E., Singh, K., Singh, B., Singh, P.K., Baral, K., Jit, S.: "2-D analytical modeling of the electrical characteristics of dual-material double gate TFETs with a SiO2/HfO2 stacked gate-oxide structure. IEEE Trans. Electron Dev. **64**, 960–968 (2017)
5. Gholizadeh, M., Hosseini, S.E.: A 2-D analytical model for double- gate tunnel FETs. IEEE Trans. Electron Dev. **61**(5), 1494–1508 (2014)
6. Kumar, M.J., Pandey, P., Vishnoi, R.: Tunnel Field-Effect Transistors (TFET) Modelling and Simulations. Wiley, Hoboken (2014)
7. Lu, H., Seabaugh, A.: Tunnel field-effect transistors: state-of-the-art. IEEE J. Electron Dev. Soc. **2**(4), 44–49 (2014)
8. Hraziia, A.V., Amara, C.V.: An analysis on the ambipolar current in Si double-gate tunnel FETs. Solid State Electron. **70**, 67–72 (2012)
9. Yadav, R., Dan, S.S., Vidhyadharan, S., Hariprasad, S.: Suppression of Ambipolar Behaviour and simulationeous improvement in RF performance of Gate overlapped tunnel field effect transistor (GOTFET) devices. Springer, Heidelberg (2020)
10. Karbalaei, M., Dideban, D., Heidari, H.: A simulation study of the influence of a high-**k** insulator and source stack on the performance of a double-gate tunnel FET. J. Comput. Electron. (2020)
11. Garg, S., Saurabh, S.: Suppression of ambipolar current in tunnel FETs using drain pocket: proposal and analysis. Superlattices Microstruct. **113**, 261–270 (2018)
12. Atlas user's manual

Investigation and Optimization of Dielectric Resonator MIMO Antenna Using Machine Learning Approach

Pinku Ranjan, Harshit Gupta, Anand Sharma, Swati Yadav, and Milka Potrebićc

Abstract A new tool, machine learning optimization (MLAO), has recently been introduced to accelerate antenna and array design. In order to improve fast response prediction, machine learning (ML) techniques, including GPR, SVM, and ANN, were used to develop antenna models. The Multiple-Input Multiple-Output (MIMO) acquainted with the 4G versatile system should be effectively tried and may work in the current recurrence of correspondence. By HFSS, the planning scheme for a dual MIMO broadband antenna has been introduced. Then the info of frequency and S parameters are collected and check out to seek out a far better algorithm that will be used for antenna designing purposes and parameter optimization. A spread of modeling techniques of AI like support vector regression, genetic algorithm has been used.

Keywords GPR · SVM · MLAO · MIMO · AI

P. Ranjan (✉) · H. Gupta
Atal Bihari Vajpayee -Indian Institute of Information Technology and Management, Gwalior, India
e-mail: pinkuranjan@iiitm.ac.in

H. Gupta
e-mail: mtcn_202005@iiitm.ac.in

A. Sharma
Motilal Nehru National Institute of Technology Allahabad, Prayagraj, India
e-mail: anandsharma@mnnit.ac.in

S. Yadav
College of Engineering Roorkee, Roorkee, Uttarakhand, India

M. Potrebićc
School of Electrical Engineering, University of Belgrade, Belgrade, Serbia
e-mail: milka_potrebic@etf.rs

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_56

1 Introduction

MIMO is a wireless technology that transfers more data simultaneously using numerous transmitters and receivers. Particularly, Multiple-Input Multiple-Output is an antenna-based technology that uses multiple transmitting and receiving to increase the capacity of RF telecommunication system. All 802.11n wireless products support MIMO. Also, the development of mobile communications is supported continually by adequate devices. An antenna is a tool that is linked to the wireless communication system. An antenna with a simple design that can work on multiple bands has been installed for wireless communication systems. Thoroughly researched and extensively developed antennas have several benefits, such as small size, low-cost manufacturing, and are of simple design. In the field of the development of Communication Systems, newer technology is always being produced.

The technology must be supported (high data rate) by high-speed data service and has a quality of service (QOS) that is reliable. Wireless technology that enables users to move around is the current emerging technology. Because of MIMO antennas, antennas that are used today can boost system capacity. There are antenna arrays on MIMO wireless systems at the terminals of both the transmitter and receiver. In MIMO, a remarkable improvement in data output is achieved using multiple data transmitting and receiving by multiple elements at the transmitter and receiver ends, and this is also the reason why wireless systems of fourth-generation (4G) support multimedia applications and videos more in real-time compared to the older generation. A natural radio-wave phenomenon, called multipath, is used by MIMO technology, letting transmitted data bounce off walls, ceilings, and other objects with multipath, reaching the receiving antenna several times at different angles and slightly different times. In past few years many MIMO antennas are designed for the wireless technology [1–5].

1.1 Machine Learning Based Modelling

Machine learning is usually described as the method where computers are programmed in order to optimize a performance criterion by analyzing examples from the past. Models are usually learned based on data collected from specific examples. The various machine-learning techniques comprised in this context are artificial neural networks, support vector regression, genetic algorithm, evolution computing, etc. [6, 7].

Artificial Neural Network

ANN is incredible assets used primarily for microwave gadget portrayal, demonstrating, and plan improvement. Up till now, its viability has been demonstrated in numerous applications depicted in the literary work.

Three-Layer Neural Network

Three layers of neurons are included, i.e., information layer, concealed layer, and yield layer. The number of neurons relies upon common sense applications. Specifically, the quantity of neurons in info layer is normally dictated by mathematical boundaries of the gadget and activity frequencies, considering that the quantity of yield layer neurons is dictated by the dissipating boundaries**.** Mathematical boundaries and the concerned yields are the amplitudes of return misfortune $|S_{11}|$ and the addition misfortune $|S_{12}|$. A neuron work should be chosen ahead of time. They are beyond what twenty competitors can be chosen as this capacity.

Preparing an ANN model is about getting familiarized with the connections among the input and output information. Furthermore, the neural organization loads can be registered. Another arrangement of information is then used to test the exhibition (e.g., forecast precision) of the ANN model. As a rule, the ANN demonstrating execution is influenced by the learning calculation and the organization structure.

Support Vector Regression

Support vector machine is mainly applied for classification. Many approaches are built on SVR due to its prediction ability. An experiment proves that RBF kernel-based SVR provide good results.

The SVR technique outperforms ANN to a certain extent by seeing of its supremacy, as described below.

- ANN uses risk minimization while SVR uses Structural risk minimization, so instead of local–global minima can occur in ANN, while in SVR, this is not possible.
- The number of parameters in SVR is much less than that in ANN.
- In general, SVM uses only four parameters whereas ANN has several layers each with a number of neurons that needs to be determined before training.
- ANN can have over fitting and under fitting. So cross-validation is used to remove this bias.

Genetic Algorithms

The genetic algorithm (GA) is one of the widely used global optimization algorithms to optimize the shape and size of the antenna, by antenna designers in order to achieve better antenna performance. By optimizing the bandwidth, multi frequency, directionality, gain, size etc., GA was used to improve the performance of MIMO antennas. The idea of GA to perform functional optimization, includes the usage of search strategies for optimization built on the Darwinian notion of natural selection and evolution. Optimization of the rectangular patch antenna parameters by the genetic algorithm and its performance analysis was carried out using CST software. The surface current in the developed antenna path meanders and the electrical antenna length is therefore increased. It implies that the total surface is reduced to a large amount by the antenna with the same resonance frequency. A huge amount of the overall surface is reduced. In this case, through a radiation pattern of typical

MIMO antenna, patch size reduction of up to 82% has been achieved compared to a conventional one resonating at the same frequency.

2 Methodology

For the ML optimization, a DR-based MIMO antenna is employed. The antenna design consists of the plus shape radiator printed on one side of the FR-4 substrate and on the opposite side two microstrip lines were etched. Further, on the plus-shaped radiator, an alumina-based ring DRA is placed. Figure 1 shows the planning of the MIMO antenna using the HFSS software.

For this study, the support vector regression technique is used. This method is additionally compared with the other regression methods. Comparison of the SVR with other algorithms in antenna designing requires the S-parameters and frequency at a specific height to seek out the best accuracy score. As MIMO antenna has both S_{11} and S_{12} parameters. Therefore, there are two scores for every algorithm.

2.1 Support Vector Regression

SVR's objective function is to minimize the coefficients, more specifically the coefficient vector, not the MSE. Thus, data between S-parameters (dB) and frequency (GHz) is taken. The results of Support Vector Regression are shown in Figs. 2 and 3.

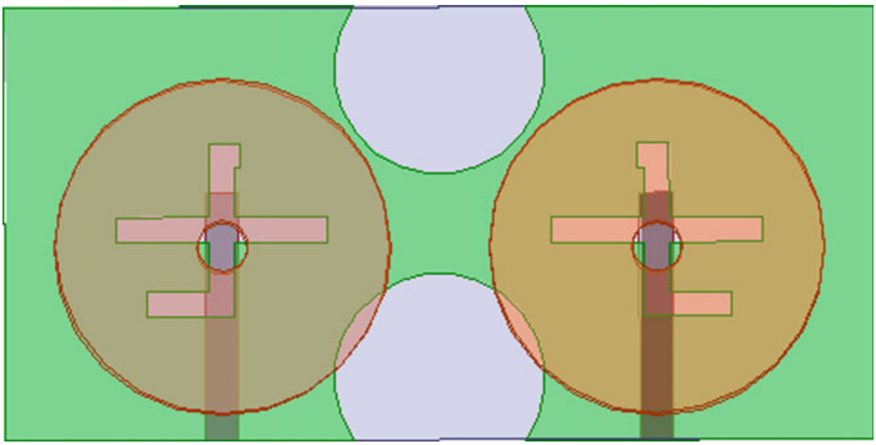


Fig. 1 Design of MIMO antenna in HFSS

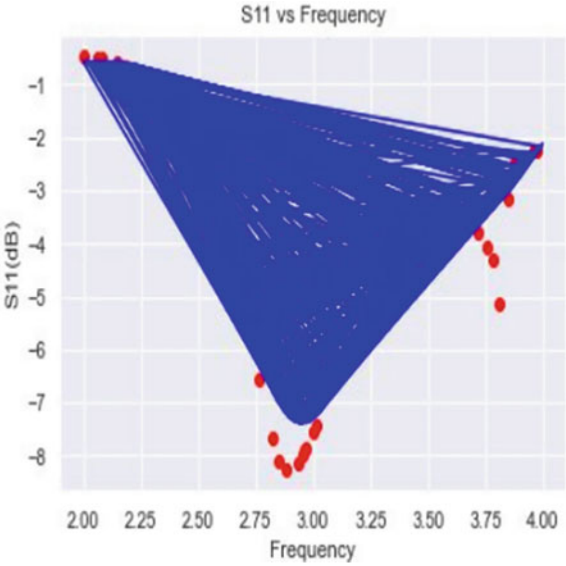


Fig. 2 SVR results for S_{11} parameter

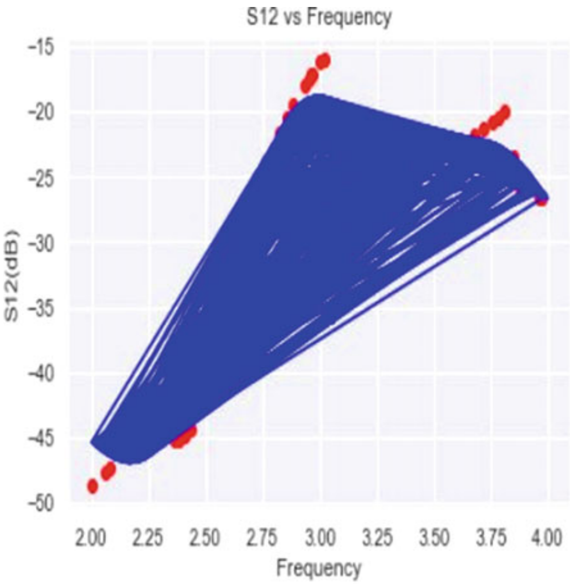


Fig. 3 SVR result for S_{12} parameter

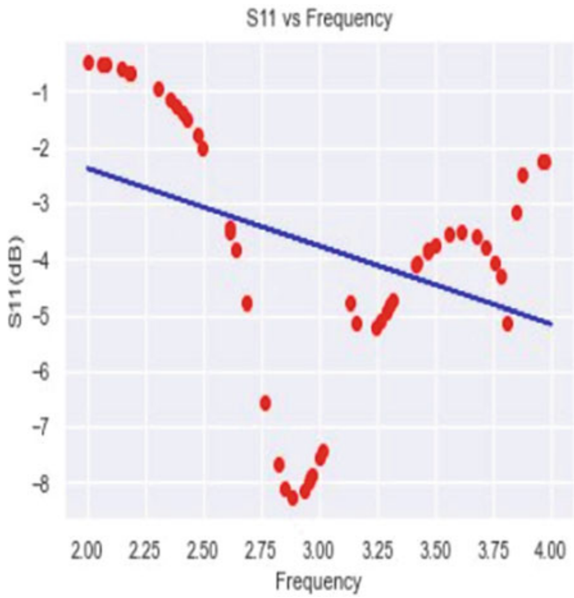


Fig. 4 SLR results for S_{11} parameter

2.2 Simple Linear Regression

SLR is a method based on statistics which encapsulates and study relationships between 2 continuous (quantitative) variables. Here also data between S-parameters (dB) and frequency (GHz) is taken. Figures 4 and 5 shows the result of simple linear regression.

- As the independent variable, one variable, denoted by x, is considered.
- The dependent variable is considered to be the other variable, denoted by y.

2.3 Decision Tree Regression

Decision tree regression is one of the most commonly used practical methods to supervise ML. DTR can solve both regression and classification problems with more practical application of the latter, and with three types of nodes, it is a tree-structured classifier.

Here data between S-parameters (dB) and frequency (GHz) is collected and the results are shown in the Figs. 6 and 7.

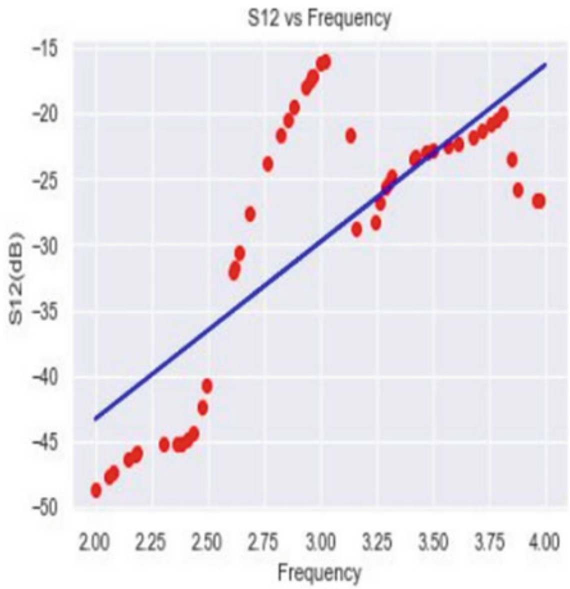


Fig. 5 SLR results for S_{12} parameter

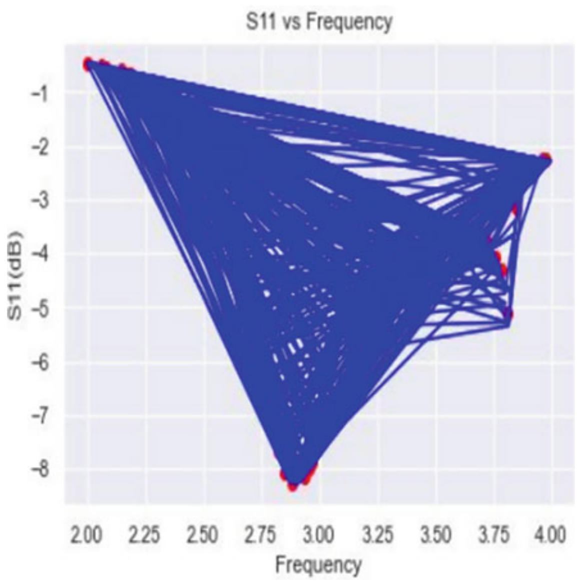


Fig. 6 DTR results for S_{11} parameter

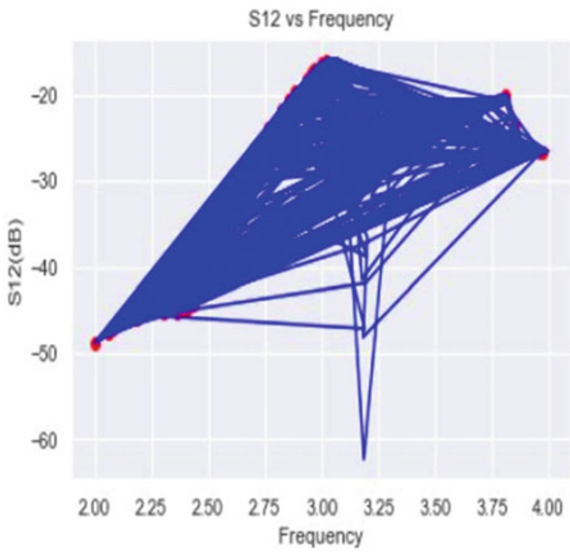


Fig. 7 DTR results for S_{12} parameter

2.4 Random Forest Regression

Random forest is a supervised learning algorithm which utilizes the classification and regression method of ensemble learning. It is a technique for bagging and not one for boosting. The trees in random forests run parallel to each other. While constructing the trees, there is no interaction between these trees. It works by constructing a multitude of decision trees at training time and producing the class i.e., the class mode (classification) or mean prediction (regression) of the individual trees and the data between S-parameters (dB) and frequency (GHz) is procured. The results achieved through this method are shown in Figs. 8 and 9.

3 Analysis of Results

3.1 Results

The results achieved from the varied regression models are collected all together to find out the best results. Tables 1 and 2 shows the results of $S_{11/22}$ and $S_{12/21}$ parameters for various models.

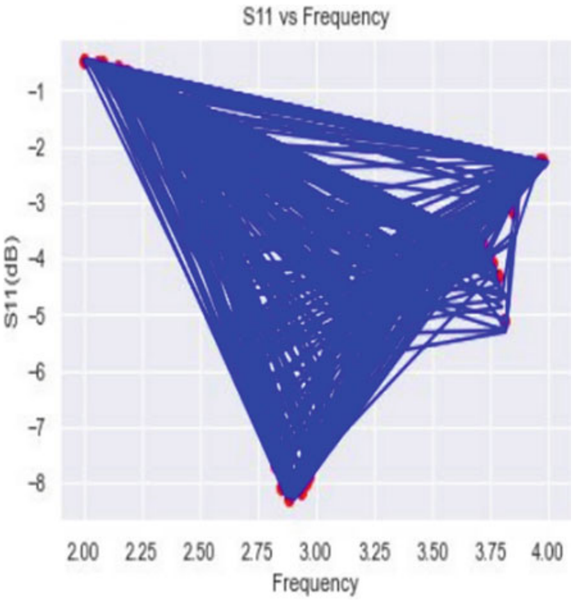


Fig. 8 RFR result for S_{11} parameter

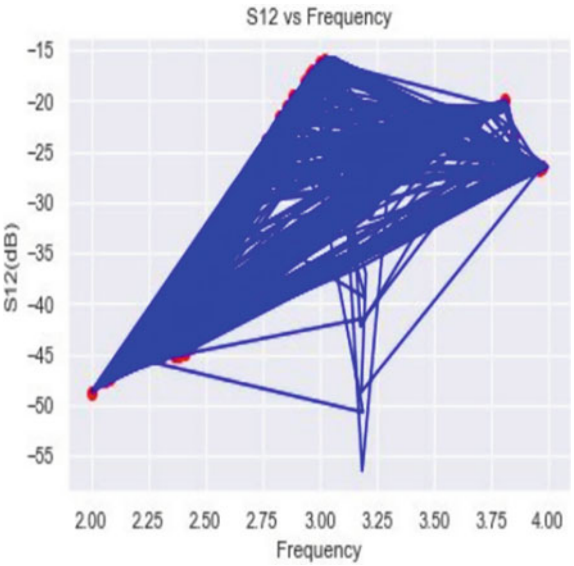


Fig. 9 RFR result for S_{12} parameter

Table 1 Results for S_{11}/S_{22} parameter

Models	SLR	SVR	DTR	RFR
MAE	5.3142	1.5259	0.15466	0.069396
RMSE	6.85513	2.2422	0.283448	0.117663
R_2 score	0.56506	0.9534	0.99925	0.9997

Table 2 Results for S_{12}/S_{21} parameter

Models	SLR	SVR	DTR	RFR
MAE	1.8015	0.33635	0.035841	0.0170925
RMSE	2.2378	0.53937	0.053838	0.0256657
R_2 score	0.10645	0.94808	0.99948	0.99988

3.2 Analysis

Comparison of four algorithms, SLR, SVR, DTR and RFR, on the given dataset is done. From the above tables, it can be clearly seen that random forest regression gives the best accurate results for both S_{11} and S_{12} (same results for S_{21} and S_{22} because $S_{11} = S_{22}$ and S_{12}=S_{21}. In early research, only SVR was the best option for this sort of research, but now Random Forest Regression is the better option which gives an accuracy of nearly 0.999 on comparing the accuracy in 3 terms MAE, RMSE and R2 score. This also provides the measure of how well the information has been trained and tested.

4 Conclusion

Due to the complexity of the EM simulators used for the design of antenna, their simulation time can change from a few seconds to few days depending on the size, frequency, and computational power of the device. Optimization is necessary optimized to maximize output and reduce scale at the same time. Techniques for rapid optimization are extremely important and desirable. An alternative to EM simulation is the application of the machine learning algorithms, which are typically used as an alternative to the general optimization techniques, such as genetic algorithm, particle optimization of the swarm and simulated annealing. For instance, the relationship between the size of the patch antenna and its resonant frequency is trained by neural networks. The calculation of the parameters for the model of antenna using ML techniques was performed. This paper combines the Machine learning (ML) with antenna architecture simulation. The inclusion of artificial intelligence (AI) may provide promising results in the field of antenna design.

References

1. Bauch, G., Alexiou, A.: MIMO technologies for the wireless future. In: IEEE 19th international symposium on personal, indoor and mobile radio communications, pp. 1–6 (2008)
2. Alexa, F., Bardeanu, B., Vatau, D.: MIMO antenna system for LTE. In: 36th international conference on telecommunications and signal processing (TSP), pp. 294–298 (2013)
3. Kumar, J.: Compact MIMO antenna. Microw. Opt. Technol. Lett. **58**, 1294–1298 (2016)
4. Zhang, J.Y., Zhang, F., Tian, W.P., Luo, Y.L.: ACS-fed UWB-MIMO antenna with shared radiator. Electron. Lett. **51**, 1301–1302 (2015)
5. Raj Kumar, R.V.S., Krishna, R., Kushwaha, N.: Design of a compact mimo/diversity antenna for UWB applications with modified TH-like structure. Microw Opt. Technol. Lett. **58**(5), 1181–1187 (2016)
6. Gunes, F., Tokan, N.T., Gurgen, F.: A knowledge-based support vector synthesis of the transmission lines for use in microwave integrated circuits. Expert Syst. Appl. **37**(10), 3302–3309 (2012)
7. Wu, Q., Cao, Y., Wang, H., Hong, W.: Machine-learning-assisted optimization and its application to antenna designs: opportunities and challenges. China Commun. **17**(4), 152–164 (2020)

Trilateration Based Localization for Underwater Sensor Networks

Pradeep Nazareth, B. R. Chandavarkar, and Priyansh Kumar Dubey

Abstract Underwater Sensor Networks (UWSNs) is one of the emerging areas in the field of communication. UWSNs are used in many applications. At the same time, it faces several challenges such as energy constraint, node mobility, routing, and localization. The sensed data is meaningful only when it is referred to with location. In underwater sensor networks (UWSNs), determining the sensor's location where the event is taking place is essential. Estimating the location of each node in a sensor network is known as localization. As global positioning system (GPS) signals are not propagated underwater, localization becomes a challenging issue. While various localization algorithms have been proposed for terrestrial sensor networks, they may not directly map them to underwater due to the variation and three-dimension nature of underwater. This paper presents a new localization technique for underwater sensor networks based on trilateration. In this technique, whenever a new node is introduced in an UWSNs, it will derive its coordinates by examining the coordinates of three of its nearest neighbors and the distance between them. Distance between them is determined by using round-trip time. The mathematical model, along with simulation, is presented in the paper. Further, a thorough analysis of results for static and mobile nodes is discussed in this paper.

1 Introduction

Underwater Wireless Sensor Networks (UWSNs) is a primary source for exploring the ocean environment. The need to sense the underwater paves the way for the development of UWSNs. The applications may have a wide variety of requirements

P. Nazareth (✉) · B. R. Chandavarkar (✉) · P. K. Dubey (✉)
Wireless Information Networking Group (WiNG), Department of Computer Science and Engineering, National Institute of Technology Karnataka, Surathkal, Mangalore, India
e-mail: pradeep.nitk2017@gmail.com

B. R. Chandavarkar
e-mail: brcnitk@gmail.com

P. K. Dubey
e-mail: priyansh.dby25@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_57

such as stationary or moving, brief or long-lived, time-sensitive or insensitive; these requirements further results in varying design paradigms [1]. The communication in UWSNs is assisted with one or more surface buoys. A surface buoy is a node placed partially above the water level and somewhat below the water level. The purpose of this surface buoy is to send the data to the satellite, which further forwards it to the receiving station situated on the land.

UWSNs are often static or semi-mobile. In static deployments, sensor nodes are attached to anchored buoys, or the seafloor, in semi-mobile underwater networks that can be suspended from buoys and deployed by a ship as shown in Fig. 1. The entire set-up is left underwater for few hours to many days as per applications. However, there may be variations in node connectivity due to water current.

UWSNs have numerous applications in the field of defense to enhance coastal area surveillance. UWSNs are used in underwater weather monitoring to detect and analyze tsunami formations and earthquakes. Further, UWSNs are used in various underwater resource exploration such as mines and changes in aquatic environment monitoring such as temperature changes, pollution monitoring.

Localization refers to determining the coordinates of a new node introduced in the network by taking the reference of the nodes already present in the network.

Localization of underwater nodes is one of the major areas of research because the location of sensed data is critical for further processing [3]. Localization underwater is much complicated than terrestrial networks. The radio waves used in Global Positioning System (GPS) will not work in water because radio waves undergo strong

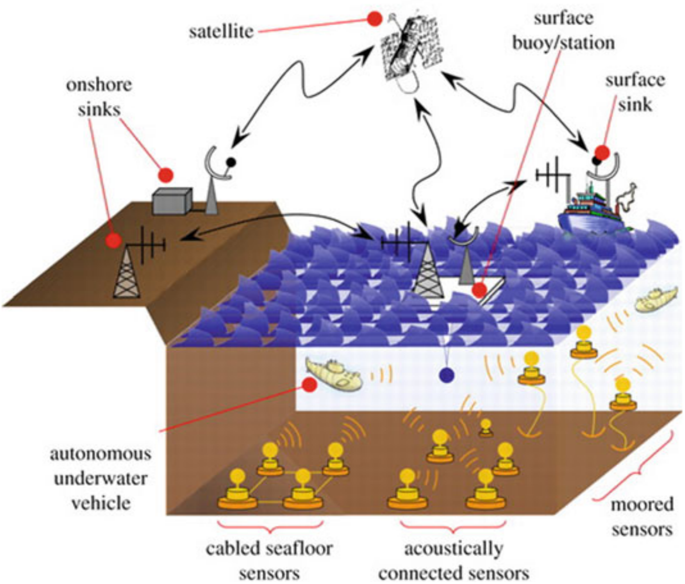


Fig. 1 Underwater node deployment [2]

attenuation in water [4]. Thus, the underwater node is not able to localize itself using GPS. Second, due to the substantial-high propagation delay in underwater acoustic signals, the overhead of maintaining clock synchronization is high. Third, the submerged sensor node's transmitting energy consumption is several times or even hundreds of times the receiving energy consumption. In terrestrial networks, the distance between the blind node and the anchor node is calculated by judging the RSSI (received signal strength indicator) value[5] of the signal. Implementing the same in the aquatic medium is not feasible due to heavy attenuation in water. In the place of radio waves, acoustic signals cannot be measured as precisely as radio waves. Thus need some alternative method to calculate the distance.

Major contributions of this paper are, design of localization techniques for underwater sensor networks based on trilateration. In this technique, whenever a new node is introduced in an UWSNs, it will derive its coordinates by examining the coordinates of three of its nearest neighbors and the distance between them. Distance between them is determined by using round-trip time. Further simulations are extended for mobile underwater nodes. Simulation is done by using the specialized underwater simulator - UnetStack. The paper is organized as follows: Sect. 2 discusses related work done in the area of underwater localization. The detailed mathematical model is presented in Sect. 3. The design is presented in Sect. 4. The Sect. 5 discusses about implementation in UnetStack. Detailed results are discussed in Sect. 6. Finally, the Sect. 7 discusses about conclusion.

2 Related Work

This section presents a brief overview of the localization techniques used in UWSNs. Underwater localization approaches can be broadly classified into range-based and range-free algorithms [6]. Range-based localization algorithms have two steps - the ranging and localization phase. In the ranging step, the angle or distance between nodes is estimated. Further, in the localization phase, actual coordinates of nodes are computed based on computed distance or angle [7]. Range-free localization does not require estimation of range to avoid computation overhead involved in determine range.

Area location scheme is a range-free underwater localization scheme [8], in which anchor nodes broadcast beacon messages at different power-level. The sensor nodes store each anchor nodes with corresponding transmitted power. The sink node obtains information from sensor nodes and detects the area of the sensor node. The significant advantage of this scheme is that it efficiently manages the sensor and anchor node's energy. This method requires low computational complexity but results in low location accuracy. The probabilistic localization method focuses on improving the accuracy of localization [9]. It consists of two phases, first determining the estimated position using the circle-based method and then determining the exact position using the measurement error's probability distribution. Uniform error distribution and normal error distribution are considered to improve the location accuracy. It is not suitable for UWSNs consists of a large number of nodes. Hyperbola-based localization approach [10] used to achieve better accuracy. It uses calibration and normal distribution to estimate computed estimation error. The significant advantages of this scheme are efficient energy consumption and lower computation complexity. On another end, it is not suitable for large UWSNs.

The silent positioning scheme is based on the Time Difference of Arrival (TDoA) approach [11]. This scheme uses four anchor nodes, from which range is estimated to underwater nodes. Multiple times range is estimated to improve efficiency, and the average range is computed and computed location. It achieves better localization accuracy and energy-efficient [12]. Energy-aware, distributed solution based on Time of Arrival (ToA) [13]. It uses broadcast ranging range estimation. There are many applications in which data is retrieved after the mission is over. The key feature of this scheme is localization obtained only after retrieving the node. Collaborative Monitoring Scheme (CMS) proposed in [14], in which two underwater nodes and a profiler are used. The profiler will deep dive underwater and move towards the depth of water. They know their depth using pressure sensors. Thus nodes collaborate to find two-dimensional coordinate estimation. Two underwater nodes follow the profiler, the range between profiler and nodes is estimated at regular intervals using ToA.

Localization with Directional Beacons (LDB) is a hybrid, anchor-free localization technique [15]. Underwater nodes estimate their depth by using pressure sensors. In LDB, Autonomous Underwater Vehicle (AUV) periodically transmits beacons. The underwater nodes receive those beacon fall in the conical beacon forms different circles with different heights. Thus they obtain 2D localization. It is energy efficient but has low localization accuracy.

3 Mathematical Modelling

In this paper, to determine the location of a sensor node, it needs to be in the range of the other three anchor nodes. The location of the anchor node is already known. The node whose location needs to be derived is called a blind node.

The concept [16] is based on the principle of the intersection of three circles: three circles will have only one unique point of intersection; there is no way that a combination of three circles will have more than one point of intersection. The idea here is that if coordinates of the neighboring three nodes and their distance from the blind node are obtained, assuming three circles are drawn with the adjacent nodes as a center, and the blind node must lie on the three circles. Hence it must satisfy the equation of all three circles as shown in Fig. 2.

$$(X - X_a)^2 + (Y - Y_a)^2 = d_a^2 \tag{1}$$

$$(X - X_b)^2 + (Y - Y_b)^2 = d_b^2 \tag{2}$$

$$(X - X_c)^2 + (Y - Y_c)^2 = d_c^2 \tag{3}$$

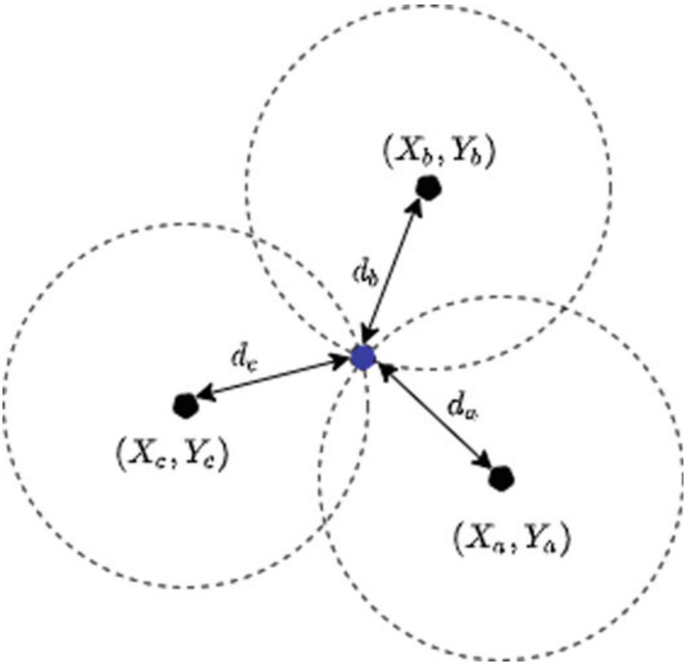


Fig. 2 Intersection of 3 circles at a unique point

These equations can be further expanded to:

$$d_a^2 = X^2 - 2X.X_a + X_a^2 + Y^2 - 2Y.Y_a + Y_a^2 \tag{4}$$

$$d_b^2 = X^2 - 2X.X_b + X_b^2 + Y^2 - 2Y.Y_b + Y_b^2 \tag{5}$$

$$d_c^2 = X^2 - 2X.X_c + X_c^2 + Y^2 - 2Y.Y_c + Y_c^2 \tag{6}$$

The three Eqs. 4, 5, and 6 are independent non-linear simultaneous equations Eq. 6 was subtracted from Eq. 5 and Eq. 4 was subtracted from Eq. 5 to get the following linear equations:

$$d_b^2 - d_c^2 = 2X(X_c - X_b) + X_b^2 - X_c^2 + 2Y(Y_c - Y_b) + Y_b^2 - Y_c^2 \tag{7}$$

$$d_b^2 - d_a^2 = 2X(X_a - X_b) + X_b^2 - X_a^2 + 2Y(Y_a - Y_b) + Y_b^2 - Y_a^2 \tag{8}$$

Rearranging and resolving the above equation, we get the intersection point ‘X’ and ‘Y’

$$Y = \frac{v_b(X_c - X_b) - v_a(X_a - X_b)}{(Y_a - Y_b)(X_c - X_b) - (Y_c - Y_b)(X_a - X_b)}$$

$$X = \frac{v_a - Y(Y_c - Y_b)}{(X_c - X_b)}$$

where

$$v_a = X(X_c - X_b) + Y(Y_c - Y_b)$$

$$v_b = X(X_a - X_b) + Y(Y_a - Y_b$$

The values for X and Y gives us an accurate position in two-dimension (2D) for the blind node.

4 Design

In the proposed approach, round trip time is used to estimate the distance from the blind node to the anchor node. The blind node broadcasts a beacon to find the existence of the anchor nodes in its range. Anchor nodes that receive the beacon respond with their coordinates. As soon as the blind node receives information about three neighboring anchor node's coordinates, it will run an agent at a specific interval to judge the round trip time and the distance value. The scheme proposed here is already popular [17] in the terrestrial wireless sensor networks, but it is not still employed in the UWSNs. This algorithm determines a specific blind node's location within the distributed nodes along the testbed area. The proposed algorithm has two phases: the initialization phase and the Coordinate computation phase.

4.1 Initialization Phase

Only the anchors will have position data before implementing the positioning algorithm, but all the nodes have identification numbers (IDs). The network is considered for this algorithm will be scalable to a considerable number of nodes spread over the testbed area. The percentage of anchor nodes will be small. It results in a situation where only a tiny percentage of nodes in the network can establish direct contact with any anchor nodes. The blind node in the direct range of the broadcast will obtain the coordinate information of the anchor node and get the distance with corresponding anchor nodes using round trip time. Through this process, all blind nodes will know the location of the anchors and their distance.

One of the significant steps in the algorithm mentioned above is estimating the distance between the anchor node and the blind node by estimating the round trip time.

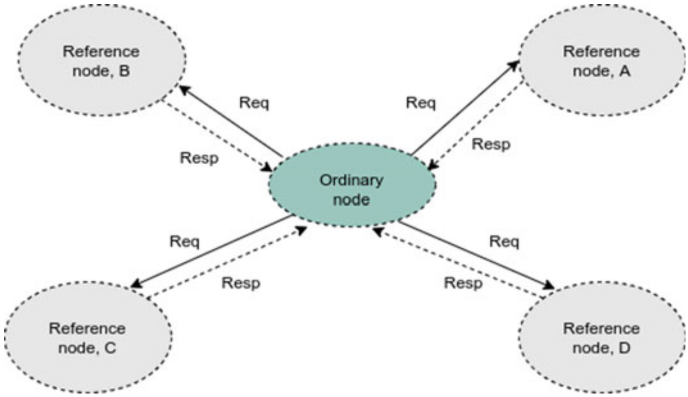


Fig. 3 Representation of estimation of distance using round trip time

Compared to the symmetric round trip-based algorithm, this procedure consumes less energy because for the localization of k blind nodes, there are four anchor nodes in the neighborhood of the blind node, which are A, B, C, and D anchor nodes. To estimate the distance [18] between them, the blind node will send a request and note the time stamp. When it receives the response, it will again mention the timestamp, and from both the timestamp, it will estimate the total time taken by the signal in one round trip Fig. 3. As the signal's speed and time are known, the distance between the nodes can be calculated.

4.2 Coordinatie Computation Phase

If a blind node can estimate its distance to at least three anchor nodes; then the blind node can perform trilateration to get its accurate location in 2D; this blind node becomes a "converted" anchor node whose position will now be sent to the sink. This process (initialization and final phases) will continue until all blind nodes become anchor nodes.

The following algorithm explains the procedure in a proper algorithmic manner.

Procedure:

1. A broadcast request is made by blind node.
2. **If** anchor node is in the range of blind node **then:** respond with its address
   endIf
3. **If** blind node gets packets from three different anchors **then:**
   for each address **do:**
   request coordinates information
   calculate the distance using round trip time

endfor
endIf
4. **else** repeat 1
5. **If** at least three different anchors identified **then:** Perform trilateration
endIf
6. **else** repeat 3
7. **If** the trilateration is successful **then:**
blind node becomes a converted anchor node
endIf
8. **else** repeat 5

End Procedure

5 Overview of UnetStack

This section of the paper presents the overview of UnetStack. UnetStack is a specialized underwater simulation tool composed of several agents. Further, users have flexibility in creating their agents as far requirement [19]. The advantage of UnetStack is simulated code can be directly ported into UnetStack enabled modem. Thus, first, required scenarios can be simulated and later tested with real-time scenarios. UnsetStack allows both discrete events and real-time simulations. UnetStack has different components; some of the components/agents used in our simulations are,

- Unet simulator: It enables writing and execution of UnetStack scripts. The simulation scripts allow the programmer to configure the underwater nodes.
- NodeInfo agent: It provides node information such as node name, address, coordinate information, and mobility information of the node.
- Ranging agent: Range estimation is a fundamental requirement in the underwater network protocol design. The range agent provides various services such as One-Way Travel Time (OWTT) and Two-Way Travel Time (TWTT) range estimation.

6 Results and Analysis

6.1 Localization for Static Blind Node

Two agents are developed to localize a node installed on blind node and anchor nodes, respectively. To localize a static blind node following Steps are followed.

- The blind node will first start broadcast requests for getting the address of the neighbor nodes.
- As soon as the anchor nodes in the neighborhood receive the request, they reply with a PDU containing their address.

- At the blind node, the agent will store the address in a list.
- As soon as the blind nodes receive the addresses from three neighbor nodes, it queries them for getting their ranges.
- In the range response packet itself node can extract the coordinates by using the *getpeerlocation()* function.
- To avoid the collision *Waker behavior* is employed in querying the ranges. It triggers the anchor nodes to respond at different times.

6.2 Localization for Mobile Blind Node

The problem with mobile nodes is that the node might move to a new position when the localization is completed. Thus, during the estimation, if the speed at which the node is moving, the distance covered by it in a given time can be obtained. Adding this distance to the estimated coordinates, it is possible to get the actual coordinates of the nodes at that time. Figure 4 represents the concept.

To localize a static blind node following algorithm is followed [20].

- The blind node will start broadcast requests for getting the address of the neighbor nodes.
- The current timestamp is saved in a variable 'start'.
- As soon as the anchor nodes in the neighborhood receive the request, they reply with a PDU containing their address.
- At the blind node, the agent will store the address in a list.
- As soon as the blind nodes receive the addresses from three neighbor nodes, it queries them for getting their ranges.
- In the range response packet itself we can extract the coordinates by using the *getpeerlocation()* function.

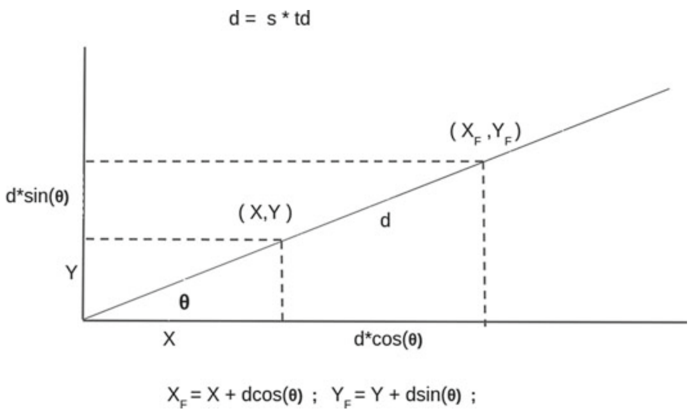


Fig. 4 Localization for mobile blind node

- To avoid the collision *Waker behavior* is employed in querying the ranges. It triggers the anchor nodes to respond at different times.
- After getting the ranges and coordinates information from the anchor nodes, the coordinates of the blind nodes are estimated.
- The estimated coordinates might be different than the current coordinates because, during all this time, the blind node has traveled some distance.
- Therefore the estimated coordinates are updated as shown in Fig. 4.

6.3 Metrics for Evaluating Algorithm

Following metrics are used to evaluate the performance of the algorithm [21].

- Localization Error: Localization Error is defined as the distance between the estimated and the actual coordinates of the node. It is computed using the equation:

$$Error = \sqrt{(x_e^i - x_a^i)^2 + (y_e^i - y_a^i)^2}$$

  where (x_e^i, y_e^i) and (x_a^i, y_a^i) are node i's estimated and actual coordinates.
- Total time taken by the algorithm: The algorithm's total time to execute should not be very high, as it will lead to low performance. The ideal situation will be that the algorithm completes its execution and updates the coordinates on the map within a reasonable time.

6.4 Simulation Setup

- The simulation is carried out using a topology, which consists of 4 nodes in the area of 3000 m × 3000 m × 3000 m. Anchor nodes are static.
- Mobile blind nodes are configured to move in a linear direction with speed of 1 m/s. The interval of response to broadcast requests varies from 1 to 4 s. I
- The mobile node moves in a straight line triggering the localization algorithm every 2 min.
- For the first 2 min, the blind node will be stationary. Next, for 10 min, it will move in a straight line, and at last, it will stop after traveling for 10 min and stays there forever.

6.5 Observations

It has been observed that for stationary blind node, negligible (0.00038 m) error is observed. Figure 5 shows the variation of error with increased mobility of blind node. It is observed that as the speed of the node increases, the localization error increases. The main reason is that distance between a blind node, and anchor node is identified with respect to one point, and later node moves. Further, there is no correspondence between x and y coordinates during node movement results in localization error. For the speed of 1mps, an average of 10 m error is observed. Figure 6 shows the time taken

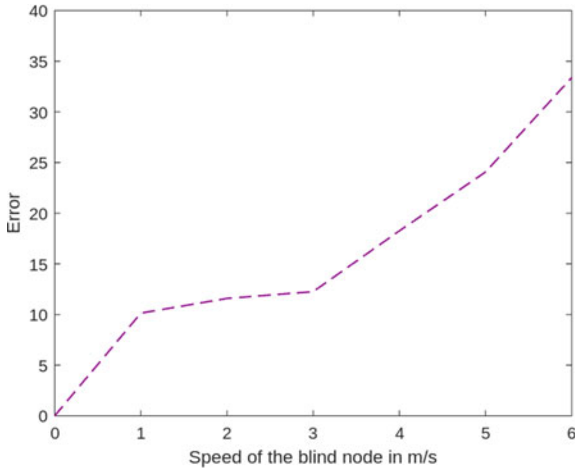


Fig. 5 Variation of error with speed of blind node

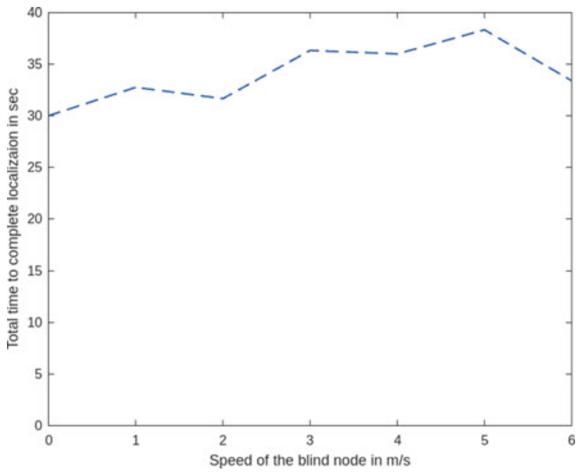


Fig. 6 Time taken for localization with speed of blind node

to complete the localization process. Speed of the node movement has no significant impact on the completing localization. It varies between 30 to 40 s, irrespective of blind nodes.

7 Conclusions and Future Work

This paper addresses the localization problem for UWSNs using the intersection of circles method. A novel approach with the mathematical model is presented, and based on the mathematical model, we proposed solutions for static and mobile blind nodes. Thorough simulations are carried out the validity of mathematical model using UnetStack simulator. Further analyzed the simulations and presented the results. It is observed that the algorithm gives very accurate results for the case of the static blind node. For mobile blind nodes, the error increases proportionally to the speed of the node. In the future, attempts can be made to make the algorithm more accurate with the increasing speed of the blind node.

References

1. Rahman, T., Yao, X., Tao, G.: Consistent data collection and assortment in the progression of continuous objects in IoT (2018)
2. Heidemann, J., Stojanovic, M., Zorzi, M.: Underwater sensor networks: applications, advances and challenges. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. **370**(1958), 158–175 (2012)
3. Hightower, J., Want, R., Borriello, G.: Spoton: an indoor 3d location sensing technology based on RF signal strength (2000)
4. Chandrasekhar, V., Seah, W.K., Choo, Y.S., Ee, H.V.: Localization in underwater sensor networks: survey and challenges. In: Proceedings of the 1st ACM International Workshop on Underwater Networks, pp. 33–40 (2006)
5. Benkic, K., Malajner, M., Planinsic, P., Cucej, Z.: Using RSSI value for distance estimation in wireless sensor networks based on ZigBee (2008)
6. Isik, M.T., Akan, O.B.: A three dimensional localization algorithm for underwater acoustic sensor networks (2009)
7. Toky, A., Singh, R.P., Das, S.: Localization schemes for underwater acoustic sensor networks-a review. Comput. Sci. Rev. **37**, 100241 (2020)
8. Chandrasekhar, V., Seah, W.: An area localization scheme for underwater sensor networks. In: OCEANS 2006–Asia Pacific, pp. 1–8 (2006). https://doi.org/10.1109/OCEANSAP.2006.4393969
9. Bian, T., Venkatesan, R., Li, C.: An improved localization method using error probability distribution for underwater sensor networks. In: 2010 IEEE International Conference on Communications, pp. 1–6 (2010). https://doi.org/10.1109/ICC.2010.5501953
10. Bian, T., Venkatesan, R., Li, C.: Design and evaluation of a new localization scheme for underwater acoustic sensor networks. In: GLOBECOM 2009–2009 IEEE Global Telecommunications Conference, pp. 1–5 (2009). https://doi.org/10.1109/GLOCOM.2009.5425366
11. Cheng, X., Shu, H.S.H., Liang, Q.: A range-difference based self-positioning scheme for underwater acoustic sensor networks. In: International Conference on Wireless Algorithms, Systems and Applications (WASA 2007), pp. 38–43 (2007). https://doi.org/10.1109/WASA.2007.40

12. Tuna, G., Gungor, V.C.: A survey on deployment techniques, localization algorithms, and research challenges for underwater acoustic sensor networks. Int. J. Commun Syst **30**(17), e3350 (2017)
13. Mirza, D., Schurgers, C.: Energy-efficient ranging for post-facto self-localization in mobile underwater networks. IEEE J. Sel. Areas Commun. **26**(9), 1697–1707 (2008). https://doi.org/10.1109/JSAC.2008.081209
14. Mirza, D., Schurgers, C.: Collaborative localization for fleets of underwater drifters. In: OCEANS 2007, pp. 1–6. IEEE (2007)
15. Luo, H., Guo, Z., Dong, W., Hong, F., Zhao, Y.: LDB: localization with directional beacons for sparse 3D underwater acoustic sensor networks. J. Netw. **5**(1), 28 (2010)
16. Schaeffer, D.G.: A mathematical model for localization in granular flow (1992)
17. Xu, J., Liu, W., Lang, F., Zhang, Y., Wang, C.: Distance measurement model based on RSSI in WSN (2010)
18. Karn, P., Partridge, C.: Improving round-trip time estimates in reliable transport protocols. ACM New York, NY, USA, vol. 17, pp. 2–7 (1987)
19. Chitre, M., Bhatnagar, R., Soh, W.S.: Unetstack: An agent-based software stack and simulator for underwater networks. In: 2014 Oceans-St. John’s, pp. 1–10. IEEE (2014)
20. Aksu, H., Aksoy, D., Korpeoglu, I.: A study of localization metrics: evaluation of position errors in wireless sensor networks (2011)
21. Zhang, S., Cao, J., Li-Jun, C., Chen, D.: Accurate and energy-efficient range-free localization for mobile sensor networks. IEEE Trans. Mob. Comput. **9**(6), 897–910 (2010)

Geometrically Robust Digital Image Watermarking Based on Zernike Moments and FAST Technique

Ranjana Dwivedi and Vinay Kumar Srivastava

Abstract Geometrically robust digital image watermarking scheme based on Zernike moments and Features from Accelerated Segment Test (FAST) technique is presented in this paper. FAST technique detect the geometrically robust feature points from the image and circular patches centered around these detected features are used to embed the given watermark data sequence. Zernike transform calculate the Zernike moment magnitudes (ZMMs) of extracted circular patches and these are modified by embedding watermark sequence. Circular patches are reconstructed by utilizing modified ZMMs and these replaces the original patches. Robustness of proposed scheme is tested by performing various operations and attacks on watermarked image. Experimental results show the proposed method performs well under various common signal processing geometrical attacks.

Keywords Zernike moments · FAST · Feature points

1 Introduction

Multimedia data such as images, audio, video, text etc. transmission over the internet has increased exponentially. With the increase of digital data transmission, it becomes possible to create, duplicate, transmit, and distribute data without the knowledge of the owner of data. To protect multimedia data from forgery, theft, illegal transmission, duplication, Digital Watermarking is one of the solution for security, data authentication and copyright protection of the digital data. Digital image watermarking is a data hiding technique to conceal information or a message into a given image. Message or information is embedded into the original image with the help of watermarking embedding algorithm and a watermarked image is produced. During transmission of watermarked image through the communication network, it may get interfered

R. Dwivedi (✉) · V. K. Srivastava
Department of Electronics and Communication Engineering, Motilal Nehru National Institute of Technology Allahabad, Prayagraj, India
e-mail: ranjana@mnnit.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_58

with some undesired noise or attacks. Attacks or noise may degrade the watermarked image quality. At the receiver end, message or information is detected or extracted using watermarking extraction algorithm. Watermarking extraction algorithm is the reverse process of embedding algorithm. While designing the watermarking algorithm, several factors must be considered, such as robustness against intentional or unintentional distortion of the proposed algorithm, imperceptibility, capacity, and security. Robustness tells us about how much it is resistant to an attack. The Watermarking scheme can be divided into three categories i.e. robust, fragile, or semi-fragile. The scheme is resistant to intentional or unintentional attacks in robust watermarking. It makes it difficult for the hacker to remove or alter the embedded information. In the case of fragile watermarking, watermark is easily destroyed on applying attacks or even slight modifications. Fragile watermarking is used in tamper detection, content authentication. The semi-fragile watermarking scheme can tolerate attacks or noise up-to some level or degree. It is robust to intentional attacks but fragile to malicious attacks. Attacks can be classified into common signal processing attacks and geometrical attacks. Common signal processing attacks mainly aims to affect the strength of embedded watermark, while geometrical attacks introduce geometric distortion. Rotation, cropping, scaling, translation are some examples of geometrical attacks. To encounter geometrical attacks, various robust image watermarking techniques [1–6] are developed.

To improve the watermarking method's robustness, various orthogonal image moments are incorporated by researchers in their methods. Image moments [7] are region-based descriptors. Image moments are named according to the type of polynomial base on which projection of image is performed. Moments have been widely used in many image analysis and processing applications. For ex. object classification, edge detection, and pattern recognition, and image retrieval. Image moments fully reconstruct the original image after the embedding of secret bits in the original image. Image moments contain the information of image properties invariant to rotation, translation and scaling operations on image. Various image moments are used in literature for watermarking methods like Zernike, Pseudo-Zernike, Legendre moments, Fourier-Mellin moments, Hu's seven moments, Tchebichef moments, Krawtchouk moments, Wavelet moments, etc.

In this paper, we propose a geometrically robust image watermarking scheme based on Zernike moments and FAST technique of feature detector. Section 2 explains the basics of Zernike moments and FAST technique which is used in our proposed scheme to detect features from the given image. Section 3 discuss the proposed watermarking embedding and extraction algorithm. Experimental results of proposed scheme and its performance under various attacks are presented in Sect. 4 and finally the conclusion is drawn.

2 Zernike Moments and Features from Accelerated Segment Test (FAST) Technique

2.1 Zernike Moments

Zernike moments are introduced in image analysis by Teague [8] in 1980. Zernike [9] proposed a complex polynomials set $\{ V_{kl}(x, y)\}$, which forms a complete orthogonal set in polar coordinate over the unit disk of $x^2 + y^2 \leq 1$. Polynomial form is defined as

$$V_{kl}(x, y) = V_{kl}(\rho, \theta) = R_{kl}(\rho)e^{il\theta} \tag{1}$$

where k is non-negative, l is integers subject to constraints k − |l| is even, and k − |l|;
ρ - length of the vector from the origin to the pixel (x, y);
θ - angle between x axis and vector ρ.
$R_{kl}(\rho)$ is Radial polynomial defined as:

$$R_{kl}(\rho) = \sum\nolimits_{s=0}^{(k-|l|)/2} (-1)^s \frac{(k-s)!}{s!(\frac{k+|l|}{2} - s)!(\frac{k-|l|}{2} - s)!}\rho^{k-2s} \tag{2}$$

The Zernike moment A_{kl} for function $f(x, y)$ is defined as:

$$A_{kl} = \frac{k+1}{\pi} \iint_{x^2+y^2 \leq 1} f(x, y)V_{kl}^{*}(x, y)dxdy \tag{3}$$

where $V_{kl}^{*}(x, y) = V_{k,-l}(x, y)$ and k is the order with repetition l.
Zernike moment of a digital image is calculated using,

$$A_{kl} = \frac{k+1}{\pi} \sum\nolimits_{x} \sum\nolimits_{y} f(x, y)V_{kl}^{*}(x, y) \tag{4}$$

satisfying $x^2 + y^2 \leq 1$

For computing the Zernike moments of an image, the image centre is considered as the origin, and pixel coordinates are mapped into the unit circle. Pixels that lie outside of the unit circle have been discarded.

Image can be reconstructed using Zernike moment A_{kl},

$$f'(x, y) = \sum\nolimits_{k=0}^{N} \sum\nolimits_{l} A_{kl}V_{kl}(x, y) \tag{5}$$

The lower order of Zernike moments captures the image's grossly shaped image, while a higher order of Zernike moments represents high-frequency details. Zernike moments have rotation invariant feature, i.e., the magnitude of Zernike moments of the original image and rotated image is equal.

Image normalization operation is performed to achieve scale and translation invariance. An image function $f(x, y)$ can be normalized by transforming it into $h(x, y)$.

$$h(x, y) = f(\frac{x}{a} + \bar{x}, \frac{y}{a} + \bar{y}) \tag{6}$$

where $(\bar{x}, \bar{y})$ is centroid of $f(x, y)$ and $a = \sqrt{\frac{\beta}{m_{00}}}$ with β a pre-determined value and m_{00} its zero order moment.

2.2 FAST Technique

Feature of an image is a distinct pattern which is different from its neighborhood present in the image. It can be a point, curve of boundaries, corner or edges of the image. FAST method [10] computes a feature point P by considering a Bresenham circle around it with radius of sixteen pixels. To check whether a point is feature point or not, its pixel P from the image is selected, and let I_P be the intensity of pixel P and Th be the threshold value. Bresenham circle with radius of 16 pixels around the point P is considered. Pixel P is considered feature point, if its n neighboring pixels in the considered circle are all brighter than $I_P + Th$ or darker than $I_p - Th$. Detection of interest points from a large set of interest point are selected by applying non-maximum suppression. FAST method is used in real time applications because it is computationally efficient.

3 Watermark Embedding and Extraction Algorithm

In our proposed scheme, Feature points are detected from a given image input using FAST technique. Around each detected features, circular patches are extracted and this extracted circular region is used to embed watermark bits. Zernike moments of circular regions are calculated and watermark bits are repeatedly embedded into their corresponding ZMM. Proposed scheme utilizes FAST technique of feature detection. There are some popularly used feature detectors in literature are Harris corner detector, Scale Invariance Feature Transform etc. For geometrically invariant features, we first detect features using FAST techniques and apply geometric attacks on the image. After applying geometric attacks such as rotation & scaling, feature points are again detected from attacked image. It is noticed that most of the detected features are same as detected before applying attacks. Only those feature points are selected further for watermark embedding which are robust against attacks for geometrically invariant digital image watermarking. After selecting features point, a circular patch is extracted centered around the selected feature point. These circular

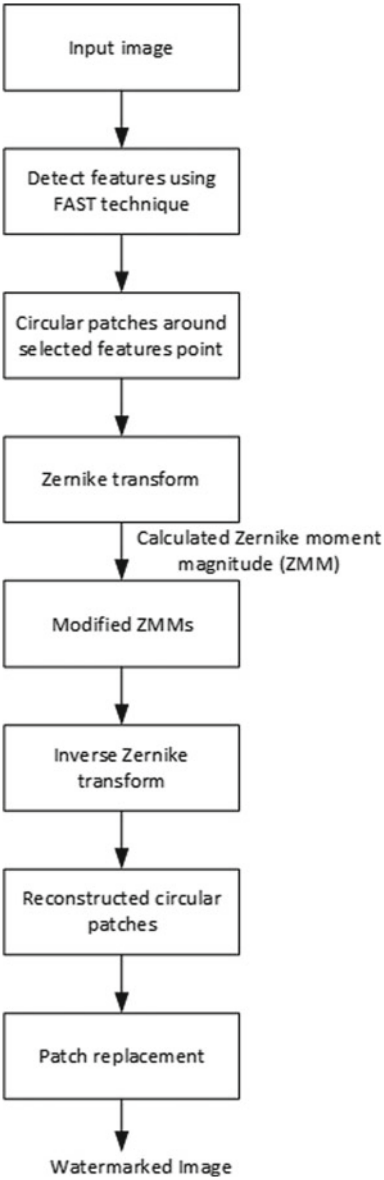


Fig. 1 Block diagram of proposed watermark embedding and extraction scheme

patches are used to embed watermark bits by modifying their calculated Zernike moment magnitudes. Figure 1 shows the block diagram of proposed watermark embedding scheme. By applying Zernike transform on each extracted circular patch, their corresponding ZMMs are calculated with order 35. Watermark data sequence of fixed length are generated and embedded into circular patches by modifying their ZMMs. These sequences are repeatedly embedded into ZMMs using Spread spectrum communication technique [11] by using Eq. (7)

$$Z^{'} = Z + \alpha.W \tag{7}$$

where Z is ZMM of extracted circular patches, α is watermark embedding strength parameter and W is watermark data sequence, $Z^{'}$ is watermarked data. α controls the watermark embedding strength, if it has large value that means watermark embedding is strong and it gives better watermark detection results. But at the same time, large α causes much distortion. Thus to keep balance between watermark detection and distortion, medium value of α is chosen for watermark embedding. After embedding watermark data sequence, inverse Zernike transform is applied to reconstruct the circular patches. These reconstructed circular patches are replaced with original circular patches and watermarked image is obtained.

For watermark extraction, FAST algorithm is applied on received watermarked image to detect its feature points. After feature points detection, circular patches centered around feature points are extracted and Zernike transform is applied on each extracted circular patches. Zernike transform calculate its corresponding ZMM with specified order of 35. To detect watermark bits which are embedded in ZMMs of watermarked image, linear correlation [12] is used. Linear correlation can be defined as,

$$C_{Linear} = \frac{1}{S} \sum w.z^{'} \tag{8}$$

where w is the same watermark data sequence used in embedding and $z^{'}$ is watermarked data. S is the size of ZMMs. Watermark detection is decided by comparing the linear correlation with a pre-defined threshold value. If linear correlation value is greater than threshold value, watermark is detected otherwise not.

4 Experimental Results

Proposed scheme is applied on various test images of size 512 × 512. Here we select 15 strong feature points for Lena image, 13 features points for Pepper image and 9 features point for Baboon image using FAST technique. Circular patches of radius 35 pixels are drawn and their corresponding ZMMs are calculated with order $O = 35$. Watermark embedding strength chosen here is 30 and experiments are performed using MATLAB. Figure 2 shows the detected features from Baboon, Pepper and Lena image by applying FAST technique and corresponding extracted circular patches. Figure 3 shows the watermarked image of Baboon, Pepper and Lena with PSNR of 45.47 dB, 45.23 dB, 46.71 dB respectively. Watermarked images are tested against various common signal processing and geometrical attacks by utilizing *Stirmark* 4.0 [13]. These attacks include median filtering, Gaussian low pass filtering, JPEG compression, cropping, scaling, rotation, shearing etc. are tested to check the robustness of proposed scheme. Table 1. shows the results of watermark detection results of proposed scheme under various attacks and its comparison with some existing feature based watermarking techniques proposed by [14, 15]. Numerator shows the number of circular patches where watermarks are detected successfully

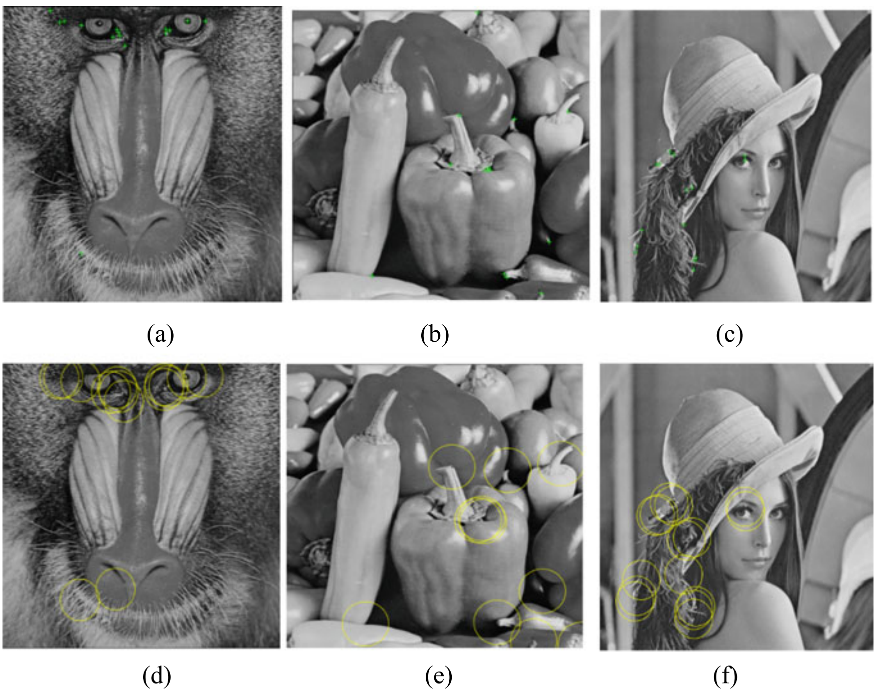


Fig. 2 Features detected using FAST technique from **a** Baboon image **b** Pepper image **c** Lena image; Extracted circular patches around features point **d** 9 circular patches of Baboon image **e** 13 circular patches of Pepper image **f** 15 circular patches of Lena image

Table 1 Watermarking detection results under various common signal processing and geometrical attacks

Attacks	Baboon			Pepper			Lena		
	[14]	[15]	Proposed	[14]	[15]	Proposed	[14]	[15]	Proposed
Median Filter 3 × 3	6/15	9/10	8/9	17/23	9/10	10/13	6/10	8/9	13/15
Gaussian Filter 3 × 3	4/15	9/10	8/9	9/23	9/10	11/13	5/10	8/9	13/15
JPEG 90	8/15	9/10	9/9	16/23	10/10	12/13	6/10	8/9	12/15
JPEG 70	8/15	8/10	9/9	14/23	10/10	12/13	4/10	8/9	13/15
Median Filter 3 × 3 + JPEG 90	5/15	9/10	7/9	16/23	9/10	11/13	3/10	8/9	13/15
Gaussian Filter 3 × 3 + JPEG 90	4/15	9/10	8/9	5/23	9/10	11/13	4/10	8/9	12/15
Cropping (10% off)	6/15	8/10	7/9	6/23	9/10	10/13	5/10	8/9	12/15
Scaling (1.5)	7/15	8/10	8/9	13/23	9/10	11/13	6/10	8/9	13/15
Rotation (5°)	5/15	9/10	9/9	9/23	9/10	12/13	4/10	8/9	12/15
Shearing (1%)	6/15	7/10	7/9	7/23	8/10	11/13	4/10	9/9	13/15
Removed 5 rows and 17 coloumns	6/15	6/10	7/9	11/23	6/10	12/13	5/10	5/9	13/15
Cropping (10% off) + JPEG 70	6/15	8/10	7/9	6/23	9/10	11/13	5/10	8/9	14/15
Rotation (5°) + Cropping + JPEG 70	5/15	7/10	7/9	7/23	9/10	12/13	5/10	7/9	13/15
Removed 5 rows and 17 columns + JPEG 70	6/15	6/10	6/9	7/23	6/10	11/13	4/10	5/9	12/15

(a) (b) (c)

Fig. 3 Watermarked images **a** Baboon image with PSNR = 45.47 dB **b** Pepper image with PSNR = 45.23 dB **c** Lena image with PSNR = 46.71 dB

and denominator denotes the total number of original watermarked circular patches. Results shows that proposed scheme performs better in comparison with existing techniques under various attacks.

5 Conclusions

Geometrically robust digital image watermarking scheme based on Zernike moment and FAST technique is proposed in this paper. Feature points are detected by employing FAST techniques of feature detection and these features are used for watermark embedding. Circular patches centered around detected feature points are extracted and their corresponding ZMMs are modifies by embedding watermark bits. Robustness of Proposed scheme is tested by applying various attacks on watermarked image and results obtained made it clear that proposed scheme well performed under attacks. Experimental results are also compared with some existing techniques.

References

1. Bas, P., Chassery, M.B.: Geometrically invariant watermarking using feature points. IEEE Trans. Image Process. **11**, 1014–1028 (2002)
2. Lee, H., Lee, C., Lee, H.: Geometrically invariant watermarking: synchronization through circular Hough transform. Multimedia Tools Appl. **34**, 337–353 (2007)
3. Seo, J., Yoo, C.: Image watermarking based on invariant regions of scale-space representation. IEEE Trans. Signal Process. **54**, 218–238 (2006)
4. Lutovac, B., Daković, M., Stanković, S., Orović, I.: An algorithm for robust image watermarking based on the DCT and Zernike moments. Multimedia Tools Appl. **76**, 23333–23352 (2017)
5. Degadwala, S.D., Kulkarni, M., Vyas, D., Mahajan, A.: Novel image watermarking approach against noise and RST attacks. Procedia Comput. Sci. **167**, 213–223 (2020)

6. Hu, R., Xiang, S.: Cover-lossless robust image watermarking against geometric deformations. IEEE Trans. Image Process. **30**, 318–331 (2020)
7. Tsougenis, E., Papakostas, G., Koulouriotis, D., Tourassis, V.: Performance evaluation of moment-based watermarking methods: a review. J. Syst. Softw. **85**, 1864–1884 (2012)
8. Teague, M.: Image analysis via the general theory of moments. Opt. Soc. Am. **70**, 920–930 (1980)
9. Zernike, F.: Diffraction theory of the cut procedure and its improved form, the phase contrast method. Physica **1**, 689–704 (1934)
10. Rosten, E., Drummond, T.: Machine learning for high speed corner detection. In: 9th Euproean Conference on Computer Vision, vol. 1, pp. 430–443 (2006)
11. Cox, M., Miller, J., Bloom, H.C.: Digital Watermarking. Morgan-Kaufmann, San Francisco (2002)
12. Zheng, D., Wang, S., Zhao, J.: RST invariant image watermarking algorithm with mathematical modeling and analysis of the watermarking processes. IEEE Trans. Image Process. **18**, 1055–1068 (2009)
13. Petitcolas, F.: Watermarking schemes evaluation. IEEE Signal Process. Mag. **17**(5), 58–64 (2000)
14. Gao, X., Deng, C., Li, X., Tao, D.: Geometric distortion insensitive image watermarking in affine covariant regions. IEEE Trans. Syst. Man Cybern. Part C: Appl. Rev. **40**, 278–286 (2010)
15. Yuan, X., Pun, C.: A geometric invariant digital image watermarking scheme based on robust feature detector and local Zernike moments. In: 9th International Conference on Computer Graphics Imaging Vision, pp. 53–56 (2012)

Extension of the ERDMOSM1 and ERDMOSM2 Reduced Group Delay Finite Impulse Response (FIR) in Digital Filtering

Sonelal Prajapati and Sanjeev Rai

Abstract The conventional digital filters may not be suitable in such DSP applications where distortion-less responses are desired along with the minimum group-delayed output. The conventional Finite Impulse Response (FIR) filters give the distortion-less output but it has an objectionable group-delay response. The Infinite Impulse Response (IIR) filters can be used to minimize the group delay but due to its recursive nature, it introduces distortion in response. In this paper, a novel modified overlap-save method-based FIR filter is proposed that improves the group delay. To design such filters, we first make the impulse response causal, and then causal impulse responses are shifted by (N-1)/2 unit for odd order impulse response and N/2 for even order impulse response. Further, we have proposed a novel algorithm for the removal of undesired samples from the convolution result. The final output response shows significant improvement in group delay over the existing techniques.

Keywords Group delay · FIR filter · IIR filter · Enhance Reduce delay modified over-lap save Method (ERDMOSM)

1 Introduction

A digital filter is classified as Finite Impulse Response (FIR) filter and Infinite Impulse Response (IIR) filter [3, 4] A digital filter performs mathematical operations ona sampled, discrete-time signal to reduce or enhance certain aspects of that signal. Close to or outside the Z plane, the join loop may be the poles of an IIR filter. A stability problem could appear in an IR filter in digital signal handling applications. A linear phase is desired, for unqualified to ability such as filter for its many benefits, absences of coefficient sensibility, and less filter phase deformation. The

S. Prajapati (✉) · S. Rai
Motilal Nehru National Institute of Technology Allahabad, Prayagraj, India
e-mail: sonelal59@gmail.com

S. Rai
e-mail: srai@mnnit.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_59

linear phase FIR indefinite application has the main losses in general group delay wherever is the filter length. For higher sequence filter this is largest by size, and in communication applications the echo of transmitting signals tends to be the main problem. Furthermore, in digital reaction device application a delay in the reaction round am in general speed. Multirate digital signal processing is used to decrease the electronic delay in the lively blast control system, therefore, this plane has been approached [11]. Mathematical notation of a sample which is digital filtered can be classified (FIR) filter and (IIR) filter. Close to or outside the Z plane, the join loop may be the poles of an IIR filter. A stability problem could affect the IIR filter, and digital signal handling applications.

In the linear phase FIR indefinite application, the main loss off to the general group delay is wherever is the filter length. For higher sequence filter, echo of transmitting signals, tends to be the and main problem. For the past few years, the design of the minimum-phase FIR filter by various authors has alerted only a quantity of interest. Linear phase FIR filter design series with [6] the low pass filter, ideal differentiator, and ideal Hilbert transformers can design a normal filter, as well as fulfill a normal-objective.

The author in, presented a fresh FIR filter design technique with a high-level passband and equiripple stopband [9, 10]. The shift of linear phase FIR filters to a minimum phase was implemented in Vaidyanathan [13], changing linear stage FIR filters to a minimum stage. A fresh set of maximally flat non-symmetric low-pass FIR filters have been expanded in sequence to improve the results used in the filter designing [8]. It offers a mildly delayed dc group than asymmetric delay. Any tiny quantity of frequency response filter. A robust response based on a discrete Hilbert transformation (DHT) was given in [2].To design optimum minimal phase digital FIR filters by change in magnitude response. Extending DHT in a complex case, the minimum stage Filter needs less computation and also lacks recall than the linear phase filter at the same moment as the delay and size response limitations immediately gathered. The size spectrum of the truncated minimum phase order shifts since the creative size band in this algorithm therefore, it is possible to find least output response but no altering the response (IR) of the filter, its display not viable according to traditional algorithms such as overlap-add method, overlap-save method OAM, OSM) [5]. Also, the convolution in,

$$y(n) = \sum_{k=0}^{N} h(k)x(n-k) \tag{1.a}$$

Whenever x (k) specifies the input signal toward the filtered,$h(k)$ is impulse reaction (IR) of the filter. The filtered signal is $y(n)$ and the filter setup is N. In [9] an optimization system iterative was given for the design by direct projection method of a low delay FIR bandpass filter with a maximum flat passband and equiripple stopband. Linear phase delay FIR filter advance N with filter arrangement progression. According to the author, this design method can accomplish a decreased delay causal.

FIR filter for duration has a transfer function,

$$\mathrm{h(z) = h(0) + h(1)_z^{-1} + h'(2)z^{-2} + + \ldots\ldots h(N-1)Z^{-(N-1)}} \tag{1.b}$$

While $h(z) = \sum_{n=0}^{N-1} h(n)z^{-n}$

Wherever $h(n)$ has the factor of the impulse reaction. While the coefficients have actual numbers. If any symmetric [i.e, $h(n) = h(N-1-n)$] or asymmetric [i.e.,$h(n) = -h(N-1-n)$] while the symmetric situation has been combined into (2), it is through the DTFT of the h(k) sequences

$$h\left(e^{j\omega}\right) = e^{-j\omega(N-2)/2} h_0\left(e^{j\omega}\right) \tag{1.c}$$

With

$$\mathrm{H_0(e^{jw})} = \begin{cases} \sum_{K=0}^{LN-1/2} d_k cos(\omega k) & if\,\mathrm{N}\,is\,odd, \\ \sum_{k=1}^{N-/2} d_k cos\left(\omega\left(k-\frac{1}{2}\right)\right) & if\,\mathrm{N}\,is\,even, \end{cases} \tag{1.d}$$

Anywhere $h_0\left(e^{j\omega}\right)$ it is real value of coefficient N range

$$\mathrm{d}_k = \begin{cases} \mathrm{h}\left(\frac{\mathrm{N}-1}{2}\right)\mathrm{fork=0}, \\ 2\mathrm{h}\left(\frac{Nj-1}{2} - \mathrm{k}\right)\mathrm{forK} \neq 0, \end{cases} \tag{1.e}$$

And stand for even, the factor is d_k known by

$$d_k = 2h(k), \mathrm{Fork} = 0, 1, 2, 3 \ldots \frac{N}{2} - 1. \tag{1.f}$$

Hence the filter's stage reaction

$$\phi(\omega) = -\left(\frac{N-1}{2}\right)\omega, \tag{1.g}$$

where is the linear phase of group delay=$\left(\frac{N-1}{2}\right)$.behalf N strange delay has an essential, where stand for N the even non -integer. If only the ranks $N-1$ has comparable size reaction this $h_0\left(e^{j\omega}\right)$ can be like achieve creating the non-causal transfer function

$$h_0\left(e^{j\omega}\right) = e^{j\omega(N-1)}\, H\left(e^{j\omega}\right) \tag{1.h}$$

2 Reduction of Group Delays: Modified Overlap-Save Technique

2.1 Zero Phase of Impulse Response

Now, therefore, It has considered that, through DFT-based circular convolution, overlap–save method produces a comparable outcome as linear convolution. Only impulse response has duration N and the information study sequence has range N, before aliasing exactly to the round convolution of two separate duration sequence is isolated by zero-padding next the end non-zero sample of the response (IR).Only N range impulse response where the least amount of zeros added will be the duration of the $L = N + M - 1$ convolution. Since the duration of the input signal was N, $N - 1$ unpredicted samples had to be isolated to the final convolution consequence

2.2 Reduction of Group Delay

According to the application, there are initially three instances where group delay reduction can be accomplished.

2.2.1 Zero delay Modified Overlap Save Method

The Zero delay Modified Overlap Save Method can be used for zero group delay. This is process, in its place of $N - 1$ zero models feel on x (n) suppliers, its improve initial $(N - 1)/2$ zero previous information filtering. Due to circulation move press $(N - 1)/2$ the filter signal of samples on separately finish. This is group delay case totally zero occurs. This procedure was termed zero delay Modified Overlap save Method (ZDMOSM). Figure 1 shows an illustration of a filter signal without group delay.

2.2.2 Reduce Delay Modified Overlap-Save Method

Zero delays modified overlap save method cannot be introduced in real-time for group delay reduction, this would, therefore, need to modify the term acquisition during processing. In the event to reduce Delay Modified Overlap-Save Method samples of $N - 1$ zeros are regarding before filtering as in the Overlap -save Method case. For processing in real-time rather than remove the comparable amount of both side models of the resulting circular convolution, $5(N - 1)/6$ models are separated since the left termination and $(N - 1)/2$ models since the right termination of the outcome. It processes safe $(N - 1)/2$ of the models, as Fig. 2 also an illustration of a filter signal poor group delay.

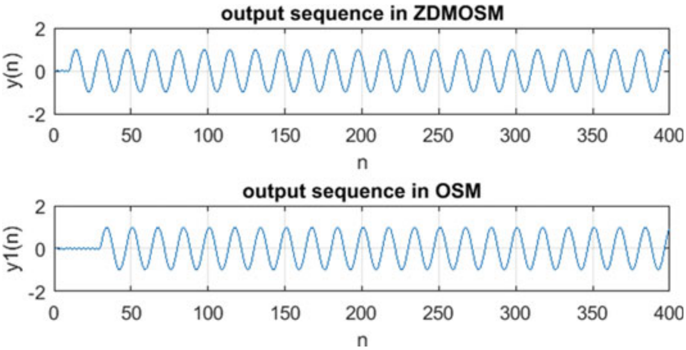


Fig. 1 Example of zero group delay filtration with ZDMOSM

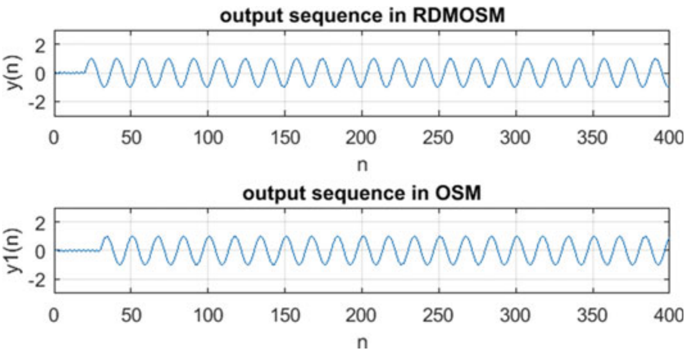


Fig. 2 Example of zero group delay filtration with RDMOSM

2.2.3 Enhanced case-: Enhance Reduce Modified Overlap-Save Method

Within this subdivision, another filtering procedure will be discussed doing the DFT-based round convolution an overlap-save method, the objective of the method is to accomplish group delay filtering. Following the zero-stage impulse response (IR), there are zero paddings.et $Y1(l)$ be the outcome of $h(l)$ and $x(l)$ circular convolution can be written as DFT

$$Y(l) = X(l)H(l) \tag{1.I}$$

$$Y_1(l) = e^{j(2\pi/K)n((l-1)/2} H(l.)X(l.)$$

We may take IDFT for certain

$$y_1(l) = \left(y\left(n + \frac{L-1}{2}\right)\right)_m \tag{1.J}$$

It is a mathematical equation to see that the outcome obtains with the zero-stage impulse (IR) i.e.,$y1(n)$ is a circular shift form of $y(n)$ moved by $N-1/2$ Specimens on the left side. Therefore,$N-1/2$ Models to be discarded outside of $N-1$ must suppress the extremity of the circular convolution result on both sides. Before it appears that the overlap save method that remove the first $N-1$ models of $y(n)$ was non appropriate meant for the routine of zero stage filters. Therefore, the Modified over save methods were

$$\left(1-\frac{\frac{N-1}{12}}{\frac{N-1}{2}}=\right)\frac{5}{6}$$

$$\left(1-\frac{\frac{N-1}{6}}{\frac{N-1}{2}}=\right)\frac{4}{6}$$

described as working theory. $N-1/2$ samples in shown Figs. 3 and 4 every end of the circular convolution. The circular convolution lead $y1(k)$ rotation does not change the group delay induced by the filter .so MOSM' provide the same result as acquired from OSM or another conventional filtering method after that, it will look, Reduction of group delay can be obtained by redefining the samples to be retained from the end rotation of the outcome is a circular convolution. The RDMOSM case reduces Group delay may behalf that detailed that technique one where group delay decrease can be obtained by more than half of the problem. When performing circular convolution based on DFT $N-1$ zero samples should be consider as previous to $x(n)$ DFT based $y1(k)$ the OSM $h(n)$,with After circular convolution of x(k), it remove $5(N-1)/6$ then $1/2(N-1)/2=$ 17 after the left end of the sample, and $(N-1)/6$ then $N-1)/12$ after the correct in of the samples, respectively, for ERDMOSM1 and ERDMOSM2. It processes on the left $(N-1)-5(N-1)/6=(N-1)/6$ samples in the event of ERDMOSM1 then $(N-1)-11(N-1)/12=(N-1)/12$ models in the ERDMOSM2 case, which in particular will have been removed according to overlap-save method passages. On the other side $5(N-1)/6$ or $11(N-1)/12$ group delay by linear phase of the in the ERDMOSM1 and ERDMOSM2.Group delay in this results decrease proscribe OSM compression factor, respectively, in the ERDMOSM1 case or ERDMOSM2.The findings obtain from two these algorithms were better than standard RDMOSM' Figs 6 and 7 Demonstrate an instance of a filtered signal in the actual period) where the group delay was reduced by factor of $5/6$ and $4/6$ for ERDMOSM1 and ERDMOSM2 respectively. It has been demand 70 the ensuing outcome signal starting since 35^{th} models into event by OSM'(i.e., group delay was $N-1/2$ samples) where as in the state of RDMOSM series starting since the sample (i.e. group delay was sample in ZDMOSM, he is resulting output series of the 18^{th}(i.e., group delay was $1/2(N-1)/2=$ 17 models). Contrary to this, the ZDMOSM output starting varies took a lot of the OSM filtered stand of comparable contribution. So, it deviated from the consequences of linear convolution. In the current's equation ERDMOSM1 the filtered output series starting to 15^{th} models (i.e., group delay was ($\frac{2}{5}\times\frac{N-1}{2}=14$ Models) Also for ERDMOSM2The resulting Filtering output Sequence Samples Starting after seven samples (i.e., since 8^{th} samples (i.e., group delay was ($\frac{1}{5} \times \frac{N-1}{2} = 7$ models). eliminate a lot of sample on the left-hand side of the circular convolution result in our current job, it has been seen that whether it is, group delay will have greatly decreased on the cost of increasing ripple amplitude. Whether it will be the $N - 1$ sample since the left side of the now it will be discussed the cause for which it is removed $5(N - 1)/6$, $11(N - 1)/12$ model numbers since the left then $(N - 1)/6$,$N - 1)/12$ sampler numbers since the right. Redefining the samples to be following Circular convolution, three necessary conditions must be followed up circular convolution consequence, what group delay will have been perfectly pressed since there will be no samples with group delay which, on the other side, will increase the deviations from the resultant filtered series since the overlap save method outcome. Therefore, the outcome will be inaccurate. Therefore, it had the outcome of circular convolution to obtain the better output to remove samples from both sides.

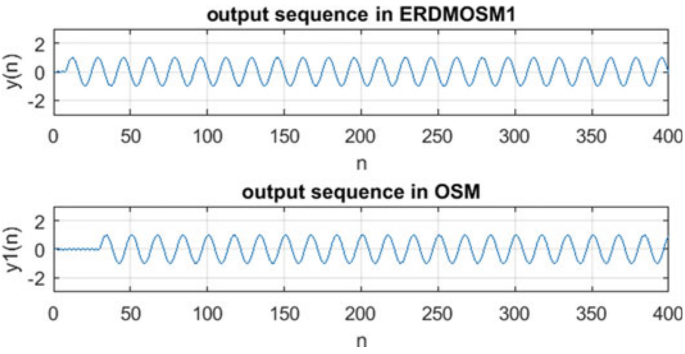


Fig. 3 Example of zero group delay filtration with ERDMOSM1

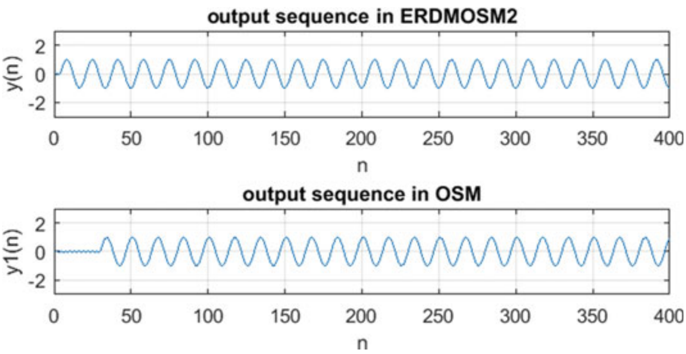


Fig. 4 Examples of reduced group delay filtration with ERDMOSM2

3 Result and Discussion

Examples will be shown from division to division to show the impacts of group delay decrease on top of the lately suggested algorithm information. The traditional OSMs comparable fresh filtering technique, the results of condensed delay are always compared to the error signal displayed by and can be written as,

$$e(n) = y1(n) - y(n) \tag{1.k}$$

where $y1(n)$ is the filtered output, it is assumed that overlap save method and $y(n)$ are the filtered output for each new filtering method. The zero-stage impulse response is shown as $h(n)$ and the filtering signal is shown as $x(n)$. For instance, as its input, audio signal and voice signal were intended to be respectively overlapped.

3.1 Examples 1

It has implemented the FIR lowpass equiripple Series 70 filter. The impulse response was shown in Fig. 5**.** Comparing The linear convolution method and OSM. Because two methodologies were same, OSM' was therefore evaluated because the references in the Fig. 6 division show that MOSM' and OSM' filtered performance in the event of processing off-line. It showed the upper to lower input, the ZERO DELAY MODIFIED OVERLAP-SAVE METHOD the result, the result of the OVERLAP-SAVE and the error signal, between the two-process outcome .A result is very different since the findings of diagram demonstrate that the(ZDMOSM')By taking into consideration comparable instances $x(n)$ and $h(n)$ because In real-time the comparable finding we rea usual Fig 7 a factor $1/2$ of another Comparable Filtering method. The aim, linear convolution must be used to approximate the addition of Modified Overlap-Save Method by reducing group delay. Nevertheless, the method needs zero-phase filters acquired via circular moves in a linear phase IR. In Figs. 8 other 9. Two Better ones Reduce delay Modified Overlap-save Methods. I e., Enhance Reduce Delay modified Overlap-Save Method1, Enhance Reduce Delay Overlap Save Method2 and the Overlap Save Method results have been plotted and comparable by it. From that figure, it can be observed that the ripple amplitude has improved the delay reduction on the cost. Here, the method Enhance Reduce Delay modified Overlap-SaveMethod1and Enhance Reduce delay Overlap-Save Method2 delays were decreased by $5/6$ and $4/6$ respectively for at the same Time, however, Ripple Amplitude Are Somewhat Higher Compare to Reduce Delay Modified Overlap- save Method.

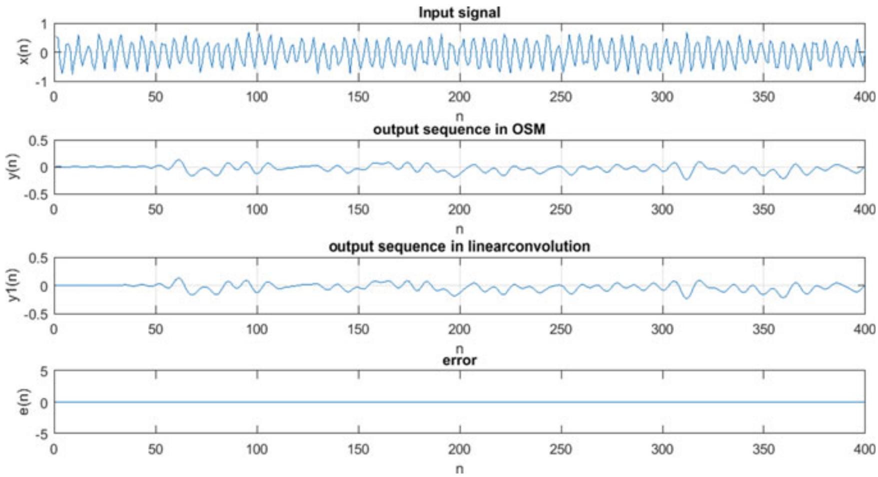


Fig. 5 Compare linear convolution with OSM technique in example

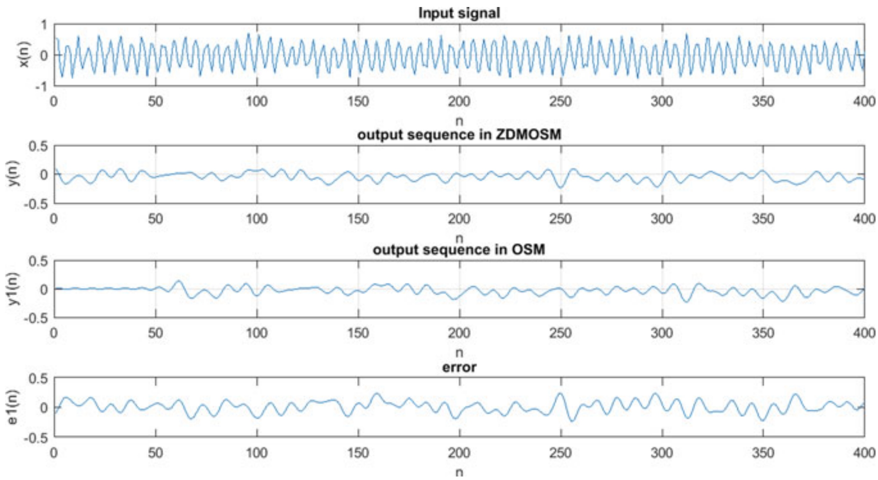


Fig. 6 Compare linear convolution with ZDOSM technique in Example 1

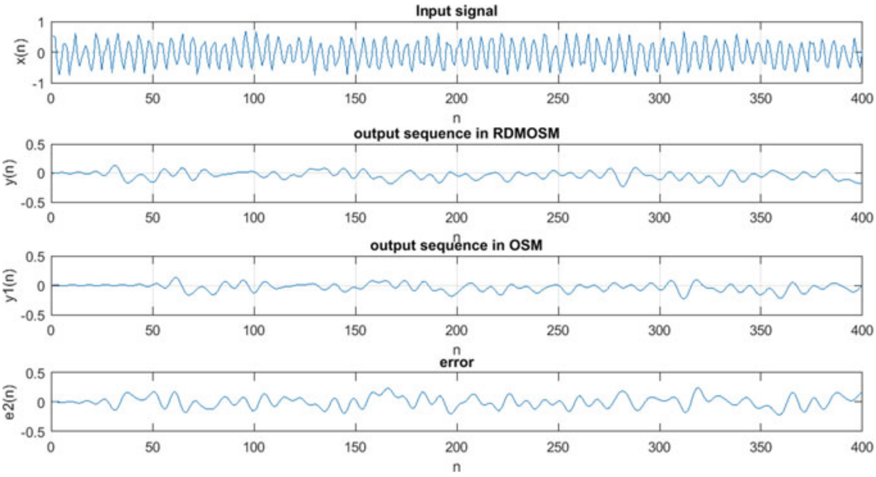


Fig. 7 Compare linear convolution with RZDOSM technique in Example 1

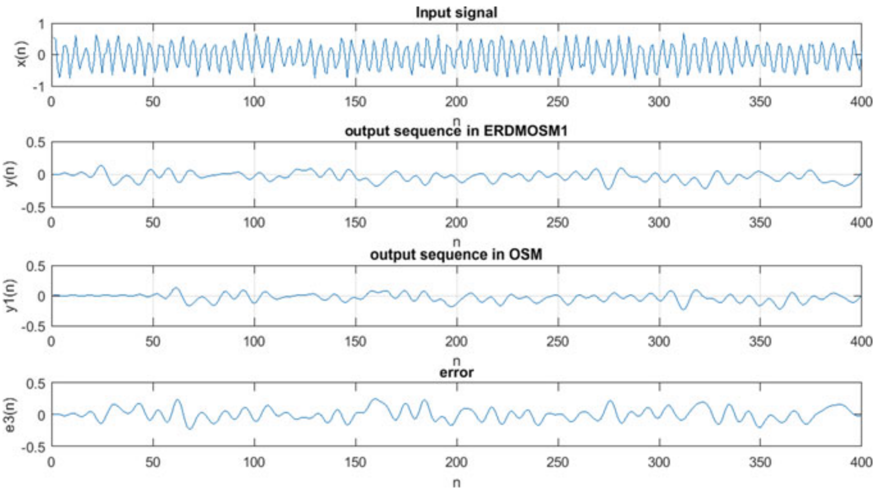


Fig. 8 Compare the OSM linear convolution with the ERDMOSM2 technique in Example 1

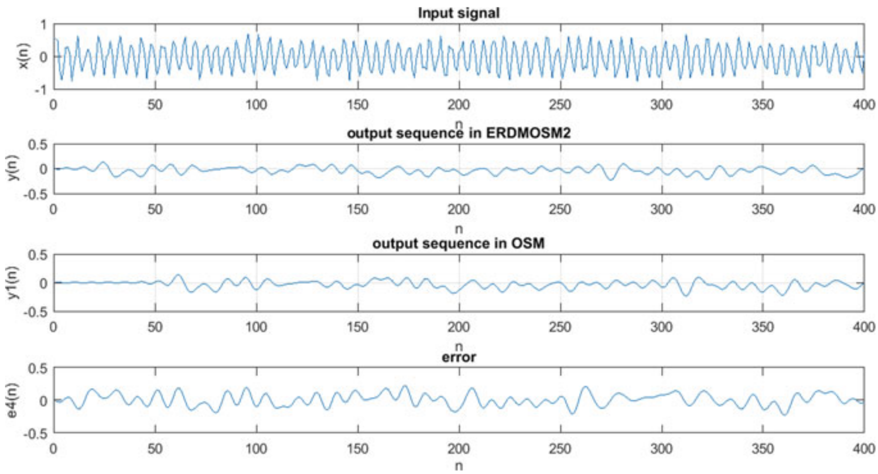


Fig. 9 Compare the OSM linear convolution with the ERDMOSM2 technique in Example 1

4 Conclusion

In this work, two new and enhanced filtering algorithms have been presented. These algorithms were only introduced with zero stage filters, the circular shift of overlap-save method. Comparison have been made with reduce delay Modified Overlap- save Method where reduction group delay has been obtained by issue 1/2 compared to overlap-save method Enhance Reduce delay modified Over-lap Save Method1 and Enhance Reduce delay Overlap-Save Method 2 which allows group delay to decrease by issue 5/6 and 4/6 respectively. The results have been verified and are found to conform to the said objectives.

References

1. Rabiner, B., McGonegal, C.: An approach to th approximation problem for no recursive digital filters. IEEE Trans. Audio Electro Acoust **18**(2), 83–106 (1970)
2. Rabiner, L.: Linear program design of finite impulse response (FIR) digital filters. IEEE Trans. Audio Electro Acoust. **20**(4), 280–288 (1972)
3. Fonda, J.S., Armand, E., et al.: Toward a group delay reduction digital filtering. Digital Signal Process. **19**(1), 22–32 (2009)
4. Kene, J.D.: Extendedoverlap-saveandoverlap-addconvolutionalgorithmsforrealsignal. In: 2007 IET-UK International Conference on Information and Communication Technology in Electrical Sciences (ICTES 2007). IET (2007)
5. Apaydin, G.: Realization of reduced-delay finite impulse response filters for audio applications. Digital Signal Process. **20**(3), 620–629 (2010)
6. Shyu, J.J., Pei, S.-C.: A generalized approach to the design of variable fractional-delay FIR digital filters. Signal Process. **88**(6), 1428–1435 (2008)

7. Damera-Venkata, N., Evans, B.L., McCaslin, S.R.: Design of optimal minimum-phase digital FIR filters using discrete Hilbert transforms. IEEE Trans. Signal Process. **48**(5), 1491–1495 (2000)
8. Bai, M.R., Lin, Y., Lai, J.: Reduction of electronic delay in active noise control systems-a multirate signal processing approach. J. Acoust. Soc. Am. **111**(2), 916–924 (2002)
9. McClellan, J., Parks, T.W., Rabiner, L.: A computer program for designing optimum FIR linear phase digital filters. IEEE Trans. Audio Electro Acoust. **21**(6), 506–526 (1973)
10. Ogata, A., Aikawa, N., Sato, M.: A design method of low delay fir band pass filters. In: 2000 IEEE International Symposium on Circuits and Systems. Emerging Technologies for the 21st Century. Proceedings (IEEE Cat No. 00CH36353), vol. 1. IEEE (2000)
11. Selesnick, I.W., Burrus, C.S.: Maximally flat low-pass FIR filters with reduced delay. IEEE Trans. Circ. Syst. II Analog Digital Signal Process. **45**(1), 53–68 (1998)
12. Vetterli, M.: Running FIR and IIR filtering using multirate filter banks. IEEE Trans. Acoust. Speech Signal Process. **36**(5), 730–738 (1988)
13. Vaidyanathan, P.P.: Multirate Systems and Filter Banks. Pearson Education, Delhi (1993)

Performance Analysis of Vedic Multiplier Using High Performance XOR-MUX Based Adder for Fast Computation

Aishita Verma, Anum Khan, and Subodh Wairya

Abstract This work presents design of a Vedic multiplier using Full adder (FA) based on XOR-MUX module. For that an energy-efficient 6T XOR cell is designed which utilizes 2:1 MUX to implement a full adder circuit. The performance of proposed full adder is compared with some existent FAs in terms of Average Power, Delay and PDP. The scalability of this adder is established by implementing and analysing 4 bit, 16 bit and 32 bit Ripple carry adder configuration. The 4 × 4 Vedic multiplier is designed using Urdhwa and Tiryakbhyam algorithm of multiplication and the performance is checked with respect to existing VM structures. The Vedic multiplier shows upto 74.26% reduction in PDP as compared with existing designs of Vedic multipliers. Furthermore, the 4 × 4 Vedic multiplier design is also extended to 8 × 8 Vedic multiplier design. All simulations are analyzed using Cadence Virtuoso using 45 nm technology over 0.8–1.2 V voltage range for Vedic multiplier and 0.8-V for proposed Full adder. Monte Carlo simulation and process corners analysis is performed of Vedic multiplier to establish its robustness.

Keywords Full adder · RCA · Monte-Carlo simulation · Corner analysis · Vedic multiplier

1 Introduction

The rapid growth of portable systems, such as laptops, has accelerated research in microelectronics. As a result, low-power design is primary design priority. Full adders are generally designed using a number of criteria. Multipliers, subtractors, differentiators, comparators, integrators, and other circuits use a 1-bit adder cell as a basic component. The number of transistors in an arithmetic circuit, such as a multiplier or an Arithmetic Logic Unit (ALU) impacts its complexity. Speed and power consumption are two other essential factors to consider when designing complete adders [1].

A. Verma (✉) · A. Khan · S. Wairya
Department of Electronics and Communication Engineering, Institute of Engineering and Technology, Lucknow, India
e-mail: aishita.verma19@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_60

They have a completely contradictory relationship. As a result, energy consumption per operation or the power delay product is introduced to achieve optimal design tradeoffs. The performance of digital circuits could be improved by selecting the suitable logic styles. Many complete adder architectures are addressed in the literature, including static CMOS, dynamic circuits, Pass Transistor Logic (PTL), transmission gates and GDI logic [2]. For sum and carry, a static CMOS adder with complementary PMOS and NMOS networks requires 28 transistors. PTL implementations are an alternative to CMOS implementations [3]. It requires fewer transistors, which reduces overall capacitances, increasing speed and lowering power, but voltage at output differs in PTL design due to threshold voltage drop across input and output. GDI logic, a low-power design technique using fewer transistors, is an alternative to CMOS logic [4]. When using cascaded operations, the switching is slow. Special attention is required to achieve full swing operations [5]. The primary component in the design of a multiplier is adder, which has notable impact on multiplier performance .As a result, full adder design is a major concern [9, 10]. A high-speed XOR-MUX based FA is presented in this paper. First XOR module is proposed in this paper. The XOR module output signals are then used to implement the carry, sum-generation circuits. The proposed FA performance has been thoroughly evaluated by comparing it to the performance of existent FAs. According to results of the simulation proposed FA and 32-bit RCA showed superior performance. The entire paper is organized as follows. In Sect. 2, Proposed adder with 6T XOR circuit and 2:1 MUX is designed. In Sect. 3 Vedic multiplier is described. The simulation results are analyzed and compared in Sect. 4 with Conclusion in Sect. 5.

2 Proposed XOR-MUX Based Full Adder

The full adder is used for arithmetic operations including multiplications, divisions [11, 12] and ALUs in digital signal processing applications [12]. This work presents full adder designed with two stages of XOR and one 2:1 MUX. Figure 1 depicts the XOR gate and 2:1 MUX schematic. Figure 2 depicts the proposed block diagram, which uses two consecutive XOR gates for Sum and a 2:1 MUX for Carry output. Table 1 depicts the truth table.

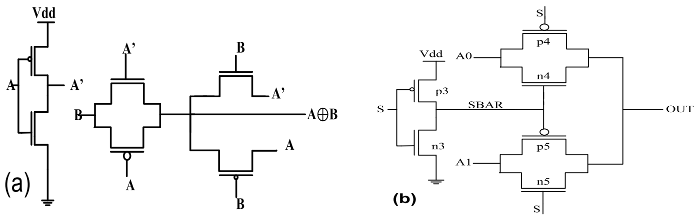


Fig. 1 **a** Full-swing XOR gate **b** 2:1 MUX

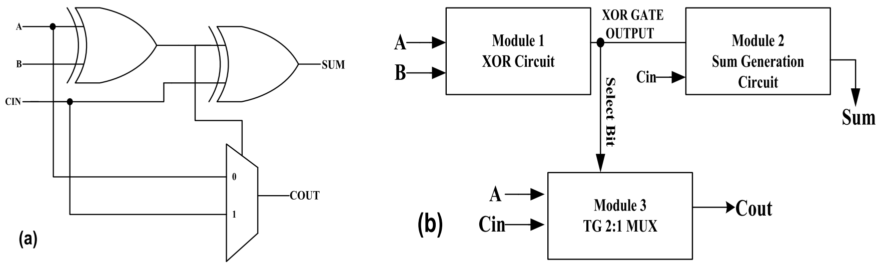


Fig. 2 Proposed XOR-MUX Full adder **a** Schematic **b** Block diagram

Table 1 Truth table of FA

Cin	A	B	Sum	Count
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

One of the most essential modules for realizing complete adder architecture is the XOR gate. To achieve complete swing in output, presented XOR gate requires 6T. The schematic is depicted in Fig. 1(a). The following is a description of how the suggested XOR gate works: For A = 1, B = 0 output is A, and for A = 1, B = 1 the $\overline{A}$. is obtained at output. The second XOR gate consists of carry input and XOR output obtained from first XOR gate. The output = A ⊕ B for Cin = 0, if Cin = 1 with A ⊕ B = 0, C in is the output, when Cin = 1, A ⊕ B = 1, $\overline{\text{Cin}}$ is the obtained output. Figure 1(b) depicts schematic of TG based 2:1 MUX [6].

The major goal is to minimize circuit complexity while increasing speed of operation. For the SUM output, which is one of the three inputs, the two XOR modules in cascade are required. The implementations of a XOR module necessitate the use of TGs and pass transistor logic. The most essential element of FA, as indicated in Fig. 2(b), is the carry generating module. The Cout of one FA is used as Cin signal for next stage in extended word length adders. As a result, the FA Cout signal requires sufficient voltage level. The Cout is determined using a TG-based MUX with A as the input and Cin as the select bit. 2:1 MUX designed using Transmission Gate (TG) for implementing the carry generation circuit. In Fig. 2(b), the Sum generation circuit is implemented with 2:1 MUX using TG1 (composed of p4 and n4) and TG2 (composed of p5 and n5). Because TG is capable of passing both strong logic 0 and 1, with full swing output.

3 Vedic Multiplier

By combining Vedic mathematics with multiplication, the computation time may be reduced [13]. In ancient Indian method of calculation the Vedic Mathematics for performing multiplication comprises sixteen sutras. "Urdhwa and Triyakbhyam" is word of Sanskrit, meaning vertically and crosswise in the order given. The algorithm of multiplication intensifies speed of computation as it performs the concurrent addition and generation of partial product. In the present work above approach is well-chosen for generating 4 × 4 multipliers [14]. The algorithm of multiplication that is Urdhwa and Triyakbhyam for multiplication of 4-bit is exemplified with diagram manifested in Fig. 3 with schematic in Fig. 4. The steps will be reduced which are requisite for computing final product in multiplication using Urdhva Tiryagbhyam Sutra. This enhances speed with reduced computation delay.

The steps mentioned below are followed in Vedic multiplier generally,

Step 1: The formation of final end result of LSB's is by multiplier vertical multiplication with multiplicand's (LSB).

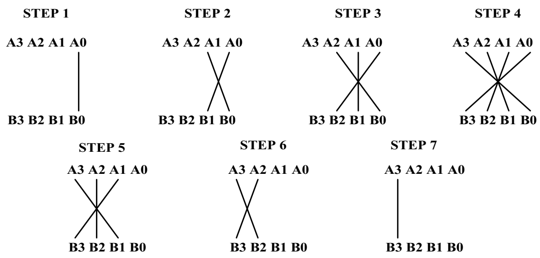


Fig. 3 Method of multiplication [14]

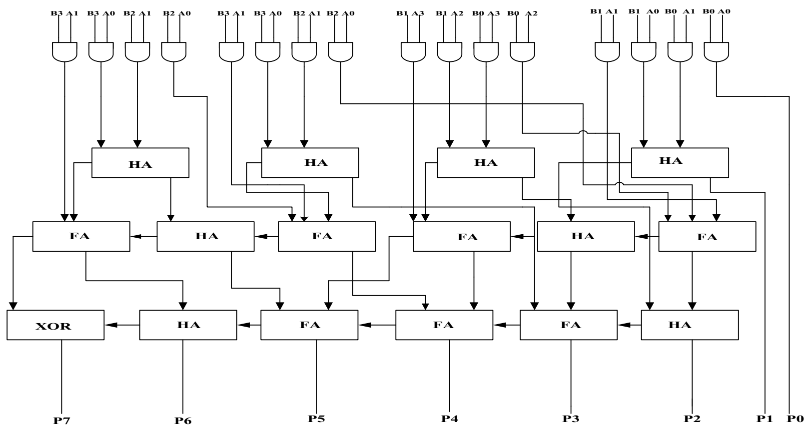


Fig. 4 4 × 4 Vedic multiplier

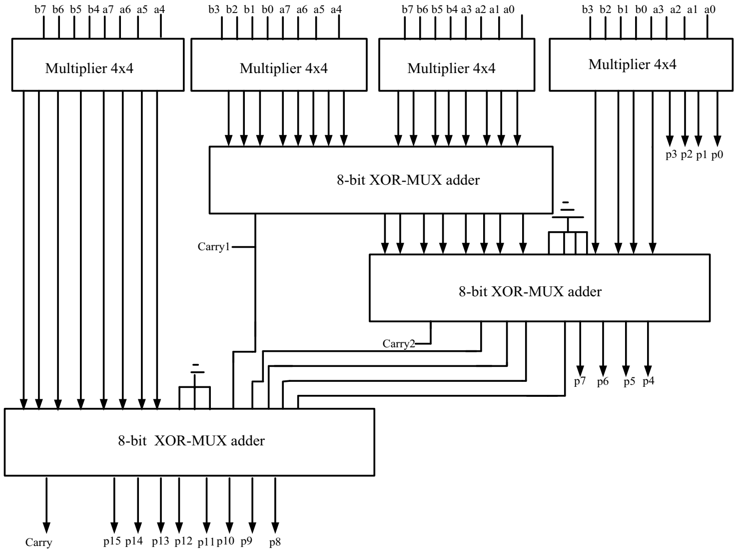


Fig. 5 8 × 8 Vedic multiplier

Step 2: There is a multiplication of LSB of multiplier with MSB of multiplicand and the multiplication of MSB of multiplier with LSB of multiplicand in the crosswise manner and addition is done of these products. With this procedure of adding second bit Final end result is stated.
Step 3: In the vertical aspect, the multiplication is performed of Multiplier and multiplicand's MSB.

8 × 8 Vedic multiplier [15] is also designed by extending 4 × 4 Vedic multiplier designs with 8-bit RCA designed with the Proposed XOR-MUX Full adder. The schematic of 8 × 8 Vedic multiplier is shown in the Fig. 5.

4 Results and Discussion

The simulation was executed utilizing Cadence virtuoso 45-nm over 0.8–1.2 V voltage range. The proposed FA circuit and designed Vedic multiplier performance is studied and compared for parameters such as power, delay, PDP, and EDP.

4.1 Performance Analysis of Adder

The XOR gate utilized in XOR-MUX full adder is fast and energy efficient. As a result, using the suggested XOR gate can improve performance metrics while also providing good driving capabilities for following stages. The total power of XOR gate at 1.0 V supply is 0.029 µW and delay is 42.024 ps, PDP is 1.25 aJ. The Simulation waveform is in Fig. 6.

The simulation waveform of 2:1 MUX is shown in Fig. 7. The total power of 2:1 MUX is 15.539 nW at 0.8-V. The full swing simulation waveform of full adder is depicted in Fig. 8 using 1 V supply to examine and validate the operation of proposed FA.

The average power consumption of FA is 0.093 µW, delay is 39.7312 ps, PDP is 3.69 aJ and EDP is 146.608 $\times$ 10^{-30} J at 1-V supply. Performance of proposed FA with existent full adders is studied with 0.8-V supply.

Table 2 depicts the obtained simulation results at 0.8-V supply and Fig. 9 depicts the Performance comparison of FA.

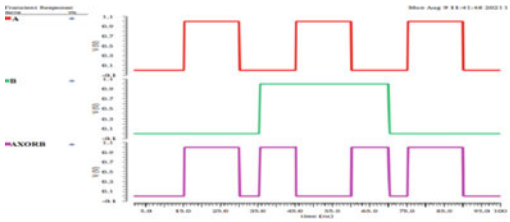


Fig. 6 XOR gate waveform

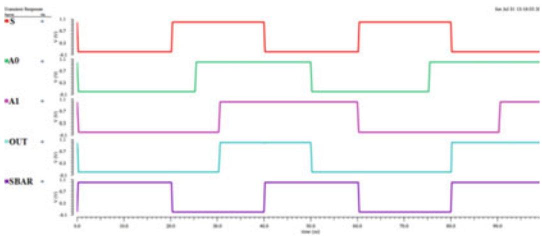


Fig. 7 Simulated waveform of 2:1 MUX

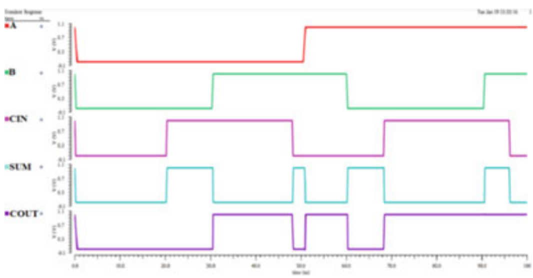


Fig. 8 Simulated waveform of full adder

Table 2 Full adder performance analysis

Full adder cell	TC	Power	Delay	PDP
CPL [16]	32	1.72	84.8	145.9
CCMOS [17]	28	0.68	125.8	85.5
TFA [18]	16	0.61	145.6	88.8
DPL [19]	22	0.87	91.9	79.9
HYBRID 1 [20]	24	0.78	189.1	145.6
HYBRID 2 [21]	16	0.35	231.4	81
GDI 1 [22]	18	0.46	98.8	43.5
GDI 2 [22]	22	0.63	77.3	48.7
PROPOSED FA	18	0.09	59.52	5.48

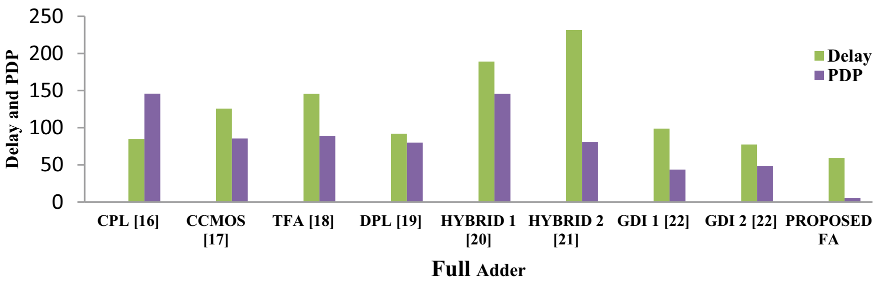


Fig. 9 PDP and delay comparison of FA

With the results of simulation presented in Table 2, it is concluded that proposed FA requires lesser amount of power, delay and PDP among single logic FAs.

The proposed FA showed excellent performance in RCA structures because its delay and PDP are lowest as shown in Table 3.When extended to 32-bit RCA, percentage improvements in 93.30% in Power, 32.95% in Delay, and 95.523% in PDP are obtained compared with CCMOS FA. In order to determine the driving capability simulation test bench is applied. Figure 10 depicts schematic of test bench for 1-bit FA. Results obtained from the circuit are summarized in Table 4.

With the drawn layout the area of Proposed 1-Bit FA adder is calculated. The calculated area is 27.11 μm^2 and Fig. 11 depicts the layout.

Table 3 Results of simulation of RCA

Full adder	Power (μW)				Delay (ps)				PDP (fJ)			
	4-bit	8-bit	16-bit	32-bit	4-bit	8-bit	16-bit	32-bit	4-bit	8-bit	16-bit	32-bit
CCMOS [17]	2.42	4.68	9.48	20.79	510.6	1042.5	2105.8	4233.2	1.24	4.88	19.96	88.01
Hybrid FA [20]	2.81	5.47	10.77	22.69	767.8	1552.5	3485.4	7057.8	2.16	8.49	37.54	160.1
GDI 1 [22]	2.08	4.38	9.4	20.16	345.9	743.6	1621.2	3550.6	3.26	15.24	71.58	345.2
Proposed FA	0.253	0.499	0.580	1.391	211.9	759.8	1463	2838	0.053	0.341	0.848	3.94

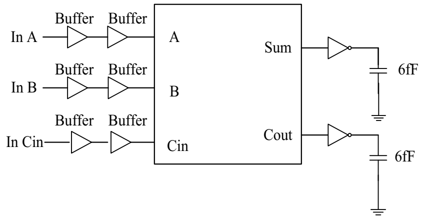


Fig. 10 Test bench for 1-bit FA driving capability analysis

Table 4 Results of simulation of full adder with load capacitance

Load capacitance (fF)	Average power (μW)	Delay (ps)	PDP (aJ)
2	0.249	196.01	48.80
4	0.365	227.6	83.07
6	0.481	257.3	123.7
10	0.695	314.6	218.6
50	2.907	815.3	2370
100	5.468	1366	7469.2

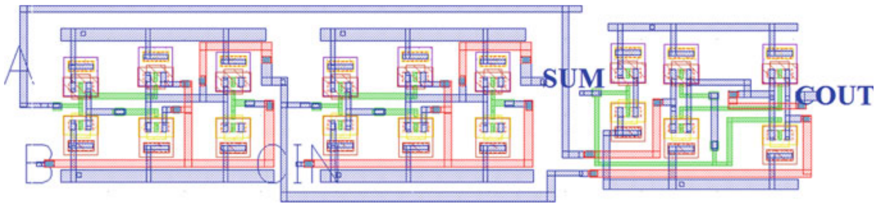


Fig. 11 Layout of 1-bit FA

4.2 Performance Analysis of Vedic Multiplier

The 4 × 4 multiplier is designed with FA and half adders. The half adder designed using AND gate and XOR gate. The simulation waveform of is depicted in Fig. 12.

Total power of Half adder is 0.0409 μW at 0.8-V supply. The product terms are P7-P0. Table 5 depicts result of simulation at different supply voltages.

The results of comparison of proposed Vedic multiplier with other designs [23] and [24] are shown in Table 6.

Table 6 shows that 4 × 4 Vedic multiplier has lowest power consumption, delay, PDP and EDP at 1-V supply indiscriminate to existing 4 × 4 multiplier designs.

The designed Vedic multiplier operates with lesser power consumption as shown in Fig. 13(a) The power saving in Vedic multiplier accomplished is 73.66, 61.61% as compared with existent Vedic multiplier in [23] and [24]. As seen from PDP comparison results given in Fig. 13(b) the Vedic multiplier depicts 74.26, 65% reduced PDP

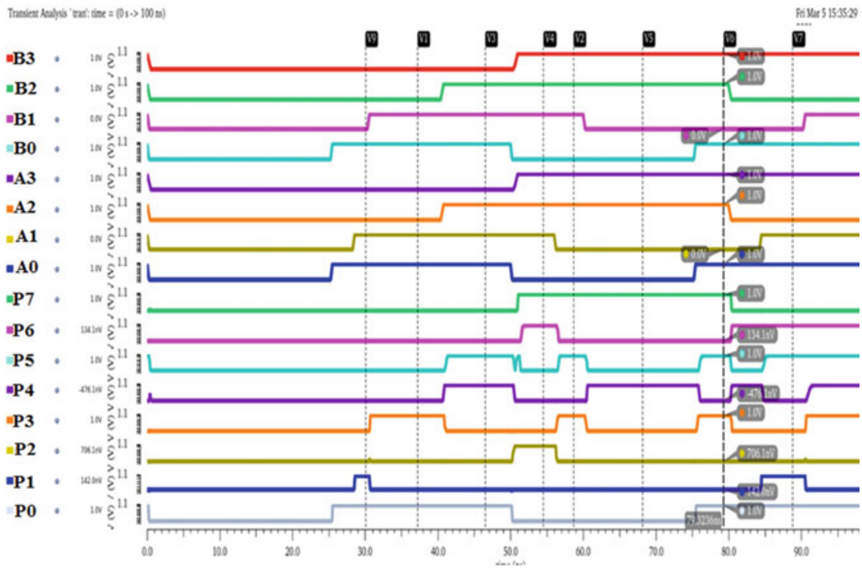


Fig. 12 Simulation waveform of Vedic multiplier for eg. 1101 × 1101 = 010,101,001

Table 5 Results of simulation of 4 × 4 Vedic multiplier at 0.8–1.2 V supply

VDD	Vedic multiplier			
	Power (μW)	Delay (ps)	PDP (aJ)	EDP (10^{-27} J)
0.8	0.627	465.31	291.75	135.75
0.9	0.801	343.53	275.16	94.52
1 V	0.998	244.34	244.07	59.63
1.1 V	1.227	154.1	189.08	29.13
1.2 V	1.540	60.07	92.50	5.55

Table 6 Results of comparison of Vedic multiplier at 1-V supply

Design	Power (μW)	Delay (ps)	PDP (aJ)
Multiplier [23]	3.79	250	947.5
Multiplier [24]	2.6	268	696.8
Vedic multiplier	0.998	244.34	244.07

as compared with design in [23] and [24]. The corner analysis results are depicted in Table 7 with nominal, minimum, maximum values observed at different corners. There is no deviation in Power, PDP and EDP but the delay alters from SS to TT corner. With miniaturization of the device dimensions, as the technology advances, it is requisite to observe the process variation analysis of circuits. Monte Carlo is the statistical analysis for determining the robustness of the circuit. Hence, Monte Carlo simulation with 200 samples is performed.

The Monte Carlo results of simulation of designed Vedic multiplier are displayed in Fig. 14 and Table 8. The results of (μ/σ) depicts that designed Vedic multiplier with XOR-MUX based adder has 6.9% performance variation in delay with full swing output.

The total power consumption of 8 × 8 Vedic Multiplier is 7.2601 μW, delay is 601.575 ps, PDP is 4.36 fJ and EDP is 2.62 × 10^{-24} J at 1-V supply.

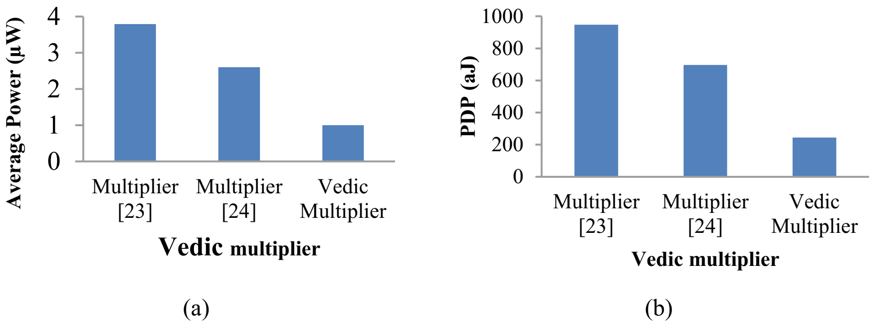


Fig. 13 Power and PDP comparison

Table 7 Performance summary at different corners of Vedic multiplier

OUTPUT	NOMINAL	MIN	MAX	SS	SF	FS	FF	TT
Power (nW)	585.3	585.3	585.3	585.3	585.3	585.3	585.3	585.3
Delay (ps)	425.2	283.9	714.2	714.2	383.2	527.7	283.9	425.2
PDP (aJ)	248.7	248.7	248.7	248.7	248.7	248.7	248.7	248.7
EDP (10^{-27} J)	105.7	105.7	105.7	105.7	105.7	105.7	105.7	105.7

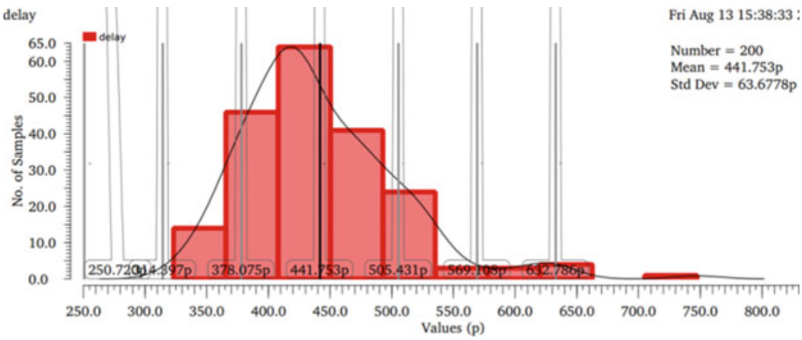


Fig. 14 Monte Carlo histogram of delay simulation with 200 samples

Table 8 Monte Carlo simulation results of Vedic multiplier

Output	Min	Max	Mean (μ)	Std. Dev. (σ)
Power (nW)	585.3	585.3	585.3	585.3
Delay (ps)	323.1	747.4	441.8	63.68
PDP (aJ)	248.7	248.7	248.7	248.7
EDP (10^{-27} J)	105.7	105.7	105.7	105.7

5 Conclusion

In this paper a full-swing, Low-Power XOR-MUX based adder is designed. For comparison and analysis Cadence Virtuoso 45-nm technology was used. The results of simulation authenticate the efficacy of design. Compared with CCMOS FA, the proposed FA shows 86.764% improvement in Power, and 52.68% improvement in Delay at 0.8-V. When extended to 32-bit RCA, percentage improvements in 93.30% in Power, 32.95% in Delay, and 95.523% in PDP are obtained compared with CCMOS FA. The calculated area of proposed 1-bit FA is 27.11 μm^2. In addition to this work 4×4 and 8×8 Vedic multiplier has been designed with proposed FA. Meticulous Monte Carlo analysis is done for understanding robustness of Vedic multiplier. The Vedic multiplier design reveals remarkable results in terms of Power, Delay and PDP. The Vedic multiplier shows 74.26, 65% reduced PDP as compared existent designs. It is notable that Vedic multiplier designed with XOR-MUX FA consumes less power and has high speed and therefore, is appropriate for VLSI circuits. In future, this design can be used in FIR filters and MAC units.

References

1. Usha, S., Ravi, T.: Design of 4-BIT ripple carry adder using hybrid 9T Full adder. In: 2015 International Conference on Circuits, Power and Computing Technologies [ICCPCT 2015], pp. 1–8 (2015). https://doi.org/10.1109/ICCPCT.2015.7159490
2. Shams, A.M., Bayoumi, M.A.: A novel high-performance CMOS 1-bit full-adder cell. IEEE Trans. Circ. Syst. II Analog Digit. Signal Process. **47**(5), 478–481 (2000). https://doi.org/10.1109/82.842117
3. Lin, J., Hwang, Y., Sheu, M., Ho, C.: A novel high-speed and energy efficient 10-transistor full adder design. IEEE Trans. Circuits Syst. I Regul. Pap. **54**(5), 1050–1059 (2007). https://doi.org/10.1109/TCSI.2007.895509
4. Bui, H.T., Wang, Y., Jiang, Y.: Design and analysis of low-power 10-transistor full adders using novel XOR-XNOR gates. IEEE Trans. Circ. Syst. II Analog Digit. Signal Process. **49**(1), 25–30 (2002). https://doi.org/10.1109/82.996055
5. Mukherjee, B., Ghosal, A.: Design & study of a low power high speed full adder using GDI multiplexer. In: 2015 IEEE 2nd International Conference on Recent Trends in Information Systems (ReTIS), pp. 465–470 (2015). https://doi.org/10.1109/ReTIS.2015.7232924

6. Murugeswari, S., Mohideen, S.K.: Design of area efficient and low power multipliers using multiplexer based full adder. In: Second International Conference on Current Trends In Engineering and Technology - ICCTET 2014, pp. 388–392 (2014). https://doi.org/10.1109/ICCTET.2014.6966322
7. Jiang, Y., Al-Sheraidah, A., Wang, Y., Sha, E., Chung, J.-G.: A novel multiplexer-based low-power full adder. IEEE Trans. Circ. Syst. II Express Briefs **51**(7), 345–348 (2004). https://doi.org/10.1109/TCSII.2004.831429
8. Bandi, V.: Performance analysis for Vedic multiplier using modified full adders. In: 2017 Innovations in Power and Advanced Computing Technologies (i-PACT), pp. 1–5 (2017). https://doi.org/10.1109/IPACT.2017.8245017
9. Yadav, A.K., Shrivatava, B.P., Dadoriya, A.K.: Low power high speed 1-bit full adder circuit design at 45nm CMOS technology. In: 2017 International Conference on Recent Innovations in Signal Processing and Embedded Systems (RISE), pp. 427–432 (2017). https://doi.org/10.1109/RISE.2017.8378203
10. Saraswat, V., Kumar, A., Pal, P.K., Nagaria, R.K.: A survey on different modules of low-power high-speed hybrid full adder circuits. In: 2017 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON), Mathura, India, pp. 323–328 (2017). https://doi.org/10.1109/UPCON.2017.8251068
11. Ahmad, N., Hasan, R.: A new design of XOR-XNOR gates for low power application. In: 2011 International Conference on Electronic Devices, Systems and Applications (ICEDSA), Kuala Lumpur, Malaysia, pp. 45–49 (2011). https://doi.org/10.1109/ICEDSA.2011.5959039
12. Basheer, B., Merin Varkey, R.: Review on various full adder circuits. In: 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), pp. 877–880 (2019). https://doi.org/10.1109/ICCMC.2019.8819680
13. Bansal, Y., Madhu, C., Kaur, P.: High speed vedic multiplier designs-a review. Recent Adv. Eng. Comput. Sci. (RAECS) **2014**, 1–6 (2014). https://doi.org/10.1109/RAECS.2014.6799502
14. Chudasama, A., Sasamal, T.N., Yadav, J.: An efficient design of Vedic multiplier using ripple carry adder in Quantum-dot Cellular Automata. Comput. Electr. Eng. **65**, 527–542 (2018). https://doi.org/10.1016/j.compeleceng.2017.09.019. ISSN 0045-7906
15. Jhamb, M., Garima, H.L.: Design, implementation and performance comparison of multiplier topologies in power-delay space. Eng. Sci. Technol. Int. J. **19**(1), 355–363. https://doi.org/10.1016/j.jestch.2015.08.006. ISSN 2215-0986
16. Zimmermann, R., Fichtner, W.: Low-power logic styles: CMOS versus pass-transistor logic. IEEE J. Solid-State Circuits **32**(7), 1079–1090 (1997). https://doi.org/10.1109/4.597298
17. Weste, N.H.E., Harris, D.M.: CMOS VLSI Design: A Circuits and Systems Perspective, 4th edn. Addison-Wesley, Boston (2010)
18. Alioto, M., Di Cataldo, G., Palumbo, G.: Mixed Full Adder topologies for high-performance low-power arithmetic circuits. Microelectron. J. **38**(1), 130–139 (2007). https://doi.org/10.1016/j.mejo.2006.09.001
19. Aguirre-Hernandez, M., Linares-Aranda, M.: CMOs full-adders for energy-efficient arithmetic applications. IEEE Trans. Very Large Scale Integr. (VLSI) Syst. **19**(4), 718–721 (2011). https://doi.org/10.1109/tvlsi.2009.2038166
20. Hassoune, I., Flandre, D., O'Connor, I., Legat, J.: ULPFA: a new efficient design of a power-aware full adder. IEEE Trans. Circuits Syst. I Regul. Pap. **57**(8), 2066–2074 (2010). https://doi.org/10.1109/TCSI.2008.2001367
21. Bhattacharyya, P., Kundu, B., Ghosh, S., Kumar, V., Dandapat, A.: Performance analysis of a low-power high-speed hybrid 1-bit full adder circuit. In: IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 23, no. 10, pp. 2001–2008, October 2015. https://doi.org/10.1109/TVLSI.2014.2357057
22. Shoba, M., Nakkeeran, R.: GDI based full adders for energy efficient arithmetic applications. Eng. Sci. Technol. Int. J. **19**(1), 485–496. https://doi.org/10.1016/j.jestch.2015.09.006. ISSN 2215-0986

23. Patel, C.R., Urankar, V., Vivek, B.A., Bharadwaj, V.K.: Vedic multiplier in 45nm technology. In: 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, pp. 21–26 (2020). https://doi.org/10.1109/ICCMC48092.2020.ICCMC-0004
24. Tripathy, S., et al.: Low power multiplier architectures using Vedic mathematics in 45nm technology for high speed computing. In: International Conference on Communication, Information and Computing Technology (2015)

Performance Estimation of Different Tunnel Field Effect Transistor Based Biosensors Used in the Biomedical and Its Future Prospective

Shilpi Gupta and Subodh Wairya

Abstract Using mathematical analysis, a comparative evaluation of several types of biosensors constructed using field impact semiconductor (FET) is conducted. A traditional back-gated device and three other nanogap-based devices are evaluated in terms of their exhibit measures, which include features such as reactivity, affectability, recognition cutoff, and dynamic reach. An electrostatic model is applied to address the detecting rule. The biochemical response is copied basically by an adjustment of the negative charge thickness and permittivity TFET constructed. Bio-sensors that are taking low force and better qualities due than Band-2-Band tunneling investigating of transporter and sharp sub edge-swing. Its original copy portrays undeniable insight concerning the TFET constructed bio-sensors directly by developing different gadget assessment to bio-sensor application that incorporates subjective and measurable boundaries examination learning like affectability boundaries with various elements influencing the affectability by looking at changed designs and the systems in question. During ongoing years, field-impact semiconductor biosensors-Bio-FE applicable in the field of biomedical has encountered various vigorous improvements in combination of advancements in different FET attributes just like the change of bioreceptor constructions. This audit at first gives thought on this hard work development by momentarily adding up astounding investigations on 2 previously stated observations. The previous research integrates manufacturing unusual nanostructures by using advanced materials for different FET transducers, though the latter fundamentally arranges minimal atomssassbio-tests (immunizer sections and aptamers). We recommend the applicable utilization of the investigating fields impact semiconductors (TFET) by way of a biosensor that further identifies bio-atoms arranged at the door oxide.

Keywords FET-biosensor · Biomedical-application · Nano transistor · FET · TFET · Tunneling · Bio-Sensors · Biomedical application field-effect transistor

S. Gupta (✉)
I.E.T., Lucknow, Lucknow, India
e-mail: shilpigupta.iet@gmail.com

S. Wairya
Dr. APJ Abdul Kalam Technical University, Uttar Pradesh, Lucknow, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_61

1 Introduction

The dangerous chimes of people are at high caution on account of the bio-assault that noticed form the most recent couple of many years' right structures the human immunodeficiency virus to introduce Corona-virus These viruses are undetectable and spreading with speedy speed deprived of the information on people & completed its life so hopeless. Along with, innovation enhancement has agreed method for the substitution of previous-style weapons with bio weapons generous extension for the bio-wars. These different bio-weapons comprise of pathogenic infection or microorganisms that escalate quietly & ended existences of guiltless individuals around forefront.

Meanwhile presented by author Clark in year 1962 [1], bio-sensors are taken generally utilized in various assorted submissions, for example, malignancy conclusion [2], poisonousness recognition [3], food investigation [4], wellbeing visualization [5], and consequently into view of Biosensors, logical devices altering over normal responses into electrical signals [6], classically contain of something like 2-essential segments consolidating organized: An organics receptor with an actual substance sensor [7]. The previous proselytes' responses from the Bio-schemical area, normally analytes fixation, interested in a substance or else real produce signal by a defined impressible, though the previous connections the symbol after the bio-acknowledgment framework's yield space, generally to the electrical area [7]. Field-impact semiconductor biosensors (Bio-FETs), for example, are one type of biosensor.

The biosensors dependent on particle touchy field impact semiconductors (ISFET) have been generally explored because of mark free, profoundly delicate and ongoing discovery of organic elements, like DNA, infection, and proteins. Different exploratory explanation is being accounted for applicable on the ISFET devices with various detecting ideas [1, 2].

Moreover, CMOS's-viable gadget constructions of sISFET instruments was gotten incredible consideration for great consistency, max reproducibility and minimal expense manufacture specially in large scale manufacturing [3, 4, 12]. These standard of electrical location in ISFET devices depends mainly at the gating impact of either negatively or positively excited bio-molecules at the entryway-oxide, i.e., being changed over straightforwardly into adjustment of electrical properties like flow, conductance & edge-voltages [5, 6].

The researcher's assembly has been modeled (SiNW) biosensors also known as Silicon-NanoWire biosensor which is entirely compatible by CMOS-based deliver circuits [7, 8]. Regular F-E-T (c-FET)-modeled bio-sensors, on the other hand, have faced theoretical limitations on the most extreme affectability due to MOSFET's inability to implement sub-threshold swing (SS)-60-mV/dec under ambient temperature conditions. TFETs (investigating Field impact semiconductors) have a low force activity and a high SS [9].

2 Device Structure with Simulation Settings

In Fig. 1 and Table 1, actual boundaries and gadget designs are shown. The suggested gadget uses a predictable p-I-architecture. Gadget's channel-Length with Width is 1 µm. Channel and source doping has been done by $2 \times 10^{\wedge}20\ cm^{-3}$. Boron & Arsenics individually in addition, fixation of channel doping are $1 \times 10^{\wedge}17\ cm^{-3}$. Additionally, Helmholtz layer having a couple of angstrom thickness is considered as dielectric layer [12]. Biomolecules are shown with certain negative charges on the door oxide. To validate the detecting components of target particles with charge, Site-restricting hypothesis is used. Negative/positive charges gets connected on the door oxide's source/channel in multi-detecting recreations. All recreations are led with Sentaurus-TM TCAD test system of synosyssInc.

2.1 Fundamental Procedure of TFET Device

Figure 2-a illustrates detecting system of a unique TFET-pH-sensor. At the point once pH increments & particles by negative charges are combined at the door oxide close to the intersection among source & channel, the channel's Energy level is elevated. It results that the possible boundary among Source & channel heavier. Subsequently, burrowing rate diminishes and burrowing-current is decreased.

Figure 2(b) illustrates the exchange bends for the reproduced TFET based pH device operated at various pH-levels. In this device, the channel inclination electrical energy is around 1 V. When the standardized channel current (Id per L/W) levels 10–7 A, the edge-voltage is separated by a steady current strategy and described as the entrance voltage.

In pH 7, the recommended TFET model has a Sub-threshold slope of roughly 45 mV per decade with a Vt around 0.23 V. As pH increases, the Vt changes in order

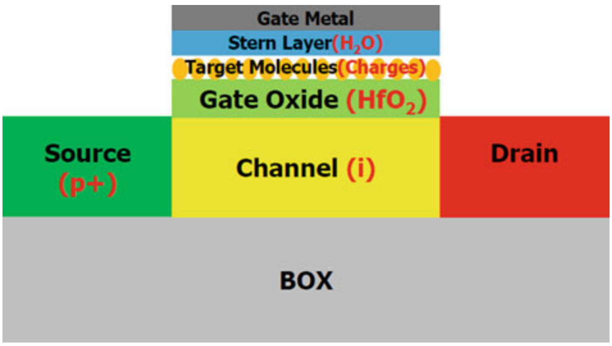
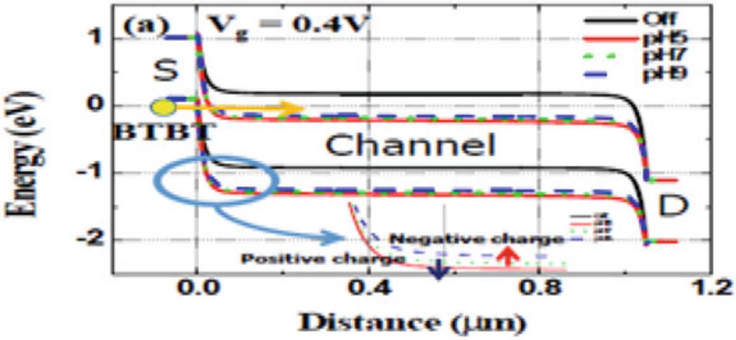


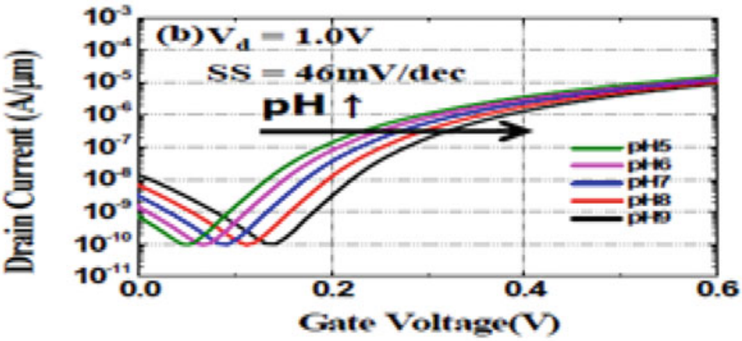
Fig. 1 Bio-Sensor-Device construction with TFET in the literature

Table 1 Parameter analysis of Various TFET with biosensor type application

Tfet types	Biosensors types	Sensitivity (s)	Improvements
(DM) designed with p-type TFET [18]	p-TFET biosensor	10ˆ9	Device optimization which enables sensing system of measurement to improved biosensor act
N+ pocket-doped vertical tunnel field-effect transistor (VTFET) [19]	Label-free biosensors	• 10ˆ6 = Sensitivity at k = 12 versus neutral biomolecules k-dielectric constant = 5, 7, 10, and 12	Label-free-Sensing with better sensitivity and minimum power consumption
Conventional TFET [20]	pH or biomolecule sensor (Multi-sensing Biosensors)	• SS = 45 mV/dec • Vt = 0.23 V	TFET sensors detects various molecules of opposite charges using single device
Si-NW-TFET [21]	Point-of-care-diagnostics-biosensor	SS = 30 mV/dec	Precise labels-frees finding of CYFRA21-1 value up to 0.5 fgml−1 or ~12.5 am, Real-time measure, skill for clinically Samples-ex-serum
DP Vertical Hetero-T FET [22]	CMOS constructed label-free biosensing	• Sensitivity = 26.78% Vs T pocket • 60.8% Vs L-pocket • 56% Vs N-pocket • and 40.6% with T ox	Improved truthful. Sensitivity estimate discloses that the proposed DP-DM-HTFET is optimistic candidate for DM-HTFET constructed label-free biosensing applications
Vertical dielectrically modulated tunnel field-effect transistor (V-DMTFET) [23]	Label-free-biosensor	• Ids sensitivity = 10ˆ • SS sensitivity = 0.50	V-DMTFET is fewer sensitive toward noise, ZrO2 has the maximum sensitivity



(a)

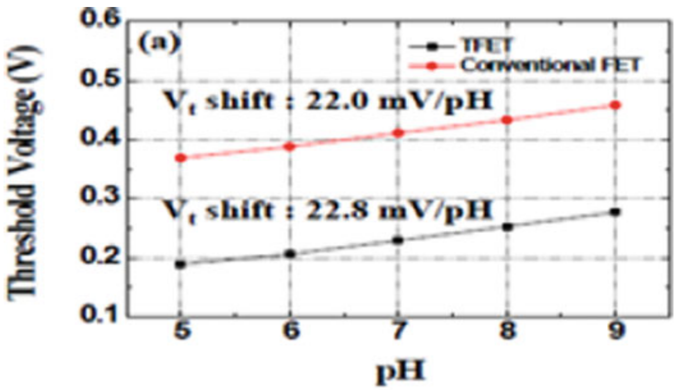


(b)

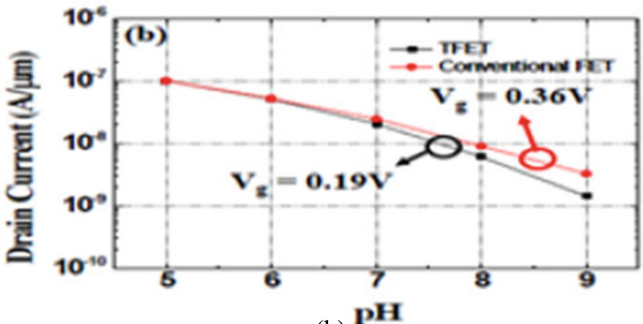
Fig. 2 Energy and drain current curve **a** Changes in the Eg-band(eV)figure of an n-type TFET based—pH sensors as a function of pH—different level **b** Transfer graphs of a n-type TFETs based—pH sensors as a function of pH—different level

of 22 mV/pH. The boundary electrical energy (Voltage) swing of the TFET device with pH variation is nearly identical to the TFET sensor (see Fig. 3(a)).

Nonetheless, detecting current Vs pH extricated up to limiting voltage with the value of pH-5 depicts that the TFET pH-device is being upgraded affectability because of the more extreme S-Slope like showed in Fig. 3(b)



(a)



(b)

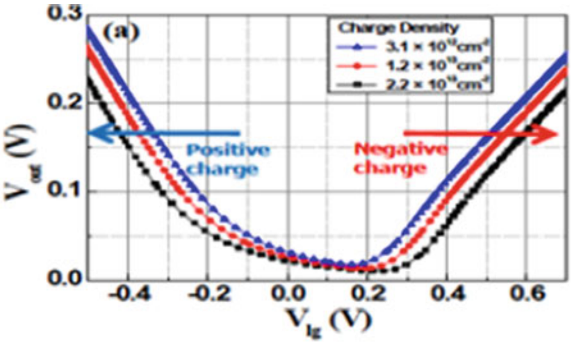
Fig. 3 Threshold voltage swing **a** and current's sensitivitys **b** on a fix Gate-voltages for TFET and conventional FET pH sensors (Vt at pH level 5)

2.2 *The TFET Sensor's Multi-sensing Capabilities*

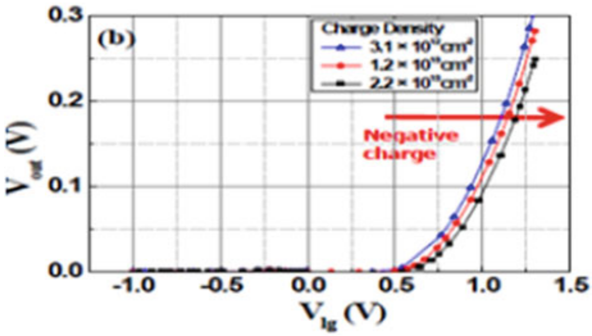
The ambipolar properties of TFET devices allow them to detect two unlike bio-particles using a single device. Figure 4(a) depicts a TFET sensor's multi-detecting behavior. The bio-atom along- with negative charges can be found in a positive entryway voltage region, although the bio-particle with positives charges might be found in a negative-entrance sV-region.

Because the planned sensor designed with a lengthy channel, the research can be easily connected and included numerous atoms in to the source-channel adjacent entrance-soxide. The energy-band (Eg) figure of the TFET sensor is publicized in Fig. 4(b) once particles along with various charges are connected to individually close of the entryway-oxides. The energy band of

Fig. 4 For varied charge densities, voltage transfer plot of CS amplifiers **a** Output voltages (V_0) of TFET sensor, **b** output voltage (V_0) of conventional FET sensor



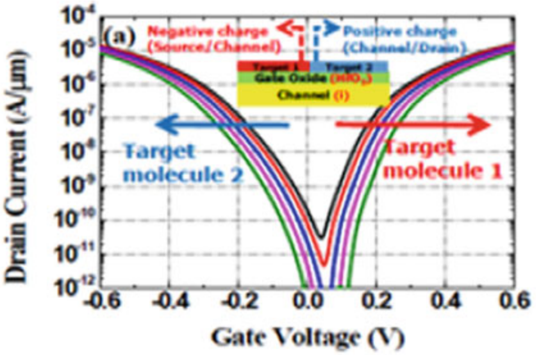
(a)



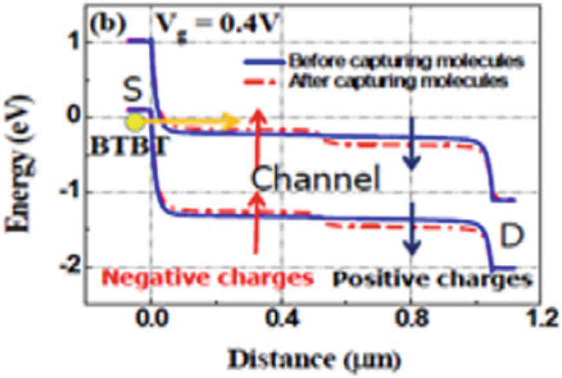
(b)

the channel is rather inaccessible. Voltage move bends in multidetecting activities should be positively moved by the negative charges joined close to the sourceintersection, although V move bends must be negatively moved through the positive joined close to the channel/channel intersection.

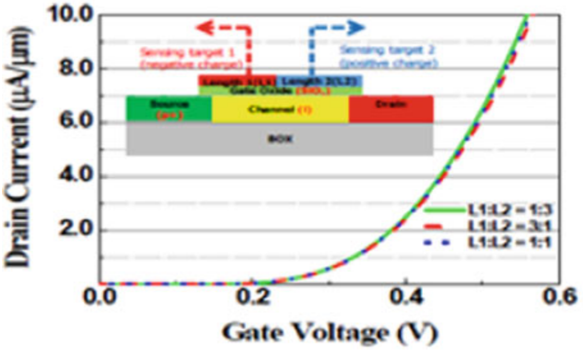
Furthermore, affectability might be changed with the correcting the conductances of the intensifier Sensors [14]. Figure 5(a) depicts yield V-variation of a TFET constructed enhancer as a function of the detecting device's door voltage range. The detection of the different aim (target) atoms might be accomplished next to both positive & negative entryway predisposition areas, comparable to the qualities of a sole sensor deprived of added hardware. Furthermore, the magnitude of yield V-variation is comparable next to the two entryway inclination locations. So, it can be suggested that multis-detecting has achieved efficiently in TFET-constructed enhancers because of ambipolar features of TFET sensors.



(a)



(b)



(c)

Fig. 5 For various charge density CS amplifiers Drain current (A/um) **a** Energy band diagram **b** Drain current (uA/um) of proposed biosensor **c** Drain current (uA/um) of proposed Biosensor

2.3 Improvement in TFET Recently for BIOSENSOR Applications

In one research methodology of Designing with Analysis of Triple Metal Vertical TFET Gate Stacked using N-type-SiGe Delta-doped layer: applying on TCAD simulation tool is published for the description of n+ delta doped layer having combining of Gate-stacking method in vertical TFET by Shilpi Gupta et al. In vertical electric field, vertical electron tunneling enhances the device's scalability in vertical structure. Upwards drive of charge carriers due to gate-electric-fields offers higher ON state current. Optimized n+ SiGe delta doped layer reduces Off-state current and enhances device operation since minimized tunneling BW amid source & channel joint. Unwanted ambipolar conduction is mitigated by using triple metal gate along with optimized work-functions next to tunneling gate, controlling gate & auxiliary gate having 4.15 eV, 4.3 eV and 4.15 eV respectively. We have discussed four different combinations with or deprived of occurrence of gate stack and n+ delta doped layer. With simulation result, it discloses 40% improvements in gate stacking method with SiGe delta doped layer in comparison to other methods. We have attained large value of current-ratios of the order (1013) and stepper subthreshold slope (9.75 mV/decade). Simulation result confirms that n+ doped TMG vertical TFET is sustainable in ultralow power applications [15].

In this paper, the author Shilpi Gupta et al. proposed and developed a novel 2-D systematic mathematical model of a 3-metal in Vertical TFET Gate stacked by n+ SiGe delta-doped layer using two-dimensional Poisson's equations. Poisson's equation was solved using parabolics approximations procedure toward explaining the channel surface-potentials and equivalent Electric-field. The most significant outcome is creation of a new channel surface potential expression that can predict the effect of source-channels and drain-channels both the biasing. The Kane's Model is utilized in estimation of the drain current by measuring the BTBT generation rate [16].

The investigation of various structures of TFET depicts that steeper subthreshold make TFET work for very low power signals but at the same time on-current and ambipolarity has been compromised. In Sect. 2, single gate SOI TFET gives subthreshold swing of 52.8 mV/decade but provides very low on-current which is enhanced by double gate TFET (DGTFET) that doubles the on-current. Later, Germanium (Ge) as a source material used replacing Silicon as Ge has low bandgap material which results in better performance and high ON/OFF current ratio. The Bilayer TFET used to improve the ambipolarity of the device which is more reduced by plasma TFET where source and drain are heavily doped which results in extremely small SS. Broken gate TFET (BGTFET) uses four gates in which 2 gates are separated which gives better performance. The doping less Vertical TFET comes with the doping less junction which formed by the induced source and the drain junction. The doping less junction less DG vertical TFET comes with greater efficiency in comparison with other mentioned devices [17] (Fig. 6).

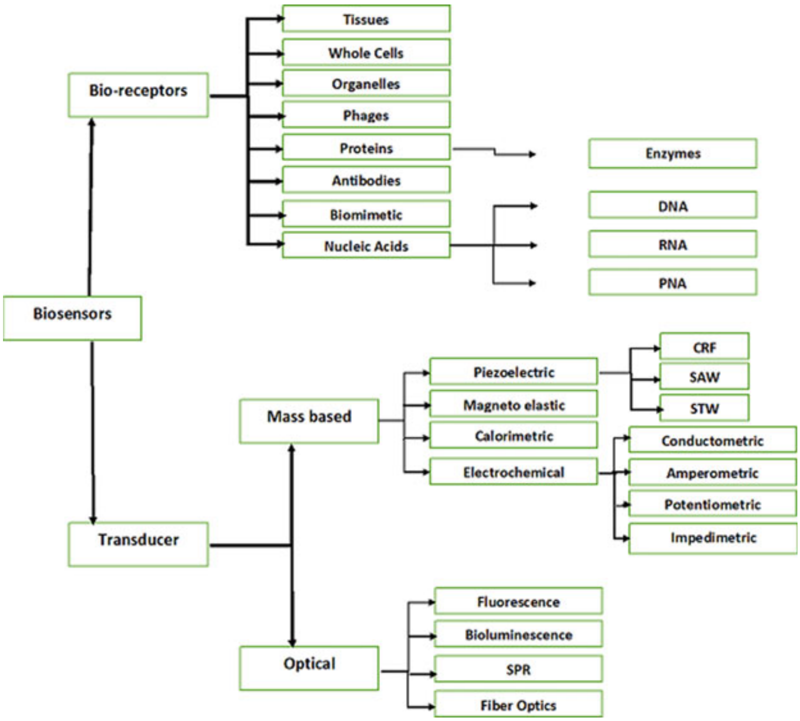


Fig. 6 Various biosensors applications in medical field

3 Results and Discussion

Following Parameter is analysis by different techniques.

4 Conclusion

We made an effort to demonstrate a TFET-based pH sensor in our research article. We made the new discovery that TFET sensors had a smaller sub-threshold swing and lower limit voltage than conventional (c)-FET sensors, thanks to TCAD replicas. Likewise, TFET based different sensors are further developed affectability because of their more extreme sub-threshold slant. Besides, TFET based-sensors is capable to detect 2-distinct atoms through inverse responsibilities in a single gadget. By utilizing two kinds of burrowing, for example, source-to-channel a lot to-deplete burrowing, the discovery of two distinctive objective particles is directed effectively at a similar detecting current-levels causes of the am-bipolar conduct of TFETs. In a typical source-speakers circuit, it is the affirmation with two distinctive objective particles could be likewise detected with the help of identical affectability. In light of these outcomes, we suppose that TFET based-Bio-sensors possibly will depiction another Vision in bio-detecting utilizations.

References

1. Clark, L.C., Jr., Lyons, C.: Electrode systems for continuous monitoring in cardiovascular surgery. Ann. N.Y. Acad. Sci. **102**, 29–45 (1962)
2. Tothill, I.E.: Biosensors for cancer markers diagnosis. Semin. Cell Dev. Biol. **20**, 55–62 (2009)
3. Ligler, F.S., Taitt, C.R., Shriver-Lake, L.C., Sapsford, K.E., Shubin, Y., Golden, J.P.: Array biosensor for detection of toxins. Anal. Bioanal. Chem. **377**, 469–477 (2003)
4. Bunney, J., Williamson, S., Atkin, D., Jeanneret, M., Cozzolino, D., Chapman, J., Power, A., Chandra, S.: The use of electrochemical biosensors in food analysis. Curr. Res. Nutr. Food. Sci. **5**, 83–95 (2017)
5. Pantelopoulos, A., Bourbakis, N.G.: A survey on wearable sensor-based systems for health monitoring and prognosis. IEEE Trans. Syst. Man Cybern. Syst. **40**, 1–12 (2010)
6. Chaplin, M.F., Bucke, C.: Biosensors. In: Enzyme Technology, 1st edn. Cambridge University Press, Cambridge (1990)
7. Thevenot, D.R., Toth, K., Durst, R.A., Wilson, G.S.: Electrochemical biosensors: recommended definitions and classification. Biosens. Bioelectron. **16**, 121–131 (2001)
8. Bergveld, P.: Development of an ion-sensitive solid-state device for neurophysiological measurements. IEEE Trans. Biomed. Eng. **BME-17**, 70–77 (1970)
9. Cui, Y., Wei, Q., Park, H., Lieber, C.M.: Nanowire: nano sensors for highly sensitive and selective detection of biological and chemical species. Science **293**, 1289–1292 (2001)
10. Avci, U.E., et al.: Comparison of performance, switching energy and process variations for the TFET and MOSFET in logic. In: VLSI symp. tech. dig., pp. 124–125 (2011)
11. Lee, J., et al.: Implementation and characterization of gate-induced drain leakage current-based multiplexed SiNW biosensor. In: The 20th Korean Conference on Semiconductors (2013)
12. Chapman, R.A., et al.: Comparison of methods tobias fully depleted SOI-based MOSFET nanoribbon pH sensors. IEEE Trans. Electron Devices **58**(6), 1752–1760 (2011)
13. Choi, B., et al.: TCAD-based simulation method for the electrolyte-insulator-semiconductor field-effect transistor. IEEE Trans. Electron Devices. **62**(3), 1072–1075 (2015)
14. Lee, J., et al.: SiNW-CMOS hybrid common-source amplifier as a voltage-readout hydrogen ion sensor. IEEE Electron Device Lett. **34**(1), 135–137 (2013)
15. Gupta, S., Wairya, S., Singh, S.: Design and analysis of triple metal vertical TFET gate stacked with N-type SiGe delta-doped layer. Silicon J. (2021). https://doi.org/10.1007/s12633-021-01211-3. Accessed 21 June 2021
16. Gupta, S., Wairya, S., Singh, S.: Analytical modeling and simulation of a triple metal vertical TFET with hetero-junction gate stack. Superlattices Microstruct. **157** (2021)
17. Gupta, S., Wairya, S.: Performance analysis of different Tunnel Field Effect Transistors (TFET) device structures with their Challenges. In: 8th International Conference on Signal Processing and Integrated Networks, SPIN 2021 (2021)
18. Dwivedi, P., Singh, R.: Investigation the impact of the gate work-function and biases on the sensing metrics of TFET based biosensors. Eng. Res. Express **2** (2020)
19. Devi, W.K., et al.: N+pocket-doped vertical TFET for enhanced sensitivity in biosensing applications: modeling and simulation. IEEE Trans. Electron Dev. **67**(5) (2020)
20. Lee, R., et al.: Investigation of feasibility of tunneling field effect transistor (TFET) as highly sensitive and multi-sensing biosensors. J. Semicond. Technol. Sci. **17**(1) (2017)
21. Gao, A., et al.: Robust ultrasensitive tunneling FET biosensor for point-of-care diagnostics. Sci. Rep. **6**, 22554. https://doi.org/10.1038/srep22554

22. Bhattacharyya, A., et al.: Performance assessment of new dual-pocket vertical heterostructure tunnel FET-based biosensor considering steric hindrance issue 0018–9383 (2019)
23. Verma, M., et al.: Performance assessment of a novel vertical dielectrically modulated TFET-based biosensor 0018–9383 (2017)

Analysis of Tunnelling Probability of Different High-K Material for Nanometer Thickness MOSFET Gate

Avijit Deb Sarkar

Abstract Presently, various mathematical and experimental research are going on for the reduction of the size of transistors for the improvement of the overall performance of electronic devices. Most of the study circumscribes around the behavior of small channel leakage current of MOSFET (metal–oxide–semiconductor field-effect transistor) which ultimately affects their respective work efficiency. The equations for drain current of the transistor shows the channel length to be the primary parameter for their calculation. Exhibiting the reduced channel length to impact the behavior of the transistor along with their respective drain current characteristics. SiO_2 has been the material of choice, since ages, of a transistor in spite of its thickness still being a limitation. Another observation is that the available mathematical models are highly complex to visualize, requiring multiple parameters for their calculation and the experimental papers, though are very limited in numbers, showing difficulty to be arranged being costly, as requires costly software for their simulations. This research has thereby proposed a simple strategy for the identification of the impact of different high-k gate material thickness based on the calculation of leakage current by tunnelling probability of electrons from channel to gate, as because the rate of flow of electron is always proportional to current in opposite direction. Hence, this theoretical research would aid in an easy identification of an appropriate gate material as per the thickness requirement.

Keywords High-k materials · Short channel MOSFET · Gate materials · Gate sizing · Leakage current

1 Introduction

The report published in International Technology Roadmap for Semiconductors (ITRS) has mentioned that there is a continuous scaling down of MOSFET size by the reduction of the channel length and gate thickness of the devices that has

A. D. Sarkar (✉)
TATA Consultancy Services, Kolkata, India
e-mail: avijitcse@gmail.com

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_62

adversely impacted their leakage current [1]. The reduced size of MOSFET evolved a new geometrical structure called FinFET (fin field-effect transistor), made of the same type of material as that of the former. But it has been experimentally shown that the geometrical reduction has been unable to handle the leakage current impact on the drain current equation [2]. This increased scaling has led to a higher leakage current along with a power dissipation. The dimension and material of the MOSFET and FinFET is thereby required to be analyzed closely for a better understanding of their exact impact on their respective leakage current behavior [2]. Thus, in this theoretical study different types of high k- gate (where "k" refers to dielectric constant) material with their respective variation in dimensions have been explored for the analysis of their effect on their leakage current based on their quantum mechanical model of tunnelling electron probability. The different choice of such materials was TiO_2, Ta_2O_5, Al_2O_3, Si_3N_4, SiO_2, La_2O_3, Y_2O_3, CeO_2, HfO_2 (Table 1). Although the high 'k' value is important for any gate material as it is one of the major factors that controls the gate leakage current but still when an ultra-thin gate section of such material is used, they have exhibited a quantum mechanical tunnelling effect leading to the generation of a charge flow through the gate dielectric [3, 4]. The tunnelling gate current is calculated using Wentzel–Kramers–Brillouin (WKB) approximation, which exponentially depends on the thickness of gate materials. Hence the gate thickness is another important physical parameter which could control the behavior of MOSFET and/or FinFET [5]. The WKB approximation formula requires multiple parameters such as threshold voltage, flatband voltage, gate voltage, voltage across dielectric, electric field in the dielectric, dielectric permittivity, physical thickness of the dielectric layer etc. to calculate the tunnelling gate current.

In nano scale design, six short-channel leakage current components have been noticed. Following are the different types of leakage current that flows inside the MOSFET and FinFET:

a. I_1 is the reverse-bias p–n junction leakage.
b. I_2 is the sub threshold leakage.
c. I_3 is the oxide tunnelling current.
d. I_4 is the gate current due to hot-carrier injection.

Table 1 Material name and list

Material name	Material symbol
Silicon dioxide	SiO_2
Silicon nitride	Si_3N_4
Aluminum oxide	Al_2O_3
Yttrium oxide	Y_2O_3
Tantalum pentoxide	Ta_2O_5
Cerium oxide	CeO_2
Hafnium oxide	HfO_2
Lanthanum oxide	La_2O_3
Titanium dioxide	TiO_2

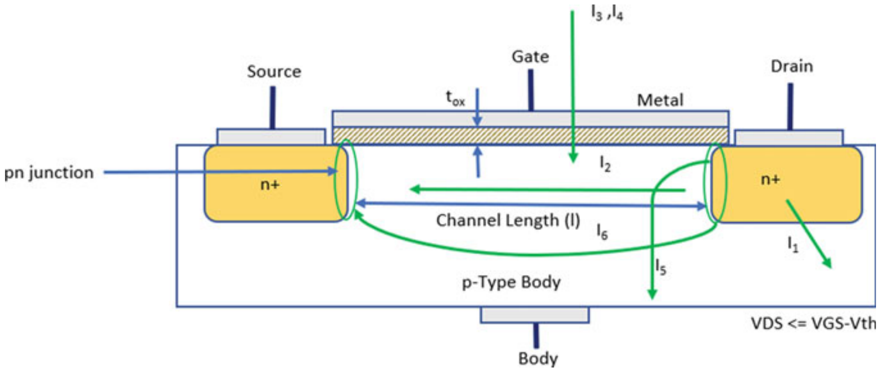


Fig. 1 Short channel leakage current components for N-Channel MOSFET

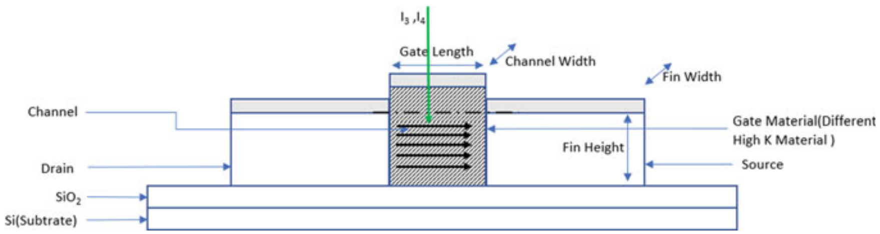


Fig. 2 Short channel leakage current components for N-Channel FinFET

e. I_5 is the Gate- induced drain leakage current.
f. I_6 is the channel punch through current.

In this study the analysis was confined to the factor I_3 which is the impacting factor showing the electron flow from channel to gate. Figures 1 and 2 are the schematic representation of the leakage current flow of MOSFET and FinFET.

2 Quantum Mechanical Analysis of Short Channel Transistor

In short channel transistor the classical behavior of the charge particles has changed in deep submicron and nanometer size MOSFET or FinFET devices. It has been reported that the mechanical behavior of the charge carriers is greatly affected by the wave-particle behavior of electrons in the semiconductor devices. In sub less than equal to 10 nm channel length with higher Drain Voltage (V_{ds}) of MOSFET or FinFET the subthreshold leakage current (I_2) flowing from drain to source without any control of gate voltage. It can be concluded that for small channel devices the electron density is proportional to V_{ds} and it impacts the conduction in channel region.

The oxide-tunnelling leakage current is another component, which depends on the thickness of oxide or dielectric layer. Smaller thickness of the dielectric layer is the main cause of energy quantization between the gate and channel. This phenomenon of energy quantization helps in the creation of an extra current flow from gate to the channel, which is called as an oxide tunnelling current (I_3) (see Figs. 1 and 2). The oxide tunnelling current density depends on the probability of density of the charge flowing from source to the drain. The following equation of conservation of probability (Eq. 1) has been used by different researchers for the analysis of the quantum mechanical behavior of such short channel MOS devices [7, 8].

$$\partial\rho/\partial \mathrm{t} = -\nabla . \mathrm{j} \tag{1}$$

where,

$\mathrm{j} = \mathrm{i} \times (\mathrm{h}/2\pi) \times (1/2\ \mathrm{m}) \times [\Psi\ \nabla\ \Psi^* - \Psi * \nabla\ \Psi]$
$\mathrm{i} = (-1)^{1/2}$
$\rho = \Psi * \Psi$ (Probability density where, Ψ is periodic wave function of the electron and Ψ^* is the complex conjugate of the wave function)
$\hbar = \mathrm{h}/2\pi = 1.\ 055{,}106 \times 10^{-34}$ J.s $= 6.582119 \times 10^{-16}$ eV.s

3 Tunnelling Electron Probability Calculation Using Schrödinger Equation

The time independent Schrödinger equation (Eq. 2) [9] for particle wave functions is solved to find the quantization of the energy levels and the variation of the surface potential only in the transverse direction i.e. along the depth of the channel or normal to the oxide/silicon interface [7, 10, 11]. Figure 3 shows the tunnelling of charge from channel to gate through a rectangular potential wall of height Eb (in eV) and width tox (in μm). This behavior of the dielectric gate layer resembles the finite depth quantum well behavior (Fig. 3). For short channel MOSFET this behavior is more prominent than long channel. It can be concluded that the tunnelling effect is inversely proportional to the length of the channel. Hence the Schrödinger equation (Eq. 2) can be considered as the best candidate for this analysis, as the wave nature of the election is more appropriate for the leakage current analysis.

$$\mathrm{d}^2\Psi/\mathrm{dx}^2 = -\,(2\mathrm{m}/\hbar)[\mathrm{E} - \mathrm{U(x)}]\ \Psi(\mathrm{x}) \tag{2}$$

where,

m = mass of the electron (0.911×10^{-27} g)
$\hbar$ = Planck's constant (1.054×10^{-34} J.S)
$\Psi(\mathrm{x})$ = The periodic wave function of the electron.
E = Energy of the electron

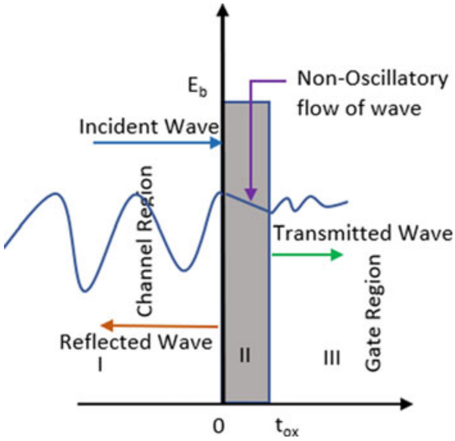
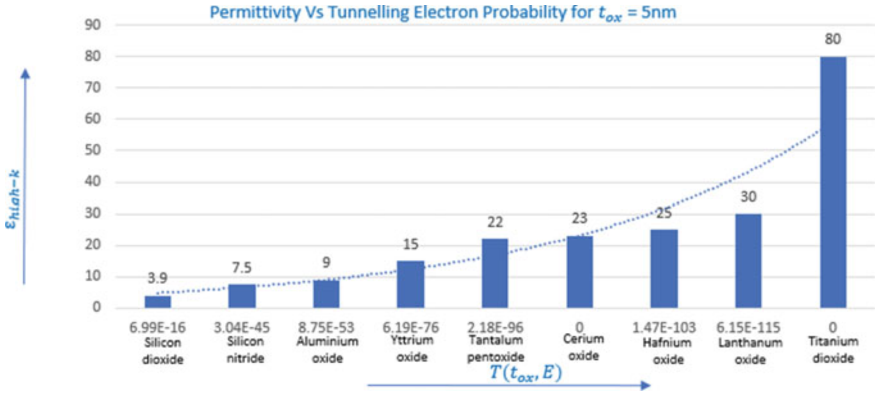


Fig. 3 Wave tunnelling through the potential barrier of height E_b and width W, which is the thickness of gate material (t_{ox})

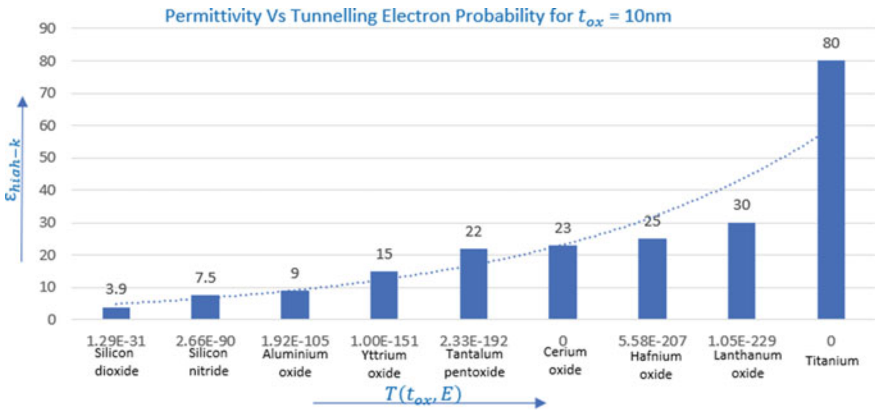


Graph 1 Dielectric constant (ε_{high-k}) vs tunnelling electron probability ($T(t_{ox}, E)$) for 5 mm thickness gate materials

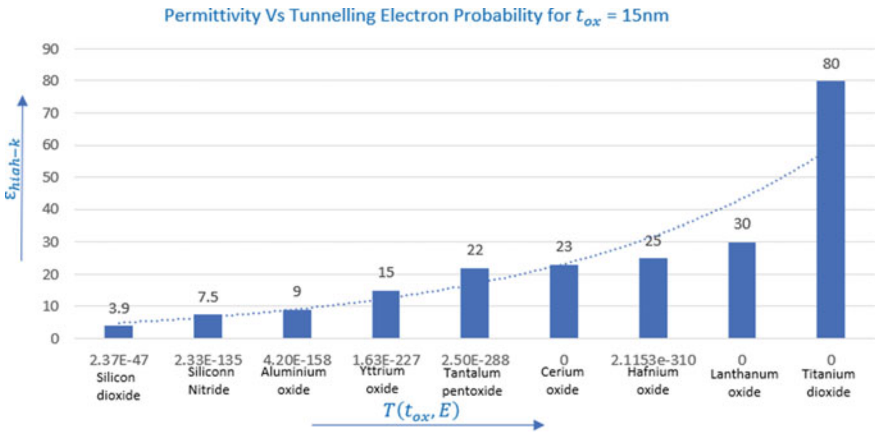
U(x) = serves the boundary condition of the wave functions with respect to position.

Where E is the allowed energies of the electrons in channel region. The de Broglie relation defines the wavelength of the electron as $\lambda = 10^{-10}$m or 0.1 nm. Thus, for small channel MOSFET the length of the channel is comparable with the wavelength of electron which in turn helps the electronic state to get quantized at discrete energy levels. Based on classical mechanics, the electron does not have enough energy to tunnel through an insulated layer and ultimately reaches the gate. However, quantum mechanically there is a finite probability that an electron can tunnel through oxide layer and then to the gate [12].

From channel to gate, the charged wave flows through three regions of different potential level.



Graph 2 Dielectric constant (ε_{high-k}) vs tunnelling electron probability ($T(t_{ox}, E)$) for 10 mm thickness gate materials

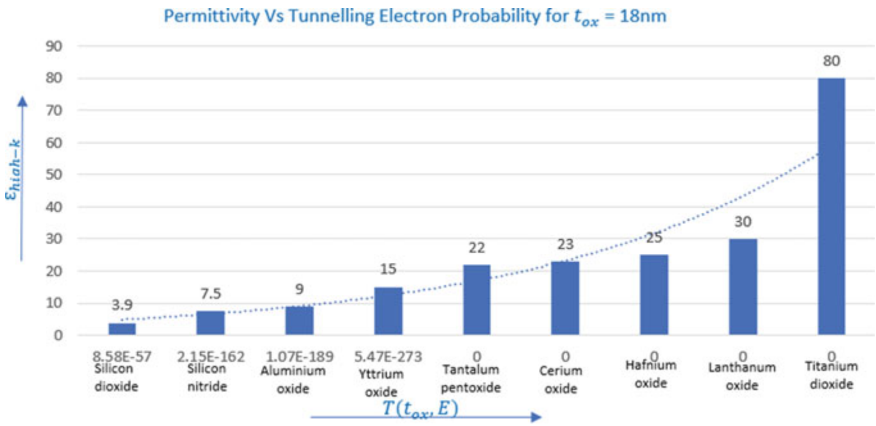


Graph 3 Dielectric constant (ε_{high-k}) vs tunnelling electron probability ($T(t_{ox}, E)$) for 15 mm thickness gate materials

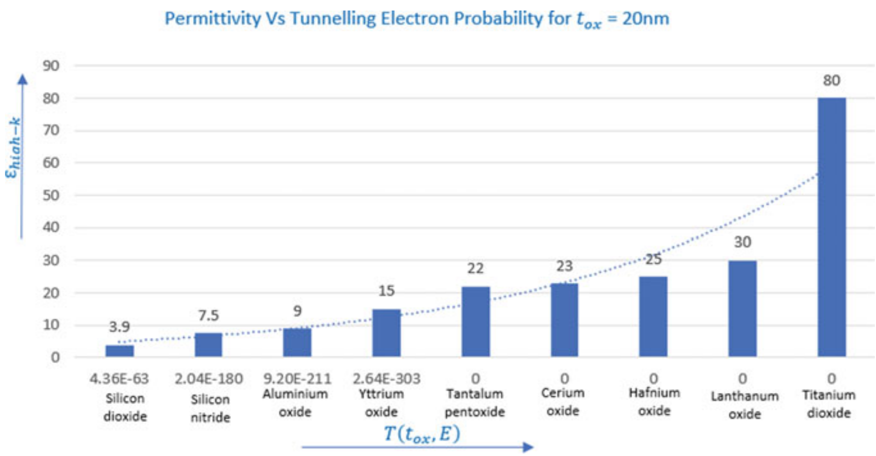
$U(x) = E_b$; $x < 0$ (channel region)
$U(x) = 0$; $0 <= x <= t_{ox}$ (dielectric region)
$U(x) = E_b$; $x > t_{ox}$ (gate region)

The region I and III are the channel and gate regions respectively. The wave nature of the election changes from sinusoidal to exponentially diminishing in form in the dielectric region. Hence the wavenumbers of the regions I and III are:

$$k_0^2 = \frac{2m(E_b - E)}{\hbar} \tag{3}$$



Graph 4 Dielectric constant (ε_{high-k}) vs tunnelling electron probability ($T(t_{ox}, E)$) for 18 mm thickness gate materials



Graph 5 Dielectric constant (ε_{high-k}) vs tunnelling electron probability ($T(t_{ox}, E)$) for 18 mm thickness gate materials

$$k_1^2 = \frac{2mE}{\hbar} \tag{4}$$

Thereby, based on the wavenumbers the Schroedinger's equation can become:

$$\frac{d^2\Psi_{I,III}}{dx^2} = \frac{2m(E_b - E)}{\hbar}\Psi_{\mathrm{I,III}} \quad (\mathrm{x} < 0 \text{ and } \mathrm{x} > \mathrm{t_{ox}}) \tag{5}$$

$$\frac{d^2\Psi_{II}}{dx^2} = \frac{2mE}{\hbar}\Psi_{\mathrm{II}} \quad (0 <= \mathrm{x} <= \mathrm{t_{ox}}) \tag{6}$$

The solution of each region of the Schroedinger's equations are [9]:

$$\text{Channel Region: } \Psi_{\mathrm{I}}(\mathrm{x}) = \mathrm{Ae}_1^{\mathrm{ikx}} + \mathrm{Be}_1^{-\mathrm{ikx}}(\mathrm{x} < 0) \tag{7}$$

$$\text{Dielectric Region: } \Psi_{\mathrm{II}}(\mathrm{x}) = \mathrm{Ce}_2^{\mathrm{ikx}} + \mathrm{De}_2^{-\mathrm{ikx}}(0 <= \mathrm{x} <= \mathrm{t_{ox}}) \tag{8}$$

$$\text{Gate Region: } \Psi_{\mathrm{III}}(\mathrm{x}) = \mathrm{Fe}_1^{\mathrm{ikx}} + \mathrm{Ge}_1^{-\mathrm{ikx}}(\mathrm{x} > \mathrm{t_{ox}}) \tag{9}$$

The Fig. 3 depicts that there is no reflected wave in the gate region, so the constant of Eq. 9 is 0 (G = 0). From the channel region some part of the wave gets transferred to the oxide layer is called an incident wave and some part gets reflected to the channel so both that both A and B will remain constant. It is also very important to know that the nature of the wave function in the insulated region is not oscillatory. The wave function in the region III (Eq. 8), is the transmitted wave function. The amplitude of the transmitted wave is given by Eq. 9. The region I (see Fig. 3) gives the amplitude of incident and reflected wave $|A|^2$ and $|B|^2$ respectively. The tunnelling probability of a wave is the ratio of incident wave to that of the transmitted wave. The following equation (Eq. 10) defines the tunnelling probability of the wave and the behavior of the waves in each region (see Fig. 3).

$$T(t_{ox}, E) = \frac{|F|^2}{|A|^2} \tag{10}$$

It was noted that for a wide and high barrier, the transmission of electrons is quite poor [9], thereby the approximated tunnelling equation can be given as follows (Eq. 11):

$$T(t_{ox}, E) = 16\frac{\mathrm{E}}{E_b}\left(1 - \frac{E}{E_b}\right)e^{-2k_0 t_{ox}} \tag{11}$$

Again, for the high-k materials, the equivalent oxide thickness can be measured by the below mentioned equation (Eq. 12)

$$EOT = \frac{\varepsilon_{high-k}}{3.9} t_{ox} \tag{12}$$

where, $\varepsilon_{\mathrm{high-k}}$ is the dielectric constant of the high-k materials.

The revised tunnelling equation is given by (Eq. 13) for the calculation of tunnelling probability of the electrons is:

$$T(t_{ox}, E) = 16\frac{\mathrm{E}}{E_b}\left(1 - \frac{E}{E_b}\right)e^{-2k_0 EOT} \tag{13}$$

The tunnelling probability is the primary cause of gate leakage current. There are three tunnelling components such as Electron Conduction Band (ECB) tunnelling [13, 14], Electron Valance Band (EVB) tunnelling [12, 13] and Hole Valance Band (HVB) tunnelling [13, 14]. ECB tunnelling is the primary property of the short channel n-MOSFET and HVB tunnelling in the primary property of short channel p-MOSFET. [13]. Table 2 provides the tunnelling probability of electrons for each high-k material mentioned in Table 1 and corresponding trend in Graphs 1, 2, 3, 4 and 5.

4 Analysis of Materials Behavior and Their Usage as a Gate Materials

Table 1 and Graphs 1, 2, 3, 4 and 5 shows the tunnelling probability of the electron through gate for different materials. The tunnelling probability of electron depends on the dielectric constant and thickness of the oxide layer of MOSFET or FinFET. For a constant thickness, tunneled electron flow could be controlled by high-k materials. However, in recent days we are still dependent on the SiO_2 as a primary gate material because of its physicochemical properties [22]. The amorphous structure and very few electronic defects make SiO_2 an excellent interfacing material with Si. It can be etched and patterned to a nanometer scale [22]. However, based on the above analysis the leakage current (I_3) is the primary concern for SiO_2 with a nanometer range gate thickness. Hence, designers need to use a new high K material or combination of high-k materials with SiO_2 as a gate material.

Any gate material or combinations of gate materials should adhere with below six properties:

- It must have a high-k so that it achieve reasonable number of years of scaling [22].
- The thermodynamic behaviour of material should not have any impact on performance during direct contact with the Si channel [22].
- It must be kinetically stable and be compatible with processing to 1000 °C for 5 s [22].
- It must act as an insulator, by having band offsets with Si of over 1 eV to minimise carrier injection into its bands [22, 23].
- It must form a good electrical interface with Si [22].
- It must have a few bulk electrically active defects and patterned to a nanometer scale [22].

Table 2 Tunnelling probability of electron for n-channel MOSFET

E (p-Type Si) in eV	t_{ox} (in nm)	Material	ε_{high-k}	E_b in eV	$T(t_{ox}, E)$
3.4	5	SiO_2	3.9	8.76	6.9939e−16
		Si_3N_4	7.5	5.0	3.0444e−45
		Al_2O_3	9	7.0	8.7541e−53
		Y_2O_3	15	5.6 [15]	6.19e−76
		Ta_2O_5	22	4 [16]	2.1818e−96
		CeO_2	23	3.19 [17]	0
		HfO_2	25	5.8 [18]	1.4715e−103
		La_2O_3	30	5.18 [19]	6.1468e−115
		TiO_2	80	3.09 [20, 21]	0
3.4	10	SiO_2	3.9	8.76	1.2873e−31
		Si_3N_4	7.5	5.0	2.6621e−90
		Al_2O_3	9	7.0	1.9174e−105
		Y_2O_3	15	5.6 [15]	1.004e−151
		Ta_2O_5	22	4 [16]	2.3335e−192
		CeO_2	23	3.19 [17]	0
		HfO_2	25	5.8 [18]	5.5791e−207
		La_2O_3	30	5.18 [19]	1.047e−229
		TiO_2	80	3.09 [20, 21]	0
3.4	15	SiO_2	3.9	8.76	2.3695e−47
		Si_3N_4	7.5	5.0	2.3278e−135
		Al_2O_3	9	7.0	4.1997e−158
		Y_2O_3	15	5.6 [15]	1.6284e−227
		Ta_2O_5	22	4 [16]	2.4958e−288
		CeO_2	23	3.19 [17]	0
		HfO_2	25	5.8 [18]	2.1153e−310
		La_2O_3	30	5.18 [19]	0
		TiO_2	80	3.09 [20, 21]	0
3.4	18	SiO_2	3.9	8.76	8.5829e−57
		Si_3N_4	7.5	5.0	2.1477e−162
		Al_2O_3	9	7.0	1.0654e−189
		Y_2O_3	15	5.6 [15]	5.4676e−273
		Ta_2O_5	22	4 [16]	0
		CeO_2	23	3.19 [17]	0
		HfO_2	25	5.8 [18]	0
		La_2O_3	30	5.18 [19]	0
		TiO_2	80	3.09 [20, 21]	0

(continued)

Table 2 (continued)

E (p-Type Si) in eV	t_{ox} (in nm)	Material	ε_{high-k}	E_b in eV	$T(t_{ox}, E)$
3.4	20	SiO_2	3.9	8.76	4.3613e−63
		Si_3N_4	7.5	5.0	2.0355e−180
		Al_2O_3	9	7.0	9.1987e−211
		Y_2O_3	15	5.6 [15]	2.6413e−303
		Ta_2O_5	22	4 [16]	0
		CeO_2	23	3.19 [17]	0
		HfO_2	25	5.8 [18]	0
		La_2O_3	30	5.18 [19]	0
		TiO_2	80	3.09 [20, 21]	0

5 Conclusion

The study thereby concludes that the Schrodinger equation can be used to calculate the tunnelling probability of the electron for ultra-thin gate materials for the identification of the flow of electron from channel to gate. The rate of flow of electron from one layer (channel) to another (gate) resulting in the leakage current flow in opposite direction that is from gate to channel. There are different mathematical models available to understand the tunnelling behavior of the MOSFET and FinFET. Each model requires four major physical parameters such as potential of the barrier, potential of the silicon, gate thickness and dielectric constant of the gate materials. Hence this technique could help the researchers to identify the gate materials easily and quickly for the construction of devices.

References

1. Arden, W., et al.: "More-than-Moore" White Paper, International Technology Roadmap for Semiconductors, ITRS 2010 update. http://www.itrs.net
2. Darbandy, G., Lime, F., Cerdeira, A., Estrada, M., Garduño, I., Iñiguez, B., et al.: Study of potential high-k dielectric for UTB SOI MOSFETs using analytical modeling of the gate tunneling leakage
3. Depas, M., Vermeire, B., Mertens, P.W., Van Meirhaeghe, R.L., Heyns, M.M.: Determination of tunneling parameters in ultra-thin oxide layer poly-Si/SiO2/Si structures. Solid-State Electron. **38**, 1465 (1995)
4. Lo, S.H., Buchanan, D.A., Taur, Y., Wang, W.: Quantum-mechanical modeling of electron tunneling current from the inversion layer of ultra-thin-oxide nMOSFET's. IEEE Electron Device Lett. **18**, 209 (1997)
5. Yeo, Y.C., Lu, Q., Lee, W.C., King, T.-J., Hu, C., Wang, X., Guo, X., Ma, T.P.: Direct Tunneling Gate Leakage Current in Transistors with Ultrathin Silicon Nitride Gate Dielectric, 11 November 2000
6. Piprek, J., et al.: "Carrier Transport" Leakage current in a TFT. http://www.stallinga.org/ElectricalCharacterization/tft/index.html

7. Chaudhry, A., Roy, J.N.: Mosfet Models, Quantum Mechanical Effects and Modeling Approaches: A Review, March 2010. http://www.jsts.org/html/journal/journal_files/2010/03/year2010volume10_01_03.pdf
8. Wang, L.: Quantum Mechanical Effects on MOSFET Scaling Limit. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.89.3422&rep=rep1&type=pdf
9. Sakurai, J.J.: Modern Quantum Mechanics (1995)
10. Ip, B.K., Brews, J.R.: Quantum effects upon drain current in a biased MOSFET. IEEE Trans. Electron. Devices **45**(10), 2213–2221 (1998)
11. Marini, E., et al.: Investigation of ALD dielectrics in silicon capacitors. In: 35th Annual Microelectronic Engineering Conference at RIT, May 2017
12. Bowman, K.A., Wang, L., Tang, X., Meindl, J.D.: A circuit-level perspective of the optimum gate oxide thickness. IEEE Trans. Electron. Devices **48**(8), 1800–1810 (2001)
13. Lee, W.C., King, T.J., Hu, C.M.: Evidence of hole direct tunneling through ultrathin gate oxide using P+ Poly-SiGe gate. Electron Device Lett. **20**(6), 268–270 (1999)
14. Lee, W.C., Hu, C.M.: Modeling CMOS tunneling currents through ultrathin gate oxide due to conduction- and valence-band electron and hole tunneling. Electron Device IEEE Trans. **48**(7), 1366–1373 (2001)
15. Roslan, Ameer F., Salehuddin, F., Zain, A.S.M., Kaharudin, K.E., Ahmad, I.: Enhanced performance of 19 single gate MOSFET with high permittivity dielectric material. Indones. J. Electr. Eng. Comput. Sci. **18**(2), 724–730 (2020)
16. Lee, J., Lu, W., Kioupakis, E.: Electronic properties of tantalum pentoxide polymorphs from first-principles calculations. Appl. Phys. Lett. **105**, 202108 (2014)
17. Ali, M.M., Mahdi, H.S., Parveen, A., Azam, A.: Optical properties of cerium oxide (CeO2) nanoparticles synthesized by hydroxide mediated method. In: 2nd International Conference on Condensed Matter and Applied Physics (ICC 2017), May 2018. AIP Conference Proceedings 1953(1):030044
18. Robertson, J.: High dielectric constant oxides. Engineering Department, Cambridge University, Cambridge CB2 1PZ, UK, December 2004. Accessed 2 Dec 2004 – © EDP Sciences
19. Reference. https://www.sciencedirect.com/topics/engineering/channel-length-modulation
20. Liu, Q.-Y., Fang, Z.-B., Ji, T., Liu, S.-Y., Tan, Y.-S., Chen, J.-J., Zhu, Y.-Y.: Band alignment and band gap characterization of La2O3 films on Si substrates grown by radio frequency magnetron sputtering. 2014 Chinese Physical Society and IOP Publishing Ltd (2014)
21. Chung, M.H., Kim, J.H.: Preparation and luminescence of europium-doped yttrium oxide thin films. Appl. Sci. Converg. Technol. **26**(2), 26–29 (2017)
22. Nagaraj, G., Dhayal Raj, A., Albert Irudayaraj, A., Josephine, R.L.: Tuning the optical band Gap of pure TiO2 via photon induced method. Optik**179**, 889–894 (2019)
23. Robertson, J.: High dielectric constant oxides. Eur. Phys. J. Appl. Phys. **28**, 265–291 (2004)
24. Robertson, J., Chen, C.W.: Appl. Phys. Lett. **74**, 1168 (1999)

Design of Tapered Vivaldi Antenna for Milli-meter Waves Applications

Ritesh Kumar Kushwaha, P. Karuppanan, Prateek Asthana, and Nand Kishore

Abstract In this paper, Tapered Vivaldi Antenna (TVA) is designed and fabricated for millimeter waves Applications. Two different TVA structures are designed and compared the radiation characteristics in the frequency band from 50 to 60 GHz. The return loss is shown below −10 dB in the operating band. Performance characteristics of the first TVA structure is improved by another TVA having an extended substrate. The modified structure shows the better directive gain and radiation efficiency. The extended substrate TVA is fabricated and measured in the laboratory and validated with the simulated results. The observed maximum gain is 9.8 dB, and radiation efficiency is 92.6%. The proposed design is suitable for 5G and WLAN applications.

Keywords Tapered Vivaldi Antenna · Extended substrate · 5G · WLAN · Milli-meter wave applications

1 Introduction

With the advancement of mm-wave technology, it has been seamless attention in this frequency regime from 30 to 300 GHz, which lies in the spectrum between microwave wave and visible light. The millimeter-wave technology has been significantly suitable in sciences, medical imaging, high data rate wireless communications, security, etc. [1]. An antenna is a key component in the millimeter-wave technology that fascinated lots to the research community. It has projected various types of antenna blow 60-GHz [2]. Nevertheless, horn-antennas was the better option for high-radiation gain due to restricted fabrication and measurement facilities in the mm-wave frequency range up to 60 GHz [3].

R. K. Kushwaha (✉)
Chandigarh University, Gharuan, Mohali, Punjab, India
e-mail: rel1713@mnnit.ac.in

P. Asthana
Bharat Institute of Engineering and Technology, Hyderabad, India

P. Karuppanan · N. Kishore
Motilal Nehru National Institute of Technology Allahabad, Allahabad, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_63

Conversely, substrate-integrated-waveguide (SIW) technology and numerous fabrication technologies, including MEMs and LTCC, reduce the conventional horn antennas [4, 5]. Remarkably, ultra-wideband (UWB) antennas substantially reduce an imprint of the system via substituting multiple and multi-band antennas conventionally required to include various bands. Therefore, Vivaldi antennas have been extensively used to obtain a wideband frequency range as an end-fire antenna in the mm-wave range. The Vivaldi was initially projected by Gibson in 1979 [6].

In the last decade, various methodologies have been adopted to improve the radiation performance of the Vivaldi antenna in the 60-GHz band. Hehnel et al. [7] designed a Broadside radiating Vivaldi antenna using In-Package Technology and obtained the gain of around 7 dBi with the back lobe of −12 dBi. Hameed et al. [8] reported a Vivaldi antenna in the operating frequency band from 50–70 GHz using an antenna on-chip (AoC) Technology which shows the maximum gain of −0.40 dBi. Similarly, 0.7 dBi gain has been reported using on-chip antenna Technology [9]. On the other hand, High gain Vivaldi antenna achieved by using Parasitic patch, Substrate Integrated waveguide (SIW), meta-surface, Balanced AVA, slots, Corrugation, and different substrate methods [11]. Remarkably, the substrate extension technique is an easier approach to achieve high gain. Wan et al. [12] designed a trapezoidal dielectric substrate for antipodal Vivaldi antenna and observed the radiation characteristics in the frequency band from 1-to-40 GHz band. Though the observed gain was around 11.1 dB at 40 GHz, but the Half Power Bandwidth (HPBW) was 13.1°. Besides, Moosazadeh [13] reported a modified tapered antipodal Vivaldi antenna with trapezoid-shaped dielectric and observed a gain around 14 dB. However, the front-to-back ratio (F-to-B) was 23 dB.

Nevertheless, the radiation characteristics of the Vivaldi antenna, such as directivity, gain, radiation efficiency, and front-to-back, are still need to be improve. In this paper, a simple Tapered Vivaldi antenna with extended substrate and micro-strip feed line is projected for 60 GHz applications. The paper is arranged as follows: Sect. 2 contains the antenna design method. Section 3 discusses the fabricated prototype and measured results. Finally, the summary and conclusion are presented in Sect. 4.

2 Antenna Design

Conventional tapered Vivaldi antenna structures having a cavity at the slotted structures. Whereas, in the proposed structure, the back cavity is removed and chosen the larger substrate structure in comparison to aperture length, as shown in Fig. 1.

The projected antenna is constructed on Rogers RT 5880 substrate (ε_r = 2.2 and $tan\delta = 0.0009$), having thickness of 0.254 mm. The dimension of the antenna is $15 \times 4 \times 0.254\,\text{mm}^3$. Further, the optimizations are performed using CST Microwave Studio. The optimum parameters of the proposed antenna is listed in Table 1. The flared structure of the metal is expressed by the exponential equation as [14]

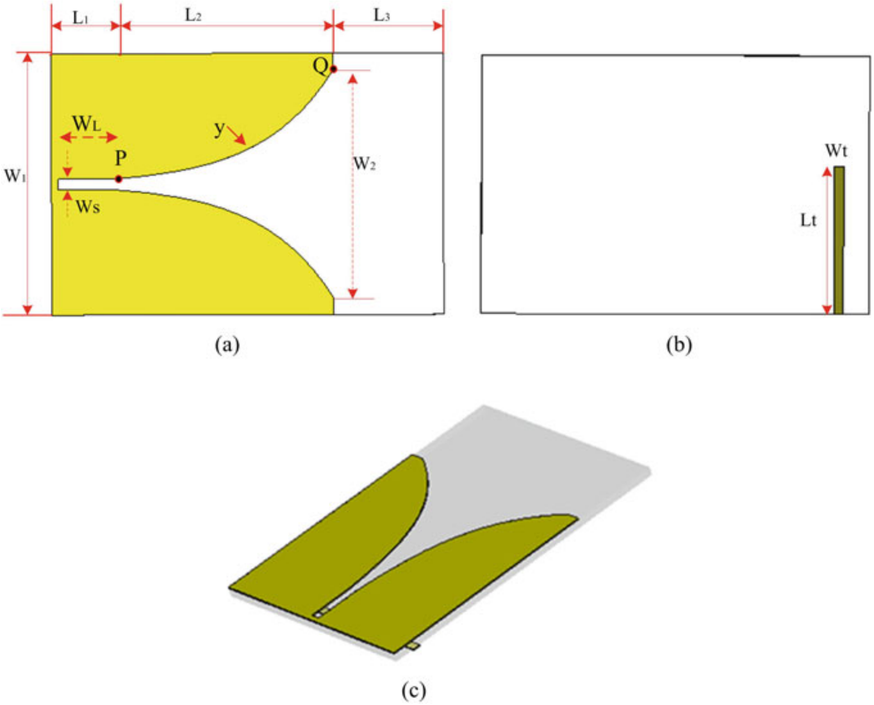


Fig. 1 Proposed geometry of Vivaldi antenna **a** Front view **b** back-view **c** Prospective view

Table 1 Dimensions of proposed antenna structure

Parameters			
L1	2.08	Ws	0.36
L2	7.24	Ls	1.85
L3	3.68	Wt	0.35
W1	8.7	Lt	5
W2	7.6	k	0.42

$$y = \pm[\frac{W_s}{2}e^{kx}] \tag{1}$$

where W_S is the slotted width and k is the tapering ratio for the metal flare structure.

In order to reduce the backward radiation, the cavity has been removed from the conventional design, and the length of the offset size is reduced. Also, the simple rectangular microstrip transmission line is employed to feed the flared arm. This provides enhanced radiation directive gain thereby reduced backward radiation and low side-lobe levels (SLLs). Further, a substrate length shall be extended to improve the directive gain, which acts as a director of produced radiations from flared arms

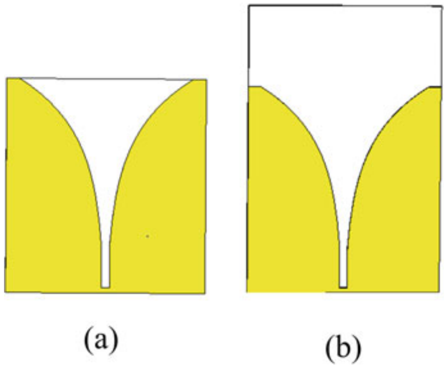


Fig. 2 Projected antenna **a** TVA-1 **b** TVA-2

Fig. 3 Proto-type design of TVA-2 and measurement setup

[4]. Here, the same rectangular structure is preferred for the director with L3 length to avoid design complexity. The proposed tapered antenna without cavity (TVA-1) and designed tapered antenna with the extended substrate (TVA-2) are shown in Fig. 2.

Moreover, the fabricated and measurement setup of tapered Vivaldi antenna with the extended substrate are shown in Fig. 3. The prototype antenna is connected to VNA-N524A for return loss measurement.

3 Results and Discussion

Tapered Vivaldi antenna, extended substrate Tapered Vivaldi antenna, and prototype Vivaldi antenna characteristics are compared in the frequency range 50–60 GHz. Simulated and measured return loss are shown in Fig. 4.

It is evident that all the proposed and fabricated antenna performance with return loss less is less than −10 dB. Also, the fabricated antenna's return loss is well-matched

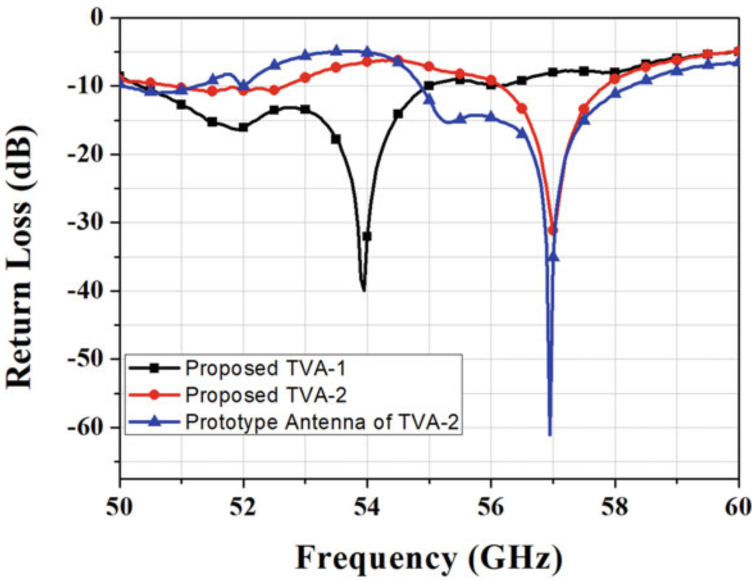


Fig. 4 Simulated and measured return loss of proposed antenna

with TVA-2. Gain and radiation efficiency performance of the TVA-1, TVA-2, and prototype antenna is shown in Fig. 5. The prototype antenna gain is around 1 dB more than TVA-1. The gain of TVA-2 is more than 9.8 dB over the frequency band of 50–60 GHz. Also, the radiation efficiency is more than 92% for TVA-2. Remarkably, gain, and radiation efficiency is well-matched with the designed prototype antenna.

The radiation patterns (E-plane and H-plane) of TVA-1 and TVA-2 at 51, 54, 55, and 56 GHz is displayed in Fig. 6. Notably, radiation patterns are nearly equal due to very less frequency variation. It is intended to integrate unique resonators structure along with changed extended substrates to improve the radiation performance in the future.

Table 2 illustrated the comparison of radiation characteristics to existing antennas in targeted band. Remarkably, the proposed antenna minimum size, along with high gain, impedance bandwidth and radiation efficiency.

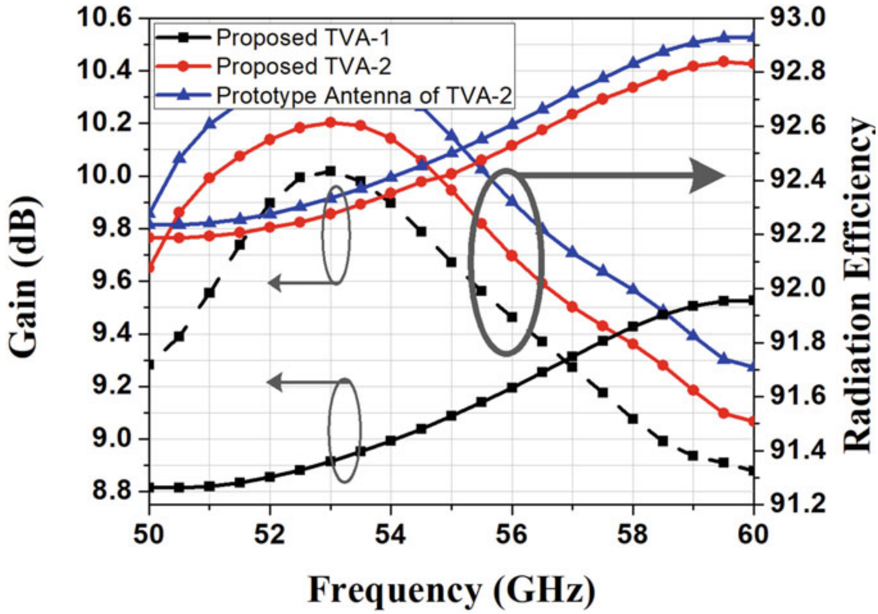


Fig. 5 Simulated/measured gain and radiation efficiency of TVA-1, TV-2 and prototype design

Table 2 Comparison of radiation characteristics of antenna with existing antenna

Ref	Frequency band (GHz)	Return loss (dB)	Gain (dB)	Radiation efficiency	Size (mm^3)
15	50–70	−44	6.3–7.5	–	$40 \times 34 \times 4.8$
16	58–64	−17	6.2–8.2	–	$20 \times 17.5 \times 0.254$
Proposed prototype TVA-2	50–60	−60	9.8–10.5	91.5–92.8	$15 \times 4 \times 0.254$

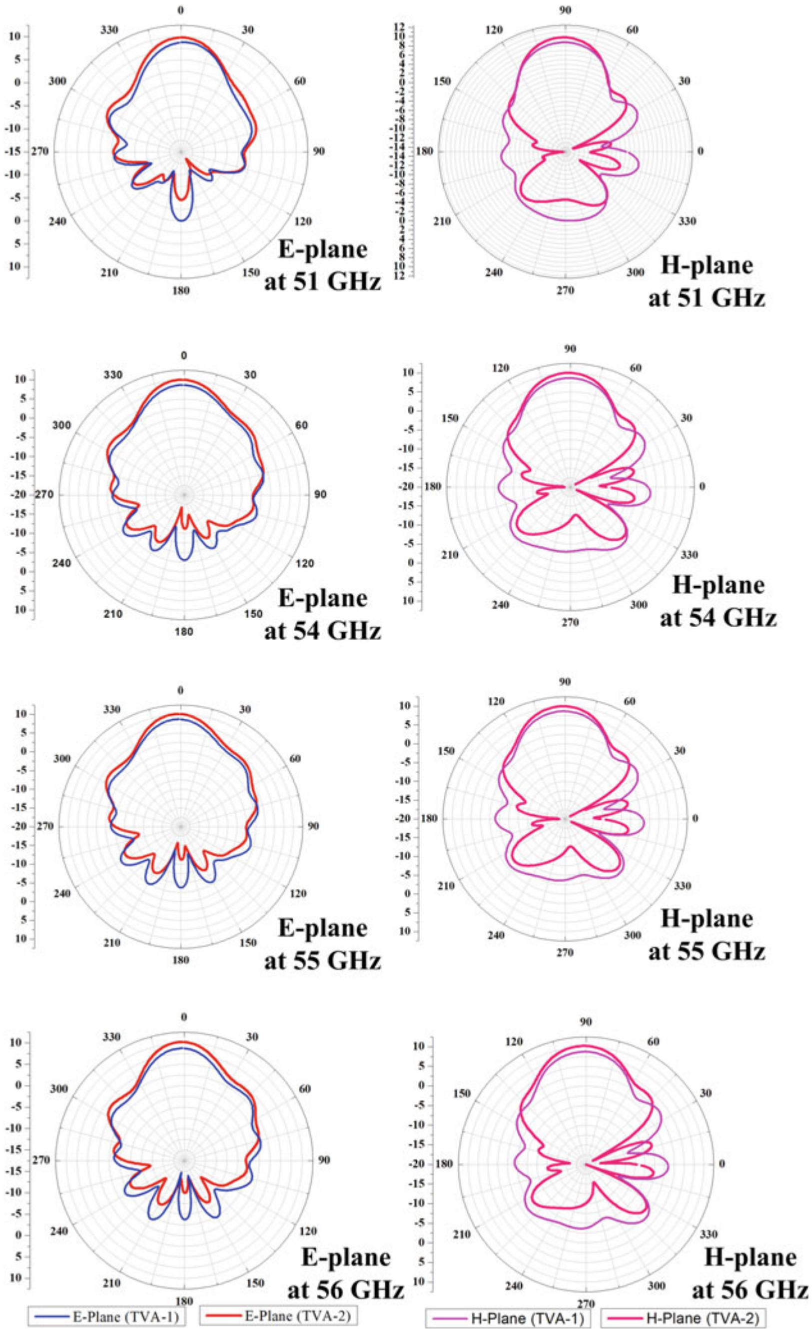


Fig. 6 Radiation pattern (E-plane and H-plane) at various operating frequencies

4 Conclusion

This paper presents Tapered Vivaldi Antennas for millimeter-wave applications. Firstly, the conventional tapered antenna is designed with some modifications, and then its radiation performances are improved by an extended dielectric substrate. Finally, the prototype TVA structure is fabricated and measured in the laboratory. The simulated and measured results are well-matched that validated by the designed TVA. The fabricated antenna provides around 1.8 dB higher gain and around 1% higher radiation efficiency in comparison to the first structure.

References

1. Wu, T., Rappaport, T.S., Collins, C.M.: Safe for generations to come: considerations of safety for millimeter waves in wireless communications. IEEE Microwave Mag. **16**(2), 65–84 (2015)
2. Zhu, F., Hong, W., Liang, W.F., et al.: A low-power low-cost 45-GHz OOK transceiver system in 90-nm CMOS for multi-Gb/s transmission. IEEE Trans. Microw. Theory Tech. **62**(9), 2105–2117 (2014)
3. Peytavit, E., Lampin, J.F., Akalin, T., et al.: Integrated terahertz TEM horn antenna. Electron. Lett. **43**, 73–75 (2007)
4. El-Nawawy, M., Allam, A.M.M.A., Ghoneima, M.: Design and fabrication of W-band SIW horn antenna using PCB process. In: 2013 1st International Conference on Communications, Signal Processing, and their Applications (ICCSPA). IEEE (2013)
5. Tajima, T., Song, H.J., Katsuhiro, A., et al.: 300-GHz step-profiled corrugated horn antennas integrated in LTCC. IEEE Trans. Antennas Propag. **62**, 5437–5444 (2014)
6. Gibson, P.J.: The Vivaldi aerial. In: Proc. 9th Eur. Microw. Conf., pp. 101–105, September 1979
7. Hahnel, R., Benedix, W.-S., Plettemeier, D.: Broadside radiating Vivaldi antenna for the 60 GHz band. In: 2013 International Workshop on Antenna Technology (iWAT). IEEE (2013)
8. El-Hameed, A., Anwer, S., et al.: A 60-GHz on-chip tapered slot Vivaldi antenna with improved radiation characteristics. In: 2016 10th European Conference on Antennas and Propagation (EuCAP). IEEE (2016)
9. Sultan, K.S., et al.: A 60-GHz gain enhanced Vivaldi antenna on-chip. In: 2018 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting. IEEE (2018)
10. Gao, C., Li, E., Zhang, Y., Guo, G.: A directivity enhanced structure for the Vivaldi antenna using coupling patches. Microw. Opt. Technol. Lett. **60**(2), 418–424 (2018)
11. Dixit, A.S., Kumar, S.: A survey of performance enhancement techniques of antipodal Vivaldi antenna. IEEE Access **8**, 45774–45796 (2020)
12. Wan, F., Chen, J., Li, B.: A novel ultra-wideband antipodal Vivaldi antenna with trapezoidal dielectric substrate. Microw. Opt. Technol. Lett. **60**(2), 449–455 (2018)
13. Moosazadeh, Kharkovsky, S.: A compact high-gain and frontto-back ratio elliptically tapered antipodal vivaldi antenna with trapezoid-shaped dielectric lens. IEEE Antennas Wireless Propag. Lett. **15**, 552–555 (2015)
14. Bourqui, J., Okoniewski, M., Fear, E.C.: Balanced antipodal Vivaldi antenna with dielectric director for near-field microwave imaging. IEEE Trans. Antennas Propag. **58**(7), 2318–2326 (2010)

15. Horst, M.J., Ghasr, M.T., Zoughi, R.: Design of a compact V-band transceiver and antenna for millimeter-wave imaging systems. IEEE Trans. Instrum. Meas. **PP**, 1–12 (2019)
16. Sharma, M., et al.: Design of a novel dual band printed antenna for future mobile applications. Procedia Comput. Sci. **171**, 917–923 (2020)

WebRTC Based Telehealth and Telecommunication

Aditya Mantri and Navjot Singh

Abstract COVID-19's somber headlines highlight the importance of remote patient monitoring as we come face to face with the pandemic every day. The increased demand on healthcare resources in an unprecedented time makes it difficult for these patients to access traditional centralized healthcare services. The condition may become worse if they are delayed in receiving care, or there may be risks associated with visiting a facility where COVID-19 patients also receive care. A decentralized, yet personalized healthcare model is certainly needed. With the technology we currently have, different telemedicine solutions have been implemented. There are often some disadvantages to these solutions, which includes: i) applications that are too particular; ii) proprietary technologies which are not compatible; iii) the need for IT personnel trained in maintaining the system; iv) the necessity of installing additional software; v) the costs involved. There is a real need for a low-cost, easy-to-use, and compatible video conferencing architecture. In this paper, we propose such a model. A model which is economical, user-friendly, and scalable at the same time. With SDKs and APIs, web realtime communication (WebRTC) is used to provide real-time telecommunication based medical service. With the help of these software development kit (SDKs), Health Monitoring devices are integrated with a patient's interface, allowing doctors to access vital signs through the screen-sharing feature powered by WebRTC.

Keywords e-Health · Mobile health · React · Remote patient monitoring · WebRTC · Wireless sensor network

A. Mantri (✉)
Motilal Nehru National Institute of Technology Allahabad, Prayagraj 211004, UP, India
e-mail: amantri7@gmail.com

N. Singh
Indian Institute of Information Technology Allahabad, Prayagraj 211015, UP, India
e-mail: navjot@iiita.ac.in

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_64

1 Introduction

Despite an increase in doctors per population, there are still not enough to monitor such a huge, ever-growing population. Unless there are significant structural and transformational changes in healthcare, it will be difficult for them to remain sustainable. Besides attracting, training, and retaining more healthcare professionals, we also must utilize their time to the greatest benefit of patients—caring for them. In order to improve the outcomes of chronic conditions and reduce the need for frequent medical visits, it is imperative that we embrace and accelerate the development of remote patient monitoring technologies. Additionally, in scenarios such as COVID-19, monitoring remotely can be a crucial tool for health care practitioners to assess their patients' conditions before they present for treatment.

The Internet of Things is currently being studied in various countries in relation to human health monitoring systems. The transfer of long-distance medical data (long-distance consultation and monitoring) has greatly advanced [1–8]. But most of the models achieving this type of advance follow a store-and-forward architecture in which the data collected from the patient is stored in a database or on a cloud server, and the data is retrieved by doctors at a later stage to analyze the patient's vitals. As a result, assistance is costly and time-consuming.

Furthermore, some of the existing models and implementations exhibit a problem of insufficient communication stability and their complexity of use. Despite receiving positive responses, video conferencing systems designed specifically for telehealth are not flawless. The purchasing price has been a concern for many small hospitals, clinics and users that are on a limited budget. In many cases, proprietary technologies are used in an ad hoc manner that is incompatible with other similar solutions. For example, Audio and video are processed using hundreds of different methods, but there is no interoperability between them. In many cases, downloading the software, understanding the setup and configuration instructions, as well as providing ongoing maintenance (upgrades), required a considerable level of IT expertise.

Because of these shortcomings, this paper presents a model for monitoring patients remotely in real-time, which will enable instant communication between stakeholders.

WebRTC emphasizes direct communication by enabling developers to build applications that can perform audio and video calling, share live video and screens as well as instant messaging. With in-browser media content support, end users no longer require downloading, installing, and manually configuring applications. Additionally, they do not require proprietary plug-ins in the browser to view media content. As a result of WebRTC, real-time communication can be accessed via a simple JavaScript API via any web application [9]. It aims to aid developers with creating browser-based cross-peer real time interaction that is often necessary for multimedia services such as video conference calls [10]. Thus, A WebRTC (web real-time communication) system was adapted to meet the model requirements of being simple, inexpensive, and interoperable.

In this study, data will be collected from a wearable device to calculate heart rate. Wireless sensors are used to collect information for a health monitoring system. React has been used to develop the interfaces for these sensors and the communication is facilitated via WebRTC as part of React integration.

Our model tends to be more effective than other existing models by-

- Providing real-time doctor-patient consultation using videoconferencing.
- Allowing interpretation of medical data using screen share facility.
- Ensuring no packet loss and reduced server overload.

2 Related Work

Lee et al. [11] discussed a device that enables remote patient monitoring via smartphone. It utilizes a smartphone app. In order to talk about the treatment process with the doctor, the patient can use a smartphone to video conference. Skype's AES-style encryption was used with this system to secure patient video data. However, real-time functionality is not available on this system. The collected vital signs need to be uploaded manually by the patient to the server.

A remote patient monitoring system utilizing RF technology was described by Roy et al. [12]. This system allows users to access medical information on mobile devices and through the web. Several components constitute the system, including sensor nodes, coordinator nodes, web and database servers, and graphical user interfaces (GUI). Data is collected by the sensor node. A central server receives the acquired data. The GUI allows users to view and analyze data. Unfortunately, in real-time mode, the system cannot function because of the delay in transmission and slow response from the server.

A new Android application designed by Mehmet [13] monitors the heart rate and heart rate variability of cardiovascular patients undergoing close and constant monitoring. The wearable sensors in the system continuously monitor the cardiovascular activity of the patient via wireless connectivity. The sensed signals are then transmitted wirelessly to an Android interface. Several parameters are configured to ensure that the system detects if the critical values of Heart rate, Heart rate variability, and Body Temperature are exceeded. When it is, the parameters and the patient's location are sent as an e-mail message and Twitter notification to the doctor and family members. The system does not offer real-time analysis of vital statistics. However, it permits patients to move around and to live a more fulfilled life in their immediate surroundings.

An e-health system based on IoT was implemented by Pap et al. [14] using Raspberry Pi devices. The data was recorded in the chart and the user could view, analyze, and download the chart using a web browser using a Node.js application. The system includes sensors for measuring blood pressure, pulse oximeters, airflow, temperature, and galvanic skin conductivity. Depending on the user's preferences, this system will record data either live or on record. Only temporary data is stored in the live session mode unlike the recording session, where the data will be stored permanently.

3 Proposed Model

Essentially, the proposed model can be broken down into two phases,

- Setting up a peer-to-peer connection between doctor and patient.
- Collection and transfer of health vitals from health monitoring devices, from patients, end to the doctor end.

In this way, instantaneous communication can be established which will make it possible for a patient to send their vitals data via webRTC as shown in Fig. 1.

In order to demonstrate the proposed model and verify its correctness a prototype was built.

3.1 Prototype Architecture

Establishing the P2P Connection Using WebRTC. This telemedicine model which we propose, is web-based. The patients connect to an HTTP server through a web browser and access the main HTML page that uses JavaScript and CSS also. The web page being accessed by the clients implements React JS code to enable the connection to the signaling server, which acts as a broker to coordinate the doctor-to-patient communication between the browsers.

PeerConnection, MediaStream, and DataChannel constitute the three core concepts of WebRTC API. Interaction between a local computer and remote peer is established using the PeerConnection interface. Connections can be made to remote peers, maintained and monitored, and closed when no longer required. When working with audio or video data, a MediaStream is used which defines the ways of working with the data, the restrictions associated with its type, and the success and error callbacks. Furthermore, DataChannel is used to represent a connection between two peers.

Primary steps for securing a connection involved were:

Fig. 1 Model overview

Exchanging SDP. An announcement or invitation is made using the Session Description Protocol, which allows the description of multimedia communication sessions. This protocol isn't used for media delivery but for negotiating between peers of different audio/video codecs, and for getting network topology and other information.

SDP offers are created using RTCPeerConnection.createOffer(). The caller sets this offer to be the local description of the connection by calling RTCPeerConnection.setLocalDescription(). This is followed by use of Signaling server by the caller to send the return offer to the intended recipient. Recipient records the remote description of the received offer using RTCPeerConnection.setRemoteDescription(). Upon this, An answer is created at the caller's end using RTCPeerConnection.createAnswer(). Passing the answer as input to RTCPeerConnection.setLocalDescription(), the recipient sets the answer's local description as the answer. Once again, the signaling server is used in order to send the answer to the caller. Finally, Response is received by the caller who Sets the answer using RTCPeerConnection.setRemoteDescription().

Exchanging ICE Candidates. A peer must also exchange information about the network connection in addition to information about the media. A candidate ICE (ICE candidate) indicates whether the peer can communicate via direct connections to the peer or via TURN servers. It is typical for peers to propose their best candidates first, working down the line to their worst candidates. Although UDP is preferred (since it's faster, and can recover from interruptions more easily), TCP candidates are also allowed under the ICE standard.

This exchange involves the callers to create and set local descriptions (SDPs). An ICE candidate is generated by STUN when the caller contacts the server. Signaling servers are used for exchanging ICE candidates, especially after STUN servers receive candidates from STUN and after setting up locally and remotely descriptors. ICE candidates are also generated by the STUN server at Callee's request. As soon as the signaling server is called, both sides call RTCPeerConnection.addIceCandidate(). An illustration of the entire process can be seen in Figs. 2 and 3.

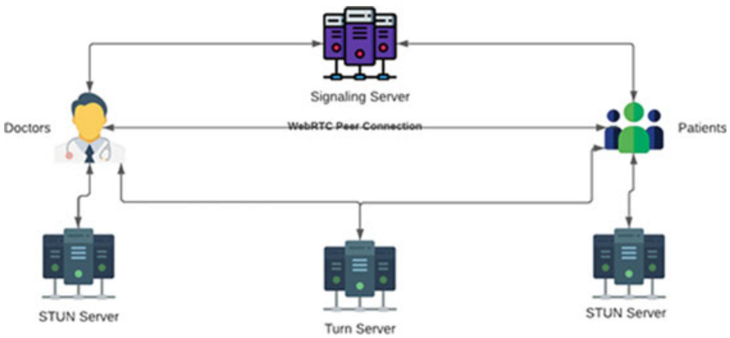


Fig. 2 Overview of WebRTC connection

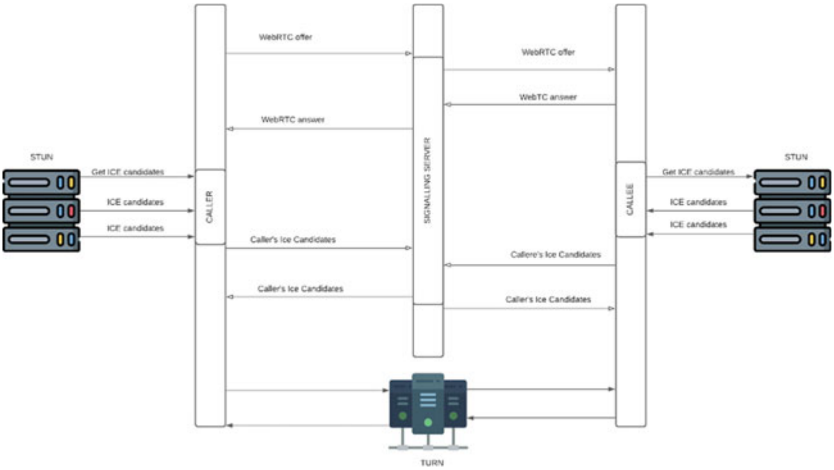


Fig. 3 Overview of exchange process to establish connection

Integrating Health Monitoring Systems. There are several different modules that make up the health monitoring system, such as the pulse rate recording module, the body temperature measurement module, the heart rate acquisition module, and the blood pressure monitoring module. The patient is connected to several sensors, which communicate with each other. In order to receive readings from the patient, sensors automatically convert the readings into signals.

The system can be configured to be able to send or retrieve data that will update an individual's record or create reports, depending on how the system is configured. Open source SDKs allow the communication between these health monitoring devices and the interface using Bluetooth/Wi-Fi.

Several methods of wireless communication are available. We have chosen Bluetooth technology since it does not require high data communication bandwidth to transmit medical data, including heart rate, that our prototype measures. Bluetooth's small power consumption, limited range of 10–100 m, along with its power and space saving characteristics makes it ideal for mobile devices and embedded systems that exchange relatively less data. Due to the frequency hopping spread spectrum (FHSS) method used by Bluetooth, the connection also has low interference. In today's world, Bluetooth technology is very attractive and can drive cost-savings and enhanced efficiency. With Bluetooth enabled devices, the wireless synchronization process is automatic. We have optimized our module to provide synchronized patient information across a variety of devices, including portable digital assistants, laptop computers, cell phones, and more [15]. Many devices and applications support Bluetooth for internet connectivity.

For testing purposes, we have only used the heart acquisition model. By measuring the difference in light transmittance in the blood vessels caused by the heartbeat, the light volume method measures the heart rate. Photoelectric sensors convert light sources into optical signals, and then an electrical signal is generated by a filter

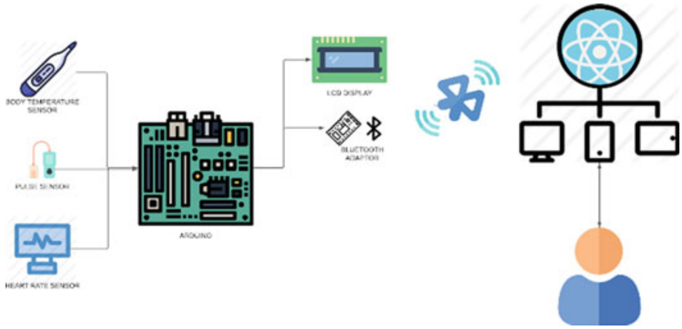


Fig. 4 Integration of telehealth devices with the interface

circuit using the optical signal. A wavelength between 650 and 750 nm was chosen. Figure 5 illustrates the heart rate sensor signal flow. The transmittance of the light source changes when a pulse of light passes through the peripheral blood vessels of the human body. During the amplifying process, a photoelectric converter reflects an optical signal through peripheral vessels of the human body is converted into electrical signals. Analog voltage is used to display the heart rate. Digital conversion is performed by Arduino, and quantities are transmitted via Bluetooth. Figure 4 depicts this integration process.

3.2 *Prototype Implementation and Interface*

The purpose of this section is to display the prototype implementation and design of a WebRTC-based video conferencing system to illustrate how the concepts described earlier were realized.

The developed application is divided into some distinct but complementary components. The core component of this application was responsible for connecting the peers with real time video and audio. In addition to its session server that identifies logged-in and available users, it also has real-time collaborative chat boxes as well as a screen sharing area.

Chat box area is used to send text messages and binary files. Care coordinators and patients can exchange data via a chat service in addition to their online meeting. This will allow prescription exchanges and the sharing of files. An online chat service requires both client and server code implementation. A chat channel is established from a socket.io server. Chat messages are sent to the client once the connection has been established.

Using the getDisplayMedia function we got the user's screen stream, inbuilt in the browser. Now, in order to send this stream, we used a package called SimplePeer in WebRTCPeerConnection. The other side of the connection will then start to fetch users screenshare streams using the getUserMedia function. Using this Screen-Share

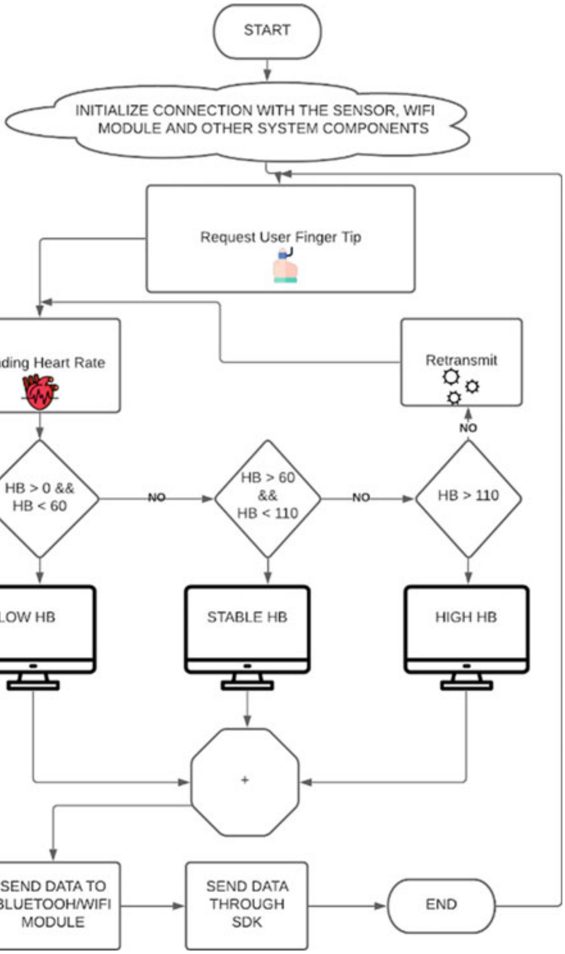


Fig. 5 Flow chart for heart rate sensing

feature, Patients can show vital statistics and data collected health monitoring devices to the doctors who will then be able to give their inferences and advice.

An open source React Native wrapper, MedM DeviceKit SDK, is used which provides a wide range of medical sensors to seamlessly connect via Bluetooth to acquire data such as heart rate, blood pressure, spirometer and temperature. A heart rate acquisition module was used to collect the data for our prototype.

In order to use the MedM DeviceKit SDK, Registering the Device Kit and Connecting it to our heart acquisition devices is done followed by vitals data collection. With this, device's data is acquired in XML format.

The React Native Web library is then used to allow use of React Native components in both web and mobile applications.

Finally, Users connected to the server will now need to identify themselves in order to access the web. Following the user's approval of the webcam and microphone control, he or she will see the interface. Our left-hand column displays the users that can initiate a WebRTC connection, and the list of common rooms created for group calls, which allows multiple participants in a single request, such as nurses, GPs and specialists. to join at the same time. Clicking on an active user starts the connection, if the callee is available, following which the patient will see the interface illustrated in Fig. 6. In the right column, the remote video appears after a WebRTC connection has been established, along with buttons to share screen, turn off local video, disable microphone, open chat box, and end the call. Upon screen-sharing by the patient, interface shown in Fig. 7 will appear at the doctor's end.

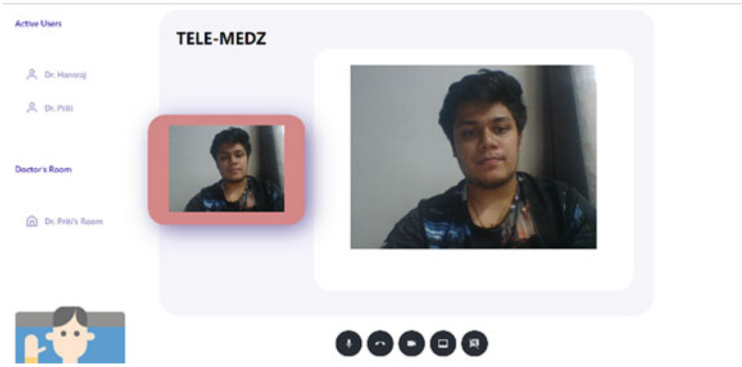


Fig. 6 Interface on connection

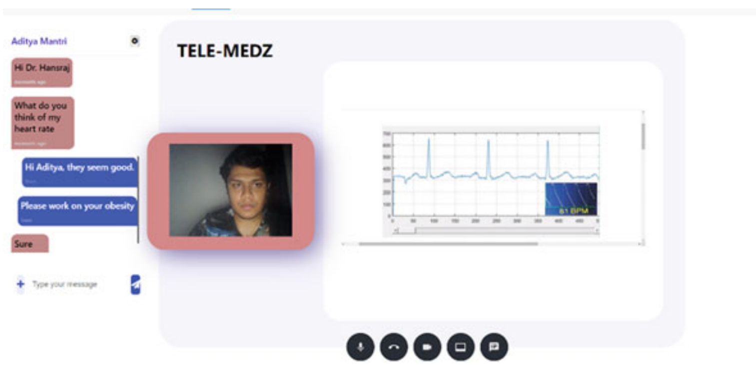


Fig. 7 Interface showing vitals data upon screen-share

Table 1 Performance of proposed model in private network

Module	Quantity of ECG packages	Latency/Delay (ms)
ECG packets sent from health monitoring device using MedM SDK	4500	3.019
ECG packets received by the mobile device	4500	

4 Results

A test was carried on the prototype using two tabs on the same browser connected to internet. There was a successful connection established as shown in Fig. 6 and we were able to transmit video (and voice), share screen, and transfer files and messages without experiencing any issues.

The Arduino microcontroller accurately received data from the heart pulse sensor. Data was then processed by the microcontroller and sent to the LCD screen for display. Wireless module then successfully transmitted the data via Bluetooth, and from there it was visualized on the interface without any errors as shown in Fig. 7.

To ensure that there are no packet anomalies, we evaluated total ECG data of a randomly selected set of 4500 packets between sender (health monitoring device) and receiver (mobile device). According to the results as shown in Table 1, the sum of the ECG values of the sender and subscriber are the same. In comparison to most existing models, the proposed model in private networks had very low latency and high performance.

5 Limitations

For our needs, we have developed and implemented a simple application that makes use of WebRTC. Few limitations are inherent in the current design and implementation. These limitations are described next, along with future plans to address them.

- Some deviations occurred during data acquisition from health monitoring devices. This was due to the sensor of the heart acquisition module being associated with the circuit design. This can be solved using sophisticated and advanced circuit hardware.
- We used Socket.io and Node.js in our web application to create a signaling server. Messages can be exchanged easily through this design choice. However, this model works well with very few users, such as the current situation. For a large number of participants, the model would not be scalable causing latency and delay issues. A signal server for this model that is capable of handling high volumes of participants can therefore be used, like SIP or Jingle.
- Bandwidth allocation while using the application deteriorated on having multiple tabs open in the browser. As a result, video conferencing reliability in browsers

was variable. By adjusting output bitrates to match the continuously fluctuating bandwidth, this issue can be solved.

6 Conclusion

In an unprecedented time of COVID 19, health needs are immediate and urgent. Now is the time to introduce remote patient monitoring standards. It is necessary to bring change to the centralized healthcare system to provide more personalized healthcare to improve outcomes and create a healthier world.

By using WebRTC technology, a standard-based interoperable, simple and cheap method of developing a video conferencing system is presented. A connection between patient and doctor was achieved by using Wireless Sensor Technology and a browser.

Based on an experimental prototype, this study was demonstrated, and a successful remote patient monitoring system was built. Our real-time communication capability enabled us to provide telehealth teleconferencing services promptly.

The paper also describes the adaptations that we made in the system architecture, as well as the component models. The prototype implementation also yielded several insights which have been discussed. Finally, we also pointed out some of the shortcomings of our current implementation.

References

1. Rezaeibagha, A.F., Mu, Y.: Practical and secure telemedicine systems for user mobility. J. Biomed. Inform. **78**, 24–32 (2018)
2. Agnisarman, S., Chalil Madathil, K., Smith, K., Ashok, A., Welch, B., McElligott, J.: Lessons learned from the usability assessment of home-based telemedicine systems. Appl. Ergon. **58**, 424–434 (2017)
3. Mugica, F., Nebot, À., Bagherpour, S., Baladón, L., Serrano-Blanco, A.: A model for continuous monitoring of patients with major depression in short and long term periods. Technol. Health Care. **25**(3), 487–511 (2017). https://doi.org/10.3233/THC-161289. PMID: 28009344
4. Tamilselvi, V., Sribalaji, S., Vigneshwaran, P., Vinu, P., Geetharamani, J.: IoT based health monitoring system. In: 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 386–389 (2020)
5. Acharya, A.D., Patil, S.: IoT based health care monitoring kit. In: 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), pp. 363–368 (2020)
6. Gregoski, M., et al.: Development and validation of a smartphone heart rate acquisition application for health promotion and wellness telehealth applications. Int. J. Telemed. Appl. **2012**, 1–7 (2012)
7. Oresko, J., et al.: A wearable smartphone-based platform for real-time cardiovascular disease detection via electrocardiogram processing. IEEE Trans. Inf Technol. Biomed. **14**(3), 734–740 (2010)
8. Trivedi, S., Cheeran, A.: Android based health parameter monitoring. In: 2017 International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 1145–1149 (2017)

9. Jennings, C., Hardie, T., Westerlund, M.: Real-time communications for the web. IEEE Commun. Mag. **51**(4), 20–26 (2013)
10. Jang-Jaccard, J., Nepal, S., Celler, B., Yan, B.: WebRTC-based video conferencing service for telehealth. Computing **98**(1–2), 169–193 (2014)
11. Lee, E.K., Wang, Y., Davis, R.A., Egan, B.: Designing a low-cost adaptable and personalized remote patient monitoring system. In: 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 1040–1046 (2017)
12. Roy, S., Rahman, A., Helal, M., Kaiser, M.S., Chowdhury, Z.I.: Low cost RF based online patient monitoring using web and mobile applications. In: 2016 5th International Conference on Informatics, Electronic Vision, ICIEV 2016, pp. 869–874 (2016)
13. Taştan, M.: IoT based wearable smart health monitoring system. Celal Bayar Üniversitesi Fen Bilimleri Dergisi, pp. 343–350 (2018)
14. Pap, I.A., Oniga, S., Orha, I., Alexan, A.: IoT-based eHealth data acquisition system. In: 2018 IEEE International Conference on Automation, Quality and Testing, Robotics, AQTR 2018 - THETA 21st Ed. Proceedings, pp. 1–5 (2018)
15. Elgharably, R., et al.: Wireless-enabled telemedicine system for remote monitoring. In: 2008 Cairo International Biomedical Engineering Conference, pp. 1–4 (2008)

Internet Security: Threats and Its Preventive Measures

Dharamdas Kumhar, Anil Kewat, and Avanish Kumar

Abstract Internet security is a broad issue that encompasses security for all transitions made over the Internet. Internet security provides authentication when data is exchanged over the Internet. Internet security is becoming a top priority, especially for financial and government institutions, as lack of security is a serious threat to the integrity and confidentiality of any organization. Since the Internet is an insecure channel for exchanging private data on messages and is susceptible to intrusion or frauds such as phishing. Therefore, it is necessary to implement certain methods to protect your data messages. In this paper, we provide an analysis of various security threats and the protective measures that can help the users to protect their data from unwanted access.

Keywords Pharming · Phishing · Internet fraud · Domain hijacking · Spamming · Malware · Scalping

1 Introduction

In today's digital environment, most of our daily activities are dependent on the Internet. Various forms of communication, entertainment, financial, and business activities are accomplished online. This means that huge quantities of sentient data and information are permanently being shared over the Internet. The Internet is typically privately owned and secure, but it can also be an insecure channel for transmitting the information. As a result of the threat of high levels of intrusion by hackers and cybercriminals, Internet security might become a top priority for businesses and individuals alike. Internet security is a set of security policies designed to protect activities and transactions that take place online on the Internet. These policies are

D. Kumhar (✉) · A. Kewat · A. Kumar
Department of Mathematical Sciences and Computer Applications, Bundelkhand University, Jhansi, India
e-mail: drprajapatibu@gmail.com

A. Kumar
CSTT, MOE, GoI, New Delhi, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_65

designed to guard users against threats like an intrusion into your computer systems, addresses of emails, or accessing websites, malware that infect and harm the systems, and fraud from hackers who steal personal information like credit card numbers and bank details [1]. Internet security can be chosen as a template for a wide range of concepts, including network security and computer security that focus on precise threats and vulnerabilities when online accessing and using the Internet. A threat could also be a threat that puts your system or practice at risk. From an Internet security perspective, a threat is a process where it causes significant damage to your computer systems. It can be divided into an activity that may occur or might not occur, but it may be forgivable enough to cause harm in this way. These threats can reveal a single computer system or a system network. Internet security is necessary because it is the process that preventing and detecting unauthorized use of your computer. Prevention measures help you to stop unauthorized users from accessing any part of your computer system. Investigating helps you to work out whether or not someone attempted to interrupt your system if they were well-turned, and what they may have done.

2 Internet Security Threats

The most common threats to Internet security come either from the use of malware or through frauds such as phishing [2]. The common threats to security are as follows:

2.1 Malware

The researcher and computer scientist Yisrael Radai have first utilized the word "Malware" in 1990 [3]. It's designed to thieve data from the computer device, damaging the computer equipment or harassing the users. Malware or malicious software (meaning software designed with the wrong intention) is software specifically designed to gain access to a computer either to disrupt its operation or to get sensitive data from it. It is a big threat to Internet security and includes computer viruses, spyware, worms, and Trojan horse. Malware is usually embedded with legitimate software that is either useful or attractive. Some common malware is listed in Fig. 1 and types of malware, what it does and their real-world example are given in Table 1.

I. Computer Virus

A computer virus is a small program that gets loaded into the computer without the user's knowledge and replicates itself repeatedly. It is dangerous because it will quickly use all available memory and finally halt the system. Even more dangerous types of the virus may corrupt or delete the file from the computer and may spread itself to another computer by using the user's email program. Resource of computer viruses includes the attachments in a message

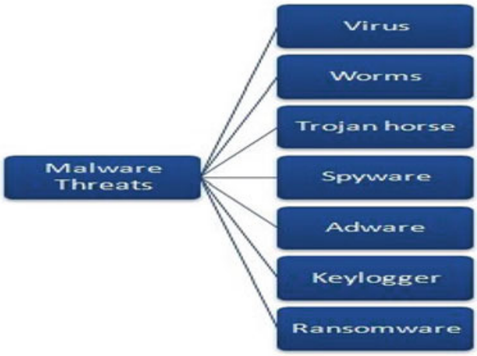


Fig. 1 Malware threats

Table 1 Type of malware, what it does and their real-world example

Type	What it does	Real-world example
Virus	Attached itself into the program and executes only when the program is run	Zeus
Worms	Loads by copying itself through a communication network	Stuxnet worm
Trojans	Disguises itself as desirable code	Emotet
Spyware	Search user data without knowing their background	DarkHotel Spyware
Adware	Provides undesirable advertising	Fireball Adware
Keyloggers	Monitors users' keystrokes	Olympic Vision
Ransomware	Disables victim's access to data until a ransom is paid	RYUK

of email or instant messaging messages, greeting card attachments that contain funny images for audio and video files as well as files that can be downloaded from the Internet [4]. Thus, a computer virus is always undesirable as they slow down the computer performance and causes erratic behavior; loss of data, and frequent crashes.

II. Worms

Like viruses, worms are programs written with malicious intention and can replicate themselves and spread across a computer network [5]. However, unlike viruses, most worms do not interface with the normal use of the computer. Moreover, they exist as separate entities and do not attach themselves to other files or programs. However, worms also take control over the computers on which they get installed and steal confidential data. Once a warm gets installed, it uses the email program of the user to send a copy of itself to everyone listed in his email address book. Then it replicates itself to send itself bent everyone cataloged in each of the recipient's address books. Therefore, it carries on with the replication procedure indefinitely. The worms use such a lot of network bandwidth and memory that causes servers of the web, services of the network, and particular computers to prevent responding. A worm once installed can connect to a remote computer over the Internet and download a more substantial piece of malicious software.

III. Trojan horse

Trojan horse is a non-self-replicating type of malicious software that pretends to be harmless so that users can easily download it on the computer. It is usually contained inside a harmless program. Once accomplished, a Trojan may decelerate the computer, causes harm or theft of data, given unprocessed access to its controller, destruction the file allocation table (FAT), or establish a virus [6].

IV. Spyware

Spyware is a malicious program that surreptitiously monitors the activity on a computer and reports that information to others without the user's knowledge. it is usually used for tracking and storing the user internet browsing patterns; gaining information about the user login, bank, and credit card information; servicing up pop-up advertisements to the Internet users; installing additional software; redirecting web browser to untrusted sides; modifying software setting; reading cookies; reducing network connections speeds, and causing slow down or even crashing of a computer system. Spyware is bundled as a hidden piece of code in a freeware or shareware program, which can be difficult to remove once downloaded from the Internet [7].

V. Keylogger

This is one of the energetic threats where a major part of the knowledge is tracked. So supported your keystrokes, the keylogger is going to be ready to keep a track of your activity. With the assistance of this program, the hackers are going to be ready to determine your personal information as password and username. They're distributing as malware and aren't finding out by software of antiviral. The keylogger inscriptions the key pressed by the user with a keyboard of the tampered device and sends the recorded data to the attacker without the knowledge of the users.

VI. Adware

Adware is also known as advertising supporting software is any software that is given to the user with advertisements embedded in this application. When users download from the Internet freeware or shareware that has adware embedded in it, the adware gets installed in the user's computer. It can also is spread through email attachments and shared files.

Adware comes under the category of malware because most of the time it is unwanted. It is a form of spyware that tracks user's Internet surfing habits and collects information about the user to display advertisements related to them. Users try to avoid adware, as they see it as a threat to their privacy and security over the Internet and because they are annoyed by the distraction caused by the adware. Adware display advertisements automatically without the user's permission. Unlike spyware, adware does not transfer the user's personal information to another location. However, it both slows down computer speed and allows constant pop-up advertisements to plague the user.

VII. Mobile ransomware

It is malware software that locks down the user's computer device and prevents them from accessing the information. The information of the infected computer system is encrypted by ransomware. When the payment is paid to the attacker, only then it decrypts the information. Both computers and cell phones are hits by ransomware. By changing the personal identification number (PIN) it locks down the cell phone and then asks for a payment to unlock the cell phone. There are two types of ransomware locker ransomware and crypto-ransomware. The locker ransomware locks down the computer devices and makes the information inaccessible to the owner of the computing device while the crypto-ransomware encrypts the file. Most of the ransomware indicates its presence within the computer devices to gather payments from the owner [8]. Ransomware is incapable of the victim's access to information while a ransom is purchased such as RYUK. Table 2 describes malware threats and their impact.

Table 2 Malware threat and their impact

S. no.	Threats	Recent attacks	Impact
1.	**Virus**	On August 15, 2012, a cyber attack against Saudi Aramco was triggered by the Cyber Shamoon virus [9]	Its 30,000 windows-based machines are infected
2.	**Worms**	On February 14, 2013, a Bizarre attack affected Linksys routers with self-replication worms [10]	1000 devices have been hit by the worm
3.	**Trojan horse**	In 2018, Emotet Trojan attract on Chilean bank Consorcio	Damages of USD 2 million
		In 2018, Emotet Trojan attracts the city of Allentown, Pennsylvania [11]	Losses of USD 1 million
4.	**Spyware**	In 2021, Pegasus Spyware attack France, India, Pakistan, Iraq, South Africa, Egypt, Morocco, Lebanon, Uganda, and Belgium [12]	Stealing of confidential information of requested user
5.	**Adware**	In 2017, Fireball Adware the highest infection rates were discovered in Indonesia, India, and Brazil [13]	According to Check Point researches, Adware called Fireball infected 250 million computers and devices
6.	**Keylogger**	In 2015, the University of California Irvine suffered incidents due to keyloggers [14]	Nearly 2,000 students had their personal and health information stolen after computers in the student. Health centers were compromised
7.	**Ransomware**	On January 14, 2021, Baltimore city was hit by RobbinHood ransomware [15]	It brought all city activities to a halt for weeks, including tax collection, transfers of property, and government emails. The attack cost the city more than $18 million

2.2 Denial of Service Attack

Denial of service (DoS) attack is an attempt to make a computer resource is unavailable to its intended users. Generally, denial of service attacks targets high-profile web services such as banks, credit card payment gateway, government organizations, media, and root name servers. The common way in which the denial of service attack can be made is as follow and listed in Fig. 2 and types of denial of service attack, what it does are given in Table 3.

I. Ping of flood

In this scheme, the attacker sends several external communication requests to the target machine so that either it is unable to respond to legitimate traffic or responds so slowly as to be rendered essentially unavailable.

Fig. 2 Denial of service attacks

Table 3 Types of denial of service and what is does

Type	What it does
Ping of flood	The attacker sends multiple external communication requests to the target machine so that it is either unable to respond to legitimate traffic or responds so slow that it is essentially unavailable
Ping of death	The ping of death occurs when an attacker crashes, destabilizes or freezes computers or services by targeting them with oversized data packets
Teardrop attack	Each packet has a sequence number so that the receiver can justify the message. In this, the attacker inserts a confusing serial number into the packet, making it difficult for the receiver to justify the message
Mail bombs	An unauthorized user may send a large number of email messages with large attachments to a particular mail server to fill its disk space so other users are denied email services
5G-based swarm attacks	A 5G swarm would allow its operators (such as cybercriminals) to rapidly find, share and correlate vulnerabilities and then change their attack methods to better exploit the vulnerabilities they discovered

II. Ping of death

Normally, a ping packet in the network should not exceed 65,535 bytes. A packet greater than this is not only difficult to handle, but also may cause the system to crash. Therefore, in this type of attack, ping packets of size greater than 65,535 bytes are deliberately sent.

III. Teardrop attack

In routing, a message is split into packets. Each packet contains a sequence number so that the receiver can reasonable the message accurately. In the teardrop method of denial of service, the attacker puts a confusion sequence number in the packets, thereby making it difficult for the receiver to reasonable the message. This may finally result in a system crash.

IV. Mail bombs

An unauthorized user may send a large number of email messages with large attachments to a particular mail server to fill its disk space so other users are denied email services. Although denial of service attacks does not cause loss of loss or theft of confidential data, they can cost the victim a great deal of time and money to handle the problem. In other types of denial of service attacks, the attacker identifies serious bugs in the target computer system and then causes the target system to crash by sending an input that takes advantage of the bugs. These bugs male lead to system crashes or may severely destabilize the system to such an extent that users are unable to access or use the system.

V. 5G based swarm attacks

With the development of the latest 5G technology and networks, higher speed displacement and huge quantities of data are often retrieved and uploaded more rapidly than ever. A fresh face of cybercrime has begun. High bandwidth-based attacks are much general than ever too; affecting the majority of technology. However, most of the attention is focused on the Internet of Things (IoT) and portable devices such as mobile devices. According to TechTarget, the number of swarm attacks has increased by over 80% in recent years and continues to grow. The swarm attacks nature collects multiples devices that are infected at the same time, which will then act on different attacking tasks, depending on their preface inside bot-coordinated attacks. This type of attack subconsciously uses Artificial Intelligence (AI) to urge new victims, change attack plans, and correlate and share information with the primary attacker. Table 4 describes denial of service attacks and their impact.

2.3 Internet Fraud

The copyright Internet fraud may be a sort of fraud that creates use of the Internet. It's not one fraud; under this, there are various frauds. Internets fraudsters are present everywhere and they use ground tricks to defraud people and withdraw money from their bank accounts. The downside of using the Internet is that it can be used for illegal activities such as theft of personal information, the activity of fraudulent transactions,

Table 4 Denial of service attack and their impact

S. no.	Threats	Recent attacks	Impact
1.	**Ping of flood attack**	Took place in February of 2020. AWS (Amazon Web Services) did not disclose which customer was targeted by the attack [16]	AWS reported mitigating a massive DDoS attack. At its peak, this attack saw incoming traffic at a rate of 2.3 terabits per second (Tbps)
2.	**Ping of death attack**	Mid 2013, ping of death attack for IPv6 packets for Microsoft windows [17]	Attack in Network layer and causes buffer overflow and system crash
3.	**Teardrop attack**	Took place in the year 2013. Windows and Linux are vulnerable to teardrop attacks, including Windows 7 and Windows Vista [18]	Attack on the network layer and generates an error in fragmentation and reassembly of packets
4.	**Mail bombs attack**	On August 18, 2016, a Massive email bombs attack the United States [19]	Hotmail account started getting flooded by several thousand subscription mails in the first 12 h
5.	**5G-based swarm attacks**	In 2019, swarm-based attacks in the U.S. [20]	Internet Speed Get Low

or passing fraudulent proceeds to financial institutions. Such fraud can take place in chat rooms, emails or newsletters, or on websites. The following are some common Internet fraud and listed in Fig. 3.

I. Purchases fraud

Purchase fraud occurs when it criminal purchases a product or service online and pays for it through fraudulent means, for example, by using a stolen or fake credit card. As a result, merchants do not get paid for the transition and lost money.

II. Online auction fraud

Online auction fraud occurs when a fraud starts an online auction of high-priced items on a website. He then accepts payment from the auction winner but either does not deliver the product or delivers a product that is less valuable than the one offered. This is the most common fraud on the Internet.

III. Online retail fraud

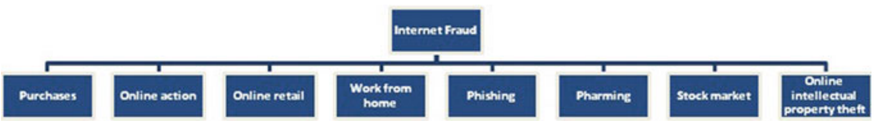


Fig. 3 Internet Fraud

Online rental fraud is similar to auction fraud in which after receiving the payment either the product is not delivered or an inferior product is delivered.

IV. Work from home

Work from home scam occurs when business opportunities are advertised on the Internet and users are asked to pay normal to substantial sums of money to get them registered. The fraudster collects the money; however, he or she never delivers the promised material or provides adequate information to subscribers but rather sends advice on how to place advertisements similar to the one that recruited him or her. In another scenario, the scammer accepts services from the victims e.g. writing directories, data entry, and reading books but then refuses to pay them by rejecting their work on grounds of substandard quickly.

V. Phishing

Phishing is a fraudulent activity in which a person or the business pretends to be trustworthy when it is not. It is done to acquire sensitive information such as passwords, account numbers, and credit card details. Phishing is usually done through official electronic notifications or messages such as email or instant messages [21]. For example, phishes will ask a potential victim to enter his or her password by displaying messages such as verifying your account, confirm building information, or enter your credit card information. Once the victim gives the details, the attacker can access the victim's information and use it for legitimate purposes, such as spamming, transferring money into his or her account, or online shopping with the victim's credit card number. While sending fraudulent messages or emails, fraudsters copy the code and graphics from legitimate websites. This is done to get the email a look and feel of legitimate communication from the trusted website so that the victims can comfortably provide them with their privacy and confidential information. Scammers may also provide links to their websites that seem to be legitimate.

VI. Pharming

Pharming occurs when a hacker exploits the vulnerabilities in the domain name system to redirect the website traffic from a legitimate website to the hacker's fraudulent website. This process of making the victim computer communicate with the wrong server is called domain hijacking. The fraudster constructs a fake website that looks similar to the legitimate site and asks for the user's personal information. For example, a fake website may ask the user to enter his password, PIN, and bank account number, credit card number, and so on to steal his information and misuse it. This combination of domain hijacking with a phishing website constitutes pharming.

VII. Stock market fraud

Stock market fraud includes an attempt to manipulate securities prices on the market for the personal profit of the scammer. The scammer usually follows any of the following two methods to do this fraud. The first method is the pump-and-dump scheme in which false information is spread to causes a dramatic increase in the prices of the thinly traded stocks through a chatroom, forums, internet boards, and email (as spam). This is called the pump. The moment the prices reach the desired level, scammers sell their stocks (called dump) to innocent victims, thereby making a substantial profit. Later, only when the prices fall to their usual level do the victims realize that it was all a fraud. In the second technique, called short-selling or scalping, the scammer spreads false information that causes and dramatic decrease in a stock price. Once the price falls to the desired level, the scammer buys the stock in bulk and then reverses the false information or waits for the company to refute the information in the media. Once the stocks regain their original price; the fraudster sells them, thereby making a high profit.

VIII. Online intellectual property theft

Individually all over the world who shares their notes and information on the Internet have an exclusive right to their material. However, many people just copy and use them without taking permission from the author. Some also copy other text, images, and multimedia data and put them on their website. Some may even copy someone else's program code or material and distribute it for free than the programmer had intended to sell it. This process of copy and using someone else's online material is called online intelligible property theft.

IX. Spam email fraud

Spam email is a common form of fraud in which the fraudster sends a bulk email to millions of email addresses to correct the receiver's computers, steal their identity, or fool them to pay for fraudulent products or services. These emails offer to recipients false dealings such as low-interest loans, lottery earnings, fancy business proposals, free credit report checks, and relationships with local singles. Sperm emails require the recipients to open the email and click on the link, which may also open up the computer to the virus, worm, and other bugs that will corrupt the computer. Table 5 describes Internet frauds and their impact.

3 Preventive Measures

Digital As more and more people are using the Internet for conducting their business, entertainment, and socializing the concern for privacy and security on the Internet is growing constantly. The following are some measures that users can take to preserve their privacy in the public domain and ensure security.

- Email filter: The good email filter will prevent many messages of this type. Employees should be trained properly to recognize spam emails and take action properly. Email saying you have when money should be ignored. If it was true, then such communication will not be sent through Yahoo or Gmail. It will be sent through the official email of that organization. Sometimes, we get sperm email from a friend climbing to have been mugged while on vacation overseas and asking for money. Ignore such email because that friend's email account may have been hijacked. In such cases, confirm the request for money through phone or

Table 5 Internet fraud and their impact

S. no.	Threats	Recent attacks	Impact
1.	**Purchases fraud**	In 2020, APP fraud on France [22]	Losses of USD 246 dollar
2.	**Online action fraud**	On 2020, Online fraud in India [23]	Losses of USD 1,614,120 dollar
3.	**Online retail fraud**	On 2021, E-commerce retailers fraud in China [24]	Loss of $20 Billion
4.	**Work from home fraud**	In September 2019 North America, Europe, Middle East & Africa, and the Asia Pacific were affected by work from home fraud [25]	These organizations experienced a total of 4,716 insider incidents during 12 months
5.	**Phishing fraud**	In 2018, a phishing attacker targeted Brazilian companies and individuals [26]	APWG member company Axur is located in Brazil. A Axur observed more than 17,600 fraud nexuses
6.	**Pharming fraud**	In September 2020, Evil Twin phishing attack on the U.S. Department of the interior's internal systems [27]	Hackers used evil twin phishing to steal unique credentials of about 30,000 emails and gain access to the department's Wi-Fi networks
7.	**Stock market fraud**	In 2021, pump and dump e-mail and social media, affinity, Ponzi & impersonation, and imposter [28]	Losses of $10 Million
8.	**Online intellectual property theft**	In 2020, attackers steal healthcare IP related to COVID-19 (Coronavirus disease) vaccines [29]	Because of these attacks, one vaccine manufacturer, Dr. Reddy's Laboratories, had to close several of their plants
9.	**Spam email fraud**	In January 2014, Fridge sends spam emails that attack smart gadgets [30]	The 100,000 devices are infected by the spam attack

any other way of communication. Do not click on the link given in emails received from an unknown sender. If there is a spam email asking for bank details, do not reply to it but rather call the bank to confirm the request. Make sure you do not call on any phone number included in the email, as it can also be a fake.

- Use the software of security: Install anti-spyware, anti-virus, anti-scamware, and anti-malware software on your computer, and update them regularly. Many software applications are now available for free on the Internet. They are used to detect and eliminate attacks of malware [31].
- Update programs and systems regularly: Updates enclose crucial security upgrades that help out protect against knowing bugs and penetrability. Make

sure that you keep your software and devices up-to-date to avoid falling prey to criminals. Do not use an obsolete operating system.

- Browser selection: Increase the browsing security settings. All browsers have their security features, but some browsers may have grave flaws that allow cybercriminals and hackers to exploit and attack. Make sure you are using a secure browser to minimize the hazard of your computer or network being compromised.
- Use strong passwords: Use strong passwords so that it is not easy for the hacker to guess them. Never select an obvious password such as your name, names of family members, or date of birth, or a simple sequence such as abcdef for 123,456. A good password should have at least 6 to 8 characters which is a good mix of an uppercase letter, lowercase letter, numbers, and punctuation symbols. Moreover, change the password every few days and do not share them even with friends and family members [32].
- Put up a firewall: Organization should also install firewalls to ensure data privacy. Firewalls are the most effective gateway protection amid your computer and the Internet and are one of the most important protections to preventing the expansion of cyber threats like malware and viruses. Make sure that your firewall devices are configured properly, and look at them regularly to make sure they have the most recent software/firmware updates installed, or they may not take full effect.
- Monitor for intrusion: Intrusion detections can be used to monitor different types of systems and customize the performance of the network. If the intrusion detection system suspects possible security breaches, it can send or generate an email or other warning based upon the nature of the detected operation.
- Server security: It monitors the protection of data and resources stored on the server. This also includes techniques and tools that to prevent intrusions as well as hacking and other malicious operations.
- Control access: Ensure that separates can only access information and services for which they are authorized. When accessing the Internet over a wireless network that is Wi-Fi, take for granted that the network is not secure even if its access is protected through the passwords. To ensure that the device does not unknowingly get connected to such networks, set devices to ask before joining networks.
- To avoid cases of non-delivery of products or services or delivery of inferior products, buy only from a trusted and well-known website that has a good reputation. Before doing any transaction, users should also read other user's feedback and comments on the product sold.
- Review your monthly credit card statement to ensure that they are accurate. In case of any discrepancy or doubt, immediately contact the issue of the credit card.
- Use pop-up blocker software to block pop-ups. Do not click on the popup alerts, not even on the cross to close the popup alert, because this may result in getting more pop-ups. Therefore, a better option, in this case, is to close the browser application.
- Do not reveal too many private details on the Internet especially on social networking sites. If it is necessary to disclose the details, then ensure that the website is a secured site. The address of the secured side begins with https, where "s" signifies secure. Moreover, remember that once some information is posted

on the Internet, it is difficult to take it back, because even after deleting it, a copy of the information can still be available with other websites or search engines.
- When accessing the Internet, do not forget to delete cookies, temporary files, or the history of the web page browsed.
- To do to a website, type its address in the address bar. Do not click on any link or cut and paste its address from unsolicited emails or web pages. Remember that links that look legitimate and genius may actually be bogus, specially designed to steal private information.
- Do not forget to sign out from email services sites or other accounts after using them.
- The organization should ensure confidentially by encrypting their message or files before transmitting them over the Internet.

4 Conclusion

Developing policies and procedures to deal with such threats is a big challenge for Internet security. Therefore, looking at the advancements in technology, we have also experienced a vast quantity of Internet security threats that an individual comes into contact with as well as affected their organizations. Therefore, one has to make sure that they understand that their personal information is important and shouldn't be compromised at any cost. Everyone should be educated and trained to maintain at least minimum security standards in their daily activities. There is no ideal solution for Internet crimes, but we must do our best to reduce them for a safe and secure future in the field of the Internet.

References

1. Rhee, M.Y.: Internet Security: Cryptographic Principles, Algorithms and Protocols. Wiley, New York (2003)
2. Data Security Overview BITUG, December 2017. https://dan-lewis-fns9.squarespace.com/s/Data-Security-BITUG-DEC-2017.pdf
3. Malware. https://searchsecurity.techtarget.com/definition/malware
4. The Internet comes down with a virus. The New York Times. https://www.nytimes.com/times-insider/2014/08/06/1988-the-internet-comes-down-with-a-virus
5. Barwise, M.: What is an internet worm? BBC (2010). http://www.bbc.co.uk/webwise/guides/internet-worms
6. Landwehr, C.E., Bull, A.R., McDermott, J.P., Choi, W.S.: A taxonomy of computer program security flaws, with examples. Naval Research Lab Washington DC (1993)
7. He, D., Chan, S., Guizani, M.: Mobile application security: malware threats and defenses. IEEE Wirel. Commun. **22**(1), 138–144 (2015)
8. Richardson, R., North, M.M.: Ransomware: evolution, mitigation and prevention. Int. Manage. Rev. **13**(1), 10 (2017)
9. International Institute for Strategy studies Survival: bi-monthly journal. https://www.iiss.org/en/publications/survival/sections/2013-94b0/survival--global-politics-and-strategy-april-may-2013-b2cc/55-2-08-bronk-and-tikk-ringas-e272

10. Bizarre attack infects Linksys routers with self-replicating malware. https://arstechnica.com/information-technology/2014/02/bizarre-attack-infects-linksys-routers-with-self-replicating-malware/
11. 11 real and famous cases of malware attacks. https://gatefy.com/blog/real-and-famous-cases-malware-attacks/
12. India Rejects use of Israeli Software Pegasus; what is Pegasus? https://youtu.be/AB7T5Ljj03s
13. FIREBALL – The Chinese Malware of 250 Million Computers Infected. https://blog.checkpoint.com/2017/06/01/fireball-chinese-malware-250-million-infection/
14. What is a keylogger? How attackers can monitor everything you type. https://www.csoonline.com/article/3326304/what-is-a-keylogger-how-attackers-can-monitor-everything-you-type.html
15. The 11 Most Common Types of Malware. https://www.crowdstrike.com/cybersecurity-101/malware/types-of-malware/
16. Famous DDoS attacks. The largest DDoS attacks of all time. https://www.cloudflare.com/learning/ddos/famous-ddos-attacks/
17. Ping of death DDoS attack. https://www.cloudflare.com/learning/ddos/ping-of-death-ddos-attack/
18. Kurose, J.: Computer Networking: A Top-Down Approach (2013). James F. Kurose, Keith W. Ross
19. Schneider, M., Shulman, H., Sidis, A., Sidis, R., Waidner, M.: Diving into email bomb attack. In: 2020 50th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), pp. 286–293. IEEE, June 2020
20. Confronting the threat of edge-enabled swarm attacks. https://internetofthingsagenda.techtarget.com/blog/IoT-Agenda/Confronting-the-threat-of-edge-enabled-swarm-attacks
21. Use a free password manager. https://web.archive.org/web/20160125015536/http://scsccbkk.org/Use%20a%20Password%20Manager%20for%20Security.pdf
22. Fraud: 5 Hot Scams In 2021, And How To Avoid Them. https://www.forbes.com/uk/advisor/personal-finance/2021/01/05/fraud-5-hot-scams-in-2021-and-how-to-avoid-them/
23. Surat: Two involved in Rs 120 crore online frauds held: The Time of India. https://timesofindia.indiatimes.com/city/surat/two-involved-in-rs-120cr-online-fraud-held/articleshow/83410934.cms
24. E-commerce fraud to surpass $20B in 2021, report finds. https://www.retaildive.com/news/e-commerce-fraud-to-hit-20-billion-2021-an-18-jump-from-prior-year/599395/
25. Working from home from invisibility to decent work. https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_765806.pdf
26. APWG, 2018. Phishing Activity Trends Report: 1st Quarter 2018. https://docs.apwg.org//reports/apwg_trends_report_q1_2018.pdf
27. 11 Types of Phishing + Real-Life Examples. https://www.pandasecurity.com/en/mediacenter/tips/types-of-phishing/
28. 5 Stock Market Fraud Schemes to Watch Out For in 2021. https://www.zamansky.com/5-stock-market-fraud-schemes-to-watch-out-for-in-2021/
29. 7 Best Practices to Prevent Intellectual Property Theft. https://www.ekransystem.com/en/blog/best-practices-to-prevent-intellectual-property-theft
30. Fridge sends spam emails as attack hits smart gadgets. https://www.bbc.com/news/technology-25780908
31. Ramzan, Z.: Phishing attacks and countermeasures. In: Stamp, M., Stavroulakis, P. (eds.) Handbook of Information and Communication Security, pp. 433–448 (2010)
32. Build Your Own Free Security Suite. https://www.pcworld.com/article/150204/free_security_suite.html

Correction to: New Idea for Characterizing of Waves Using Indian MST Radar

Rajendra Prasad Rao, Prakhar Yadav, and R. S. Yadav

Correction to: Chapter "New Idea for Characterizing of Waves Using Indian MST Radar" in: A. Dhawan et al. (eds.) *Advances in VLSI, Communication, and Signal Processing*, Lecture Notes in Electrical Engineering 911, https://doi.org/10.1007/978-981-19-2631-0_11

In the original version of the book, the following belated corrections have been incorporated: In Chapter 11 "New Idea for Characterizing of Waves Using Indian MST Radar", for the authors, Rajendra Prasad Rao and Prakhar Yadav, affiliation has been changed from "JK Institute of Applied Physics and Technology, University of Allahabad, Gadanki, India" to "JK Institute of Applied Physics and Technology, University of Allahabad, Allahabad, India", and for the author, R. S. Yadav, affiliation has been changed from "Electronics and Communication Engineering, JK Institute of Applied Physics and Technology, University of Allahabad, Gadanki, India" to "Electronics and Communication Engineering, JK Institute of Applied Physics and Technology, University of Allahabad, Allahabad, India". The book has been updated with the changes.

The updated original version of this chapter can be found at
https://doi.org/10.1007/978-981-19-2631-0_11

R. P. Rao (✉) · P. Yadav
JK Institute of Applied Physics and Technology, University of Allahabad, Allahabad, India
e-mail: rajendra.rao434@gmail.com

R. S. Yadav
Electronics and Communication Engineering, JK Institute of Applied Physics and Technology, University of Allahabad, Allahabad, India

A. Dhawan et al. (eds.), *Advances in VLSI, Communication, and Signal Processing*,
Lecture Notes in Electrical Engineering 911,
https://doi.org/10.1007/978-981-19-2631-0_66