# SCHOLARLY PUBLICATIONS

*A CURRENT AWARENESS BULLETIN*
*OF RESEARCH OUTPUT*

## @DTU

(77th Edition)

***MAY 2019***

## BY: CENTRAL LIBRARY


DELHI TECHNOLOGICAL UNIVERSITY
(FORMERLY *DELHI COLLEGE OF ENGINEERING)*
GOVT. OF N.C.T. OF DELHI
SHAHBAD DAULATPUR, MAIN BAWANA ROAD
DELHI 110042

# PREFACE

This is the **Seventy Seventh** Issue of Current Awareness Bulletin started by Delhi Technological University, Central Library. The aim of the bulletin is to compile, preserve and disseminate information published by the faculty, students and alumni for mutual benefits. The bulletin also aims to propagate the intellectual contribution of Delhi Technological University (DTU) as a whole to the academia.

The bulletin contains information resources available in the internet in the form of articles, reports, presentations published in international journals, websites, etc. by the faculty and students of DTU. The publications of faculty and student which are not covered in this bulletin may be because of the reason that the full text either was not accessible or could not be searched by the search engine used by the library for this purpose.

The learned faculty and students are requested to provide their uncovered publications to the library either through email or in CD, etc. to make the bulletin more comprehensive.

This issue contains the information published during **May, 2019**. The arrangement of the contents is alphabetical. The full text of the article which is either subscribed by the university or available in the web is provided in this bulletin.

**Central Library**

# CONTENTS


1. A State-of-the-Arts and Prospective in Neural Style Transfer, ***6.Sagar*** and ***3.Dinesh Kumar Vishwakarma***, IT, DTU

2. A Visual Cognizance Based Multi-Resolution Descriptor for Human Action Recognition using Key Pose, ***3.Dinesh Kumar Vishwakarma[1]*** and ***6.Tej Singh[2]***, [1]IT and [2]ECE, DTU

3. Analysis of Signals via Non-Maximally Decimated Non-Uniform Filter Banks, Sandeep Patel , Ravindra Dhuli, and ***7.1.Brejesh Lall***, Electronics and Communications, DTU

4. Color tunable photoluminescence properties in Eu3+ doped calcium bismuth vanadate phosphors for luminescent devices, ***6.Harpreet Kaur*** and ***3.M. Jayasimhadri***, Applied Physics, DTU

5. Design and analysis of an optimal ECC algorithm with effective access control mechanism for big data, ***3.Om Prakash Verma[1]***, ***6.Nitin Jain[2]*** and Saibal Kumar Pal, [1]ECE and [2]IT, DTU

6. Design of Low Power ECRL based Power Gated 4:2 Compressor, ***8.Naman Malhotra***, ***8.Aarushi Mann*** and ***3.Neeta Pandey***, Electronics and Communications, DTU

7. Effects of Flyash Addition on the Dry Sliding Tribological Behavior of Aluminium Composites, ***3.Vipin Kumar Sharma***, ***3.Ramesh Chandra Singh***, ***3.Rajiv Chaudhary***, ***8.Mayank Saxena*** and ***6.Mukul Anand***, Mechanical, DTU

8. Electrochemical properties of Na0:66V4O10 nanostructures as cathode material in rechargeable batteries for energy storage applications, Rakesh Saroha, Tuhin S. Khan, Mahesh Chandra, Rishabh Shukla, ***3.Amrish K. Panwar***, Amit Gupta, M. Ali Haider, Suddhasatwa Basu and Rajendra S. Dhaka, Applied Physics, DTU

9. Experimental Investigation on the Thermal Behavior of Untreated and Alkali-Treated Pineapple Leaf and Coconut Husk Fibers, ***7.Mohit Mittal*** and ***3.Rajiv Chaudhary***, Mechanical, Production and Industrial, DTU

**10.** Facile one-pot synthesis of 5-amino-1H-pyrazole-4-carbonitriles using alumina–silica-supported MnO2 as recyclable catalyst in water, ***3.Poonam*** and ***3.Ram Singh***, Applied Chemistry, DTU

**11.** High Isolation EBG Based MIMO Antenna for XB and Applications, Gaurav Saxena, ***3.Priyanka Jain*** and Y K Awasthi, Electronics & Communication, DTU

**12.** Inverse Document Frequency-Weighted Word2Vec model to recommend apparels, ***3.Priyanka Meel*** and ***7.Agniva Goswami***, IT, DTU

**13.** LECTOR incorporated differential cascode voltage swing logic (L-DCVSL), ***7.Pratibha Bajpai • 3.Neeta Pandey***, Kirti Gupta and ***3.Jeebananda Panda***, Electronics and Communication, DTU

**14.** Low Noise Amplifier in Smart Healthcare Applications, ***3.Malti Bansal*** and ***6.Jyoti***, Electronics and Communication, DTU

**15.** Management and Disposal of Mercury and Amalgam in the Dental Clinics of South India: A Cross-Sectional Study, ***8.Karthik Krishna Ramesh***, Maya Ramesh and Ramesh Krishnan, Environmental, DTU

**16.** Performance Evaluation of QAM for Improvement of BER, Balwant Singh, ***7.1.Vinod Shokeen*** and Rishi Asthana, Electronics and Communication, DTU

**17.** Performance of energy-efficient cooperative cognitive radio system over erroneous Nakagami-m and Weibull fading channels, Kiran Kumar Godugu, ***6.1.Ashok Kumar Nallagonda*** and Srinivas Nallagonda, Electronics and communications, DTU

**18.** Polycyclic aromatic hydrocarbons: soil pollution and remediation, ***6.Sakshi***, ***3.S. K. Singh*** and ***3.A. K. Haritash***, Environmental, DTU

**19.** Realisation of CBTA Based Current Mode Frequency Agile Filter, ***8.Kshitij Chhabra***, ***8.Sanjhi Singhal*** and ***3.Neeta Pandey***, Electrical and Electronics and Communications, DTU

**20.** Stability of a delayed SIR epidemic model by introducing two explicit treatment classes along with nonlinear incidence rate and Holling type treatment, ***6.Abhishek Kumar*** and ***3.Nilam***, Applied Mathematics, DTU

**21.** State-of-the-Arts Person Re-Identification using Deep Learning, ***7.Shradha Jaiswal*** and ***3.Dinesh Kumar Vishwakarma***, IT, DTU

**22.** The burden of prostate cancer is associated with human development index: evidence from 87 countries, 1990–2016, ***3.Rajesh Sharma***, Management, DTU

**23.** Variation of flow parameters and Magnetic flux in Viscoelastic Magnetic Abrasive Finishing Process, ***3.K. Srinivas***, ***3.Q. Murtaza*** and ***3.A.K Aggarwal***, Mechanical, DTU

**1.** *Chancellor*

| | |
|---|---|
| ***2.*** *Pro Vice Chancellor* | ***2.1.*** *Ex Pro Vice Chancellor* |
| ***3.*** *Faculty* | ***3.1.*** *Ex Faculty* |
| ***4.*** *Teaching-cum-Research Fellow* | ***4.1.****Asst. Librarian* |
| ***5.*** *Alumni* | |
| ***6.*** *Research Scholar* | ***6.1.*** *Ex Research Scholar* |
| ***7.*** *PG Scholar* | ***7.1.*** *Ex PG Scholar* |
| ***8.*** *Undergraduate Student* | ***8.1.*** *Ex Undergraduate Student* |

2019 6th International Conference on Signal Processing and Integrated Networks (SPIN)

# A State-of-the-Arts and Prospective in Neural Style Transfer

Sagar[1], Dinesh Kumar Vishwakarma[2]
Department of Information System
Delhi Technological University
Delhi, India
pawarsgr18@gmail.com[1], dinesh@dtu.ac.in[2]

***Abstract*—Artistic style transfer is one the enormous research fields in computer vision. It is a technique to synthesize a content image and the style together to form an stylized image. Various popular applications like Prisma, Deepart, Pikazoapp provide a platform to synthesis images in order to produce very exciting artistic style. These applications use Convolutional Neural Networks to transform the images in an artistic style. Prisma photo editor has more than 250 modern art filters and 110 million users. Hence, the amount of images being synthesized is huge and computation requires to perform the task efficiently is large. Abundant CNN architectures like VGG16, VGG19, GoogleNet etc are introduced to solve this problem of heavy computations. The system helps users to transform their images into artworks using style of famous artists for stirring image filters, and it might be useful for artists to have an indication about the art. This paper presents the fundamental concepts, classification of Style Transfer and major progress towards Neural Style Transfer.***

***Keywords*—Neural Networks, Style Transfer.***

## I. Introduction

Human imagination and their creative skills are generally described by art which is the most fascinating activity since Ancient time. Art in visual form can be referred as painting or sculpture which generates works to be appreciated primarily for their beauty and emotional power. We all have seen the work of famous artists like Pablo Picasso, Vincent Van Gogh, Frida Kahlo and so on, these are outstandingly created artworks where the high-level artistic skills and huge amount of time had to incorporate to get such results [1]. The evolution of paintings had been very revolutionary which is differentiated periodically: Medieval, Early Renaissance, Northern Renaissance, High Renaissance, Mannerism, Baroque, Rococo, Neoclassicism, Romanticism and Realism.

These art movements affected the world throughout human evolution but the most comprehensive discovery has ever been made was digital art which came into picture in 1950's. Digital art helped in various research fields which are well established now a day. To enable collection of expressive style for digital art, we can use Non-Photorealism Rendering (NPR) [1],[18]. Unlike, computer graphics which focuses on photorealism it is initiated by artistic styles and has wide variety of applications like video games, scientific visualization illustrations and many more. NPR for most part works or builds for particular artistic styles and difficult to extend to other styles [3]. Therefore, the style transfer has become the separate research field which is a study of re-composition images in the style of other images. The first study [2],[16],[12] on Neural Style transfer has been done by Gatys et al. in 2016, they produced very exciting results which catches eyes in both academics and industry. Convolutional Neural network is the key behind style transfer where there are various architectures with which Neural Style Transfer can be implemented. After Gatys paper there were extensive amount of researches took place and research papers published. It opened a new field of study that is exciting to work and visual to comprehend [2],[5].

In Neural Style Transfer there are broadly two input images [2] namely "Content image" and "Style image" where content of an image is defined as the feature response from a pre-trained CNN and style is summary feature statistics. The task is to convert an image into artistic style by changing the style however the key challenge is to preserve the desired semantic content. Researchers are able to achieve it using weights form pre-trained CNN models [10]. In the given Fig. 1, [2] the two input images have been taken namely content image and style image to produce an artistic output image where the style image can be any artistic image like "New York at night" and content image can be any image (i.e. London by day).



Style Image | Content Image

Transformed Image

Fig. 1 Neural Style Transfer (NST) of New York at night onto London by day illustrated by Gatys et al. [2]

Jing et al. [16] proposed a method in which they broadly classified Style transfer as given below:

- Rendering based on strokes which is a rendering procedure [13] in which virtual strokes are placed on a digital canvas.
- Image Analogy based on supervised technique in which we map between set of source image and stylized images.
- Neural Network which is an optimization technique to blend content image and style image.
- Image Filtering where we perform image simplification and abstraction
- Emerging homogeneous texture referring as Texture Synthesis.

In this paper, we will focus on the classification (i.e. figure 1) and Neural Networks optimization fundamental techniques and progress on style transfer.

In [16] on style transfer in which they classified the neural style transfer techniques in the following categories:

### *A. Slow Neural Method Based On Online Image Optimization*

It is the technique to transfer the style by iterative optimization whereas based upon the slow image reconstruction techniques. Out of so many researchers the first to attempt on slow image reconstruction technique using reverse CNN was carried out by DeepDream [17].

The notion behind it is to build the model and extract the relevant information from the respective images and then form the stylized output iteratively. The slow neural methods are categories in two categories, according to the way they model the style. The classification of neural style transfer is given in figure 2, NST has two majorly category: Slow Neural Method depending on Online Image Optimization and Fast Neural Method depending on Offline Image Optimization.

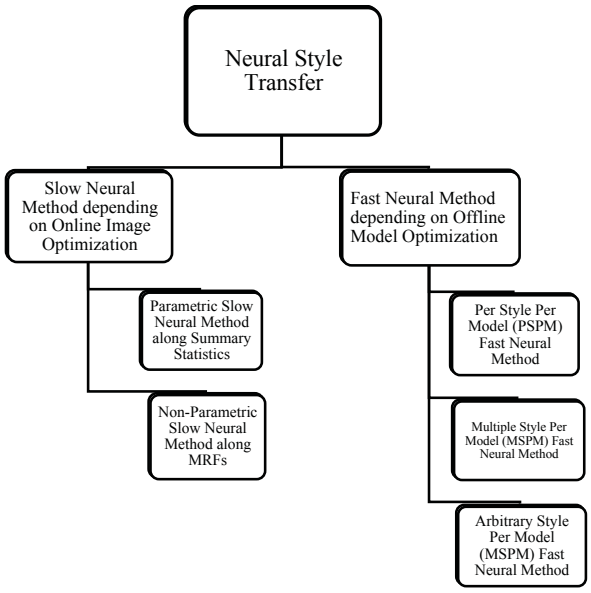



Fig. 2 Classification of Neural Style Transfer. [16]

1) Parametric slow neural method with summary statistics is the category of slow neural method in which style is described as a set of spatial summary statistics [2],[22].

2) Non-Parametric slow neural method with MRF's is slow neural method which transfer style by operating on patches compare style [4],[25].

### *B. Fast Neural Method Based On Offline Model Optimization.*

Due to few shortcomings and limitations like efficiency in Slow neural method the another category came into picture called Fast neural method which has the improved computational cost and speed [16]. Feed-forward network '$g$' can be optimized over the large set of images '$I_c$' with many styles images $I_s$, we can express it by the following equation 1:

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \mathcal{L}_{total}\left(I_c, I_s, g_\theta * (I_c)\right), I^* = g_\theta * (Ic) \quad (1)$$

On the basis of the amount of produced by individual $g$ it is further categorized as follows:

1) Per Style per model (PSPM) fast Neural method which at the testing stage pre-trains a feed-forward style specific network and generate a stylized output in single forward pass [19].

2) Multiple Style per model (MSPM) fast Neural method is proposed to improve the flexibility of the PSPM [20],[23].

3) Arbitrary Style per model (ASPM) fast Neural method is able to transfer arbitrary artistic styles with just one trainable model [21],[24].

## II. REPRESENTATION OF NST

To procure a visually pleasant image that is artistic stylized image one needs to provide the input images which can be called as content input image '$c$' and style input image '$s$'. Any ordinary image that needs to be filtered known as content image whereas the style image is the one whose style gets transferred onto the content image [2]. One can produce images in the style of high skilled art-work like the Starry Night by Vincent van Gogh and innumerable art work. If we are able to capture content and style of input images separately then this can be achieved.

Convolution Neural networks are capable of extracting low-level features(i.e., color, texture) from shallow layers and high level features from deeper layers to apply these features onto the semantic high level features of content image. Texture synthesis is constrained while synthesizing only style to preserve the semantic content in order get an artistic stylized image [2]. Content and style of an image are explained in the following sections.

### *A. Content of the input image*

Various aspect of the image content is learnt by each layer therefore it is rational to expect that the set of images having same content would have same feature maps. Building images having feature maps on a particular convolution layer matches the corresponding feature maps of the first input image that is content image [11][29]. Accordingly we can say two images

TABLE I. Classification of NST with respect to the recent researches.

| S. No. | Paper Title | Model/Network | Class | Publication Year |
|---|---|---|---|---|
| 1. | Yang et al. [2] | VGG19 | Parametric slow neural method along summary statistics | 2016 |
| 2. | Li et al. [4] | MRF | Non-Parametric slow neural method along MRFs | 2016 |
| 3. | Johnson et al. [19] | VGG16 | Per Style Per Model (PSPM) | 2016 |
| 4. | Domoulin et al. [20] | VGG16 | Multi Style Per Model (MSPM) | 2017 |
| 5. | Chen et al. [21] | VGG19 | Arbitrary Style Per Model (ASPM) | 2016 |
| 6. | Berger et al. [22] | VGG-19 network [Simonyan and Zisserman] | Parametric slow neural method with summary statistics | 2016 |
| 7. | Li et al. [9] | VGG-19 network [Simonyan and Zisserman] | Parametric slow neural method with summary statistics | 2017 |
| 8. | Li et al. [23] | VGG19 | Multi Style Per Model (MSPM) | 2017 |
| 9. | Li et al. [24] | VGG19 | Arbitrary Style Per Model (ASPM) | 2017 |
| 10. | Gu et al. [25] | VGG19 | Non Parametric slow Neural along MRFs | 2018 |

having same content need not to have same style and texture. Gatys et al. represented *l* as content layer and the content loss is the mean squared error of the feature map *F* where $\vec{p}$ and $\vec{x}$ are the original image and the generated image respectively [2]. $P^l$ and $F^l$ are the respective feature representation. Thus the formulation is given by:

$$\mathcal{L}_{content}(\vec{p},\vec{x},l) = \frac{1}{2}\sum_{ij}\left(F_{ij}^l - P_{ij}^l\right)^2 \quad (2)$$

Derivation of the loss with their respective activation results the gradient with the help of back-propagation. Minimizing the content loss, we obtain the close feature activation of generated imaged to the content image. Provided the layer we choose it transfers the outline from content image to generated image.

### B. Style of the input image

Style of the image is defined by the correlation between activations across channels. We formulate these correlations by Gram matrix $G_{ij}^l \in R^{N1 \times N2}$ where $G_{ij}^l$ is the inner product between the feature maps *I* and *j* in layer *l:*

$$G_{ij}^l = \sum_k F_{ik}^l F_{jk}^l \quad (3)$$

The dot-products of the feature activation results the Gram-Matrix of the style layer. Values in Gram-Matrix is a mean to decide whether the activations activate simultaneously or not. These activation patterns then help in creating the stylized output image [12],[15].

Now to calculate the total loss at the layer *l* of the original image $\vec{a}$ and the $\vec{x}$ generated image having the $A^l$ and $G^l$ are the style representations, the expression is give below:

$$E_l = \frac{1}{4N_l^2 M_l^2}\sum_{ij}\left(G_{ij}^l - A_{ij}^l\right)^2 \quad (4)$$

from the equation 4, The style loss is given as follows:

$$G_{ij}^l = \sum_{i=0}^{L} F_{ik}^l F_{jk}^l \quad (5)$$

The contribution to the total loss is assess by the weighting factors *w1* and the gradients are calculated by standard error back-propagation [2].

### C. Style Transfer

In order to transfer the style of one image onto another image one needs to synthesize the images in a way that they simultaneously compare the content and style representations [14]. The purpose behind this is that we minimizes the distance of feature representation of a white noise from content image and style representation from the stylized image on various layers of CNN [11], therefore the total over all loss function can be written as:

$$\mathcal{L}_{total}(\vec{p},\vec{a},\vec{x}) = \alpha\mathcal{L}_{content}(\vec{p},\vec{x}) + \beta\mathcal{L}_{style}(\vec{a},\vec{x}) \quad (6)$$

where α and β are the weighting factors for content and style representation [29] with respect to the pixel value cost gets minimized using Gradient Descent [2].

## III. Recent Progress

Gatys et al. proposed a revolutionary method of extraction semantic content of from the target image in 2015 since then it became salient research topic [2]. The algorithm worked quite well on pre-trained VGG-19 layered architecture [14]. They were able to generate high perceptual quality results via parametric slow Neural method [16] but still their model had some technical limitations. Resolution of the synthesized image was the major limiting factor. Berger et al. [22] presented a state of the art approach to satisfying long-range consistency constraints in the image synthesis and used VGG19 network Simonyan and Zisserman (2014) as the pre trained CNN inspired by Gatys. et al.

Li and Wand [4] used non-parametric slow methods using MRFs. They were able to transfer photo-realistic styles with some plausibility and were able to preserve the mesostructures in the synthesized image whereas the stricter control of mesostructures was the major limitation. Many researchers worked towards the fast Neural method based on offline model optimization and discovered new results. Johnson et al [19] worked with respect to Per Style Per Model (PSPM) in which to generate output image they merged the concept of feed-forward image transformation task and optimization based methods by feed-forward transformation network with perceptual loss function. Dumoulin et al [20] worked on Multiple Style Per Model (MSPM) and to model multiple styles at the same time they demonstrated that conditional instance normalization forms a simple, efficient and scalable modification. Whereas Tian Qi Chen and Mark Schmidt [21][24] presented a new CNN-based method that aims to be fast and adaptable to arbitrary styles using Arbitrary Style Per Model (ASPM). Their method concatenates the content and style in the form of single CNN layer, despite of their one layer restriction they were able to generate pleasant results. Their simplified solution sacrifices quality over efficiency, when it

comes to video frame-by-frame it leads to a local flickering effect.

## IV. Colnclusion

Artistic style transfer has been evolving field of study but from recently few years Neural style transfer is getting all of the attention by not only researchers but from the industries as well [6],[7],[27],[26]. In this literature survey we discussed about the current advancement in Neural Style transfer and also discussed about the fundamentals. Most used network by researchers is pre-trained VGG Network architecture by Oxford which achieve very satisfactory results on datasets like ImageNet. NST sometimes also referred as texture transfer in which image synthesis takes place. In order to synthesize the images the content and style cost must be minimized, this is carried out by backpropagation. Many industries coming up with the exciting and new demanding NST inbuilt applications like Prisma [6], Deepart [27] and Pikazo. In a very short span of time these applications got popular, millions of users are using these applications though there are still a lot to uncover in the field of NST. Hence Neural Style Transfer has great scope of advancement.

## References


[1] Y. Chen, Y. Lai and Y. Liu, "Transforming photos to comics using convolutional neural networks," *IEEE International Conference on Image Processing (ICIP)*, pp. 2010-2014, 2017.

[2] L. A. Gatys, A. S. Ecker and M. Bethge, "Image Style Transfer Using Convolutional Neural Networks," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2414-2423, 2016.

[3] L. A. Gatys, A. S. Ecker, M. Bethge, A. Hertzmann and E. Shechtman, "Controlling Perceptual Factors in Neural Style Transfer," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3730-3738, 2017.

[4] C. Li and M. Wand, "Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2479-2486, 2016.

[5] D. Ulyanov, V. Lebedev, A. Vedaldi, and V. Lempitsky, "Texture Networks: Feed-forward Synthesis of Textures and Stylized Images," *In ICML,* pp. *1349-1357*, 2016.

[6] Prisma Labs, "Prisma: Turn memories into art using artificial intelligence," [Online], Available: http://prisma-ai.com, 2016.

[7] A. J. Champandard, "Deep forger: Paint photos in the style of famous artists," [Online], Available: http://deepforger.com, 2015.

[8] M. Elad and P. Milanfar, "Style Transfer Via Texture Synthesis," in *IEEE Transactions on Image Processing*, vol. 26, no. 5, pp. 2338-2351, 2017.

[9] Y. Li, N. Wang, J. Liu, and X. Hou, "Demystifying neural style transfer," *arXiv preprint arXiv:1701.01036,* 2017.

[10] E. Risser, P. Wilmot, and C. Barnes, "Stable and controllable neural texture synthesis and style transfer using histogram losses," *arXiv preprint arXiv:1701.08893,* 2017.

[11] S. Li, X. Xu, L. Nie, and T. Chua, "Laplacian-steered neural style transfer," *In Proceedings of the ACM on Multimedia Conference*, pp. 1716-1724. ACM, pp. 1716-1724, 2017.

[12] D. Holden, I. Habibie, I. Kusajima and T. Komura, "Fast Neural Style Transfer for Motion Data," in *IEEE Computer Graphics and Applications*, vol. 37, no. 4, pp. 42-49, 2017.

[13] W. Chen, G. Sun, H. Li, Q. Sun and M. Shi, "A Style Transfer Algorithm of Caricature," *International Conference on Multimedia Information Networking and Security*, Nanjing, Jiangsu, pp. 49-53, *2010*.

[14] X. Huang and S. Belongie, "Arbitrary Style Transfer in Real-Time with Adaptive Instance Normalization," *IEEE International Conference on Computer Vision (ICCV)*, pp. 1510-1519, 2017.

[15] S. Prasad and B. N. Keshavamurthy, "An efficient artistic image-video transfer approach," *International Conference on Mechatronics and Machine Vision in Practice (M2VIP)*, pp. 1-4, 2016.

[16] Y. Jing, Y. Yang, Z. Feng, J. Ye, and M. Song, "Neural Style Transfer: A Review," *CoRR, abs/1705.04058,* pp 1-21, 2017.

[17] A. Mordvintsev, C. Olah, and M. Tyka. "Inceptionism: Going deeper into neural networks," pp. 1-17, 2015.

[18] J. E. Kyprianidis, J. Collomosse, T. Wang and T. Isenberg, "State of the "Art": A Taxonomy of Artistic Stylization Techniques for Images and Video," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 5, pp. 866-885, 2013.

[19] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual Losses for Real-Time Style Transfer and Super-Resolution," *ECCV*, pp. 619-711, 2016

[20] V. Dumoulin, J. Shlens, and M. Kudlur, "A Learned Representation For Artistic Style," *CoRR, abs/1610.07629*, 2016.

[21] T. Q. Chen, and M. Schmidt, "Fast patch-based style transfer of arbitrary style," *arXiv preprint arXiv:1612.04337*, 2016.

[22] G. Berger, and R. Memisevic, "Incorporating long-range consistency in CNN-based texture generation," *CoRR, abs/1606.01286,* pp. 1-17, 2016.

[23] Y. Li, C. Fang, J. Yang, Z. Wang, X. Lu and M. Yang, "Diversified Texture Synthesis with Feed-Forward Networks," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1-11, 2017.

[24] Y. Li, C. Fang, J. Yang, Z. Wang, X. Lu, and M. Yang, "Universal style transfer via feature transforms." In *Advances in Neural Information Processing Systems*, pp. 386-396. 2017.

[25] S. Gu, C. Chen, J. Liao, and L. Yuan, "Arbitrary Style Transfer with Deep Feature Reshuffle," *CoRR, abs/1805.04103*, pp. 1-10, 2018.

[26] R. Sreeraman, "Neural styler: Turn your videos/photos/gif into art," [Online], Available: http://neuralstyler.com/ 16, 2016

[27] "DeepArt," [Online], Available: https://deepart.io/, 2016

[28] K. Zeng, R. Zhang, X. Lan, Y. Pan and L. Lin, "Color style transfer by constraint locally linear embedding," *IEEE International Conference on Image Processing*, pp. 1121-1124, 2011.

[29] F. Luan, S. Paris, E. Shechtman and K. Bala, "Deep Photo Style Transfer," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6997-7005, 2017.

[30] E. Risser, P. Wilmot, and C. Barnes, "Stable and controllable neural texture synthesis and style transfer using histogram losses," *arXiv preprint arXiv:1701.08893,* 2017.

[31] B. J. Joshi, K. Stewart, and D. Shapiro, "Bringing impressionism to life with neural style transfer in come swim," *In Proceedings of the ACM SIGGRAPH Digital Production Symposium*, p. 5, 2017.

# Accepted Manuscript




Regular paper

A Visual Cognizance Based Multi-Resolution Descriptor for Human Action Recognition using Key Pose

Dinesh Kumar Vishwakarma, Tej Singh

PII: S1434-8411(19)30394-2
DOI: https://doi.org/10.1016/j.aeue.2019.05.023
Reference: AEUE 52756

To appear in: *International Journal of Electronics and Communications*

Received Date: 12 February 2019
Revised Date: 12 April 2019
Accepted Date: 15 May 2019

Please cite this article as: D. Kumar Vishwakarma, T. Singh, A Visual Cognizance Based Multi-Resolution Descriptor for Human Action Recognition using Key Pose, *International Journal of Electronics and Communications* (2019), doi: https://doi.org/10.1016/j.aeue.2019.05.023

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

# A Visual Cognizance Based Multi-Resolution Descriptor for Human Action Recognition using Key Pose

Dinesh Kumar Vishwakarma[1], Tej Singh[2]
[1]Department of Information Technology, [2]Department of Electronics and Communication Engineering
Delhi Technological University, New Delhi, India
[1]dvishwakarma@gmail.com, [2]ttomar07@gmail.com

**Abstract**

Human activity recognition using videos sequences is a well-known phenomenon which has many real-life applications such as daily assistive living, secuirity and surveillance, patient monitoring, robotics, and sports analysis. Recently, single or still images based action recognition is becoming very popular due to spatial cues present in an image and required less computation. Hence, a robust framework is constructed by computation of textural and spatial cues of still images at multi-resolution. A fuzzy inference model is used to select the single key image from action video sequences using maximum histogram distance between stacks of frames. To represent, these key pose images the textural traits at various orientations and scales are extracted using Gabor wavelet while shape traits are computed through a multilevel approach called Spatial Edge Distribution of Gradients (SEDGs). Finally, a hybrid model of action descriptor is developed using shape and textural evidence, which is known as Extended Multi-Resolution Features (EMRFs) model. The highest classification accuracy is achieved through SVM classifier on various human action datasets: Weizmann Action (100 %), KTH (95.35%), Ballet (92.75%), and UCF YouTube (96.36%). The highest accuracy achieved on these datasets are compared with similar state-of-the-art approaches and EMRFs shows superior performance.

**Keywords:** Human Action and Activity Recognition; Key poses; Fuzzy Logic; SVM, k-NN.

## 1. Introduction

Since the last few decades, vision-based human action recognition (HAR) is a lively area of research in the field of computer vision and pattern recognition. It has countless applications such as security and surveillance, assistive healthcare, human-computer interaction, robotics, user-interface design, video browsing, sports analysis, human interaction, human object tracking, robotics, and prevention of terrorist activities [1] [2] [3]. However, the task of recognising

ACCEPTED MANUSCRIPT

human action is challenging and complex due to the cluttered background, occlusion, and variation in viewpoint and lightening conditions [4] [5] [6] [7] [8] [9].

Recently, still images based action recognition becomes an evolving research topic in the area of computer vision. In comparison with video-based action recognition, the still image-based recognition tries to find out a person's behaviour or action using an only a single image. Although, most of the present literature regarding behaviour analysis uses videos analysis in which temporal and spatial cues are used. It can be considered more challenging to recognised action in still image than video analysis because it offers greater advantages as it does not involve the temporal information, clothing variation, environmental change, scaling, segmentation problems, and alignment of the images. Further, it reduces the computation time and complexity of the system. The action recognition based on poselets approaches [10] [11] [12] introduced in earlier works of still images but labelling and training was a complex task for these approaches. Later on, Delaitre et al. [13] used a local feature descriptor approach for action localisation in still images. A fuzzy logic model-based method for the prediction of human behaviour and recognition is presented in [14], and line pose based action detection [15] model motivates our work toward this research. By considering these advantages of still images and human physiological point of view, most of the human action can be easily discriminated and interpreted by observing a single still image without having a full view of the action image. Therefore, it raises a query in our cognizance that why it is not done automatically?

Moreover, in an attempt to answer these problems, we propose an EMRFs technique for action recognition in a still image by computing multiple spatial features rather than single features and these spatial features are extracted directly on the still image without segmentation of cluttered background. In still images, action recognition is based on the cue of the human body pose because it does not carry temporal information. Hence, no spatiotemporal based techniques can be applied for action recognition, and only spatial information based techniques can be used to estimate the action.

We have reviewed previous works for HAR in still images and there are numerous spatial-temporal information-based techniques have been proposed for action recognition, but very few works are reported based on still images [12] [13] [16].

Wang et al. [17] proposed a contour based human pose model to recognized action in still images with the help of a canny edge detector and a spectral clustering technique for clustering

ACCEPTED MANUSCRIPT

similar poses. However, they deal with the difficulty of action recognition only in the unsupervised manner and for single sports scene. Li and Ma [18] introduced a visual information based feature descriptor 'Exemplarlet' and showed that action recognition using still images dominated over spatiotemporal features. Further, to improve the recognition accuracy of an integrated model [19] which consist of appearance information with the occurrence of action scenes. Thurau and Hlavac [20] introduced a human pose model feature descriptor for action recognition based histogram of the gradient (HOG) on a selected region of interest (ROI) and represented a feature vector using non-negative matrix factorisation. Raja et al. [21] proposed a subspaces graphical information approach using connecting images for HAR in still frames. Bosch et al. [22] proposed a pyramid kernel descriptor based on local image shape and spatial layout of the objects. The local shape is computed via the spatial distribution of edges in a particular region and spatial layout by tiling the images into multiple resolutions regions. Liu et al. [23] proposed an approach to recognized actions from the unconstrained realistic video. They utilized both motion and static feature and AdaBoost for final classification. The work presented in Zheng et al. [24], gives a still image based hybrid technique combining poselet with contextual information. The poselet information is computed using poselet activation vectors, and contextual information is computed using sparse coding of foreground and background. A sequential approach for detection of human action in the non-controlled environment is presented by Conde et al. [25]. In this they use Gabor filter for preprocessing and enlightening the inherent information of the human pose and a further histogram of oriented gradient (HOG) descriptor is applied for extraction of shape and appearance information. They also observed that the cascaded model gives the improved performance of the HOG. In [26], used a deformable model using conditional random filed to extract the human body pose. For representing these poses, they used the rectangular patches on human silhouettes and determined the histogram of the oriented gradient to form the feature vector. In [27] they used non-negative matrix factorisation for high-level cues to recognise an action in still single images from a video frame of Weizmann dataset, and Google downloaded images dataset. Zhang et al. [16] proposed a systematic approached to detect the shape of human interaction regions and a product quantisation approached was used for action labelling to obtained feature from the HOI parts. Zhao et al. [28] proposed the Riemannian projection model in which each video is considered as an image set and Grassmannian point is extracted for every six frames and projected through into a subspace using

ACCEPTED MANUSCRIPT

SVD. Han et al. [29] proposed the dis-ordered multi-layer deep convolutional network, and they developed high-level features through transfer learning for action recognition in videos. In [30] they split each clip of different activities into image subsets and constructed Grassmann points for clustering analysis. They assumed that each action image set is noise and disturbance-free but practically it not possible. Ming et al. [31] proposed a non-linear dynamic system for action recognition using trajectories of the skeleton graph. However, they fail to classify 'turning' action properly. Safaei and Foroosh [32] introduced a CNN model based on a prediction of future motion of action in still images. They recognized shape and location feature in images with the help of the saliency map.

It has been observed from earlier works that still image based action recognition approach is better in processing time with fair accuracy as compared to [33] videos based action recognition and almost all works reveals that recognition accuracy has been tried to improve by incorporating key features of the scene by their own way. In more recent, some techniques have been presented that shows the performance of the system increases when multiple features [34] are incorporated for the representation of human action rather than the single feature. Therefore, we also trust that human body estimation can be accomplished accurately by extracting more than one features which describe the originality of the pose.

The primary objective of this work is to improve the recognition accuracy and view independency of the human actions in still images by incorporating the visual information that the human visual system perceives from the scene. In general, the human visual system is more sensitive towards the change in appearance, colour, and texture of the scene. The appearance-based shape and rotation information of the action is determined by using spatial edge distribution and Gabor wavelet respectively. The main contribution of the works are as follows:

- A fuzzy model-based approach used to select single key pose action images from input video sequences.
- The textural features at various orientations and scales are extracted through Gabor wavelet while shape features are computed through a multilevel approach called Spatial Edge Distribution of Gradients (SEDGs).
- An Extended Multi-Resolution Features (EMRFs) model is developed by concanetation of shape and textural evidences.

ACCEPTED MANUSCRIPT

- The performance of the EMRFs is measured in terms of accuracy on four publically available datasets such as Weizmann, KTH, Ballet, and UCF YouTube. The accuracy of EMRFs is compared with earlier state-of-the-arts and demonstrate superior performance.

The organisation of this work as follows: Section 2, gives details of the proposed framework of action recognition which comprehends selection of key pose from video sequences, feature extraction using SEDGs, and action recognition. In Section 3, the experimental setup, performance evaluation and result are discussed details. Finally, in Section 4 work is concluded.

## 2. Proposed Framework of Human Activity Recognition (HAR)

The proposed framework consists of three principal stages: a) Selection single key pose from the video sequences using fuzzy logic approach b) Pose representation by extracting extended multi-resolution features (EMRFs) c) Action recognition. The block diagram of the proposed work is depicted in Fig. 1. The details descriptions of each block are discussed in following sub-sections.

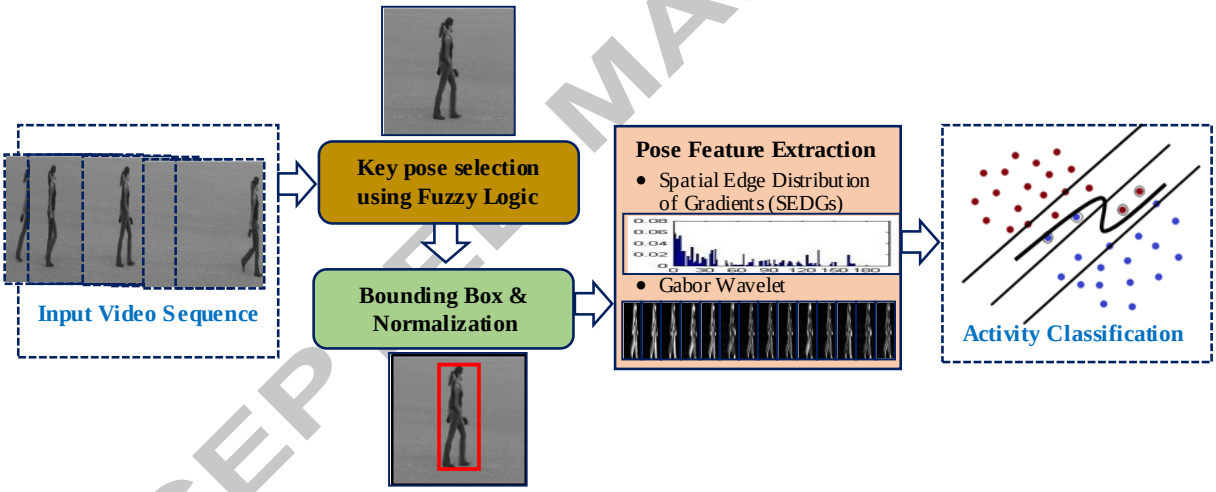



**Fig. 1.** Block diagram of the proposed framework based on Multi-Resolution Descriptor for HAR using key pose.

### 2.1 Selection of Single Key Pose using Fuzzy Logic Model

In the following sub-sections, we extract the stacks of key poses frames from input action video sequences using histogram distance between adjacent frames. Further, we select the single key still image based on Fuzzy logic inference model.

ACCEPTED MANUSCRIPT

### 2.1.1 Key Poses extraction from Action Video Sequence

Human activity can be effectively recognized with the help of key poses extracted from the video sequence and provides an explicit representation of human body posture. These selected key poses are chosen as a single prime key frame by using histogram distance which shows a spatial representation of 2D posture of human body motion. For choosing the key poses, the stacks of ten frames are selected at a regular interval from the input video sequences. Further, these stacks of frames are transformed into a different CIELab colour space. These converted stacks of frames are device invariant [35] as compared with other colour spaces. The histogram distances are calculated for three parameters: Luminance(L), and hue angle axis at 0º 'a' and hue angle at 90º 'b' of CIELab color space. The histogram distance (Ɗ) between the frames can be devised using Eq. 1.

$$Ɗ = \left\| \sum_{i=1}^{M} \sum_{j=1}^{N} S_{ij}^{t} - S_{ij}^{t+1} \right\| \quad (1)$$

Where $S$ is the stacks of frames of size $M \times N$ and $t$ represents the number of frames. The histogram distance is useful for measure instance changes in the frames of input video sequences. The computed distances for adjacent frames are utilized for the selection of key poses in the following sub-section.

### 2.1.2 Fuzzy Inference model

It can be visualized that abrupt scene change is a common phenomenon in videos. Some videos sequence shows little variation from frame to frame while other too fast. Therefore, it is not a good idea to extract key poses frames using normal distance function or some fixed threshold because of the risk of high information lost. Therefore, the fuzzy logic model is utilized to extract key pose images from video sequences. Hence, the histogram distances are calculated for adjacent frames to form the fuzzy rule-based model for selecting optimised keyframes.

The fuzzy logic models are based on the degree of membership and proposed by L. Zadeh [36]. This theory is the extension of Crips Boolean logic. Crips theory takes a hard decision whether a particular class belongs to a group or not such that:

$$\mu_A(x) = \begin{cases} 1 & if and only if\ x \epsilon A \\ 0 & if\ and\ only\ if\ x \notin A \end{cases} \quad (2)$$

On the other hand, Fuzzy logic is simple and deals with rule-based IF X AND Y THEN Z approach to assign the degree of membership to a particular class rather than to model a system

ACCEPTED MANUSCRIPT

mathematically. Fuzzy logic assigns a flexible membership between 0 and 1 to a variable class for a particular group. Fig.2. depicts the Fuzzy trapezoidal membership function which is used for selecting the key frames from video sequences.

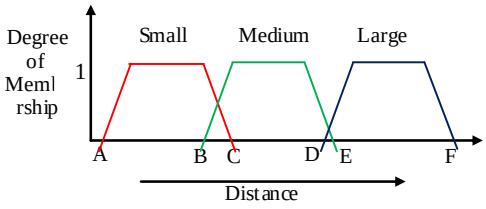



**Fig. 2.** Fuzzy Trapezoidal Membership function

In this approach, the keyframes are selected using the fuzzy model as shown in Fig. 3. The key pose frames are internally compared and ranked according to histogram distance metrics. It can be observed that the higher value of histogram distance frames shows maximum variations as compared to other frames. If the extracted key frames are denoted as $f_1 f_2 f_3 f_4 \ldots . f_n$, then the histogram distance is calculated for adjacent frames. the highest distance key frames that is having higher pixel difference as compared with other and selected as a single key still frame as illustrated in Fig. 3.

$$key\ pose\ frame = \arg\max\ (\mathcal{D}_i), \quad i \in n \tag{3}$$

The proposed algorithm for single key pose extraction is listed in Algorithm 1.

**Algorithm 1: Fuzzy Inference Model for Selection of Single Key Pose**

Step 1: The image frames are selected at the interval of 10 frames from the input video sequence.

Step 2: Selected frames are converted into new colour space 'CIELab'.

Step 3: Histogram distances are computed for $'\boldsymbol{L}'(\boldsymbol{lightness})$ , $'\boldsymbol{b}'(\boldsymbol{hue\ angle\ axis\ at}\ \mathbf{90°})$ and $'\boldsymbol{a}'(\boldsymbol{hue\ angle\ axis\ at}\mathbf{0°})$, parameters using Eq.(1)

Step 4: The means $\boldsymbol{\mu_\mathfrak{d}}$ are computed for all adjacent frames differences as:

$$\boldsymbol{\mu_\mathfrak{d}} = \left[\frac{\boldsymbol{Count\ L + Count\ b + Count\ a}}{\mathbf{3}}\right]$$

Step 5: Find the values of endpoints components of the membership function shown as in Fig.2 $\boldsymbol{A} = (\boldsymbol{\mu_\mathfrak{d}} - \boldsymbol{\mu_\mathfrak{d}} * \mathbf{0.4})$, $\boldsymbol{B} = (\boldsymbol{\mu_\mathfrak{d}} - \boldsymbol{\mu_\mathfrak{d}} * \mathbf{0.3})$, $\boldsymbol{C} = (\boldsymbol{\mu_\mathfrak{d}} - \boldsymbol{\mu_\mathfrak{d}} * \mathbf{0.2})$, $\boldsymbol{D} = (\boldsymbol{\mu_\mathfrak{d}} + \boldsymbol{\mu_\mathfrak{d}} * \mathbf{0.4})$, $\boldsymbol{E} = (\boldsymbol{\mu_\mathfrak{d}} + \boldsymbol{\mu_\mathfrak{d}} * \mathbf{0.5})$, $\boldsymbol{F} = (\boldsymbol{\mu_\mathfrak{d}} + \boldsymbol{\mu_\mathfrak{d}} * \mathbf{0.8})$.

Step 6: Create the trapezoidal fuzzy membership function for computed mean $\boldsymbol{\mu_\mathfrak{d}}$ as depicted in Fig. 2, where linguistic parameters are defined as: small, medium, and large.

- Rule 1: IF the distance between a segment frame and its neighboring segment frame is "medium" THEN it is a key frame.

ACCEPTED MANUSCRIPT

- Rule 2: IF the distance between a segment frame and its neighbouring segment frame is "large" THEN it is a key frame.
- Rule 3: IF the distance between a segment frame and its neighboring segment frame is "small" THEN it is NOT a key frame.

Step 7: Set the fuzzy rules based on neighbouring small or large distances frames to extract key pose frames.

Step 8: Compared to the selected frames with internal frames and ranked according to a histogram distances difference between them is a single still key pose and denoted as $\boldsymbol{f_s(x, y)}$ of size $\boldsymbol{M \times N}$ as depicted in Fig. 3.

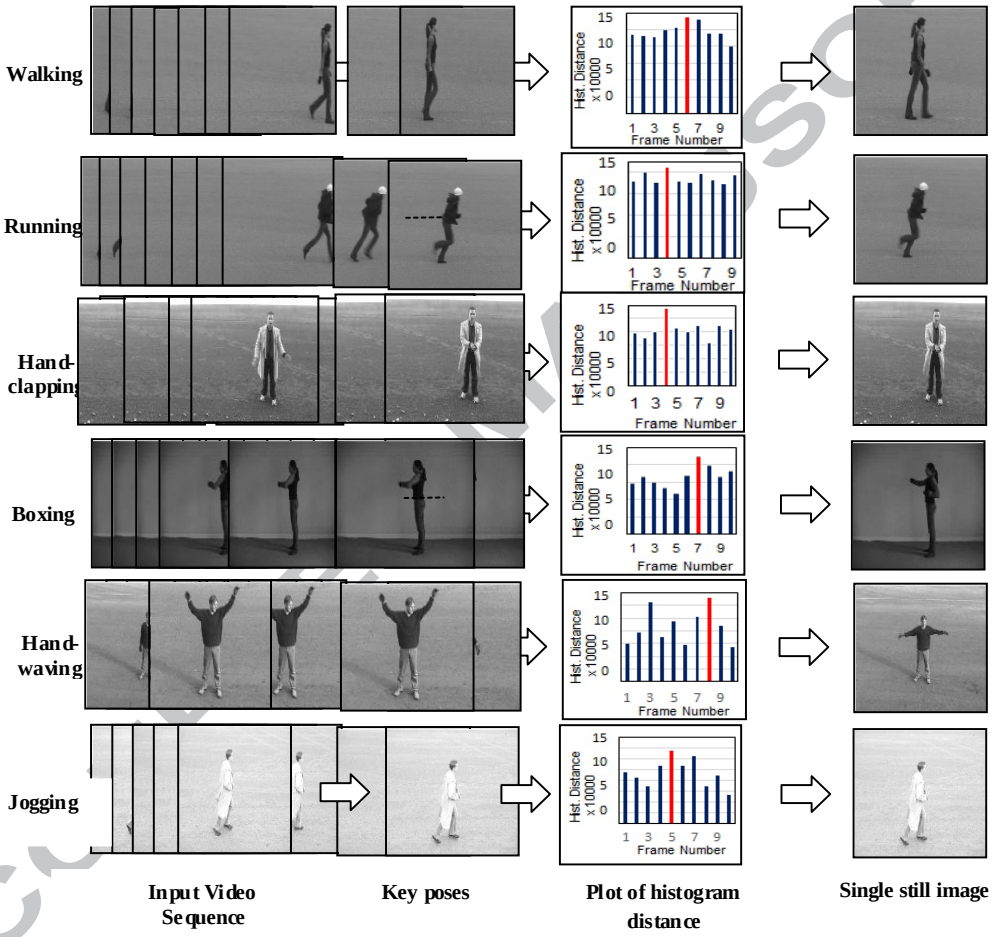



**Fig. 3.** The illustration of workflow for selecting a single still image from input video sequences of KTH dataset.

## 2.2 Extended Multi-Resolution Features (EMRFs)

It can be observed that actions are characterized by articulation and movement of the different body parts. The movement of the body parts reflects regarding pose, and most of the poses can be easily distinguished by the visual information in the scene, whether the scene is in hard printed copy or a soft copy of digital images.

ACCEPTED MANUSCRIPT

Based on the human visual perceptions of an action in a still image, the shape and orientation information of the human pose is used as a cue, to identify the kind of action pose. To extract out this evidence, we have applied Gabor wavelet and Spatial Edge Distribution of Gradients (SEDGs) to obtain the EMRFs for the action pose representation.

### 2.2.1 SEDGs Feature Map

The body posture performed by human contains information about body motion. The 2-D representation of these images gives spatial information about human motion. Such spatial distribution of postures provides the attitude of action behaviour of persons [37]. The single still images offer great advantages as they do not involve the temporal information, clothing variation, environmental change, scaling, segmentation problems, and alignment of the images. The still image approaches are less complex and time efficient. Edges detection are the most important task for efficient shape feature extraction from an image. In our approach we have efficiently utilized the Canny edge detector to obtain the edges and a threshold mechanism based on pixel variation are employed to remove unnecessary edges present in an image. To extract shape feature, the region of interest (ROI) is chosen which further divide into sub-regions at different sub-levels. The proposed algorithm of SEDGs is shown in Algorithm 2.

**Algorithm 2: SEDGs Feature Map**

Step1: Select 't' set of frames from input video sequences $\mathrm{F}(\mathbf{x},\mathbf{y},\mathbf{t})$.

Step 2: Select a single key pose frame from as described in section 3.1.

Step 3: Choose ROI and normalized it to the fixed spatial dimension 50×50, and denoted as:

$$\mathcal{R}(\mathbf{x},\mathbf{y},\boldsymbol{\phi}), \quad \textbf{for all } \mathbf{0} \leq \mathbf{x},\mathbf{y} \leq \mathbf{50}.$$

Step 4: Apply the canny edge detector to detect edges of selected ROI and given as:

$$\boldsymbol{\epsilon}(\mathbf{x},\mathbf{y},\boldsymbol{\phi}) = \textbf{canny}\langle\, \mathcal{R}(\mathbf{x},\mathbf{y},\boldsymbol{\phi})\rangle$$

Step 5: Find the spatial edge distribution vector at any point $(\mathbf{x},\mathbf{y})$ of the entire image $\boldsymbol{\epsilon}(\mathbf{x},\mathbf{y},\boldsymbol{\phi})$ different sub levels as:

**i.** At level-0, the Magnitude $\boldsymbol{M}(\mathbf{x},\mathbf{y}) = \sqrt{\mathfrak{N}_{\mathbf{x}}(\mathbf{x},\mathbf{y})^{\mathbf{2}} + \mathfrak{N}_{\mathbf{y}}(\mathbf{x},\mathbf{y})^{\mathbf{2}}}$ and Orientation $\boldsymbol{\phi}(\mathbf{x},\mathbf{y}) = \textbf{arctan}\frac{\mathfrak{N}_{\mathbf{x}}(\mathbf{x},\mathbf{y})}{\mathfrak{N}_{\mathbf{y}}(\mathbf{x},\mathbf{y})}$. Where $\mathfrak{N}_{\mathbf{x}}(\mathbf{x},\mathbf{y})$, $\mathfrak{N}_{\mathbf{x}}(\mathbf{x},\mathbf{y})$ are $\mathbf{x}$ and $\mathbf{y}$ direction gradients of image respectively. Each sub region is quantized into the 8 orientation bins evenly distributed between $\mathbf{0}°$ to $\mathbf{360}°$. The resulted feature vector for selected ROI is of $\mathbf{8} \times \mathbf{1}$ dimension.

**ii.** At level-1, the total region of an image $\boldsymbol{\epsilon}(\mathbf{x},\mathbf{y},\boldsymbol{\phi})$ is sub-divided into 4 sub-image regions, and represented as: $\boldsymbol{\epsilon}(\mathbf{x},\mathbf{y},\boldsymbol{\phi}) = \{\boldsymbol{\mathcal{S}}_{\mathbf{1}}(\mathbf{x},\mathbf{y},\boldsymbol{\phi}), \boldsymbol{\mathcal{S}}_{\mathbf{2}}(\mathbf{x},\mathbf{y},\boldsymbol{\phi}), \boldsymbol{\mathcal{S}}_{\mathbf{3}}(\mathbf{x},\mathbf{y},\boldsymbol{\phi}), \boldsymbol{\mathcal{S}}_{\mathbf{4}}(\mathbf{x},\mathbf{y},\boldsymbol{\phi})\}$. The feature vector of dimension $\mathbf{8} \times [\mathbf{1}+\mathbf{4}]$ is formed using (step 5-i).

**iii.** At level-2, Each of these sub-blocks (Step 5-ii) are further divided into four sub-blocks. A feature vector is of dimension $\mathbf{8} \times [\mathbf{1}+\mathbf{4}+\mathbf{16}]$ is obtained from 16 sub-block as in (step 5-i).

Step 6: A final feature vector based on spatial edge distribution is formed and summing all vectors of sub-levels and represented as: $\boldsymbol{\mathcal{F}}_{\boldsymbol{\epsilon}} = \mathbf{8} \times [\mathbf{1}] + \mathbf{8} \times [\mathbf{1}+\mathbf{4}] + \mathbf{8} \times [\mathbf{1}+\mathbf{4}+\mathbf{16}] = \mathbf{216}$. The results of the final feature vector is

ACCEPTED MANUSCRIPT

depicted in Fig. 4.

Fig. 4 and 5 depict the results obtained at a different level using SEDGs feature extractor. It can be observed from Fig. 5 that the extracted shape features are discriminative due to the variation of the histogram for different activities. Therefore, these features are robust to represents human activities. It is a fast and straightforward approach to calculate the features based on spatial shape information.

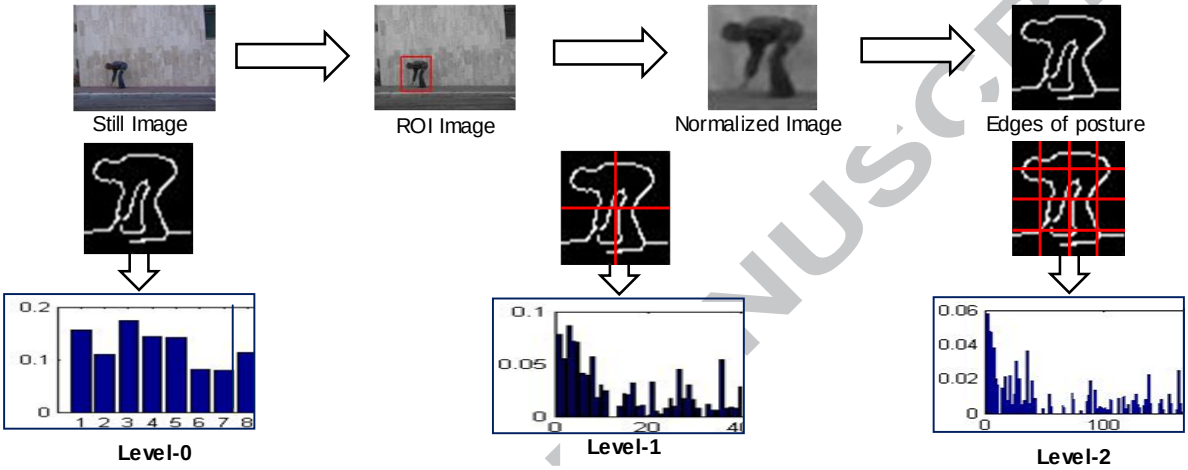



**Fig. 4.** Simulation Results of Proposed Algorithm 2.

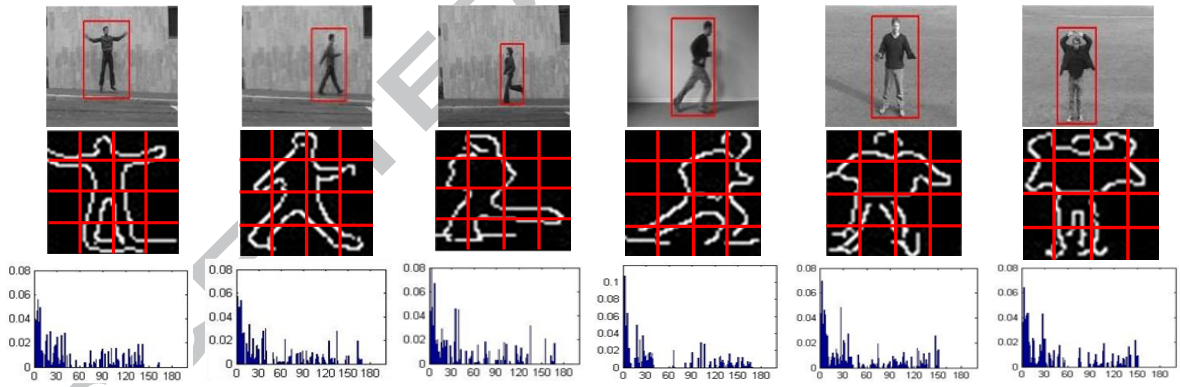


**Fig. 5.** First row shows Region of Interest (ROI) on various activities images, Second row shows edges computed on diffrent postures, and third row represents histogram of SEDGs at level '2'.

### 2.2.2 Orientation Feature Map

The orientation information of the action pose is extracted by Gabor filter, which is one of the widely used techniques for orientation and texture in the image. Arivazhagan et al. [38] introduced an invariant rotation approach for texture classification based on wavelet and defined as per Eq. (4).

ACCEPTED MANUSCRIPT

$$\psi(x,y) = (2\pi\sigma_x\sigma_y)^{-1}\exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)\right)\exp\,(2\pi j\omega x) \quad (4)$$

Where $\sigma, \omega$ are sacle and modulation frequency respectively. For a given still image $f_s(x,y)$ of size $M \times N$, the Gabor wavelet transform (GWT) at scale $p$ and orientation $q$, is obtained by convolving $f_s(x,y)$ with ${\psi^*}_{pq}(x,y)$ as per Eq. (5).

$$\mathcal{G}_{pq} = \sum\sum f_s(x,y) * {\psi^*}_{pq}(x,y) \quad (5)$$

Where ${\psi^*}_{pq}$ is complex conjugate of mother wavelet as given in Eq.(4). The Gabor wavelets are formed by exciting function as given in Eq. (6).

$$\psi(x,y) = a^{-p}\psi(X,Y) \quad (6)$$

Where $X = a^{-p}(xcos\theta + ysin\theta)$ , $Y = a^{-p}(xcos\theta - ysin\theta)$, $\theta$ is orenation parameter, $a$ is the scaling parameter, for $a > 1$ , $\theta = \frac{q\pi}{Q}$. $p = 0,1 \ldots \mathcal{P}-1, q = 0,1 \ldots\ldots Q-1$. $\mathcal{P}$ and $Q$ are the total number of scale and orienations respectivily.

The still image at various scales and orientations can be represented through the convolution of the image with Gabor wavelet transformed as shown in Fig.6, which have three scale and eight orientations. The obtained images at different orientations and scales are arranged according to the energy content. The orientation with the highest energy image is called the dominant orientations. Hence, the feature extracted from these images has to place first in the feature vector. The energy is computed using Eq. 7.

$$\mathcal{U}(p,q) = \sum_x \sum_y \|\mathcal{G}(x,y)\| \quad (7)$$

The mean and standard deviation of all the transformed coefficient is computed using Eq. 8. These values represent the region of homogenous texture in the image.

$$\mu_{pq} = \frac{\mathcal{U}(p,q)}{MN} \text{ and } \sigma_{pq} = \sqrt{\frac{\sum_x \sum_y \|\mathcal{G}(x,y)\| - \mu_{pq}}{MN}} \quad (8)$$

A Gabor feature map $\mathcal{F}_G$ are formed for $\mathcal{P}$: Scales and $Q$: Oreinations as in Eq.9.

$$\mathcal{F}_G = [\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \mu_{02}, \sigma_{02}, \ldots\ldots\ldots, \mu_{\mathcal{P}-1Q-1}, \sigma_{\mathcal{P}-1Q-1}] \quad (9)$$

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

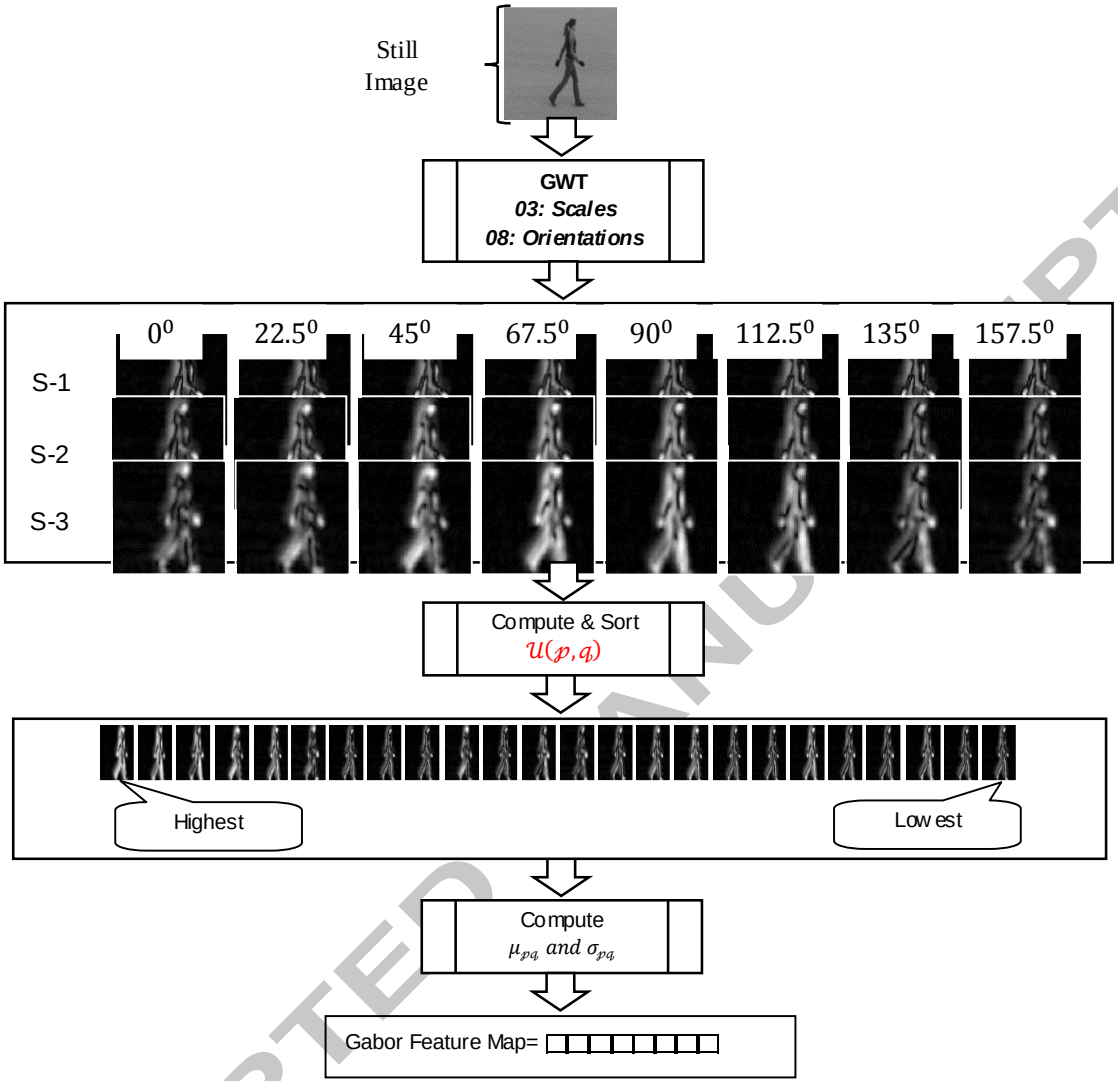



**Fig.6.** Procedure for extraction of orientation feature map

### 2.3 EMRFs Representation

The EMRFs model is inspired with the human visual cognizance [39], and based on human visual perception and cognizance a model is introduced in Fig. 7 for concatenation of texture and shape evidence. The EMRFs representation is achieved by combining the Gabor feature vector with spatial edge distribution feature vectors. The EMRFs provide an informative representation of the human body pose by capturing the multiple features at multiple resolutions, which includes the orientational texture, and shape.

ACCEPTED MANUSCRIPT

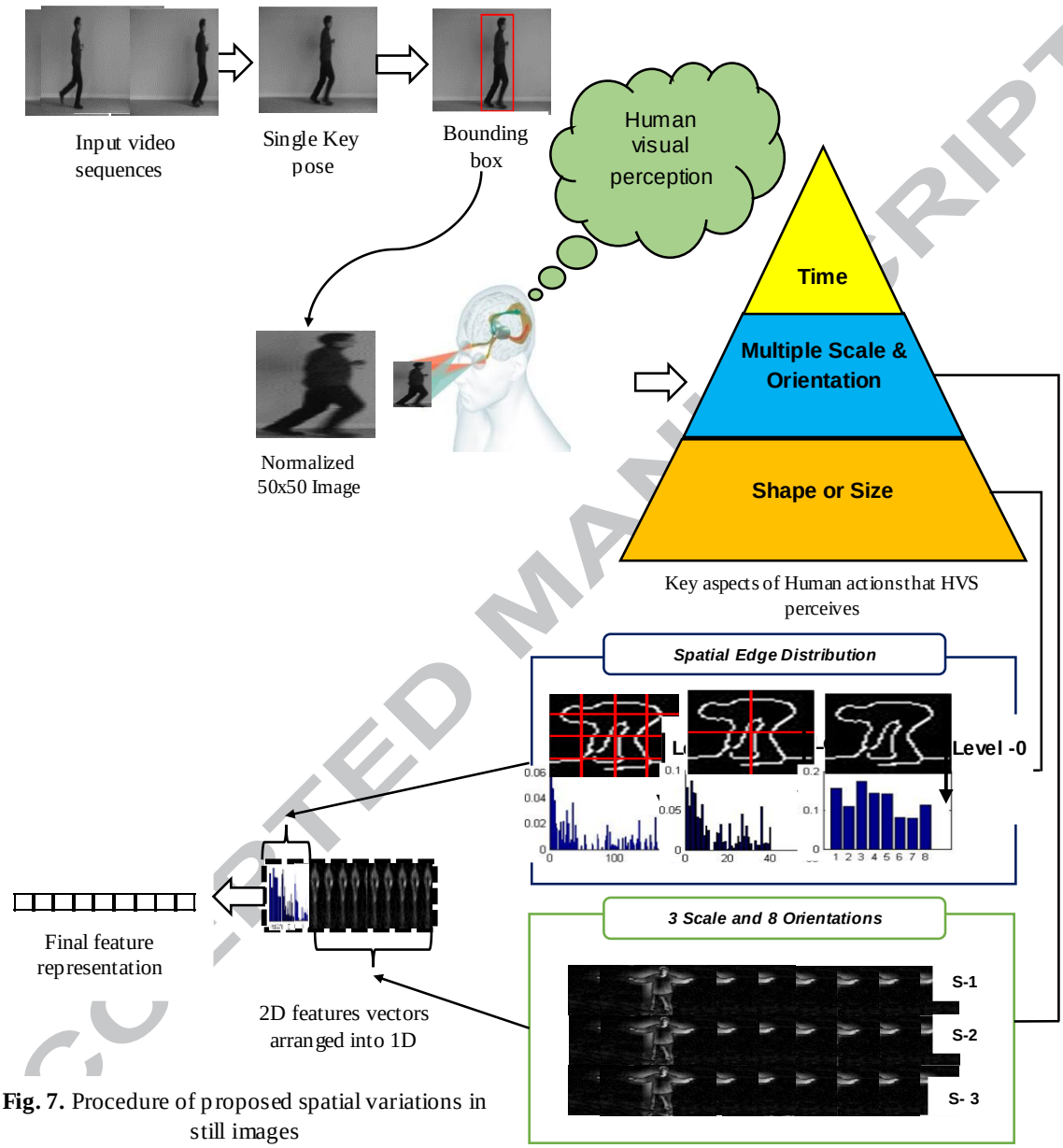



**Fig. 7.** Procedure of proposed spatial variations in still images

The features are computed at multiple scale and resolutions, hence EMRF relish the property of invariance to scaling, rotation and translation. The effectiveness of EMRFs representation is measured by experimenting with human actions datasets.

ACCEPTED MANUSCRIPT

## 3. Performance Evaluation and Result Analysis

In this section, the performance of our proposed approach is evaluated on four publically available human action datasets such as Weizmann Action [40], KTH [41], Ballet Movement [42] and UCF YouTube [23]. These human action datasets are challenging regarding viewpoint variation, occlusion, cluttered background and varying illuminations conditions. In the following sub-section, a brief description of all four datasets and different experimental conditions are presented. Further, the recognition accuracy of EMRFs is computed and compared with the techniques of others.

### 3.1 Human Activity Datasets

#### 3.1.1 Weizmann Action Dataset

Blank et al. [40] proposed this simple human action dataset. It consists of 10 action class performed by 9 people such as Run, Side, Skips, Jump, jumping -jack, Bend, Jack, Walk, Wave-1, and Wave-2, as shown in Fig. 8. There are a total of 90 videos clips recorded with 15 fps and have spatial resolution 144×180 pixels. The key challenge in this data set is in the interclass similarity among few actions like running, walking and jumping, but by exploiting the fine spatial shape distribution and textural features the action pose in still images is distinguished.

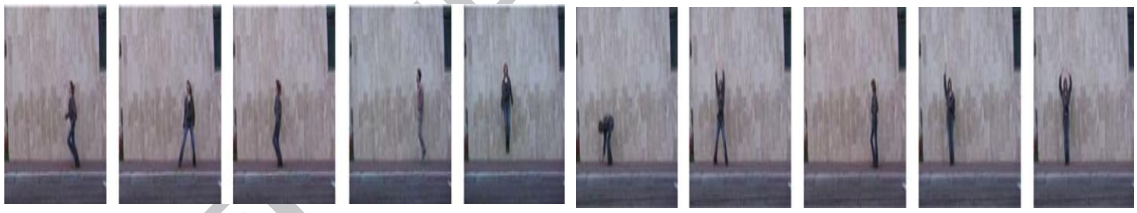


**Fig. 8.** Example frames from Weizmann Action Dataset.

#### 3.1.2 KTH Action Dataset

KTH Action is the most widely used data set for human action analysis. It was proposed by Schuldt et al. [41]. This dataset is more challenging as compared with Weizmann because changing illumination conditions and outdoor environments. It consists of 25 subjects performing six activities such as jumping, hand-clapping, jogging, walking, hand-waving, and running. There is a total of 600 videos recorded with 25 fps with spatial resolution 160×120 pixels, in 4 different scenarios. The example frames from KTH action dataset is depicted in Figure 9.

ACCEPTED MANUSCRIPT



**Fig. 9.** Sample frames of KTH Action Dataset.

In these (Weizmann and KTH) datasets, actions are recorded as a video sequence, hence for the extraction and selection of key still images; an automatic fuzzy-based model is adopted. This model gives the most discriminative frame of the video sequence, which has the highest visual appearance information about the human.

### 3.1.3 The Ballet Dataset

Fathi and Mori [42] introduced and collected this dataset from ballet DVD videos. This dataset is challenging regarding large intra-class dissimilarity and inter-class similarity, clothing variations, the speed of activity, and spatiotemporal variations. It consists of 8 movement activities performed by 1 woman and 2 men actors such as: turning, jumping, left to the right-hand opening, standing still, right to the left-hand opening, leg swinging, and hopping. There are total four annotated video sequences. Each actor is performing in each video for a fixed interval of time. The examples frames from ballet movement dataset are shown in figure 10.



**Fig. 10.** Samples frames from Ballet Movement Dataset.

### 3.1.4 UCF YouTube Action

Liu et al. [23] introduced this dataset to recognised complex action in a realistic environment. It consists of 11 action classes as tennis swinging, volleyball spiking, soccer juggling, walking with a dog, basketball shooting, cycling, horseback riding, swinging, golf swinging, trampoline jumping, and diving. There are twenty-five video groups each having four video clips sharing common features. These dataset videos are challenging due to view-points, occlusions, varying

ACCEPTED MANUSCRIPT

illumination conditions, and clutter backgrounds. The samples image frames from UCF YouTube dataset is shown in Fig. 11.



**Fig.11.** Sample frames from UCF YouTube Action Dataset.

### 3.2 Experimental Results and Comparison

In our approach, the EMRFs are computed on still images; hence in some cases where the action is recorded in video dataset, first the still images are obtained as explained in section 2 and then the representation is done by extracting and clubbing the spatial edge distribution feature map and Gabor feature map. The spatial edge distribution feature map is quantized into 8 orientation bins in the range of $[0,360^0]$ and Gabor feature map is computed with 5 scales and 8 orientations. Finally, all still images are represented as EMRFs by fusing the Gabor feature map and spatial edge feature map. The action classification is carried through a Support Vector Machine (SVM) [43] and K-nearest neighbour (k-NN) [44] classifiers. The accuracy of our EMRFs is measured as average recognition accuracy (ARA) in leave-one-out-cross-validation (LOOCV) evaluation protocol. The ARA is calculated using Eq. 10.

$$\text{Accuracy} = \frac{\text{True Postive +True Negative}}{\text{True Postive +TrueNegative+False Positive+FalsePositve}} \times 100\% \quad (10)$$

The classification accuracy is computed by dividing the correct number of predicted images by the number of tested images. The classification result of K-NN and SVM on various datasets are as shown in Fig. 12 and Fig.13 respectively.

The confusion matrix of the proposed method with a k-NN classifier for all dataset is shown in Fig. 12. The confusion matrix of Weizmann activity dataset is depicted in Fig.12(a), and it can be observed that the accuracy achieved on this dataset is very high. The activities in video dataset are classified with fewer ambiguities, and also the same can be inferred from the dataset too. Similarly, the confusion matrixes of KTH, Ballet, and UCF YouTube datasets are shown in

ACCEPTED MANUSCRIPT

Fig.12(b), (c) and (d) respectively. The less accuracy is achieved on Ballet dataset as compared with other datasets because this dataset has challenges such as self-occlusion and complex activities and also have high intra-class similarity.

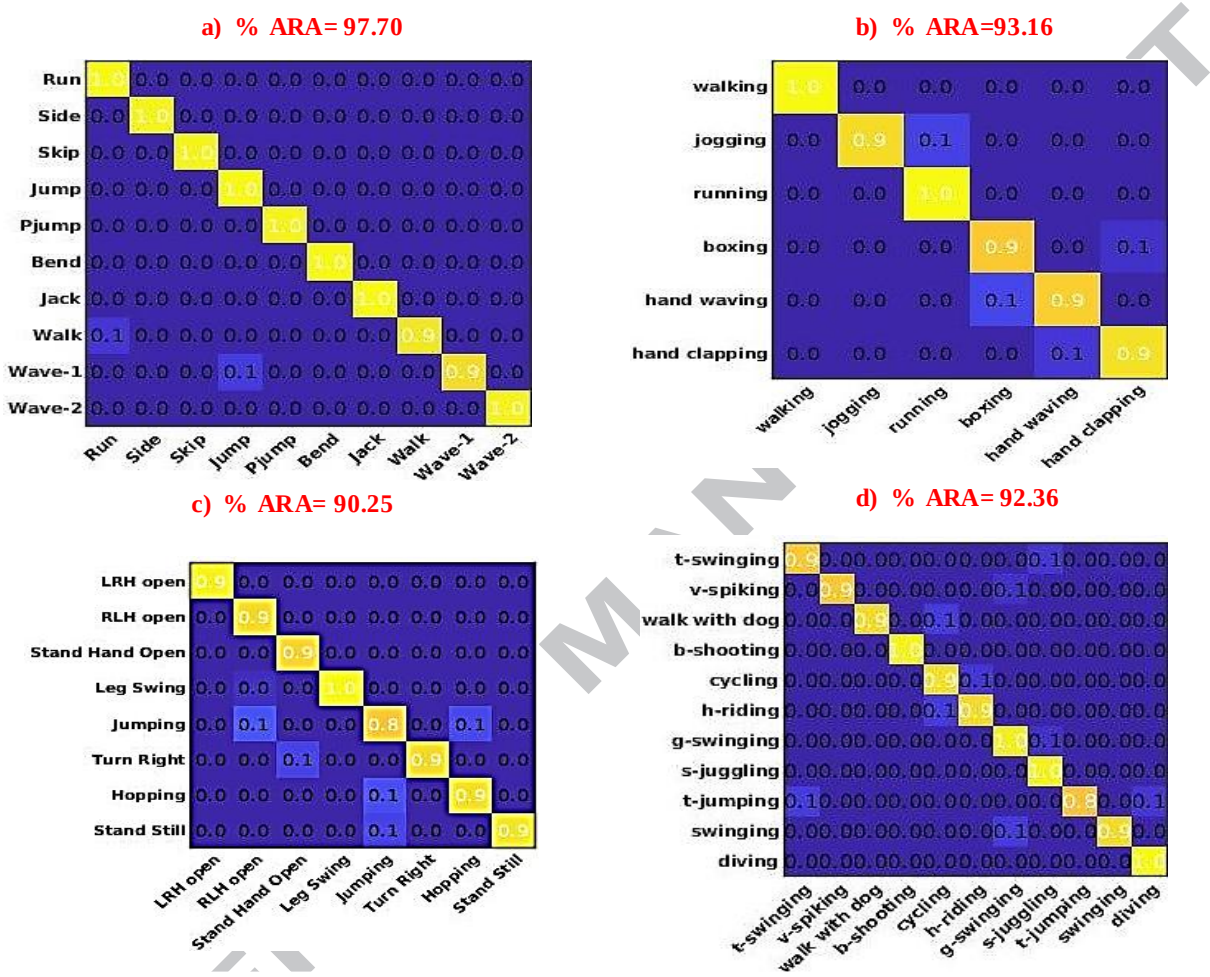



Fig. 12 Classification result of k-NN classifier on (a) Weizmann Action (b) KTH (c) Ballet Movement (d) UCF YouTube, of human action datasets.

The classification result is presented in the form of a confusion matrix with SVM classifier for each dataset is as shown in Fig. 13. There are various challenges present in the video because of the high interclass similarity in still key poses of running, walking, and jumping but our model outperforms in such kind of actions and each action is discriminated with highest recognition accuracy. Fig. 13 (a) shows the confusion matrix of the Weizmann dataset and cross-validation results of different actions. There is slightly less confusion in the case of 'jack hand' and 'wave-2' due to similar key poses. In Fig. 13(b) the confusion matrix for KTH dataset shows much satisfactorily results with average recognition accuracy 95.85 % and much lesser

ACCEPTED MANUSCRIPT

misclassification is found only three classes: ‘running’, ’walking’, and ‘jogging’ due to similar key still key poses. Further, our model easily classified with 100 % accuracy on the rest three action classes.

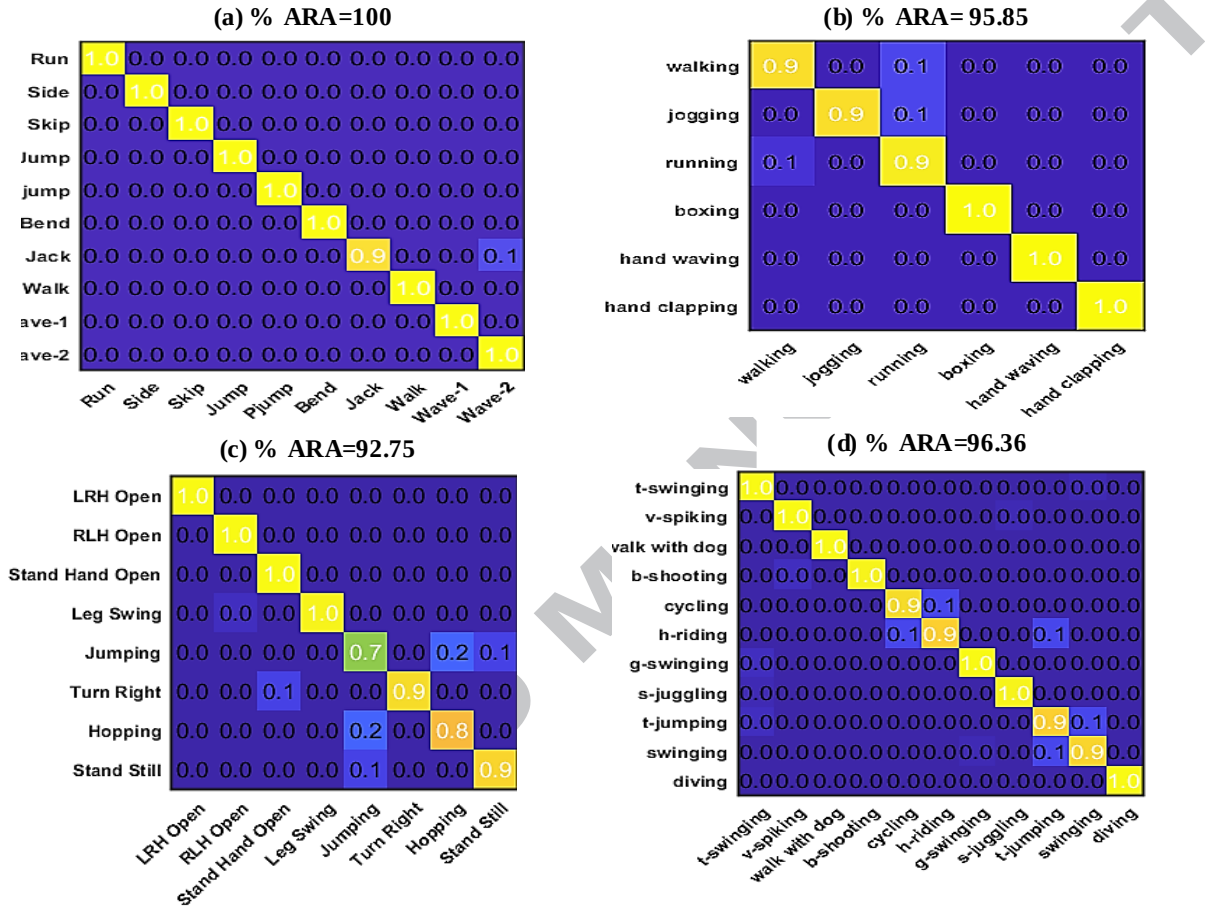



Fig. 13 Classification result of SVM classifier on (a) Weizmann Action (b) KTH (c) Ballet Movement (d) UCF YouTube, of human action datasets.

The ARA of 92.75 % computed through our method on Ballet dataset is shown in the confusion matrix in Fig. 13(c). The action recognition in this dataset is complex due to clothing, gender, and size variations. Our feature extractor EMRFs is insensitive to these variations and complexity of actions. It can be observed from the confusion matrix Fig.13(c) that there are little bit confused about still key pose of action pair such as ‘hopping’ and’ jumping’, ‘leg swing’ and ‘Right to left-hand opening’, ‘turning right’ and ‘stand hand opening’, and ‘jumping and standing still’ besides this our model gives comparable accuracy with the existing state-of-the-art methods [29], [45].

ACCEPTED MANUSCRIPT

In Fig. 13(d) shown the confusion matrix of UCF YouTube dataset. Some similar action is creating the confusion in motion feature such as 'cycling' and 'horseback riding' and 'jogging' and 'running'. Our method gives the highest accuracy on this challenging realistic video dataset.

The accuracy achieved through k-NN and SVM has been compared and given in Fig. 14. It can be seen that the performance of SVM is better than k-NN for all dataset because the k-NN classifier shows best results with higher dimensional training data and a larger value of k, but the large values of k lead to high computation.

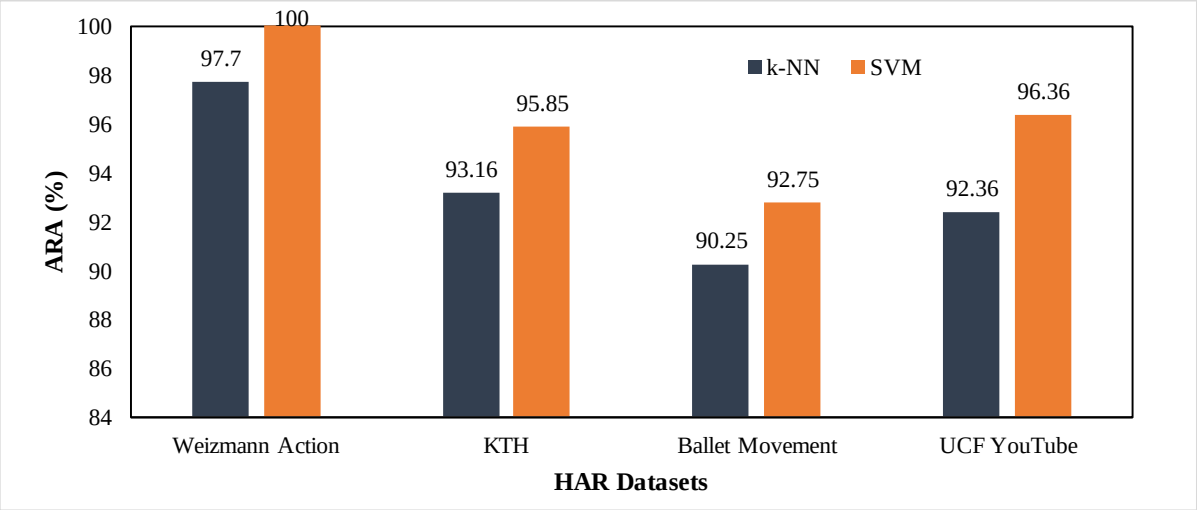



**Fig. 14** Comparison of k-NN and SVM accuracy on HAR datasets

However, the ARA is varying from one data set to another because of the recording conditions and environment setting of the dataset. The highest accuracy achieved for all datasets through SVM has been compared with the similar state-of-the-art.

### 3.3 Comparison of EMRFs with similar State-of-the-Art Approaches

The state-of-the-art comparison of proposed EMRFs is illustrated in Tables 1-4. The comparison is carried through in terms of earlier works, various form of input data to the feature extractors, techniques, evaluation protocol, types of classifiers used for classification of action recognition, and ARA. It can be observed from tables that most of the approaches for action recognition in video sequences relied on spatial-temporal features [20] [28] [29] [50] [51] [53] and very few on approaches using a single image or still image data information [21] [30] [49]. Because it is a challenging task to extracts robust features from spatial cue due to the absence of temporal information. The proposed EMRFs approach with SVM classifier achieved the highest *ARA* on the action datasets such as Weizmann, KTH, and UCF YouTube is compared with other

ACCEPTED MANUSCRIPT

state-of-the-arts. The *ARA* achieved on Weizmann human action dataset is 100%. The main reason for achieving such a high recognition rate is due to non-varying environmental conditions and the effectiveness of our proposed EMRFs. When we compare the obtained result with the techniques of others, then it shows the dominance. In the literature, one may find the similar accuracy on this dataset by the methods of others, but their way of representation is mostly based on the video based or set of the sequences of the frames.

**Table 1** Result comparison with state-of-the-art on Weizmann Action Dataset

| Methods | Input | Feature | Classifier | Test scheme | ARA (%) |
|---|---|---|---|---|---|
| Niebles and Fei [46] | Spatiotemporal | BoF Model | SVM | LOOCV | 55.00 |
| Thrurau [47] | Temporal | HOG | SVM | LOOCV | 57.45 |
| Eweiwi et al. [48] | Still Image | NMF | Bayesian | - | 55.20 |
| Thurau and Hlavac [20] | Spatiotemporal | HOG | 1-NN classifier | LOOCV | 74.40 |
| Chaaraoui et al. [49] | Still Image | Silhouettes | SVM | LOSO | 92.80 |
| Baysal and Duygulu [15] | Temporal | GPB, DTW | K-NN | LOOC | 95.10 |
| Guan et al. [27] | Still Image | TNMF | - | CV | 91.70 |
| Batchuluun et al. [14] | Temporal | Silhouettes | Fuzzy Logic | - | 99.20 |
| **Ours** | Still Image | EMRFs | SVM | LOOCV | **100** |

The result shown in Table 1 is for the Weizmann dataset while the result depicted in Table 2 is for the KTH dataset, which is recorded under varying ecological conditions. In the KTH dataset, the proposed approach based action classification leads to the *ARA* equal to 95.83%. The drop of recognition rate in this dataset as compared to Weizmann dataset is due to the more varying illumination and challenging environmental conditions. The comparative analysis of the similar state-of-the-art of EMRFs on KTH dataset is depicted in Table 2. It is also observed that EMRFs gives a significant amount of increase in recognition accuracy.

**Table 2** Result comparison with state-of-the-art on KTH Dataset

| Methods | Input | Feature | Classifier | Test scheme | ARA (%) |
|---|---|---|---|---|---|
| Raja et al. [21] | Still Image | HOG | LSVM | - | 86.58 |
| Baysal and Duygulu [15] | Temporal | GPB | k-NN | LOOC | 81.30 |
| Saghafi and Rajan [50] | Spatiotemporal | PDE | - | LOOC | 92.60 |
| Han et al. [29] | Spatiotemporal | Deeper spatial ConvNets | | Splits | 61.11 |
| Zheng et al. [45] | Temporal | Fisher vector | LSVM | Splits | 94.58 |
| **Ours** | Still Image | EMRF | SVM | LOOCV | **95.83** |

**Table 3** Result comparison with state-of-the-art on Ballet Dataset

| Methods | Input | Feature | Classifiers | Test Scheme | ARA (%) |
|---|---|---|---|---|---|
| Fathi & Mori [42] | Temporal | Optical flow | Adaboost | LOOCV | 51.00 |
| Wang and Mori [33] | Temporal | BoWs | S-CTM | LOO | 91.30 |

ACCEPTED MANUSCRIPT

| | | | | | |
|---|---|---|---|---|---|
| Guha and Ward [51] | Spatiotemporal | Cuboids+ LMP | RSR | LOO | 91.10 |
| Iosifidious et al. [52] | Temporal | BoWs | SVM | LOO | 91.10 |
| Zhao et al. [28] | Spatiotemporal | RKHS | K-means | - | 79.78 |
| Wang et al. [30] | Still Image | LGLRR | K-means | - | 60.87 |
| Vishwakarma et al. [53] | Spatiotemporal | LDA | SVM-NN | LOOCV | **94.00** |
| **Ours** | Still Image | **EMRFs** | **SVM** | **LOOCV** | 92.75 |

The comparison of the various state-of-the-art methods on Ballet dataset is listed in Table 3. It is considered a challenging dataset regarding the complexity of human activities performed such as speed of action and low illumination, but due to enclosed setup conditions, and EMRFs features extraction results for the dataset is practically satisfactory. The experimental setup used for this dataset is similar to work [53]. The average recognition accuracy achieved is 92.75% which is better than [30] [28] [52] but less than [53]. The main reason for the slightly less accuracy is using of spatial shape feature only in still image of complex motion action dataset.

**Table 4** Result comparison with state-of-the-art on UCF YouTube dataset

| **Methods** | **Input** | **Feature** | **Classifiers** | **Test Scheme** | **ARA (%)** |
|---|---|---|---|---|---|
| Liu et al. [23] | Spatiotemporal | Hybrid features | Adaboost | LOOCV | 71.20 |
| Cinbis and Sclaroff [54] | Spatiotemporal | MIL | SVM | LOOCV | 75.20 |
| Le et al. [55] | Spatiotemporal | Independent sub space analysis | K-Means | - | 75.80 |
| Yi and Lin [56] | Spatiotemporal | Spatio-temporal graph | - | LOO | 84.63 |
| Wang et al. [57] | Spatiotemporal | Dense trajectories | - | LOOCV | 85.40 |
| Shao et al. [58] | Spatiotemporal | Kernelized multi-view projection | Naïve Bayes | 5-fold CV | 87.60 |
| Jung and Hong [59] | Temporal | Bag of Sequence lets | SVM | LOOCV | 89.90 |
| Nazir et al. [60] | Spatio-Temporal | Bag of Expression(BoE) | KNN | - | 96.68 |
| **Ours** | Still Image | **EMRFs** | **SVM** | **LOOCV** | **96.36** |

Table 4 consist of various state-of-the-art approaches on UCF YouTube datasets. It is more challenging activity dataset because most of the video was recorded in free environment and complex background conditions. Our method gives 96.36% recognition accuracy with spatial cues only which outperforms over existing spatiotemporal approaches on this complex dataset.

## 4. Conclusion

A multiresolution based feature descriptor model is developed for the recognition of human actions in still images and videos too. It is more challenging to recognized action in still image than video analysis because it deals with the absence of temporal information, clothing variation,

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

environmental change, scaling, segmentation problems, and alignment of the images. In this work,

the still key poses images are selected from action videos using fuzzy inference model based on maximum histogram distances between adjacent frames. Further, the Gabor wavelet transform is used to make these key poses frames invariant for different orientations and scales. It has seen that there is a number of the parameters such as numbers of bins, a number of the scale and orientations used for development of the EMRFs and these parameters are chosen empirically, and these parameters affect the performance.

The performance of the EMRFs is measured on publically available datasets, and these datasets are challenging in respect of lightening variations, and zoom in and out. Due to the variation of lightening the texture of the scene was affected. There are few cases where the texture variations were due to the change of clothes of the actor; however, our EMRFs performs well. Our feature extractor shows best results with two best discriminative classification techniques such as supervised SVM and non-parametric k-nearest neighbour (k-NN). However, the classification accuracy of k-NN classifier somewhat less as compared with SVM classifier because k-NN required higher dimensional data for training the model. The proposed method outperforms in terms of recognition accuracies on Weizmann, KTH, UCF YouTube datasets as compared with other state-of-the-art. There are comparable recognition accuracy is achieved on Ballet dataset due to variations of clothes, speed and changing illumination conditions.

In future, a more realistic study may be conducted on the unconstrained dataset, and EMRFs can be used for many other applications such as visual sentiment representation and analysis, movie analysis, content-based recommender systems etc.

## ACKNOWLEDGEMENTS

The authors would like to acknowledge the computation of the work was supported by Biometric Research Laboratory, Department of Information Technology, Delhi Technological University, New Delhi, India.

## Reference

[1] Vishwakarma D, Kapoor R, Maheshwari R, Kapoor V. Raman S, Recognition of abnormal human activity using the changes in orientation of silhouette in keyframes. International Conference on Computing for Sustainable Global Development, New Delhi. 2015.

ACCEPTED MANUSCRIPT

[2] Tripathi G, Singh K, Vishwakarma D. Convolutional neural networks for crowd behaviour analysis: a survey. The Visual Computer. 2018; 1-24.

[3] Liu Y, Yang F, Zhong C, et al. Visual tracking via salient feature extraction and sparse collaborative model. AEU - International Journal of Electronics and Communications. 2018; 87: 134-143.

[4] Aggarwal J, Ryoo M. Human activity analysis: A review. ACM Computing Survey 2011; 43(3): 1-43.

[5] Poppe R. A survey on vision-based human action recognition. Image and Vision Computing. 2010; 28(6): 976-990.

[6] Guo G, Lai A. A survey on still image based human action recognition. Pattern Recognition 2014; 47(10): 3343-3361.

[7] Nguyen D, Li W, Ogunbona P. Human detection from images and videos: A survey. Pattern Recognition 2016; 51: 148–175.

[8] Singh T, Vishwakarma D. Video benchmarks of human action datasets: a review. Artificial Intelligence Review 2018; 1-48.

[9] Vishwakarma D, Kapoor R, Dhiman A. Unified framework for human activity recognition: An approach using spatial edge distribution and ℜ-transform. AEU - International Journal of Electronics and Communications 2016; 70(3): 341-353.

[10] Bourdev L, Malik J. Poselets: body part detectors trained using 3d human pose annotations. In ICCV 2009.

[11] Maji S, Malik J. Action recognition from a distributed representation of pose and appearance. In CVPR 2011.

[12] Yang W, Wang Y, Mori G. Recognizing human actions from still images with latent poses. In CVPR 2010.

[13] Delaitre V, Laptev I, Sivic J. Recognizing human actions in still images: a study of bag-of-features and part-based representations. In Proceedings of the British Machine Vision Conference 2010.

[14] Batchuluun G, Kim J, Hong H, et al. Fuzzy system based human behavior recognition by combining behavior prediction and recognition. Expert Systems with Applications 2017; 81: 108–133.

[15] Baysal S, Duygulu P. A line based pose representation for human action recognition. Signal processing: Image Communication 2013; 28: 458-471.

[16] Zhang Y, Cheng L, Wu J, et al. Action Recognition in Still Images with Minimum Annotation Efforts. IEEE Transactions on Image Processing 2016; 25(11): 5479-5490.

[17] Wang Y, Jiang H, Drew M, et al. Unsupervised discovery of action classes. In IEEE Conference on Computer Vision and Pattern Recognition 2006.

[18] Li P, Ma J. What happening in a still picture? In IEEE Asian conference on pattern Recognition 2011.

ACCEPTED MANUSCRIPT

[19] Li J, Fei-Fei L. What, where and who? classifying events by scene and object recognition. In IEEE Conference on Computer Vision, Rio de Janeiro, Brazil 2007.

[20] Thurau C, Hlavac V. Pose primitive based human action recognition in videos or still images. In IEEE Conference on Computer Vision and Pattern Recognition 2008.

[21] Raja K, Laptev I, Pérez P, et al. Joint pose estimation and action recognition in image graphs. In IEEE International Conference on Image Processing, Brussels. 2011.

[22] Bosch A, Zisserman A, Munoz X. Representing shape with a spatial pyramid kernel. ACM International conference on Image and Video Retrieval 2007; 58: 401-408. Amsterdam.

[23] Liu J, Luo J, Shah M. Recognizing Realistic Actions from Videos "in the Wild". In IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) 2009.

[24] Zheng Y, Yao H, Sun X, et al. Distinctive action sketch for human action recognition. Signal Processing 2018; 144: 323–332.

[25] Conde C, Moctezuma D, Diego I, et al. HoGG: Gabor and HoG-based human detection for surveillance in non-controlled environments. Neurocomputing 2013; 19-30.

[26] Ikizler N, Cinbis R, Pehlivan S, et al. Recognizing actions from still images. In IEEE International Conference Pattern Recognition, Tampa, FL, 2008.

[27] Guan N, Tao D, Lan L, et al. Activity Recognition from Still Images with Transductive Non-negative Matrix Factorization. In ECCV 2014.

[28] Zhao K, Alavi A, Wiliem A, et al. Efficient clustering on Riemannian manifolds: A kernelised random projection approach. Pattern Recognition 2015; 51: 333-345.

[29] Han Y, Zhang P, Zhuo T, et al. Going deeper with two-stream ConvNets for action recognition in video surveillance. Pattern Recognition Letters 2017.

[30] Wang B, Hu Y, Gao J, et al. Localized LRR on Grassmann Manifold: An Extrinsic View. IEEE Transactions on Circuits and Systems for Video Technology 2017.

[31] Ming X, Xia J, Zheng L. Human action recognition based on chaotic invariants. Journal of South Central University. 2014; 20: 3171-3179.

[32] Safaei M, Foroosh H. Single Image Action Recognition by Predicting Space-Time Saliency. arXiv:1705.04641v1 [cs.CV]. 2017; 1-9.

[33] Wang Y, Mori G. Human Action Recognition Using Semi-Latent Topic Model. IEEE Transactions on Pattern Analysis and Machine Intelligence 2009; 31(10): 1762-1764.

[34] Mahapatra A, Mishra T, Sa P, et al. Human recognition system for outdoor videos using Hidden Markov model. AEU - International Journal of Electronics and Communications 2014; 66(3): 227-236.

ACCEPTED MANUSCRIPT

[35] Zeng P, Chen Z. Perceptual quality measure using JND model of the human visual system. In IEEE International Conference on Electric Information and Control Engineering 2011.

[36] Zadeh L. Fuzzy sets. Information Control 1965; 8(3): 338-353.

[37] Vishwakarma D, Singh K. Human Activity Recognition Based on Spatial Distribution of Gradients at Sublevels of Average Energy Silhouette Images. IEEE Transactions on Cognitive and Developmental Systems 2017; 9(4): 316-327.

[38] Arivazhagan S, Ganesan L, Priyal S. Texture classification using Gabor wavelets based rotation invariant features. Pattern Recognition Letters 2006; 1976-1982.

[39] Pasupathy A, El-Shamayleh Y, Popovkina D. Visual Shape and Object Perception. Neuroscience 2018.

[40] Blank M, Gorelick L, Shechtman E, et al. Actions as space-time shapes. In tenth IEEE International Conference on Computer Vision (ICCV), Beijing 2005.

[41] Schuldt C, Laptev I, Caputo B. Recognizing human actions: a local SVM approach. Proc. of the International conference on Pattern Recognition 2004.

[42] Fathi A, Mori G. Action recognition by learning mid-level motion features. In IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA 2008.

[43] Vapnik V. An overview of statistical learning theory. IEEE Transaction Neural Network 1999; 10(5): 989-99.

[44] Cover T, Hart P. Nearest neighbour pattern classification. IEEE Transactions on Information Theory 1967; 13(1): 21-27.

[45] Zheng Y, Yao H, Sun X, et al. Distinctive action sketch for human action recognition. Signal Processing 2018; 144: 323–332.

[46] Niebles J, Fei-Fei L. A Hierarchical Model of Shape and Appearance for Human Action Classification. In IEEE Conference on Computer Vision and Pattern Recognition 2007.

[47] Thurau C. Behavior Histograms for Action Recognition and Human Detection. Lecture Notes in Computer Science, Springer 2007; 4814: 299-312.

[48] Eweiwia A, Cheema M, Bauckhage C. Action recognition in still images by learning spatial interest regions from videos. Pattern Recognition Letters 2015; 51: 8-15.

[49] Chaaraoui A, Pérez P, Revuelta F. Silhouette-based human action recognition using sequences of key poses. Pattern Recognition Letters 2013; 34(15): 1799-1807.

[50] Saghafi B, Rajan D. Human action recognition using Pose-based discriminant embedding. Signal Processing: Image Communication 2012; 27: 96-111.

[51] Guha T, Ward R. Learning sparse representations for human action recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 2012; 34(8): 1576-1588.

ACCEPTED MANUSCRIPT

[52] Iosifidis A, Tefas A, Pitas I. Discriminant bag of words based representation for human action recognition. Pattern Recognition Letters 2014; 49: 185-192.

[53] Vishwakarma D, Kapoor R. Hybrid classifier based human action recognition using silhouettes and cells. Expert Systems with Applications 2015; 42(20): 6957-6965.

[54] Ikizler C, Sclaroff S. Object, scene and actions: combining multiple features for human action recognition. In ECCV 2010.

[55] Le Q, Zou W, Yeung S, et al. Learning hierarchical invariant spatio-temporal features for action recognition with independent subspace analysis. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2011.

[56] Yi Y, Lin M. Human action recognition with graph-based multiple-instance learning. Pattern Recognition 2016; 53:148-162.

[57] Wang H, Klaeser A, Schmid C, et al. Dense trajectories and motion boundary descriptors for action recognition. IJCV 2013.

[58] Shao L, Liu L, Yu M. Kernelized multi-view projection for robust action recognition. Int. J. Comput. Vis. 2015; 1-15.

[59] Jung H, Hong K. Modeling temporal structure of complex actions using Bag-of-Sequencelets. Pattern Recognition Letters 2017; 85: 21-28.

[60] Nazir S, Yousaf M, Nebel J, et al. A Bag of Expression framework for improved human action recognition. Pattern Recognition Letters. 2018; 103: 39–45.



**Dinesh Kumar Vishwakarma (M'16, SM'19)** received the B.Tech. degree from Dr. Ram Manohar Lohia Avadh University, Faizabad, India, in 2002, the M.Tech. degree from the Motilal Nehru National Institute of Technology, Allahabad, India, in 2005, and the Ph.D. degree from Delhi Technological University, New Delhi, India, in 2016. He is currently an Associate Professor with the Department of Information Technology, Delhi Technological University, New Delhi. His current research interests include Computer Vision, Machine Learning, Deep Learning, Sentiment Analysis, Fake News and Rumor Analysis, Crowd Behaviour Analysis, Person Re-Identification, Human Action and Activity Recognition. He is a reviewer of various Journals/Transactions of IEEE, Elsevier, and Springer. He has been awarded with "Premium Research Award" by Delhi Technological University, Delhi, India in 2018.



Tej Singh received the B.Tech. degree from Madan Mohan Malaviya University of Technology, Gorakhpur, India, in 2010, the M.E degree

ACCEPTED MANUSCRIPT

from the Thapar University, Patiala, India, in 2014. Currently, he is a student of Doctoral of Philosophy (Ph.D.) in the Department of Electronics and Communication Engineering, Delhi Technological University, New Delhi, India. His research interests include human action and activity recognition, image processing, pattern analysis and machine learning.

ACCEPTED MANUSCRIPT

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



# Analysis of Signals via Non-Maximally Decimated Non-Uniform Filter Banks

Sandeep Patel, Ravindra Dhuli, and Brejesh Lall, *Member, IEEE*

***Abstract*— This paper addresses the important problem of reconstructing a signal from multiple multirate observations. The observations are modeled as the output of an analysis bank, and time-domain analysis is carried out to design an optimal FIR synthesis bank. We pose this as a minimizing the mean-square problem and prove that at least one optimal solution is always possible. A parametric form for all optimal solutions is obtained for a non-maximally decimated filter bank. The necessary and sufficient conditions for an optimal solution that results in perfect reconstruction (PR), are derived as time-domain pseudocirculant conditions. This represents a novel theoretical contribution in multirate filter bank theory. We explore PR in a more general setting. This results in the ability to design a synthesis bank with a particular delay in the reconstruction. Using this delay, one can achieve PR in cases where it might not have been possible otherwise. Further, we extend the design and analysis to non-uniform filter banks and carry out simulations to verify the derived results.**

***Index Terms*— Matrix Wiener filter, multirate filter bank, perfect reconstruction, time-domain pseudocirculant conditions.**

## I. Introduction

MULTIRATE filter banks have been used extensively for applications in the fields of image and video processing [1]–[3], communications [4], [5], subband adaptive filtering [6], radar systems [7], etc. The filter bank theory evolved from the need to decompose a signal spectrum into multiple frequency subbands in order to efficiently perform a number of operations like compression, coding, noise cancellation, etc. Initially, two channel quadrature mirror filter bank (QMF) was proposed in the literature [8], and later, it was extended to arbitrary $M$-channel case [9], [10]. To better exploit the signal characteristics in a number of applications like audio coding, spectral analysis, subband adaptive filtering, etc., non-uniform filter banks (NUFBs) having different channel bandwidth were introduced [11]–[13]. In [14], filter banks with different multirate factors were used, to realize the filter transfer functions for reducing the multiplicative complexity. NUFBs with rational sampling factors were considered in [12], [15]. However, they were found to be equivalent to integer decimated NUFBs [16]. A number of design methods like direct method [15], recombination method [11], modulation based method [17], etc. exist in the literature. Application specific design methods have also been proposed in the literature. For example, in [18], an NUFB was designed for acoustic echo-cancellation by maximizing the signal-to-alias ratio. All these methods design both the analysis and synthesis banks of an NUFB. The analysis and synthesis filters in a filter bank can be IIR or FIR. FIR filters have advantage over IIR in respect of stability and linear-phase property. In the literature, various design methods using transform like Lapped Orthogonal Transform [19], Generalized Lapped Orthogonal Transform [20], Extended Lapped Transform [21], etc. have been proposed to design a perfect reconstruction (PR) FIR maximally decimated uniform filter bank (UFB). A comprehensive analysis of FIR perfect reconstruction filter banks can be found in [22]. It has been shown that oversampled filter banks (OFBs) offer additional design freedom as compared to critically sampled filter banks [23], [24]. The inbuilt redundancy present in them has been exploited for a variety of applications [25]–[28]. A frame theoretic analysis of OFBs can be found in [29]. Given an oversampled analysis bank, a parameterization for all synthesis banks resulting in PR was developed in [30]. Undersampled filter banks which provide underdetermined representation of signals have found application in compression of signals [31]. Despite the filter banks have been around for more than two decades, they continue to be applied to new and emerging areas. Graph signal processing is one such interesting area where filter banks techniques find applications in solving diverse problems [32].

The necessity in any scenario to estimate a signal or its properties from its subband components can be addressed well using the filter bank based processing. Expectedly, many research problems appear posing such requirement. Synthesis from subband components has been utilized in applications like subband coding [33], [34], analog/digital conversion [35], transmultiplexer [33], ultrawideband receiver [36], TV white space transceiver [37], channel coding [38], etc. To get better results, processing on subband components, rather than on full-band signal, have been carried out in problems like speech enhancement [39], ECG processing [40], subband adaptive filtering [41], seismic signals processing [42], etc. The processed subband signals are then combined using a synthesis bank to get the final output. In multirate sensors networks, sensor outputs are modeled as an analysis bank output, and these outputs are fused to get a unified measurement [43].

Manuscript received January 9, 2019; revised April 15, 2019; accepted April 26, 2019. This paper was recommended by Associate Editor D. Comminiello. *(Corresponding author: Sandeep Patel.)*

S. Patel and B. Lall are with the Bharti School of Telecom Technology and Management, IIT Delhi, New Delhi 110 016, India (e-mail: sanpatel90@gmail.com; brejesh@ee.iitd.ac.in).

R. Dhuli is with the Department of Electronics and Communications Engineering, Vellore Institute of Technology, Amaravati 522 237, India (e-mail: ravindra.d@vitap.ac.in).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSI.2019.2914302

1549-8328 © 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



In [44], construction of a high-resolution signal from multiple low-resolution signals has been carried out. The problem of linear prediction from multirate observations has been examined in [45]. In many applications, rather than a signal, a property of interest is to be estimated from multirate observations. Time delay estimation from low rate samples is an example of that, and it finds application in speech localization, processing using microphone arrays, radar, underwater acoustics, wireless communications and other areas [46], [47]. Direction of arrival estimation is another example which finds application in wireless communications, medical imaging, seismology, etc. [48]. In [31], estimation of the spectrum of a signal from its low resolution signals was carried out. This is required when a signal can be observed only indirectly using low-rate sensors.

To address the problem of estimating a signal from its subbands, various solutions using Wiener filtering, Kalman filtering, $\mathcal{H}^\infty$ filtering, compressive sensing, adaptive filtering, etc. have been proposed in the literature. Wiener filtering based methods require joint stationarity of the subband input and desired signals along with the knowledge of second order statistics of them. Kuchler and Therrien [49] used Wiener filtering to estimate a signal from multiple multirate observations. Scrofani and Therrien [44] extended the analysis to two-dimensional signals. Vaidyanathan and Chen [50] designed a statistically optimal synthesis bank for a subband coder by using the biorthogonal solution followed by a matrix Wiener filter in the synthesis stage. In case of Kalman filtering, the statistics of the input and noise signals are required to be known even if that vary with time. Multirate Kalman filter based methods can be found in [51]–[53]. $\mathcal{H}^\infty$ optimization based methods, on the other hand, do not require any statistical assumption and perform worst-case design [54]. As $\mathcal{H}^\infty$ solution is generally not unique, mixed $\mathcal{H}^2/\mathcal{H}^\infty$ optimization methods [55], [56] have also been proposed in the literature. The complexity of $\mathcal{H}^2$ and $\mathcal{H}^2/\mathcal{H}^\infty$ solutions are very high since they are solved as linear matrix inequalities (LMIs) [56]. A least square based solution for the problem was presented in [57]. However, the solution presented was for two channels only. In [58], Chai *et al.* formulated the problem using structural similarity criterion and obtained closed-form solution under certain assumptions like the filter length is equal to the decimation factor, the mean of the source is zero, etc. Compresive sensing based method was used in [59] to estimate a signal from multirate observations. However, the signal has to be sparse. In [60] and [61], adaptive filters were used to estimate a signal from two observation sequences. An adaptive least square lattice filter was used in [62]. However, the filter obtained was periodically time-varying. Statistical analysis of the estimation from the subbands has been performed by a number of authors. For example, in [63]–[65], power spectral density of a signal was estimated from multiple observations. In [66], a measure for information contained in a multirate observations was defined.

### A. Contribution

In this paper, the problem of synthesizing a signal from its multirate observations is addressed by performing time-domain analysis. This approach has led to some additional novel contributions to the field which are now briefly presented. Under time-domain analysis, the matrix representation for various operations like filtering, downsampling, upsampling, etc. are developed and used to parameterize synthesis bank solutions. This parameterization is consistent and at least one optimum synthesis filter bank is guaranteed. We design the synthesis bank for a more general set up in comparison to existing methods [57]–[61]. The synthesis filter length can be arbitrary but given. Further, we have used delay in the reconstruction as a parameter to have flexibility in the synthesis filter design, and using that, we obtain better reconstruction error than possible with a zero delay design. Another advantage of our analysis is that we obtain a linear time-invariant (LTI) optimal matrix filter unlike the linear periodically time-varying (LPTV) optimal filters obtained in earlier works [49], [62]. Our method always results in a solution in comparison to LMI based methods [55], [56] where a solution for an LMI may not be feasible [67]. The necessary and sufficient condition on a synthesis bank solution to result in PR is derived. Under this condition, the polyphase matrices for the analysis and synthesis banks satisfy the general pseudocirculant property [68], [69], unlike the limited identity matrix condition in the earlier works [23], [29]. The necessary and sufficient conditions on the analysis bank and the delay for a PR synthesis bank solution to exist are derived. This analysis forms an alternative to the frame theory based approach in previous works [29], [30]. Time-domain pseudocirculant result derived in this paper form an important theoretical contribution to the multirate theory. The derivation of a range of delay values for which PR is possible is another contribution. We have also extended the design methodology for any arbitrary NUFB. Under experimental results, a number of experiments are presented to highlight contributions and to show application of our work.

### B. Organization

This paper is organized as follows: Section II discusses the various matrix representations for time-domain signals and operations. In Section III, we introduce the framework used to design the synthesis filter for a given analysis bank. In Section IV, we derive the pseudocirculant condition in time-domain and present the necessary and sufficient conditions on an analysis bank and on the delay to result in PR. We provide the experimental results in Section V and finally, concluding remarks are given in Section VI.

### C. Notation

The $(i, j)$-th entry of a matrix $\mathbf{B}$ is given by $B_{i,j}$ or $B(i, j)$. The transpose and conjugate-transpose of a matrix or vector quantity are denoted by the superscripts $(.)^T$ and $(.)^H$, respectively. An identity matrix of size $P$ is represented by $\mathbf{I}$ when the size is implicit and by $\mathbf{I}_P$ otherwise. The pseudoinverse of a matrix $\mathbf{A}$ is denoted by $\mathbf{A}^\dagger$. The superscript $(.)^*$ stands for conjugation of a quantity. Convolution of two sequences is denoted by $*$. The ceiling function is represented by $\lceil . \rceil$.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



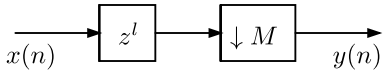



Fig. 1. A decimator preceded by time-advance.

## II. Matrix Representations

In this section, we develop the matrix representations for various multirate operations, which are required for an effective time-domain analysis of multirate filter banks.

### A. Time-Domain Sequence Representation

A vector signal is developed from the given signal $v(n)$ by stacking the past $P$ samples as:

$$\mathbf{v}(n) = \begin{bmatrix} v(n-P+1) & v(n-P+2) & \dots & v(n) \end{bmatrix}^T. \quad (1)$$

It is termed as an observation vector for the signal $v(n)$ [70]. Its flipped version or reversal can be obtained as

$$\bar{\mathbf{v}}(n) = \mathbf{J}_P \mathbf{v}(n), \quad (2)$$

where $\mathbf{J}_P$ is a counter-identity or reverse-identity matrix. For a matrix $\mathbf{B}$, reversal operation results in reversing of order along both rows and columns. Mathematically, the same can be expressed as

$$\bar{\mathbf{B}} = \mathbf{J}_M \mathbf{B} \mathbf{J}_N \quad (3)$$

if the size of the matrix is $M \times N$. For two matrices $\mathbf{A}$ and $\mathbf{B}$ whose product is defined, we have the following relation:

$$\overline{\mathbf{AB}} = \bar{\mathbf{A}}\bar{\mathbf{B}}. \quad (4)$$

### B. Decimation

Consider a system, as shown in Fig. 1, with a $M$-fold decimator preceded by a time-advance by $l$. The input and output of it are related as

$$y(n) = x(Mn + l). \quad (5)$$

Such a system can be used to obtain polyphase components of a signal when $l$ is in the range $[0, M-1]$. Two matrix representations are possible for a system. If an observation vector

$$\mathbf{y} = \begin{bmatrix} y(0) & y(1) & \dots & y(P-1) \end{bmatrix}^T \quad (6)$$

for the output is to be evaluated, then the same can be done as follows:

$$\begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(P-1) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} x(l) \\ x(l+1) \\ \vdots \\ x(M+l) \\ \vdots \\ x(M(P-1)+l) \end{bmatrix},$$

$$\mathbf{y} = \mathbf{D}^0_{P,M(P-1)+1} \mathbf{x}, \quad (7)$$

where $\mathbf{D}^0_{P,M(P-1)+1}$ is known as a decimation matrix. On the other hand, if an input sample vector

$$\mathbf{x} = \begin{bmatrix} x(0) & x(1) & \dots & x(P-1) \end{bmatrix}^T \quad (8)$$

is given, then the corresponding output samples can be obtained as

$$\mathbf{y} = \mathbf{D}^l_{\lceil \frac{P}{M} \rceil, P} \mathbf{x}, \quad (9)$$

where the elements of the decimation matrix $\mathbf{D}^l_{\lceil P/M \rceil, P}$ are given by

$$D^l_{\lceil \frac{P}{M} \rceil, P}(i,j) = \begin{cases} 1, & \text{if } j = Mi + l \\ 0, & \text{otherwise.} \end{cases} \quad (10)$$

Here $l$ is in the range $0 \le l \le M-1$. For $l \ge M$, the same matrix representations can be extended.

### C. Expansion

If an input observation vector

$$\mathbf{x} = \begin{bmatrix} x(0) & x(1) & \dots & x(P-1) \end{bmatrix}^T \quad (11)$$

is available, then the output of a $M$-fold expander can be obtained as

$$\mathbf{y} = \mathbf{U}_{M(P-1)+1,P} \mathbf{x}, \quad (12)$$

where $\mathbf{U}_{M(P-1)+1,P}$ is an expansion matrix. This matrix is related to a decimation matrix as

$$\mathbf{U}_{M(P-1)+1,P} = (\mathbf{D}^0_{P,M(P-1)+1})^T. \quad (13)$$

It can be easily verified that

$$\begin{aligned} &\mathbf{D}^0_{P,M(P-1)+1} \mathbf{U}_{M(P-1)+1,P} = \mathbf{I}_P, \\ &\mathbf{U}_{M(P-1)+1,P} \mathbf{D}^0_{P,M(P-1)+1} \neq \mathbf{I}_{M(P-1)+1}. \end{aligned} \quad (14)$$

### D. Convolution

Consider a causal linear time-invariant (LTI) filter with impulse response $\{h(0), h(1), \dots, h(Q-1)\}$. Its input and output are related as

$$y(n) = \sum_{k=0}^{Q-1} h(k) x(n-k). \quad (15)$$

We present the matrix representations for the above relation for two use cases. Under first case, consider an observation vector

$$\mathbf{y}(n) = \begin{bmatrix} y(n-P+1) & y(n-P+2) & \dots & y(n) \end{bmatrix}^T \quad (16)$$

for the output which needs to be evaluated. The same can be achieved using

$$\mathbf{y}(n) = \mathbf{H}\mathbf{x}(n), \quad (17)$$

where

$$\mathbf{x}(n) = \begin{bmatrix} x(n-P+2-Q) & x(n-P+3-Q) & \dots & x(n) \end{bmatrix}^T, \quad (18)$$

and $\mathbf{H}$ is a convolution matrix given by

$$\mathbf{H} = \begin{bmatrix} h(Q-1) & h(Q-2) & \dots & h(0) & \dots & 0 \\ 0 & h(Q-1) & \dots & h(1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & h(0) \end{bmatrix}. \quad (19)$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



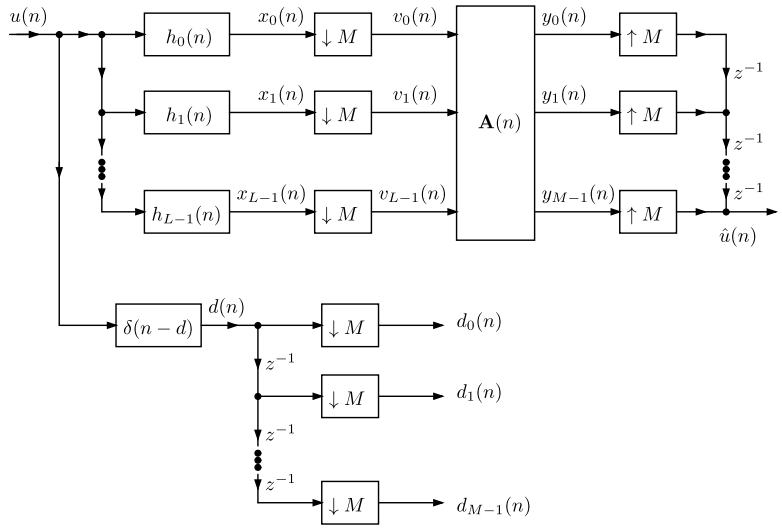



Fig. 2. Matrix Wiener filter based synthesis stage for a UFB

The second representation is obtained for a scenario where an input observation vector is given, and the corresponding output samples need to be evaluated. Consider an input observation vector

$$\mathbf{x} = \begin{bmatrix} x(0) & x(1) & \dots & x(P-1) \end{bmatrix}^T, \tag{20}$$

then the complete output due to it can be expressed as

$$\begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(Q-1) \\ \vdots \\ y(P+Q-2) \end{bmatrix} = \begin{bmatrix} h(0) & 0 & \dots & 0 \\ h(1) & h(0) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ h(Q-1) & h(Q-2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & h(Q-1) \end{bmatrix} \times \begin{bmatrix} x(0) \\ x(1) \\ \vdots \\ x(P-1) \end{bmatrix} \tag{21}$$

or

$$\mathbf{y} = \tilde{\mathbf{H}}\mathbf{x}. \tag{22}$$

It is implicit in the above representation that $x(n) = 0$ for $n \notin [0, P-1]$. The utility of the representation is in evaluation of convolution of two filters. The two convolution matrices, so far obtained, are related with each other as:

$$\tilde{\mathbf{H}} = \overline{\mathbf{H}}^T, \tag{23}$$

where $\overline{\mathbf{H}}$ denotes the reversal of the matrix.

In the following section, we address the design of a synthesis filter bank using the time-domain representations discussed so far.

## III. Design Framework

We have multirate observations, derived from various sub bands, and we aim to reconstruct or estimate the original high resolution signal. Initially, we consider observations to be at the same low rate and model them as output of a uniform analysis bank. The case when the observations are at different rates is addressed later in this paper. The subbands are applied to a matrix synthesis filter, as shown in Fig. 2, where reconstruction happens at a low rate to reduce computational cost. The output of the synthesis filter is then unblocked [71] to get the final high resolution signal. We model our problem of designing the matrix synthesis filter bank in the framework of Wiener filtering. The Wiener filter is an optimum filter to minimize the mean-square error (MSE). For a multi-input multi-output system, the solution is known as a matrix Wiener filter. The details of the design setup are explained next.

Consider the filter bank shown in Fig. 2. A wide-sense stationary (WSS) input signal $u(n)$ passes through a given analysis bank consisting of the filters $\{h_i(n)\}_{i=0}^{L-1}$. The analysis filters are assumed to be FIR and causal. The output signals $\{x_i(n)\}_{i=0}^{L-1}$ of the analysis bank are uniformly decimated by the factor $M$ to obtain the subband signals $\{v_i(n)\}_{i=0}^{L-1}$ which are then applied to a matrix synthesis filter $\mathbf{A}(n)$. We choose the synthesis filter to be FIR and causal, and formulate a Wiener filtering problem for a given length of the synthesis filter. The desired signal for the problem is derived from the input signal $u(n)$ by adding a given delay $d$ and then performing blocking operation [71]. The delay is added to compensate the time lag that appears in the input signal when it passes through the filter bank. The blocking operation ensures that the desired signal $\mathbf{d}(n)$ is jointly stationary with the subband signal $\mathbf{v}(n)$ [72]—a necessary condition required for application of Wiener filtering [73].

### A. FIR Matrix Synthesis Filter Expression

In Appendix, we obtain the following expression for an FIR matrix Wiener filter with length $P$:

$$\mathbf{r}_{d_i v} = \mathbf{a}_i^T \mathbf{R}_{vv}, \quad 0 \le i \le M-1, \tag{24}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



where $\mathbf{R}_{vv}$ is the auto-correlation matrix of the input signal $\mathbf{v}(n)$ to the filter, $\mathbf{r}_{d_i v}$ is the cross-correlation vector between the desired signal component $d_i(n)$ and the signal $\mathbf{v}(n)$, and $\mathbf{a}_i$ is the $i$-th row of the filter. This expression can be simplified in the context of the UFB given in Fig. 2. In order to do that, a vector $\mathbf{v}_s(n)$ is created by stacking the observation vectors of the subband signals as

$$\mathbf{v}_s(n) = \begin{bmatrix} \bar{\mathbf{v}}_0^T(n) & \bar{\mathbf{v}}_1^T(n) & \dots & \bar{\mathbf{v}}_{L-1}^T(n) \end{bmatrix}^T, \tag{25}$$

where $\mathbf{v}_j(n)$ is an observation vector representing the $j$-th input to the Wiener filter at time $n$ and is given by

$$\mathbf{v}_j(n) = \begin{bmatrix} v_j(n-P+1) & v_j(n-P+2) & \dots & v_j(n) \end{bmatrix}^T. \tag{26}$$

The quantities $\mathbf{R}_{vv}$ and $\mathbf{r}_{d_i v}$ can be expressed in terms of $\mathbf{v}_s(n)$ as

$$\begin{aligned} \mathbf{R}_{vv} &= \mathbf{E}[\mathbf{v}_s(n)\mathbf{v}_s^H(n)], \\ \mathbf{r}_{d_i v} &= \mathbf{E}[d_i(n)\mathbf{v}_s^H(n)]. \end{aligned} \tag{27}$$

To make use of the Wiener filter in (24), we derive an expression of the vector $\mathbf{v}_s(n)$ in terms of the input to the filter bank. We can note from Fig. 2 that $v_j(n) = x_j(Mn)$ and this relationship can also be expressed in matrix form as

$$\mathbf{v}_j(n) = \mathbf{D}^0_{P,M(P-1)+1}\mathbf{x}_j(Mn), \tag{28}$$

where

$$\begin{aligned} &\mathbf{x}_j(Mn) \\ &= \begin{bmatrix} x_j(M(n-P+1)) & x_j(M(n-P+1)+1) & \dots & x_j(Mn) \end{bmatrix}^T. \end{aligned} \tag{29}$$

Further, the vector $\mathbf{x}_j(Mn)$ can be expressed in terms of the input $u(n)$. We start with an assumption that all the analysis filters are causal and have equal length $Q$. It can be assured by appropriate zero padding. Now we can obtain $\mathbf{x}_j(Mn)$ as

$$\mathbf{x}_j(Mn) = \mathbf{H}_j\mathbf{u}(Mn), \tag{30}$$

where $\mathbf{H}_j$ is a convolution matrix. The entire input to the Wiener filter can be expressed as follows:

$$\begin{aligned} \mathbf{v}_s(n) = \begin{bmatrix} \bar{\mathbf{v}}_0(n) \\ \bar{\mathbf{v}}_1(n) \\ \vdots \\ \bar{\mathbf{v}}_{L-1}(n) \end{bmatrix} &= \begin{bmatrix} \mathbf{D}^0_{P,M(P-1)+1}\overline{\mathbf{H}}_0 \\ \mathbf{D}^0_{P,M(P-1)+1}\overline{\mathbf{H}}_1 \\ \vdots \\ \mathbf{D}^0_{P,M(P-1)+1}\overline{\mathbf{H}}_{L-1} \end{bmatrix} \bar{\mathbf{u}}(Mn), \\ &= \mathbf{K}\bar{\mathbf{u}}(Mn), \end{aligned} \tag{31}$$

where $\mathbf{K}$ is a matrix of size $LP \times [M(P-1)+Q]$. Using this result, the auto-correlation matrix $\mathbf{R}_{vv}$ can be written as

$$\begin{aligned} \mathbf{R}_{vv} &= \mathbf{K}\mathbf{E}[\bar{\mathbf{u}}(Mn)\bar{\mathbf{u}}^H(Mn)]\mathbf{K}^H, \\ &= \mathbf{K}\mathbf{R}_{uu}\mathbf{K}^H. \end{aligned} \tag{32}$$

Similarly the cross-correlation vector $\mathbf{r}_{d_i v}$ becomes

$$\begin{aligned} \mathbf{r}_{d_i v} &= \mathbf{E}[u(Mn-i-d)\bar{\mathbf{u}}^H(Mn)]\mathbf{K}^H, \\ &= \mathbf{r}_{uu}^{i+d}\mathbf{K}^H, \end{aligned} \tag{33}$$

where $\mathbf{r}_{uu}^{i+d}$ is a row vector given by

$$\mathbf{r}_{uu}^{i+d} = \begin{bmatrix} r_{uu}(-i-d) \\ r_{uu}(-i-d+1) \\ \vdots \\ r_{uu}(-i-d+M(P-1)+Q-1) \end{bmatrix}^T. \tag{34}$$

By substituting these results in (24), we obtain

$$\mathbf{r}_{uu}^{i+d}\mathbf{K}^H = \mathbf{a}_i^T\mathbf{K}\mathbf{R}_{uu}\mathbf{K}^H, \qquad 0 \le i \le M-1. \tag{35}$$

The resulting equation is of form

$$\mathbf{A}\mathbf{a}_i = \mathbf{b}_i, \tag{36}$$

where

$$\begin{aligned} \mathbf{A} &= \left(\mathbf{K}\mathbf{R}_{uu}\mathbf{K}^H\right)^T, \\ \mathbf{b}_i &= \left(\mathbf{r}_{uu}^{i+d}\mathbf{K}^H\right)^T. \end{aligned} \tag{37}$$

If $\mathbf{A}$ is invertible, then we have a unique solution. Otherwise, the existence of a solution depends on the nature of the system of equations. If the same is consistent, then all solutions can be expressed as

$$\mathbf{a}_i = \mathbf{A}^\dagger\mathbf{b}_i + \left(\mathbf{I} - \mathbf{A}^\dagger\mathbf{A}\right)\mathbf{w}, \tag{38}$$

where $\mathbf{w}$ is an arbitrary vector. In the next theorem, we prove that the equation always results in a solution.

*Theorem 1: For an FIR uniform analysis bank, there will be at least one optimal synthesis bank solution that minimizes the mean-square error.*

*Proof:* We have to prove that the solution given by (38) is always an exact solution of the equation. Substituting the solution in the equation, we obtain

$$\begin{aligned} \mathbf{A}\mathbf{a}_i &= \mathbf{A}\mathbf{A}^\dagger\mathbf{b}_i + \mathbf{A}\left(\mathbf{I} - \mathbf{A}^\dagger\mathbf{A}\right)\mathbf{w}, \\ &= \mathbf{A}\mathbf{A}^\dagger\mathbf{b}_i. \end{aligned} \tag{39}$$

The matrix $\mathbf{A}$ can be expressed as

$$\begin{aligned} \mathbf{A} &= \left(\mathbf{K}\mathbf{U}\Lambda\mathbf{U}^H\mathbf{K}^H\right)^T, \\ &= \left(\mathbf{K}\mathbf{U}\Lambda^{1/2}\right)^*\left(\mathbf{K}\mathbf{U}\Lambda^{1/2}\right)^T, \\ &= \mathbf{C}\mathbf{C}^H, \end{aligned} \tag{40}$$

by performing eigendecomposition of the auto-correlation matrix $\mathbf{R}_{uu}$. With this,

$$\begin{aligned} \mathbf{A}\mathbf{a}_i &= \mathbf{C}\mathbf{C}^H(\mathbf{C}\mathbf{C}^H)^\dagger\mathbf{b}_i, \\ &= \mathbf{C}\mathbf{C}^H(\mathbf{C}^H)^\dagger\mathbf{C}^\dagger\mathbf{b}_i. \end{aligned} \tag{41}$$

Using the pseudoinverse identities [74], the equation reduces to

$$\begin{aligned} \mathbf{A}\mathbf{a}_i &= \mathbf{C}\mathbf{C}^\dagger\mathbf{b}_i, \\ &= (\mathbf{K}\mathbf{U}\Lambda^{1/2})^*\left((\mathbf{K}\mathbf{U}\Lambda^{1/2})^*\right)^\dagger\mathbf{K}^*(\mathbf{r}_{uu}^{i+d})^T. \end{aligned} \tag{42}$$

The auto-correlation matrix $\mathbf{R}_{uu}$ is in general positive definite but always positive semi-definite. We consider the solution of equation under these two cases:

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



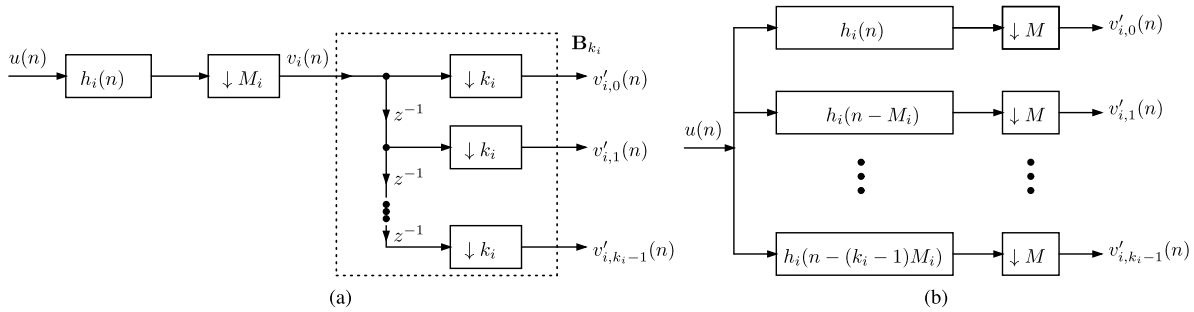



Fig. 3. Conversion of an NUFB into an equivalent UFB. (a)Blocking of a subband. (b)Equivalent representation.

1) Case I: Positive Definite Correlation Matrix
In this case, the auto-correlation matrix is invertible, so is the diagonal matrix $\Lambda$. Hence we can process the equation as follows:

$$\begin{aligned}\mathbf{A}\mathbf{a}_i &= (\mathbf{K}\mathbf{U}\Lambda^{1/2})^*((\mathbf{K}\mathbf{U}\Lambda^{1/2})^*)^\dagger(\mathbf{K}\mathbf{U}\Lambda^{1/2})^* \\ &\quad \cdot \left((\mathbf{U}\Lambda^{1/2})^*\right)^{-1}(\mathbf{r}_{uu}^{i+d})^T, \\ &= (\mathbf{K}\mathbf{U}\Lambda^{1/2})^*\left((\mathbf{U}\Lambda^{1/2})^*\right)^{-1}(\mathbf{r}_{uu}^{i+d})^T, \\ &= \mathbf{b}_i. \end{aligned} \tag{43}$$

2) Case II: Positive Semi-Definite Correlation Matrix
The diagonal matrix $\Lambda$ is not invertible under this case. However we exploit the relation between the vector $\mathbf{r}_{uu}^{i+d}$ and the matrix $\mathbf{R}_{uu}$. For $(i+d) < (M(P-1)+Q)$, the vector $\mathbf{r}_{uu}^{i+d}$ is one of the row of the correlation matrix. If $(i+d) \geq (M(P-1)+Q)$, then the vector $\mathbf{r}_{uu}^{i+d}$ can still be expressed as a linear combination of the rows of the auto-correlation matrix [75]. Thus, we can always express the vector $\mathbf{r}_{uu}^{i+d}$ as

$$\mathbf{r}_{uu}^{i+d} = \sum_j c_j \mathbf{R}_{uu}(j,:) = \sum_j c_j \mathbf{U}(j,:)\Lambda\mathbf{U}^H, \tag{44}$$

where $\mathbf{R}_{uu}(j,:)$ is the $j$-th row of the correlation matrix. With this,

$$\begin{aligned}\mathbf{A}\mathbf{a}_i &= (\mathbf{K}\mathbf{U}\Lambda^{1/2})^*\big((\mathbf{K}\mathbf{U}\Lambda^{1/2})^*\big)^\dagger(\mathbf{K}\mathbf{U}\Lambda^{1/2})^* \\ &\quad \cdot\Lambda^{1/2}\Big(\sum_j c_j\mathbf{U}(j,:)\Big)^T, \\ &= (\mathbf{K}\mathbf{U}\Lambda^{1/2})^*\Lambda^{1/2}\Big(\sum_j c_j\mathbf{U}(j,:)\Big)^T, \\ &= \mathbf{b}_i. \end{aligned} \tag{45}$$

We can observe that we have

$$\mathbf{A}\mathbf{a}_i = \mathbf{A}\mathbf{A}^\dagger\mathbf{b}_i = \mathbf{b}_i \tag{46}$$

under both the cases and hence the solution given by (38) is exact. □

From the above theorem, we can conclude that if the matrix $\mathbf{A}$ is non-invertible, we have infinite number of optimum solutions possible. This gives an opportunity to optimize the synthesis bank for a desired property.

*1) Computational Complexity:* The computational complexity can be evaluated from various algebraic manipulations required for the solution in (38). To start with, the matrix $\mathbf{A}$ evaluation needs $O\big(LP(M(P-1)+Q)^2\big)+O\big((LP)^2\big)$ calculations and its pseudoinverse requires $O\big((LP)^3\big)$ calculations. The term $(\mathbf{I}-\mathbf{A}^\dagger\mathbf{A})$ can then be evaluated with $O\big((LP)^3\big)$ cost. As the vector $\mathbf{b}_i$ needs to be evaluated for each $i \in [0, M-1]$, the total cost of obtaining $\mathbf{b}_i$'s is $O\big(MLP(M(P-1)+Q)\big)$. All the products $\mathbf{A}^\dagger\mathbf{b}_i$ and $(\mathbf{I}-\mathbf{A}^\dagger\mathbf{A})\mathbf{w}$ can then be obtained with complexity $O\big(M(LP)^2\big)$. Therefore, the complexity is quadratic with respect to $M$ and $Q$, and cubic with respect to $L$ and $P$.

Our synthesis filter is multiple-input multiple-output (MIMO) FIR system, thus enabling the operations at the lowest sampling rate. The MIMO implementation brings down the complexity drastically, thereby for each reconstructed output sample, the number of multiplication and additions required are only $LP$ and $(LP-1)$, respectively.

### B. Extension to NUFB

The results developed so far for a UFB can be extended for an NUFB case. Before applying Wiener filtering on an NUFB, we convert it into an equivalent UFB with the help of blocking operation. This conversion is needed as Wiener filtering requires joint stationarity of the input and the desired signals [72]. Each subband of the NUFB with a decimation factor $M_i$ is blocked with a factor

$$k_i = M/M_i, \qquad 0 \leq i < L \tag{47}$$

where $L$ is the number of subbands and $M$ is the LCM of all decimation factors. The process is shown in Fig. 3a. An equivalent representation of the process, as shown in Fig. 3, can be derived using the noble identities [68], [72]. Thus with blocking, the NUFB behaves like a UFB with decimation factor equal to $M$ and with analysis filters given by

$$h'_{i,l}(n) = h_i(n - lM_i), \quad 0 \leq i < L;\ 0 \leq l < k_i. \tag{48}$$

Now the Wiener synthesis filter for the NUFB can be obtained using (38).

### C. Minimum Mean-square Error

We now determine the MSE for a given length of the Wiener filter. The result can be used to choose the length of the

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



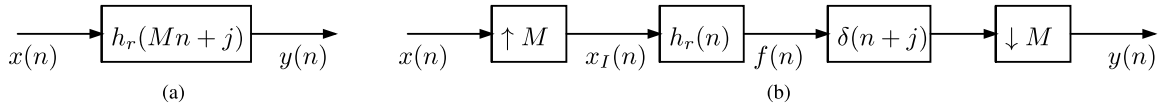



Fig. 4. The polyphase identity. (a)A polyphase component. (b)Equivalent multirate circuit.

synthesis filter based on the desired level of the error. The MSE is defined as

$$J = \mathbf{E}[\mathbf{e}^H(n)\mathbf{e}(n)] = \sum_{i=0}^{M-1} \mathbf{E}[e_i^*(n)e_i(n)]. \tag{49}$$

Its minimum value is achieved using the Wiener filter and is denoted by $J_{\min}$. Further, $J_{\min}$ is the sum of channel-wise MSEs which are given by

$$\begin{aligned} J_{\min}^i &= \mathbf{E}[e_i^*(n)e_i(n)], \\ &= \mathbf{E}[e_i^*(n)d_i(n)] - \mathbf{E}[e_i^*(n)y_i(n)]. \end{aligned} \tag{50}$$

Using the principle of orthogonality [76], the above equation reduces to

$$\begin{aligned} J_{\min}^i &= \mathbf{E}[e_i^*(n)d_i(n)], \\ &= \mathbf{E}[d_i^*(n)d_i(n)] - \mathbf{E}[y_i^*(n)d_i(n)], \\ &= r_{uu}(0) - \mathbf{E}[\mathbf{v}_s^H(n)\mathbf{a}_i^* d_i(n)], \\ &= r_{uu}(0) - \mathbf{r}_{d_i v}\mathbf{a}_i^*. \end{aligned} \tag{51}$$

Substituting the expressions for $\mathbf{r}_{d_i v}$ and $\mathbf{a}_i$ from (33) and (38), respectively, we obtain

$$\begin{aligned} J_{\min}^i &= r_{uu}(0) - \mathbf{b}_i^T \mathbf{A}^{\dagger *}\mathbf{b}_i^* - \left(\mathbf{b}_i - \mathbf{A}^H(\mathbf{A}^\dagger)^H\mathbf{b}_i\right)^T \mathbf{w}^*, \\ &= r_{uu}(0) - \mathbf{b}_i^T \mathbf{A}^{\dagger *}\mathbf{b}_i^* - \left(\mathbf{b}_i - \mathbf{A}\mathbf{A}^\dagger\mathbf{b}_i\right)^T \mathbf{w}^*. \end{aligned} \tag{52}$$

The term with the arbitrary vector is zero due to (46) and the equation reduces to

$$J_{\min}^i = r_{uu}(0) - \mathbf{b}_i^T \mathbf{A}^{\dagger *}\mathbf{b}_i^*. \tag{53}$$

This proves that all the solutions for the synthesis bank results in the same MSE, which is intuitive.

*1) Optimal Delay:* The channel MSE $J_{\min}^i$ depends on the delay $d$. The same is also observed in our simulation results. The exact form of the relation can be obtained from (37) as

$$\begin{aligned} J_{\min}^i &= r_{uu}(0) - \mathbf{r}_{uu}^{i+d}\mathbf{K}^H\left(\mathbf{K}\mathbf{R}_{uu}\mathbf{K}^H\right)^\dagger \mathbf{K}\left(\mathbf{r}_{uu}^{i+d}\right)^H \\ &= r_{uu}(0) - \mathbf{r}_{uu}^{i+d}\mathbf{B}\left(\mathbf{r}_{uu}^{i+d}\right)^H \end{aligned} \tag{54}$$

$J_{\min}^i$ can be observed as a quadratic form with the matrix $\mathbf{B}$ being a Hermitian matrix. If we are interested to obtain an optimal delay value, we can use the existing results [77], [78] on the quadratic forms.

In the following section, we derive the necessary and sufficient conditions on a filter bank for the MSE to be zero.

## IV. FIR Perfect Reconstruction Filter Bank

In the following analysis, we prove that the designed synthesis bank can achieve PR provided that the analysis bank satisfies certain condition. To justify PR property for any arbitrary input signal on a strong analytical background, we verify the pseudocirculant property [68], [69] for the solution.

### A. Pseudocirculant Property

We evaluate the product $\mathbf{P}(z) = \mathbf{A}(z)\mathbf{E}(z)$ in time-domain where $\mathbf{A}(z)$ and $\mathbf{E}(z)$ are the polyphase component matrices for the synthesis bank and the analysis bank, respectively [69]. The matrix $\mathbf{P}(z)$ is designed as an identity matrix in previous works [23], [29]. Our design approach is more general where $\mathbf{P}(z)$ satisfies the pseudocirculant property. In time-domain, we obtain the matrix $\mathbf{P}(z)$ from convolution of the two matrix filters as

$$p_{i,j}(n) = \sum_{r=0}^{L-1} a_{i,r}(n) * e_{r,j}(n), \qquad 0 \le i, j \le M-1. \tag{55}$$

In the above equation, the filter $e_{r,j}(n)$ can be obtained from the analysis filters as [68]

$$e_{r,j}(n) = h_r(Mn + j). \tag{56}$$

As $Q$ is the length of the analysis filters, the length of the filter $e_{r,j}(n)$ will be at max

$$T = \left\lceil \frac{Q}{M} \right\rceil. \tag{57}$$

By zero padding, we ensure that all the constituent filters of the polyphase filter $\mathbf{E}(n)$ have the same length $T$. With that, the convolution given in (55) can be expressed in matrix form using (22) as

$$\mathbf{p}_{i,j} = \sum_{r=0}^{L-1} \tilde{\mathbf{E}}_{r,j}\mathbf{a}_{i,r}, \tag{58}$$

where $\tilde{\mathbf{E}}_{r,j}$ is a convolution matrix of size $(P+T-1)\times P$. We like to express this convolution matrix in terms of the analysis filter $h_r(n)$. For that, we use a polyphase identity, as shown in Fig. 4, obtained from the existing result [68] by finding the 0-th polyphase component of a time-advanced filter $h_r(n+j)$. Using this identity, we obtain the following relation

$$\begin{aligned} \tilde{\mathbf{E}}_{r,j} &= \mathbf{D}^j_{P+T-1,M(P-1)+Q}\tilde{\mathbf{H}}_r(\mathbf{D}^0_{P,M(P-1)+1})^T, \\ &= \mathbf{D}^j_{P+T-1,M(P-1)+Q}\overline{\mathbf{H}}_r^T(\mathbf{D}^0_{P,M(P-1)+1})^T, \\ &= \mathbf{D}^j_{P+T-1,M(P-1)+Q}\left(\mathbf{D}^0_{P,M(P-1)+1}\overline{\mathbf{H}}_r\right)^T. \end{aligned} \tag{59}$$

It can be seen that Fig. 4b contains a combination of upsampling, filtering, time-advance and downsampling. Even then the matrix representations developed in this paper for various multirate operations have been successfully applied to deal with it. Continuing forward, we use the above relation in (58) and obtain

$$\begin{aligned} \mathbf{p}_{i,j} &= \mathbf{D}^j_{P+T-1,M(P-1)+Q}\sum_{r=0}^{L-1}\left(\mathbf{D}^0_{P,M(P-1)+1}\overline{\mathbf{H}}_r\right)^T\mathbf{a}_{i,r}, \\ &= \mathbf{D}^j_{P+T-1,M(P-1)+Q}\mathbf{K}^T\mathbf{a}_i. \end{aligned} \tag{60}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



Next we derive a necessary and sufficient condition, in the form of a theorem, for the analysis and synthesis banks of a filter bank to result in PR.

*Theorem 2: An FIR uniform filter bank is PR if and only if*

$$\mathbf{K}^T\mathbf{a}_i = c\mathbf{e}_{i+d}, \tag{61}$$

*where $c$ is a non-zero constant and the column vector $\mathbf{e}_{i+d}$ is defined as follows:*

$$e_{i+d}(l) = \begin{cases} 1, & \text{if } l = i+d \\ 0, & \text{otherwise.} \end{cases} \tag{62}$$

*Proof:* The proof follows the Vaidyanathan's well-known pseudocirculant property. A PR filter bank must satisfies the generalized pseudocirculant condition given by [68]

$$\mathbf{P}(z) = cz^{-n_0}\begin{bmatrix} \mathbf{0} & \mathbf{I}_{M-r} \\ z^{-1}\mathbf{I}_r & \mathbf{0} \end{bmatrix}. \tag{63}$$

We have evaluated $\mathbf{P}(z)$ in time-domain and next we find conditions for it to be pseudocirculant. Let $\mathbf{K}^T\mathbf{a}_i = \mathbf{s}_i$, then (60) becomes

$$\mathbf{p}_{i,j} = \begin{bmatrix} s_i(j) & s_i(j+M) & \dots & s_i(j+M(P+T-2)) \end{bmatrix}^T. \tag{64}$$

If the matrix filter $\mathbf{P}(z)$ is pseudocirculant, then the following properties are true in time-domain:

1) Property 1: In a row, only one $\mathbf{p}_{i,j}$ is non-zero i.e. only one out of $\mathbf{p}_{i,0}, \mathbf{p}_{i,1}, \dots, \mathbf{p}_{i,M-1}$ is non-zero.
2) Property 2: In a column, only one $\mathbf{p}_{i,j}$ is non-zero i.e. only one out of $\mathbf{p}_{0,j}, \mathbf{p}_{1,j}, \dots, \mathbf{p}_{M-1,j}$ is non-zero.
3) Property 3: For a vector $\mathbf{p}_{i,j}$, at most one entry is non-zero.
4) Property 4: All the entries on upper-left to lower-right diagonals are the same i.e.

$$\mathbf{p}_{i+1,j+1} = \mathbf{p}_{i,j}, \quad 0 \le i, j \le M-2. \tag{65}$$

5) Property 5: The first column will have

$$p_{i+1,0}(l) = \begin{cases} 0, & \text{if } l = 0 \\ p_{i,M-1}(l-1), & \text{otherwise} \end{cases}, \quad 0 \le i \le M-2. \tag{66}$$

The above properties will result in the following constraints on the vector $\mathbf{s}_i$:

1) Property 1 implies that only one vector

$$\mathbf{p}_{i,j} = \begin{bmatrix} s_i(j) & s_i(j+M) & \dots & s_i(j+M(P+T-2)) \end{bmatrix}^T, \quad 0 \le j < M \tag{67}$$

is non-zero in a row. Further, Property 3 means that only one element of the non-zero vector is non-zero. Together these two conditions imply that only one element of the vector $\mathbf{s}_i$ is non-zero. Assume that $\mathbf{p}_{i,j} \neq 0$ for $j = j_i$ and $p_{i,j_i}(l_i) = s_i(j_i + Ml_i) \neq 0$, then we can express

$$\mathbf{s}_i = c_i\mathbf{e}_{j_i+Ml_i}. \tag{68}$$

2) Property 4 results in

$$s_{i+1}(Ml+j+1) = s_i(Ml+j), \quad 0 \le j \le M-2; \quad 0 \le l \le P+T-2. \tag{69}$$

From the above equation, all the elements of the vector $\mathbf{s}_{i+1}$, except at indices that are multiples of $M$, can be obtained from the vector $\mathbf{s}_i$.

3) Property 5 implies

$$p_{i+1,0}(0) = s_{i+1}(0) = 0, \quad 0 \le i \le M-2. \tag{70}$$

Another implication is that

$$s_{i+1}(Ml) = s_i(Ml-1). \tag{71}$$

Combining the two results (69) and (71), we get

$$s_{i+1}(l) = \begin{cases} 0, & l = 0, \\ s_i(l-1), & \text{otherwise.} \end{cases} \tag{72}$$

Thus the vector $\mathbf{s}_{i+1}$ is just a shifted version of the vector $\mathbf{s}_i$.

The last constraint implies that all the vectors $\mathbf{s}_i$ can be obtained from a single vector $\mathbf{s}_0$. If the vector $\mathbf{s}_0$ has form

$$\mathbf{s}_0 = c\mathbf{e}_{Ml_0+j_0} = c\mathbf{e}_{d_0}, \tag{73}$$

then all the vectors $\mathbf{s}_i$ can be expressed as

$$\mathbf{s}_i = c\mathbf{e}_{i+d_0}, \qquad 0 \le i \le M-1. \tag{74}$$

Due to the above equation, we have

$$\mathbf{p}_{i,j} = c\mathbf{D}^j_{P+T-1,M(P-1)+Q}\mathbf{e}_{i+d_0}. \tag{75}$$

In z-domain, this means that

$$\mathbf{P}(z) = cz^{-l_0}\begin{bmatrix} \mathbf{0} & \mathbf{I}_{M-j_0} \\ z^{-1}\mathbf{I}_{j_0} & \mathbf{0} \end{bmatrix} \tag{76}$$

and the reconstructed output is $\hat{u}(n) = cu(n-d_0-M+1)$. One can easily see that the variable $d_0$ is the same as the delay $d$ in the desired signal path. □

It is to be noted that the above theorem puts restriction on both the analysis and synthesis banks. It may happen that a solution may not exist for the synthesis bank to result in PR. We will first obtain a parameterization of all PR solutions for the synthesis bank and later derive necessary & sufficient conditions for the same to exist.

### *B. Parametric Form for PR Solutions*

In the previous subsection, we obtained the following condition on the analysis and synthesis banks to have PR:

$$\mathbf{K}^T\mathbf{a}_i = c\mathbf{e}_{i+d}, \quad 0 \le i < M. \tag{77}$$

The above equation is of form

$$\mathbf{A}\mathbf{a}_i = \mathbf{b}_i \tag{78}$$

and thus represents a system of equations. If the same is consistent, then all the PR solutions for the synthesis bank are given by

$$\begin{aligned} \mathbf{a}_i &= \mathbf{A}^\dagger\mathbf{b}_i + (\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{w}, \\ &= c(\mathbf{K}^\dagger)^T\mathbf{e}_{i+d} + (\mathbf{I} - \mathbf{K}\mathbf{K}^\dagger)^T\mathbf{w}, \end{aligned} \tag{79}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



where $\mathbf{w}$ is an arbitrary vector. We can observe that the solutions have turned out to be independent of the input signal.

### C. Necessary and Sufficient Conditions on Analysis Bank and Delay

In order for a PR solution to exist, the analysis bank has to satisfy certain condition. We present the same in the form of a theorem about the matrix $\mathbf{K}$ which represents the analysis bank.

*Theorem 3: For an FIR uniform analysis bank, at least one PR synthesis bank solution is possible if the matrix* $\mathbf{K}$ *satisfies the relation*

$$\mathbf{K}^{\dagger}\mathbf{K} = \mathbf{I} + \begin{bmatrix} \mathbf{B}_{p\times p} & \mathbf{0}_{p\times r} & \mathbf{C}_{p\times(q-p-r)} \\ \mathbf{0}_{r\times p} & \mathbf{0}_{r\times r} & \mathbf{0}_{r\times(q-p-r)} \\ \mathbf{C}^{H}_{p\times(q-p-r)} & \mathbf{0}_{(q-p-r)\times r} & \mathbf{D}_{(q-p-r)\times(q-p-r)} \end{bmatrix}, \tag{80}$$

*where* $\mathbf{0}_{r\times r}$ *is the largest zero sub-block possible with* $r \geq M$, *and* $\mathbf{B}_{p\times p}$, $\mathbf{D}_{(q-p-r)\times(q-p-r)}$ *and* $\mathbf{C}_{p\times(q-p-r)}$ *are arbitrary matrices with the first two being Hermitian.*

*Proof:* We require the system of linear equations given by

$$\mathbf{K}^{T}\mathbf{a}_i = c\mathbf{e}_{i+d} \tag{81}$$

to be consistent to have an exact solution. In order to obtain the condition for that, we substitute a PR solution in the above equation and obtain

$$c\left(\mathbf{K}^{\dagger}\mathbf{K}\right)^{T}\mathbf{e}_{i+d} + \left(\mathbf{K} - \mathbf{K}\mathbf{K}^{\dagger}\mathbf{K}\right)^{T}\mathbf{w} = c\mathbf{e}_{i+d}. \tag{82}$$

Using the fact that $\mathbf{K}\mathbf{K}^{\dagger}\mathbf{K} = \mathbf{K}$, the equation simplifies to

$$\left(\left(\mathbf{K}^{\dagger}\mathbf{K}\right)^{T} - \mathbf{I}\right)\mathbf{e}_{i+d} = \mathbf{0}, \quad 0 \leq i \leq M-1. \tag{83}$$

If the matrix $\mathbf{K}^{\dagger}\mathbf{K}$ is of dimension $q \times q$, then in order for the above equation to be satisfied, we require

$$\left(\mathbf{K}^{\dagger}\mathbf{K}\right)^{H} = \mathbf{I} + \begin{bmatrix} \mathbf{X}_{q\times p} & \mathbf{0}_{q\times r} & \mathbf{Y}_{q\times(q-p-r)} \end{bmatrix}, \tag{84}$$

where $p \leq d$, $r \geq M$, and $\mathbf{X}_{q\times p}$ and $\mathbf{Y}_{q\times(q-p-r)}$ are arbitrary matrices of dimensions $q \times p$ and $q \times (q-p-r)$, respectively. As $\mathbf{K}^{\dagger}\mathbf{K}$ is a Hermitian matrix, we have

$$\begin{bmatrix} \mathbf{X}_{q\times p} & \mathbf{0}_{q\times r} & \mathbf{Y}_{q\times(q-p-r)} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^{H}_{q\times p} \\ \mathbf{0}_{r\times q} \\ \mathbf{Y}^{H}_{q\times(q-p-r)} \end{bmatrix}. \tag{85}$$

Thus the two matrices $\mathbf{X}_{q\times p}$ and $\mathbf{Y}_{q\times(q-p-r)}$ have following form:

$$\mathbf{X}_{q\times p} = \begin{bmatrix} \mathbf{B}_{p\times p} \\ \mathbf{0}_{r\times p} \\ \mathbf{C}^{H}_{p\times(q-p-r)} \end{bmatrix} \tag{86}$$

and

$$\mathbf{Y}_{q\times(q-p-r)} = \begin{bmatrix} \mathbf{C}_{p\times(q-p-r)} \\ \mathbf{0}_{r\times(q-p-r)} \\ \mathbf{D}_{(q-p-r)\times(q-p-r)} \end{bmatrix}, \tag{87}$$

where matrices $\mathbf{B}_{p\times p}$ and $\mathbf{D}_{(q-p-r)\times(q-p-r)}$ are Hermitian matrices. Hence the condition on the matrix $\mathbf{K}$ becomes

$$\mathbf{K}^{\dagger}\mathbf{K} = \mathbf{I} + \begin{bmatrix} \mathbf{B}_{p\times p} & \mathbf{0}_{p\times r} & \mathbf{C}_{p\times(q-p-r)} \\ \mathbf{0}_{r\times p} & \mathbf{0}_{r\times r} & \mathbf{0}_{r\times(q-p-r)} \\ \mathbf{C}^{H}_{p\times(q-p-r)} & \mathbf{0}_{(q-p-r)\times r} & \mathbf{D}_{(q-p-r)\times(q-p-r)} \end{bmatrix}. \tag{88}$$

□

The PR requirement imposes the following condition on the desired signal delay $d$.

*Lemma 1: With an analysis bank matrix* $\mathbf{K}$ *satisfying the above condition, PR is possible only when the delay* $d$ *is in the range:*

$$p \leq d \leq p + r - M. \tag{89}$$

*Proof:* The proof easily follows from the equations (83) and (84). □

The above lemma together with Theorem 3 constitute the necessary and sufficient conditions for a PR synthesis solution to exist. The lemma points that with suitable delay, PR is achievable for an analysis bank which is otherwise not possible. This way we have attempted to expand the space of PR FIR filter banks. Further, we can have PR for a number of delay values which gives an opportunity to use delay as an optimization parameter, based on various design criterion like Hermitian symmetry, time or frequency localization [23], coding gain [79], etc. For this, the number of channels required is $L > M$ and the channels should satisfy the PR condition given by Theorem 3. For a fixed delay, it is an unconstrained problem with respect to the arbitrary vector $\mathbf{w}$ [23], [30]. The required optimized solution can be selected among the solutions for various delays by imposing an additional constraint. Design specific constraints and their solutions can be considered for future work.

It can be seen that the matrix $\mathbf{K}$ easily satisfy the PR condition if it has full column rank. The PR condition induces certain restriction on its null space which is given by the following lemma.

*Lemma 2: For a PR UFB, the analysis bank matrix* $\mathbf{K}$ *has a null space spanned by the following set of vectors:*

$$\mathbf{z} = \begin{bmatrix} \mathbf{f}_{1\times p} & \mathbf{0}_{1\times r} & \mathbf{g}_{1\times(q-p-r)} \end{bmatrix}^{T}, \tag{90}$$

*where* $\mathbf{f}_{1\times p}$ *and* $\mathbf{g}_{1\times(q-p-r)}$ *are arbitrary vectors.*

*Proof:* The matrix $\mathbf{K}$ represents the analysis bank of a filter bank. Hence, when an input signal $u(n)$ is applied to the analysis bank, the resulting subband vector can be obtained as

$$\mathbf{v}_s(n) = \mathbf{K}\bar{\mathbf{u}}(Mn). \tag{91}$$

Consider a case where the vector $\bar{\mathbf{u}}(Mn)$ lies in the null space of the matrix $\mathbf{K}$. As a result, the vector $\mathbf{v}_s(n)$ as well as the output of the synthesis bank, both are zero. The reconstructed signal is then

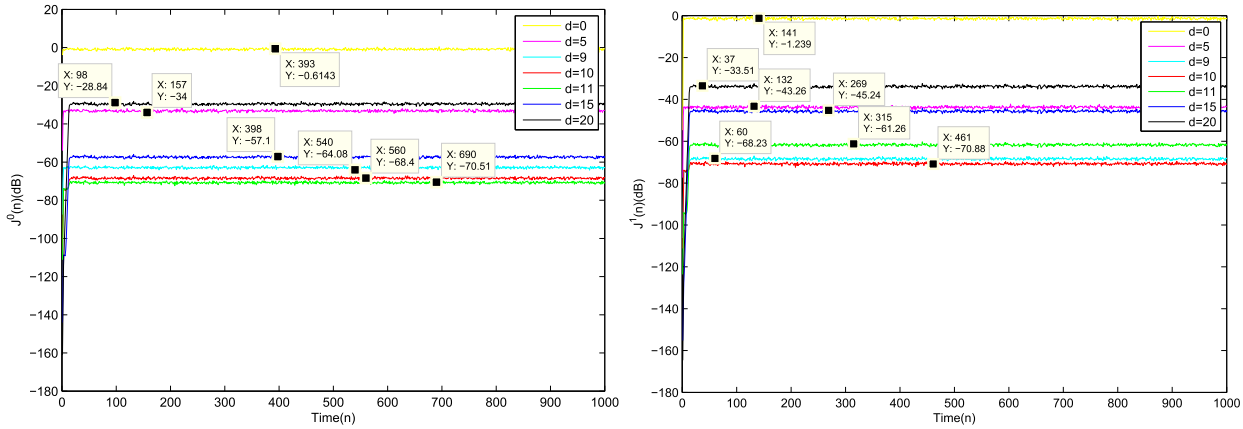
$$\hat{u}(Mn+i) = y_{M-1-i}(n) = 0, \quad 0 \leq i < M. \tag{92}$$

For PR, we require

$$\hat{u}(n) = u(n-d-M+1). \tag{93}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.






(a) (b)

Fig. 5. Experiment 1: Ensemble-averaged squared error for various delay values. (a)Channel 0. (b)Channel 1.

Combining the two, we get

$$\hat{u}(Mn+i)=u(Mn+i-d-M+1)=0, \quad 0\le i\le M-1. \tag{94}$$

Thus the vector $\bar{\mathbf{u}}(Mn)$ should be of form

$$\bar{\mathbf{u}}(Mn) = \begin{bmatrix} u(Mn) \\ u(Mn-1) \\ \vdots \\ u(Mn-d+1) \\ \mathbf{0}_{M\times 1} \\ u(Mn-d-M) \\ \vdots \\ u(Mn-q+1) \end{bmatrix}^T . \tag{95}$$

A basis vector for the null space of the matrix $\mathbf{K}$ should satisfy the above condition for all possible values of $d$ and hence should have the form

$$\mathbf{z} = \begin{bmatrix} \mathbf{f}_{1\times p} & \mathbf{0}_{1\times r} & \mathbf{g}_{1\times(q-p-r)} \end{bmatrix}^T , \tag{96}$$

where $\mathbf{f}_{1\times p}$ and $\mathbf{g}_{1\times(q-p-r)}$ are arbitrary vectors. □

In this section, we presented various necessary and sufficient conditions for a PR synthesis bank solution to exist. In the next section, we present some experiments which validates our results.

## V. Experimental Results

In this section, we verify the derived results with a number of experiments. These experiments are chosen to demonstrate the wide scope of application of our design. In Experiment 1, we obtain the MSE for various delay values in the desired signal and compare the results. We present the improvement in reconstruction, in terms of the MSE, with addition of a new subband in Experiment 2. We also examine the impact of delay on PR. In Experiment 3, a PR non-uniform filter bank obtained through our design is verified with a non-stationary input signal.

### A. Experiment 1

We consider a 2-channel maximally decimated uniform analysis bank designed using fir1 routine of Matlab with the following parameters: the analysis filter $h_0(n)$ is low-pass filter with cut-off 0.6 and length 9, and $h_1(n)$ is high-pass filter with cut-off 0.4 and length 10. The input to the filter bank is a zero mean, unit variance second-order autoregressive (AR(2)) process with coefficients 0.7 and 0.1. We choose the length of the constituent filters of the matrix synthesis filter to be 11 and design the filter for various delay values using (38). We obtain ensembled-averages of the channel-wise squared errors, shown in Figs. 5a and 5b, for these synthesis filters by performing 200 independent runs. The corresponding theoretical values, obtained using (53), are provided in Table I. It is to be mentioned here that the analysis bank does not satisfy Theorem 3. We can see that the observed error values are closer to the theoretical values for both channels. Further, we can notice that the best result among the selected delay values, in terms of the total MSE, is with delay 10. This shows that a better reconstruction can be obtained if some delay in the reconstructed output is acceptable.

TABLE I
Experiment 1: MSE for various Delay Values

| $d$ | $J^0_{\min}$(dB) | $J^1_{\min}$(dB) | $J_{\min}$(dB) |
|---|---|---|---|
| 0 | -0.8673 | -1.3434 | 1.9115 |
| 5 | -33.1717 | -43.7244 | -32.8052 |
| 9 | -62.8689 | -68.3970 | -61.7967 |
| 10 | -68.3970 | -70.7269 | -66.3972 |
| 11 | -70.7269 | -61.6691 | -61.1606 |
| 15 | -57.3320 | -45.5881 | -45.3068 |
| 20 | -29.5288 | -33.7826 | -28.1442 |

### B. Experiment 2

In this experiment, we examine the effect of number of subbands on the MSE along with the role of delay in achieving PR. We consider an Extended Lapped Transform (ELT) filter bank with decimation factor $M = 4$ and overlapping factor $K' = 2$. The prototype filter for ELT is designed using the optimized butterflies angles mentioned in [21]. The input signal to the filter bank is a realization of an AR(1) process with coefficient 0.95, zero mean and unit variance. The length of the synthesis filter is 4 and the delay is 10. We vary the number of subbands ($L$) available to the synthesis filter and obtain the MSE for each channel. On analyzing the

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



TABLE II
EXPERIMENT 2: VARIATION IN MSE WITH NUMBER OF AVAILABLE SUBBANDS

| $L$ | Available analysis filters | $d$ | $J^0_{\min}$(dB) | $J^1_{\min}$(dB) | $J^2_{\min}$(dB) | $J^3_{\min}$(dB) | $J_{\min}$(dB) |
|---|---|---|---|---|---|---|---|
| 1 | $h_0(n)$ | 10 | -12.9106 | -12.9711 | -15.0069 | -14.9634 | -7.8231 |
| 1 | $h_1(n)$ | 10 | -0.2141 | -0.2360 | -0.0490 | -0.1161 | 5.8675 |
| 1 | $h_2(n)$ | 10 | -0.0826 | -0.0577 | -0.1747 | -0.0823 | 5.9215 |
| 1 | $h_3(n)$ | 10 | -0.0865 | -0.0499 | -0.0612 | -0.0241 | 5.9652 |
| 2 | $h_0(n)$, $h_1(n)$ | 10 | -16.7631 | -19.0161 | -16.7292 | -19.1518 | -11.7388 |
| 2 | $h_0(n)$, $h_2(n)$ | 10 | -14.3354 | -13.1462 | -18.6208 | -15.8517 | -9.0295 |
| 2 | $h_0(n)$, $h_3(n)$ | 10 | -13.7442 | -13.8427 | -15.6931 | -15.6224 | -8.6055 |
| 2 | $h_1(n)$, $h_2(n)$ | 10 | -0.2899 | -0.2923 | -0.2217 | -0.2139 | 5.7663 |
| 2 | $h_1(n)$, $h_3(n)$ | 10 | -0.3044 | -0.2881 | -0.1109 | -0.1410 | 5.8104 |
| 2 | $h_2(n)$, $h_3(n)$ | 10 | -0.1676 | -0.1071 | -0.2355 | -0.1088 | 5.8662 |
| 3 | $h_0(n)$, $h_1(n)$, $h_2(n)$ | 10 | -20.5040 | -20.3481 | -23.1342 | -23.1368 | -15.5518 |
| 3 | $h_0(n)$, $h_1(n)$, $h_3(n)$ | 10 | -19.2065 | -24.3326 | -17.8184 | -21.2263 | -14.0058 |
| 3 | $h_0(n)$, $h_2(n)$, $h_3(n)$ | 10 | -15.5108 | -14.0397 | -20.4628 | -16.7167 | -10.0944 |
| 3 | $h_1(n)$, $h_2(n)$, $h_3(n)$ | 10 | -0.3786 | -0.3432 | -0.2834 | -0.2416 | 5.7092 |
| 4 | $h_0(n)$, $h_1(n)$, $h_2(n)$, $h_3(n)$ | 10 | -45.1648 | -32.0844 | -145.2323 | -143.7481 | -31.8758 |
| 4 | $h_0(n)$, $h_1(n)$, $h_2(n)$, $h_3(n)$ | 11 | -32.0844 | -145.2323 | -143.7481 | -143.5253 | -32.0844 |
| 4 | $h_0(n)$, $h_1(n)$, $h_2(n)$, $h_3(n)$ | 12 | -145.2323 | -143.7481 | -143.5253 | -145.3962 | -138.3732 |
| 4 | $h_0(n)$, $h_1(n)$, $h_2(n)$, $h_3(n)$ | 13 | -143.7481 | -143.5253 | -145.3962 | -30.2194 | -30.2194 |

obtained results, as given in Table II, we can conclude that for number of subbands $L < 4$, there is a variation in the MSE with subbands chosen. However independent of the choice of subbands, if we add additional subband to the system, then we always see an improvement in the total MSE as well as in the channel-wise MSEs. This is possible since we have exploited the correlation between the input signal components while designing the matrix synthesis filter. With the number of subbands $L = 4$, the analysis bank satisfies the PR condition given by Theorem 3. We vary the delay and find out that PR is possible only with $d = 12$. This is consistent with the range of delay values provided by (89).

### C. Experiment 3

In this experiment, we design a PR synthesis bank for a non-uniform analysis bank and test it with a non-stationary input signal. We have the following setup:

$$\begin{aligned} M_0 = 2, \quad & h_0(n) = \{-0.1295, -0.12, 0.3695, 0.5018, \\ & 0.3695, -0.12, -0.1295\}, \\ M_1 = 3, \quad & h_1(n) = \{0.1308, 0.1728, -0.3775, 0.2117, \\ & 0.2117, -0.3775, 0.1728, 0.1308\}, \\ M_2 = 6, \quad & h_2(n) = \{0.0717, 0.0749, -0.1148, 0.1659, \\ & -0.2069, 0.2224, -0.2069, 0.1659, -0.1148, \\ & 0.0749, 0.0717\}, \\ M_3 = 6, \quad & h_3(n) = \{0.0881, 0.1617, -0.1686, -0.1538, \\ & 0.1752, 0.1752, -0.1538, -0.1686, \\ & 0.1617, 0.0881\}. \end{aligned} \tag{97}$$

The analysis filters are designed using Parks-McClellan algorithm. We convert the NUFB into an equivalent UFB with the help of blocking operation. The resulting UFB has decimation factor equal to 6 and the following analysis filters:

$$\begin{aligned} & g_0(n) = h_0(n), \quad g_1(n) = h_0(n-2), \quad g_2(n) = h_0(n-4), \\ & g_3(n) = h_1(n), \quad g_4(n) = h_1(n-3), \quad g_5(n) = h_2(n), \\ & g_6(n) = h_3(n). \end{aligned} \tag{98}$$

We choose the length of the synthesis filter to be 7 and delay $d = 0$. The resulting matrix $\mathbf{K}$ satisfies the PR condition (80)

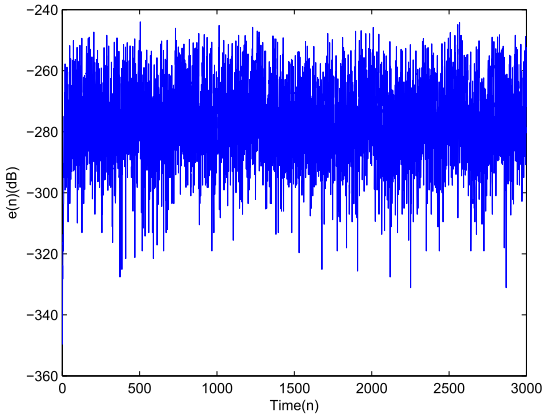



Fig. 6. Experiment 3: Error versus time curve.

and using (79), we obtain a PR solution with $\mathbf{w} = \mathbf{0}$. We apply a non-stationary signal generated using $\texttt{randn}(n)\sin(0.1n^2)$ to the filter bank and evaluate the reconstruction error as (refer Fig. 2):

$$e(n) = |\hat{u}(n) - d(n - M + 1)|. \tag{99}$$

Fig. 6 shows the resulting error versus time curve. The small error which we obtain is due to finite precision in MATLAB.

## VI. CONCLUSION

We have addressed an important problem of designing an optimal FIR synthesis bank for a given non-maximally decimated analysis bank. To design such a synthesis bank, a framework was developed to apply matrix Wiener filtering to the problem and solution is obtained for a given length of synthesis filter and delay in the reconstructed output. We experimentally showed that better MSE can be obtained in some cases by adding appropriate delay in the reconstructed output. We found that the reconstruction error keeps decreasing as more and more multirate observations are added for a signal. The matrix representation developed in this paper can be utilized to carry out time-domain analysis of any

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



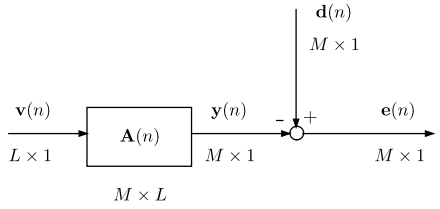



Fig. 7. A general matrix Wiener filtering setup.

multirate circuit. The infinite number of PR or non-PR synthesis bank solutions possible with our design can be optimized for a desired property in a future work.

## APPENDIX
## FIR MATRIX WIENER FILTER DERIVATION

Consider an optimum matrix filtering problem as shown in Fig. 7 where the vector signal $\mathbf{v}(n)$ is the input and the signal $\mathbf{d}(n)$ is an estimate of the desired signal. In order to obtain the Wiener solution, the input and desired signals are assumed to be jointly stationary [73]. The Wiener solution can be derived for either FIR or IIR case. Here we derive the FIR Wiener solution for a matrix filter whose constituent filters are causal and of length $P$. We begin by applying the principle of orthogonality [76] to the setup and obtain

$$\mathbf{E}[e_i(n)v_j^*(n-k)] = 0, \quad 0 \le i < M; \quad 0 \le j < L; \ 0 \le k < P. \quad (100)$$

The above equation can be written as

$$\begin{aligned}\mathbf{E}[d_i(n)v_j^*(n-k)] &= \mathbf{E}[y_i(n)v_j^*(n-k)] \\ &= \sum_{r=0}^{L-1}\sum_{l=0}^{P-1} a_{i,r}(l)\mathbf{E}[v_r(n-l)v_j^*(n-k)].\end{aligned} \quad (101)$$

As the input and desired signals are jointly stationary, we can write

$$r_{d_iv_j}(k) = \sum_{r=0}^{L-1}\sum_{l=0}^{P-1} a_{i,r}(l) r_{v_rv_j}(k-l). \quad (102)$$

The above equation can be expressed in matrix form as

$$\mathbf{r}_{d_iv_j} = \sum_{r=0}^{L-1} \mathbf{a}_{i,r}^T \mathbf{R}_{v_rv_j}, \quad 0 \le i < M; \ 0 \le j < L, \quad (103)$$

where

$$\begin{aligned}\mathbf{r}_{d_iv_j} &= \begin{bmatrix} r_{d_iv_j}(0) & r_{d_iv_j}(1) & \dots & r_{d_iv_j}(P-1)\end{bmatrix}, \\ \mathbf{a}_{i,r} &= \begin{bmatrix} a_{i,r}(0) & a_{i,r}(1) & \dots & a_{i,r}(P-1)\end{bmatrix}^T, \\ \mathbf{R}_{v_rv_j} &= \begin{bmatrix} r_{v_rv_j}(0) & r_{v_rv_j}(1) & \dots & r_{v_rv_j}(P-1) \\ r_{v_rv_j}(-1) & r_{v_rv_j}(0) & \dots & r_{v_rv_j}(P-2) \\ \vdots & \vdots & \ddots & \vdots \\ r_{v_rv_j}(-P+1) & r_{v_rv_j}(-P+2) & \dots & r_{v_rv_j}(0)\end{bmatrix}.\end{aligned} \quad (104)$$

A further compact form is

$$\mathbf{r}_{d_iv} = \mathbf{a}_i^T \mathbf{R}_{vv}, \quad 0 \le i \le M-1, \quad (105)$$

where

$$\begin{aligned}\mathbf{r}_{d_iv} &= \begin{bmatrix}\mathbf{r}_{d_iv_0} & \mathbf{r}_{d_iv_1} & \dots & \mathbf{r}_{d_iv_{L-1}}\end{bmatrix}, \\ \mathbf{a}_i &= \begin{bmatrix}\mathbf{a}_{i,0}^T & \mathbf{a}_{i,1}^T & \dots & \mathbf{a}_{i,L-1}^T\end{bmatrix}^T, \\ \mathbf{R}_{vv} &= \begin{bmatrix} \mathbf{R}_{v_0\,v_0} & \mathbf{R}_{v_0\,v_1} & \dots & \mathbf{R}_{v_0\,v_{L-1}} \\ \mathbf{R}_{v_1\,v_0} & \mathbf{R}_{v_1\,v_1} & \dots & \mathbf{R}_{v_1\,v_{L-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{v_{L-1}v_0} & \mathbf{R}_{v_{L-1}v_1} & \dots & \mathbf{R}_{v_{L-1}v_{L-1}}\end{bmatrix}.\end{aligned} \quad (106)$$

The obtained result is a system of equations containing $LP$ equations in terms of $LP$ unknowns. These are the Wiener-Hopf equations for the matrix Wiener filter. Using them, we can determine the $i$-th row of the matrix Wiener filter.

## REFERENCES

[1] V. Strela, P. N. Heller, G. Strang, P. Topiwala, and C. Heil, "The application of multiwavelet filterbanks to image processing," *IEEE Trans. Image Process.*, vol. 8, no. 4, pp. 548–563, Apr. 1999.

[2] K.-O. Cheng, N.-F. Law, and W.-C. Siu, "Multiscale directional filter bank with applications to structured and random texture retrieval," *Pattern Recognit.*, vol. 40, no. 4, pp. 1182–1194, 2007.

[3] S.-T. Hsiang and J. W. Woods, "Embedded video coding using invertible motion compensated 3-D subband/wavelet filter bank," *Signal Process., Image Commun.*, vol. 16, no. 8, pp. 705–724, 2001.

[4] B. Farhang-Boroujeny, "Filter bank spectrum sensing for cognitive radios," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1801–1811, May 2008.

[5] F. Cruz-Roldán and M. Blanco-Velasco, "Joint use of DFT filter banks and modulated transmultiplexers for multicarrier communications," *Signal Process.*, vol. 91, no. 7, pp. 1622–1635, 2011.

[6] J. J. Shynk, "Frequency-domain and multirate adaptive filtering," *IEEE Signal Process. Mag.*, vol. 9, no. 1, pp. 14–37, Jan. 1992.

[7] J. Xu, J. Yu, Y.-N. Peng, and X.-G. Xia, "Radon-Fourier transform for radar target detection, I: Generalized Doppler filter bank," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 2, pp. 1186–1202, Apr. 2011.

[8] A. Croisier, D. Esteban, and C. Galand, "Perfect channel splitting by use of interpolation/decimation/tree decomposition techniques," in *Proc. Int. Symp. Inf. Circuis Syst.*, vol. 2, Patras, Greece, 1976, pp. 443–446.

[9] P. Vaidyanathan, "Theory and design of M-channel maximally decimated quadrature mirror filters with arbitrary M, having the perfect-reconstruction property," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 4, pp. 476–492, Apr. 1987.

[10] P. Vaidyanathan, "Quadrature mirror filter banks, M-band extensions and perfect-reconstruction techniques," *IEEE ASSP Mag.*, vol. 4, no. 3, pp. 4–20, Jul. 1987.

[11] X. M. Xie, S. C. Chan, and T. I. Yuk, "Design of perfect-reconstruction nonuniform recombination filter banks with flexible rational sampling factors," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, no. 9, pp. 1965–1981, Sep. 2005.

[12] P.-Q. Hoang and P. P. Vaidyanathan, "Non-uniform multirate filter banks: Theory and design," in *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 1, May 1989, pp. 371–374.

[13] K. Nayebi, T. P. Barnwell, III, and M. J. T. Smith, "The design of perfect reconstruction nonuniform band filter banks," in *Proc. ICASSP*, vol. 3, Apr. 1991, pp. 1781–1784.

[14] V. M. Gadre and R. K. Patney, "Some novel multirate architectures for filter realization with reduced multiplicative complexity," *IEEE Trans. Signal Process.*, vol. 42, no. 9, pp. 2492–2495, Sep. 1994.

[15] J. Kovacevic and M. Vetterli, "Perfect reconstruction filter banks with rational sampling factors," *IEEE Trans. Signal Process.*, vol. 41, no. 6, pp. 2047–2066, Jun. 1993.

[16] S. Akkarakaran and P. P. Vaidyanathan, "New results and open problems on nonuniform filter-banks," in *Proc. ICASSP*, vol. 3, Mar. 1999, pp. 1501–1504.

[17] B. W.-K. Ling, C. Y.-F. Ho, K.-L. Teo, W.-C. Siu, J. Cao, and Q. Dai, "Optimal design of cosine modulated nonuniform linear phase FIR filter bank via both stretching and shifting frequency response of single prototype filter," *IEEE Trans. Signal Process.*, vol. 62, no. 10, pp. 2517–2530, May 2014.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.




[18] R. C. Nongpiur and D. J. Shpak, "Maximizing the signal-to-alias ratio in non-uniform filter banks for acoustic echo cancellation," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 59, no. 10, pp. 2315–2325, Oct. 2012.

[19] H. S. Malvar, "Lapped transforms for efficient transform/subband coding," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 6, pp. 969–978, Jun. 1990.

[20] R. L. D. Queiroz, T. Q. Nguyen, and K. R. Rao, "The GenLOT: Generalized linear-phase lapped orthogonal transform," *IEEE Trans. Signal Process.*, vol. 44, no. 3, pp. 497–507, Mar. 1996.

[21] H. S. Malvar, "Extended lapped transforms: Properties, applications, and fast algorithms," *IEEE Trans. Signal Process.*, vol. 40, no. 11, pp. 2703–2714, Nov. 1992.

[22] M. Vetterli and D. Le Gall, "Perfect reconstruction FIR filter banks: Some properties and factorizations," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 37, no. 7, pp. 1057–1071, Jul. 1989.

[23] J. Gauthier, L. Duval, and J. C. Pesquet, "Optimization of synthesis oversampled complex filter banks," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3827–3843, Oct. 2009.

[24] H. Bölcskei and F. Hlawatsch, "Oversampled cosine modulated filter banks with perfect reconstruction," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 45, no. 8, pp. 1057–1071, Aug. 1998.

[25] F. Labeau, J.-C. Chiang, M. Kieffer, P. Duhamel, L. Vandendorpe, and B. Macq, "Oversampled filter banks as error correcting codes: Theory and impulse noise correction," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4619–4630, Dec. 2005.

[26] H. Johansson and P. Löwenborg, "Flexible frequency-band reallocation networks using variable oversampled complex-modulated filter banks," *EURASIP J. Adv. Signal Process.*, vol. 2007, no. 1, 2006, Art. no. 063714.

[27] Q.-G. Liu, B. Champagne, and D. K. C. Ho, "Simple design of oversampled uniform DFT filter banks with applications to subband acoustic echo cancellation," *Signal Process.*, vol. 80, no. 5, pp. 831–847, 2000.

[28] S. Rahimi and B. Champagne, "Oversampled perfect reconstruction DFT modulated filter banks for multi-carrier transceiver systems," *Signal Process.*, vol. 93, no. 11, pp. 2942–2955, 2013.

[29] Z. Cvetković and M. Vetterli, "Oversampled filter banks," *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1245–1255, May 1998.

[30] H. Bölcskei, F. Hlawatsch, and H. G. Feichtinger, "Frame-theoretic analysis of oversampled filter banks," *IEEE Trans. Signal Process.*, vol. 46, no. 12, pp. 3256–3268, Dec. 1998.

[31] L. Gan, "Block compressed sensing of natural images," in *Proc. Int. Conf. Digit. Signal Process.*, Jul. 2007, pp. 403–406.

[32] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, Apr. 2013.

[33] M. Vetterli, "A theory of multirate filter banks," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 3, pp. 356–372, Mar. 1987.

[34] J. Rothweiler, "Polyphase quadrature filters—A new subband coding technique," in *Proc. ICASSP*, vol. 8, Apr. 1983, pp. 1280–1283.

[35] S. R. Velazquez, T. Q. Nguyen, and S. R. Broadstone, "Design of hybrid filter banks for analog/digital conversion," *IEEE Trans. Signal Process.*, vol. 46, no. 4, pp. 956–967, Apr. 1998.

[36] W. Namgoong, "A channelized digital ultrawideband receiver," *IEEE Trans. Wireless Commun.*, vol. 2, no. 3, pp. 502–510, May 2003.

[37] R. A. Elliot, M. A. Enderwitz, F. Darbari, L. H. Crockett, S. Weiss, and R. W. Stewart, "Efficient TV white space filter bank transceiver," in *Proc. 20th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2012, pp. 1079–1083.

[38] S. Weiss, S. Redif, T. Cooper, C. Liu, P. D. Baxter, and J. G. McWhirter, "Paraunitary oversampled filter bank design for channel coding," *EURASIP J. Appl. Signal Process.*, vol. 2006, Dec. 2006, Art. no. 031346.

[39] W.-R. Wu and P.-C. Chen, "Subband Kalman filtering for speech enhancement," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 45, no. 8, pp. 1072–1083, Aug. 1998.

[40] V. X. Afonso, W. J. Tompkins, T. Q. Nguyen, and S. Luo, "Multirate processing of the ECG using filter banks," in *Proc. Comput. Cardiol.*, Sep. 1996, pp. 245–248.

[41] K. A. Lee and W. S. Gan, "Improving convergence of the NLMS algorithm using constrained subband updates," *IEEE Signal Process. Lett.*, vol. 11, no. 9, pp. 736–739, Sep. 2004.

[42] L. Duval and T. Tanaka, "Denoising of seismic signals with oversampled filter banks," in *Proc. ICASSP*, vol. 6, Apr. 2003, p. VI-189.

[43] O. S. Jahromi and P. Aarabi, "Theory and design of multirate sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1739–1753, May 2005.

[44] J. W. Scrofani and C. W. Therrien, "A stochastic multirate signal processing approach to high-resolution signal reconstruction," in *Proc. IEEE ICASSP*, vol. 4, Mar. 2005, pp. 561–564.

[45] C. W. Therrien, "Multirate linear prediction," in *Proc. 13th Eur. Signal Process. Conf.*, Sep. 2005, pp. 1–4.

[46] K. Gedalyahu and Y. C. Eldar, "Time-delay estimation from low-rate samples: A union of subspaces approach," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3017–3031, Jun. 2010.

[47] O. Jahromi and P. Aarabi, "Time delay estimation and signal reconstruction using multi-rate measurements," in *Proc. ICME*, vol. 2, Jul. 2003, p. II-597.

[48] Y. Wang, G. Leus, and A. Pandharipande, "Direction estimation using compressive sampling array processing," in *Proc. IEEE/SP 15th Workshop Stat. Signal Process.*, Aug. 2009, pp. 626–629.

[49] R. J. Kuchler and C. W. Therrien, "Optimal filtering with multirate observations," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, vol. 1, Nov. 2003, pp. 1208–1212.

[50] P. P. Vaidyanathan and T. Chen, "Statistically optimal synthesis banks for subband coders," in *Proc. 28th Asilomar Conf. Signals, Syst. Comput.*, vol. 2, Oct./Nov. 1994, pp. 986–990.

[51] B.-S. Chen, C.-W. Lin, and Y.-L. Chen, "Optimal signal reconstruction in noisy filter bank systems: Multirate Kalman synthesis filtering approach," *IEEE Trans. Signal Process.*, vol. 43, no. 11, pp. 2496–2504, Nov. 1995.

[52] Y.-M. Cheng and B. Chen, "Nonuniform filter bank design with noises," *IEEE Trans. Signal Process.*, vol. 46, no. 9, pp. 2326–2344, Sep. 1998.

[53] J. Q. Ni, K. L. Ho, K. W. Tse, J. S. Ni, and M. H. Shen, "Multirate Kalman filtering approach for optimal two-dimensional signal reconstruction from noisy subband systems," in *Proc. Int. Conf. Image Process.*, vol. 1, Oct. 1997, pp. 157–160.

[54] H. Vikalo, B. Hassibi, and T. Kailath, "On $H^\infty$ optimal signal reconstruction in noisy filter banks," in *Proc. IEEE ICASSP*, vol. 3, Mar. 1999, pp. 1493–1496.

[55] H. Vikalo, B. Hassibi, and T. Kailath, "Mixed $H^2/H^\infty$ optimal signal reconstruction in noisy filter banks," in *Proc. IEEE ICASSP*, vol. 1, Jun. 2000, pp. 500–503.

[56] S. S. Bora, R. Dhuli, and B. Lall, "Mixed $H^2/H^\infty$ design of non-uniform filter banks using an efficient realisation of synthesis stage," in *Proc. Annu. IEEE India Conf. (INDICON)*, Dec. 2010, pp. 1–4.

[57] C. W. Therrien and A. H. Hawes, "Least squares optimal filtering with multirate observations," in *Proc. Conf. Rec. 36th Asilomar Conf. Signals, Syst. Comput.*, vol. 2, Nov. 2002, pp. 1782–1786.

[58] L. Chai, Y. Sheng, and J. Zhang, "Design of synthesis filter banks for the structural similarity index," in *Proc. 20th ICIP*, Sep. 2013, pp. 849–853.

[59] E. M. Ekşioˇğlu, A. K. Tanc, and A. H. Kayran, "A compressive sensing framework for multirate signal estimation," in *Proc. 10th ISSPA*, May 2010, pp. 716–719.

[60] A. H. Hawes and C. W. Therrien, "LMS adaptive filtering with multirate observations," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, vol. 1, Nov. 2003, pp. 567–570.

[61] A. N. Delopoulos and S. D. Kollias, "Optimal filter banks for signal reconstruction from noisy subband components," *IEEE Trans. Signal Process.*, vol. 44, no. 2, pp. 212–224, Feb. 1996.

[62] A. K. Tanc, E. M. Eksioglu, and A. H. Kayran, "Adaptive multirate signal estimation with lattice orthogonalization," in *Proc. 16th ICECS*, Dec. 2009, pp. 117–119.

[63] O. S. Jahromi, B. A. Francis, and R. H. Kwong, "Spectrum estimation using multirate observations," *IEEE Trans. Signal Process.*, vol. 52, no. 7, pp. 1878–1890, Jul. 2004.

[64] A. K. Tanc and A. H. Kayran, "Iterative maximum entropy power spectrum estimation for multirate systems," *AEU-Int. J. Electron. Commun.*, vol. 64, no. 2, pp. 93–98, 2010.

[65] A. N. Amini, M. S. Takyar, and T. T. Georgiou, "A homotopy approach for multirate spectrum estimation," in *Proc. IEEE ICASSP*, vol. 3, May 2006, pp. 532–535.

[66] O. Jahromi, R. H. Kwong, and B. A. Francis, "Information theory of multirate systems," in *Proc. IEEE ISIT*, Washington, DC, USA, Jun. 2001, p. 188.

[67] J. G. Van Antwerp and R. D. Braatz, "A tutorial on linear and bilinear matrix inequalities," *J. Process Control*, vol. 10, no. 4, pp. 363–385, 2000.

[68] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



[69] P. P. Vaidyanathan and S. K. Mitra, "Polyphase networks, block digital filtering, LPTV systems, and alias-free QMF banks: A unified approach based on pseudocirculants," *IEEE Trans. Acoust., Speech Signal Process.*, vol. ASSP-36, no. 3, pp. 381–391, Mar. 1988.
[70] D. A. Koupatsiaris, "Analysis of multirate random signals," M.S. thesis, Dept. Elect. Comput. Eng., Naval Postgraduate School, Monterey, CA, USA, Dec. 2000.
[71] R. Dhuli and B. Lall, "Cyclostationary signals in multirate linear systems," in *Proc. Nat. Conf. Commun. (NCC)*, Jan. 2010, pp. 1–5.
[72] S. Patel, R. Dhuli, and B. Lall, "Design and analysis of matrix Wiener synthesis filter for multirate filter bank," *Signal Process.*, vol. 102, pp. 256–264, Sep. 2014.
[73] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. Hoboken, NJ, USA: Wiley, 2002.
[74] J. Barata and M. Hussein, "The Moore–penrose pseudoinverse: A tutorial review of the theory," *Brazilian J. Phys.*, vol. 42, nos. 1–2, pp. 146–165, 2012.
[75] K.-R. Koch, *Parameter Estimation and Hypothesis Testing in Linear Models*. New York, NY, USA: Springer-Verlag, 1999.
[76] S. S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 1996.
[77] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 2000.
[78] C. A. Floudas and V. Visweswaran, "Quadratic optimization," *Handbook of Global Optimization*. Dordrecht, The Netherlands: Kluwer, 1995.
[79] F. Labeau, "Synthesis filters design for coding gain in oversampled filter banks," *IEEE Signal Process. Lett.*, vol. 12, no. 10, pp. 697–700, Oct. 2005.



**Sandeep Patel** received the B.Tech. degree in electronics and communications engineering from IIT Roorkee, Roorkee, in 2007, and the MS(R) degree from IIT Delhi, New Delhi, in 2013, where he is currently pursuing the Ph.D. degree from the Bharti School of Telecom Technology and Management. He joined Nvidia Graphics and worked on video technologies until 2008. From 2008 to 2010, he was with the Flash Media Server Team, Adobe Systems.

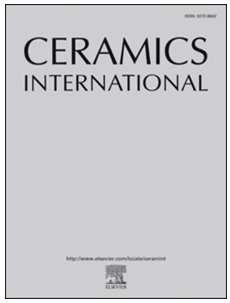


**Ravindra Dhuli** received the Ph.D. degree in signal processing from the Department of Electrical Engineering, IIT Delhi, in 2010. He is currently an Associate Professor with the Department of Electronics and Communications Engineering, Vellore Institute of Technology, India. His research interests include multirate signal processing, statistical signal processing, image processing, and mathematical modeling.



**Brejesh Lall** (M'05) received the B.E. and M.E. degrees in electronics and communications engineering from the Delhi College of Engineering, Delhi University, in 1991 and 1992, respectively, and the Ph.D. degree in multirate signal processing from IIT Delhi, New Delhi, in 1999. From 1997 to 2005, he was with the Digital Signal Processing Group, Hughes Software Systems. Since 2005, he has been with the Faculty of the Department of Electrical Engineering, IIT Delhi, New Delhi, where he is currently a Professor and the Head of the Bharti School of Telecom Technology and Management and also the Coordinator of the Ericsson 5G Center of Excellence. He has more than 150 papers in international journals and refereed conferences. His research interests are in signal processing with application to image processing and communication systems.

# Accepted Manuscript




Color tunable photoluminescence properties in $Eu^{3+}$ doped calcium bismuth vanadate phosphors for luminescent devices

Harpreet Kaur, M. Jayasimhadri

PII: S0272-8842(19)31132-0

DOI: https://doi.org/10.1016/j.ceramint.2019.05.034

Reference: CERI 21486

To appear in: *Ceramics International*

Received Date: 18 April 2019

Revised Date: 4 May 2019

Accepted Date: 5 May 2019

Please cite this article as: H. Kaur, M. Jayasimhadri, Color tunable photoluminescence properties in $Eu^{3+}$ doped calcium bismuth vanadate phosphors for luminescent devices, *Ceramics International* (2019), doi: https://doi.org/10.1016/j.ceramint.2019.05.034.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

# Color tunable photoluminescence properties in $Eu^{3+}$ doped calcium bismuth vanadate phosphors for luminescent devices

**Harpreet Kaur, M. Jayasimhadri***

*Luminescent Materials Research Lab, Department of Applied Physics, Delhi Technological University, Bawana Road, New Delhi-110 042, India*

## ABSTRACT

Single phase $Eu^{3+}$ activated calcium bismuth vanadate ($CaBiVO_5$) phosphors have been successfully synthesized using solid state reaction method. X-ray diffraction (XRD) analysis confirms the pure phase formation and scanning electron microscope (SEM) micrographs exhibit inhomogeneous particle formation with irregular morphology of $Eu^{3+}$ doped $CaBiVO_5$ (CBV) phosphor. The photoluminescence excitation (PLE) spectrum indicates significant absorption in the ultraviolet (UV) and near ultraviolet (n-UV) spectral region for un-doped CBV sample by monitored at 613 nm wavelength, whereas $Eu^{3+}$ doped CBV phosphors reveal various sharp absorption bands in n-UV and blue region along with host excitation bands. Trivalent europium activated CBV phosphors under 342 nm excitation exhibit dominant red emission peak at 613 nm wavelength accompanied by weak broadband originating from $VO_4$ groups, whereas the phosphors under 464 nm excitation, exhibit similar emission peaks with most intense one centered at 613 nm without any host emission bands. The energy transfer mechanisms and the probable cause for concentration quenching beyond 4.0 mol% of $Eu^{3+}$ ions concentration have been discussed in detail. The CIE chromaticity coordinates for the optimized phosphor, (0.551, 0.398) and (0.639, 0.358), situated in the reddish-orange and red region under 342 and 464 nm excitations, respectively. The CIE coordinates calculated based on the emission spectra measured under 464 nm excitation are close to the commercial phosphor $Y_2O_2S$: $Eu^{3+}$ (x= 0.622 and y= 0.351). All the above mentioned results support the utilization of $Eu^{3+}$ doped CBV phosphor as a potential red emitting component for luminescent devices.

*Keywords*: Phosphors; Photoluminescence; Color tunabilty; Energy transfer; w-LEDs

ACCEPTED MANUSCRIPT

*Corresponding author: ***E-mail:*** jayaphysics@yahoo.com, Tel: +91-9013553360

## 1. Introduction

In recent years, the escalating demand for the annual energy consumption and other environmental issues force the population to shift towards the alternatives for conventional lighting sources [1, 2]. In this regard, white light emitting diode (w-LED) is currently serving as a fourth generation lighting device since it has remarkably reduced annual energy consumption [3, 4]. Nowadays, w-LEDs are acting potentially as a substitute for traditional incandescent and fluorescent lamps on account of their unprecedented luminous efficiency, low energy consumption, less emission of harmful gases, high brightness, compact design, fast switching and long working lifetime [5–8]. Predominantly, there are two approaches which have been availed to fabricate w-LEDs. The first one includes using phosphor free RGB-LEDs in which monochromatic light from red, green and blue LEDs are mixed directly to get white light emission [9, 10]. These integrated LEDs need to have different driving currents for the individual LED to produce constant white light emission which is difficult to obtain and made it less accessible. Another approach includes the accomplishment of white light emission via combination of blue light emitting InGaN chip and yellow phosphor (YAG: $Ce^{3+}$), which is presently used in most of the commercial w-LEDs [11, 12]. Despite of its several advantages over RGB-LEDs like low cost, simplicity in fabrication and high brightness, such w-LEDs have some drawbacks like low color rendering index (CRI < 70), high correlated color temperature (CCT >7765 K), halo effect and low efficiency due to lack of the red component [13]. Therefore, researchers over the globe proposed alternative approaches to deal with this obnoxious problem such as “blue LED coated with green and red phosphors” and “n-UV LED coated with an appropriate mixture of RGB phosphor” [14, 15]. Apparently, both the above mentioned approaches obligate the synthesis of a promising red emitting phosphor.

Trivalent rare earth (RE) ion like europium ($Eu^{3+}$) doped appropriate host lattice may be utilized as an efficient red phosphor due to its characteristic f-f transitions [16]. Owing to their *f-f* intra-configurational characteristics, $Eu^{3+}$ doped phosphors have also been gaining a noteworthy attention for applications in luminescent devices such as plasma panel displays (PDPs), lasers, photovoltaic cells as down-converters, and backlight displays [17, 18]. Currently, commercially available red phosphor ($Y_2O_2S$: $Eu^{3+}$) gives excellent red emission but lacks on account of chemical and thermal stability in comparison with green (ZnS: $Cu^{+}$, $Al^{3+}$) and blue phosphor ($BaMgAl_{10}O_{17}$: $Eu^{2+}$) [19, 20]. Moreover, the sulphide based phosphor is not environment friendly phosphor that necessitates exploring new environmental friendly red phosphor with excellent thermal and chemical stability.

Nowadays, inorganic oxides are gaining excellent attention by virtue of their synthesis procedure, non-toxic in nature, excellent thermal and chemical stability [21–23]. Amongst various inorganic oxides, vanadates with binary, ternary and quaternary compositions are being explored to accomplish their utility as a potential host lattice for luminescent materials [24–26]. Vanadates exhibit excellent properties such as wide optical band gap, significant absorption in UV/n-UV region, excellent thermal, optical and chemical stability. In addition, the vanadate materials require low temperatures for the preparation and are cost effective [27–30].

In the present work, an investigation has been made on a new calcium bismuth vanadate ($CaBiVO_5$) host lattice doped with different concentrations of $Eu^{3+}$ ions synthesized via solid-state reaction method to harvest the above-mentioned precedence for vanadate based hosts. The prepared phosphors exhibit excitation and dopant concentration dependent tunable emission and intense red emission with high purity. Therefore, structural, morphological and luminescent properties of the $Eu^{3+}$ doped CBV phosphors have been studied extensively to delve into the possibility of utilising it as a new red phosphor in luminescent devices.

## 2. Experimental procedures

### *2.1. Sample preparation*

A series of $CaBi_{1-x}VO_5$: $xEu^{3+}$ (1.0 mol% ≤ x ≤ 6.0 mol%) phosphors with different doping concentration of $Eu^{3+}$ ions have been synthesized using conventional solid state reaction technique. Starting materials, $CaCO_3$ (Fisher scientific, 99.9%), $NH_4VO_3$ (Fisher scientific, 99.9%), $Bi_2O_3$ (Fisher scientific, 99.0 %), and $Eu_2O_3$ (CDH, 99.99%) have been taken into stoichiometric ratio. All the chemicals have been dispersed in acetone and ground thoroughly for an hour in an agate mortar. The mixture has been taken in an alumina crucible and calcined at 950 °C for 3 hours in a programmable muffle furnace. Then, naturally cooled samples to room temperature (RT) have been powdered finely for further characterizations.

### *2.2. Sample characterizations*

The CBV sample has been characterised for thermo-gravimetric studies using Star-e system under $N_2$ atmosphere and heating rate of 10 °C/min. The XRD patterns have been recorded on Rizaku Miniflex 300 XRD analyzer with nickel-filtered Cu-Kα radiation (λ = 1.5406 Å) in the angular range 20° ≤ θ ≤ 60°. The accelerating voltage and tube current of the XRD machine have been set at 30.0 kV and 15 mA, respectively. SEM micrographs have been collected using the JSM-7100F SEM machine to reveal the morphological structure of as-synthesized samples. The PLE and emission spectral measurements have been recorded using Shimadzu spectrofluorophotometer (Model No. RF-5301 PC) provided with xenon arc lamp for excitation. Also, time resolved photoluminescence (TR-PL) measurements have been carried out using Edinburgh FLSP 900 spectrofluorophotometer with a spectral resolution of 0.1 nm, where xenon lamp acts as a source of excitation, to estimate the lifetime of the as-synthesized phosphors.

## 3. Result and Discussion

### *3.1. Thermogravimetric studies*

Fig. 1 displays the thermogravimetric analysis (TGA) of the $CaBiVO_5$ precursors ($CaCO_3$, $NH_4VO_3$, and $Bi_2O_3$) taken in a stoichiometric ratio and heated in the temperature range of 90-900 °C. The TGA curve depicts that weight loss during the formation of $CaBiVO_5$ sample proceeds through two distinct steps. The initial weight loss of 5.05 % occurs in the range of 90-234 °C, has been attributed to the evaporation of surface moisture and decomposition of residual organic species such as acetone used for homogeneous mixing. Then, a significant weight loss of 10.866 % takes place with an increase in temperature up to 726 °C, is ascribed to the degradation of organic material from the synthesized sample, which could possibly be due to the elimination of $CO_2$, $NH_3$, and $H_2O$ molecules. As the temperature is further increased, a solid state reaction occurs amongst the chemicals characterized by a nominal weight loss. The mixture has been sintered at various temperatures beyond 750 °C to identify the pure phase forming temperature for the $CaBiVO_5$ phosphors.

*3.2. Structural studies*

Fig. 2(a) represents the XRD patterns for $CaBiVO_5$ (CBV) powders sintered at various temperatures by taking reference of TGA results. It is evident from Fig. 2(a) that the phase formation begins at 750 °C with the presence of some impurity peaks and mismatch of relative peak intensities. The impurity peaks gradually diminish with increasing the calcination temperature and pure phase of CBV sample obtained at 950 °C. All the peaks in the diffraction pattern of CBV sample sintered at 950 °C for 3 h are well matched to the standard data of $CaBiVO_5$ (JCPDS card no. 81-1775), which confirms the monophasic formation of $CaBiVO_5$ as shown in Fig. 2 (a). As-synthesized phosphors crystallize in an orthorhombic unit cell belong to space group $D_{2h}^{15}$ −Pbca with lattice parameters a = 11.2022 Å, b = 5.4283 Å, c = 15.5605 Å and cell volume 946.22 $Å^3$. Based on the observed XRD patterns, the CBV structure has been refined by adopting the Rietveld powder diffraction profile fitting technique employing FullPROF suite software. For the Rietveld refinement of the as-prepared sample, we have utilized the initial values of lattice parameters, space group

and Wyckoff positions of the atoms reported by A. Barth et. al. [31]. A pseudo-Voigt function, to simulate the peak shape and linear interpolation between a set background points with refinable heights parameters in the software have been adopted. Fig. 2(b) represents the refined and experimental XRD pattern along with an acceptable deviation from refined data for the CBV samples. The refinement results as listed in Table 1, confirm the pure phase formation of CBV sample as the lattice parameter values do not vary much from the initial values and also from residual factors ($R_{wp}$, $R_{p,}$ and $\chi^2$) values. The crystal structure of CBV is characterized by layers of the edge linked $CaO_7$ polyhedra, isolated $VO_4$ tetrahedra and an asymmetric surrounding of a $Bi^{3+}$ ion by oxygen atoms [31]. The absence of any diffraction peaks corresponding to $Eu_2O_3$ and complete agreement with the standard XRD data also supports the successful doping as depicted in Fig. 2(c). The values of ionic radii for $Ca^{2+}$ and $Bi^{3+}$ ions are 1.06 Å and 0.96 Å. respectively. So, when the crystal structure is doped with $Eu^{3+}$ ions ($R_{Eu3+}$ = 0.947 Å), it can favorably replace $Bi^{3+}$ ions in the lattice due to comparable radii and same valence state. Upon keen observation, it is found that the diffraction peaks of CBV: $xEu^{3+}$ phosphors shift towards higher angle with increasing $Eu^{3+}$ concentration in comparison with standard data. The most intense peak present in XRD pattern is separately plotted and shown on the right side of Fig. 2(c) to reveal the peak shift clearly. It may be explained that the substitution of smaller $Eu^{3+}$ ions in larger $Bi^{3+}$ ions leads to the decrease in interatomic spacing and finally the contraction of the unit cell. It will cause strain in the host lattice and eventually yields to shifting of diffraction peaks towards larger angles [32]. Further, the crystallite size (D) and strain (ε) for CBV :$xEu^{3+}$ (where x = 0.0, 2.0, 4.0 and 6.0 mol%) have been calculated employing Williamson Hall (W-H) equation, which is given by [33]

$$\beta \cos\theta/\lambda = (k/D) + 4\varepsilon \sin\theta/\lambda \quad (1)$$

where λ is the wavelength of Cu-Kα radiation used in XRD, β is full width half maximum (FWHM), k is the shape factor, and θ is the diffraction angle in radians of the observed peak.

Thus, the computed average crystallite size and strain values for $CaBi_{1-x}VO_5$: $xEu^{3+}$ (x = 0.0, 2.0, 4.0 and 6.0 mol%) from W-H plot (shown in Fig. 2(d)) are 38.02, 37.82, 32.32, and 28.85 nm and 0.00028, 0.00033, 0.00035 and 0.00047 respectively. The decrease in crystallite size with increasing $Eu^{3+}$ ion concentration ascribed to the increase in strain owing to the replacement of $Bi^{3+}$ ions by $Eu^{3+}$ ions of comparatively smaller radii [34, 35]. The obtained average crystallite size values are in good agreement with that of the values calculated using Debye-Scherrer formula ($D = k\lambda/\beta\cos\theta$) and the values for $CaBi_{1-x}VO_5$ :$xEu^{3+}$ (x = 0.0, 2.0, 4.0 and 6.0 mol%) are found to be 45.10, 39.24, 35.78, and 31.05 nm respectively [36].

*3.3. SEM analysis*

SEM images recorded for $CaBi_{1-x}VO_5$: $xEu^{3+}$ (x= 4.0 mol%) at different resolutions have been displayed in the Fig. 3 (a & b). The SEM micrographs reveal the microcrystalline structure formation due to agglomeration of grains obtained by sintering the sample at high temperatures. The particle size is irregular and ranging from 1 - 20 µm. Particle size in micrometer range is adequate to utilize CBV phosphors in w-LEDs, since other commercial phosphor's particle size currently available in the market are also of few micrometers [37].

*3.4. Luminescent properties*

The photoluminescence excitation (PLE) and emission spectra of undoped CBV sample have been shown in Fig. 4. The PLE spectral measurements have been carried out by monitoring the emission at 542 nm wavelength. The excitation band can be resolved into two broad peaks located at 277 and 342 nm corresponding to $^1A_1 \rightarrow {}^1T_2$ and $^1A_1 \rightarrow {}^1T_1$ transitions of $[VO_4]^{3-}$ groups, respectively [38]. The emission spectrum for CBV excited under 342 nm wavelength, reveals a weak broad emission band ranging from 400 nm to 680 nm corresponding to $^3T_1 \rightarrow {}^1A_1$ transition of $[VO_4]^{3-}$ groups [4]. In comparison to various other vanadate based lattices, the compound under present investigation shows weak emission broadband ranging from 400-680 nm centered at 542 nm [39]. Z. Pei et al. have studied host

luminescence for $CaBiVO_5$ and observed strong luminescence at liquid helium temperature in comparison to the luminescence at RT due to thermal quenching [40]. Weak luminescence observed in $CaBiVO_5$ due to the presence of small amount of Ca/Bi disorder in $CaBiVO_5$ similar to $Ca_4BiV_3O_{13}$ and $Sr_2Bi_3V_3O_{14}$.

Fig. 5 represents the PLE spectra of $Eu^{3+}$ doped CBV phosphors at various doping concentrations monitored at intense $Eu^{3+}$ emission wavelength ($\lambda_{em}$= 613 nm). It exhibits intense broadband at 342 nm similar to the undoped sample. This intense band might be due to an overlap of $O^{2-} \rightarrow Eu^{3+}$ charge transfer (CT) transitions and transitions of the $VO_4$ group. Along with the broadband in the PLE spectra, five sharp excitation bands observed at 382, 393, 415, 464 and 488 nm. These bands have been attributed to the transitions of $Eu^{3+}$ ion from its ground state $^7F_0$ to the different excited states $^5L_7$, $^5L_6$, $^5D_3$, $^5D_2$ and $^5D_1$, respectively [41]. These results circumstantiate that the $Eu^{3+}$ doped CBV phosphors can be efficiently excited by UV, n-UV and blue light, which further support the potentiality of CBV phosphors. Amongst these excitation bands, the broad CT band centered at 342 nm is most intense, and another located at 464 nm wavelength also exists with an appreciable intensity. Fig. 6 shows the comparison of emission spectra for 4 mol% $Eu^{3+}$ ions doped CBV phosphors excited under 342 and 464 nm wavelengths. The comparison of emission spectra communicates that the $Eu^{3+}$ doped CBV phosphors exhibit characteristic 4f-4f transitions of $Eu^{3+}$ under 342 nm and 464 nm excitations. Moreover, the CBV phosphor gives maximum intensity for 4f-4f transitions and weak broad band host emission under 342 nm excitation.

Fig. 7(a) represents the emission spectra for varying doping concentrations of $Eu^{3+}$ ions in CBV phosphors excited under 342 nm wavelength. The emission spectra as shown in the Fig. 7(b) is normalized at the most intense peak of 613 nm to reveal energy transfer from host to $Eu^{3+}$ ions. It exists in two parts, one is weak broad emission band similar to undoped sample has been observed due to the transition of $[VO_4]^{3-}$ groups and second part is that sharp

bands located at 580, 593, 613, 653 and 704 nm wavelengths attributed to the transitions between $Eu^{3+}$ ion energy levels viz $^5D_0 \rightarrow {}^7F_J$ (where J = 0, 1, 2, 3, 4), respectively [41]. Along with these emission bands, two more peaks located at 537 and 558 nm wavelengths overlapped with host emission are corresponding to the bands originated from $^5D_1 \rightarrow {}^7F_J$ transitions (where J = 1 and 2 respectively) of $Eu^{3+}$ ions [42]. The effect of $Eu^{3+}$ ions concentration on intensity of host emission as well as $Eu^{3+}$ characteristic emission ($^5D_0 \rightarrow {}^7F_2$) is shown as an inset of Fig. 7(a). It can be observed from the inset that with increasing the concentration of $Eu^{3+}$ ions, emission intensity of the transition $^5D_0 \rightarrow {}^7F_2$ enhances appreciably up to 4 mol% $Eu^{3+}$ ions concentration and becomes dominant while host emission remains almost unaffected. It indicates that energy transfer from $[VO_4]^{3-}$ groups to $Eu^{3+}$ ions may occur under 342 nm excitation wavelength. When the phosphors are excited at this wavelength, firstly $[VO_4]^{3-}$ groups may relax non radiatively to $^1T_1$ or $^3T_1$ state, and eventually, energy transfer from $[VO_4]^{3-}$ groups to the $^5D_0$ energy level of $Eu^{3+}$ is expected to happen. However, if and only if the aforementioned process is involved, emission corresponding to $Eu^{3+}$ ions should not enhance significantly with increasing concentration. Therefore, it can be concluded that the intense excitation band centred at 342 nm wavelength is an overlap of CT band ($O^{2-} \rightarrow Eu^{3+}$) and transition originated from $VO_4$ groups [43].

Fig. 8(a) represents the emission spectra for CBV: $xEu^{3+}$ (x = 1.0, 2.0, 3.0, 4.0, 5.0 and 6.0 mol%) phosphors excited under 464 nm wavelength. The emission spectra of all $Eu^{3+}$ doped CBV phosphors represent only peaks corresponding to the transitions between 4f-4f energy levels of $Eu^{3+}$ ions which have been previously discussed. The emission peak located at 593 nm in orange (O) region corresponds to the magnetic dipole (MD) allowed transition ($^5D_0 \rightarrow {}^7F_1$) and the other located at 613 nm in red (R) region corresponds to the forced electric dipole (ED) allowed transition ($^5D_0 \rightarrow {}^7F_2$). The former MD allowed transition follows the selection rule $\Delta J = 1$, while latter follows the selection rule $|\Delta L| \leq 2$; $|\Delta J| \leq 2$, which is a hypersensitive transition and highly influenced by the crystal field environment

[44]. The dominance of emission energy distribution in the red region conveys that $Eu^{3+}$ ions have been located at low symmetry sites without an inversion center [45]. The inversion symmetry distortion degree of the environment in the vicinity of $Eu^{3+}$ ions can be represented by the integrated intensity ratio value of ED (R) and MD (O) transitions, hence termed as asymmetric ratio. The asymmetric ratio (R/O) has been calculated for optimized (4 mol%) $Eu^{3+}$ doped CBV phosphor under 464 nm excitation, which comes out to be 6.67. The high value of asymmetric ratio for 4.0 mol% of $Eu^{3+}$ doped CBV phosphor suggests high purity of red color emission. The emission intensity of the peaks corresponding to $Eu^{3+}$ ions at $\lambda_{ex}$ = 342 and 464 nm and that of host at $\lambda_{ex}$ = 342 nm with varying $Eu^{3+}$ ions concentration is shown as Fig. 8(b). It can be easily seen that the emission intensity of the characteristic peaks of $Eu^{3+}$ rises with the increase in $Eu^{3+}$ ions concentration up to 4.0 mol%, after which the emission intensity of $Eu^{3+}$ ions is quenched. It provides that the optimized $Eu^{3+}$ ions concentration in CBV phosphors comes out to be 4.0 mol%. The decrease in the luminescence intensity with increasing concentration of $Eu^{3+}$ ions beyond 4.0 mol% is due to non-radiative energy transfer among $Eu^{3+}$ ions and is termed as concentration quenching. Blasse et al. propounded that the fluorescence quenching at higher concentration of activator occurs as the distance between luminescent centers decreases with an increase in activator ion concentration. It leads to possible non-radiative energy transfer among activator ions instead of emitting the excitation energy radiatively [46]. These non-radiative energy transfer may be explained in terms of either resonance energy transfer (RET) or cross relaxation among $Eu^{3+}$-$Eu^{3+}$ ions [47]. The non-radiative energy transfer mainly depends on the critical distance between $Eu^{3+}$-$Eu^{3+}$ ions, which can be calculated according to the formula given by Blasse [48]:

$$R_c \sim 2\left[\frac{3V}{4\pi X_c N}\right]^{\frac{1}{3}} \quad (2)$$

where, $R_c$ represents the critical distance, V is the volume of unit cell of host lattice, $X_c$ is the critical concentration beyond which quenching occurs and N is the effective number of cations in a unit cell. The critical distance for the CBV host lattice is found to be 17.81 Å by considering the parameters V = 0.94622 $nm^3$, $X_c$ = 0.04 and N = 8. According to Van Uitert, energy transfer through exchange interaction generally takes place when $R_c$ is typically close to 5 Å [49]. In the present system, the non-radiative energy transfer takes places through multipolar interaction, since the critical distance is far from the value (5Å). The nature of the multipolar interaction can be determined using Dexter's theory using the following equation [50]

$$\log\left(\frac{I}{x}\right) = -\frac{Q}{3}logx + A \quad (3)$$

which represents the relationship between luminescence intensity (I) and concentration of dopant ions (x), where A is a constant and Q represents the type of multipolar interaction. The different values of Q such as Q=6, 8 and 10 represent electric dipole-dipole, dipole-quadruple and quadruple-quadruple interaction respectively. A graph of log (I/x) vs. log x for concentrations higher than 1.0 mol% is shown in Fig. 9, based on emission spectra of $CaBi_{1-x}VO_5$: $Eu^{3+}$ excited under 464 nm wavelength. The slope of the graph is found to be 1.834 and the calculated value of Q is 5.50, which is close to 6, confirms the energy transfer process in the present case occurs due to the electric dipole-dipole interaction [51].

The possible transitions between energy levels of $Eu^{3+}$ ions and partial energy transfer between $VO_4$ and $Eu^{3+}$ ions in a schematic manner have been demonstrated in Fig. 10. In the energy level scheme of $VO_4$ groups, $^1A_1$ represents the ground state while $^1T_1$, $^1T_2$, and $^3T_1$ represent the excited states. The excitation band of CBV host can be resolved into two broad bands assigned to partly spin allowed $^1A_1 \rightarrow {}^1T_2$ and $^1A_1 \rightarrow {}^1T_1$ transitions [52]. When CBV host is doped with $Eu^{3+}$ ions, the partial energy transfer may take place either from $^1T_1$ level of $VO_4$ group to $Eu^{3+}$ CT band or from $^3T_1$ to $^5D_0$ energy level of $Eu^{3+}$ ion. These energy

transfers are shown as dashed horizontal arrows. The $Eu^{3+}$ ions may excite by absorbing energy to CTB or various higher excited states such as $^5L_7$, $^5L_6$, $^5D_3$, $^5D_2$ and $^5D_1$ of $Eu^{3+}$ ions as shown in upward arrows in Fig. 10. From the excited states ($^1T_2$ and $^1T_1$), the $VO_4$ groups can relax non-radiatively to $^3T_1$ and then radiatively to ground level ($^1A_1$) finally. Likewise, $Eu^{3+}$ ions may relax non-radiatively from CTB and other excited states to $^5D_0$ or $^5D_1$ levels, followed by de-excitation of $Eu^{3+}$ ions through radiative transitions from $^5D_0$ or $^5D_1$ level to ground state multiplet ($^7F_J$, where J = 0, 1, 2, 3 and 4 ) [24]. The non-radiative and radiative transitions are represented as curly and straight downward arrows respectively in Fig. 10.

*3.5. CIE chromaticity coordinates*

Using color matching functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$ defined by Commission International de I' Eclairage (CIE) 1931, the chromaticity coordinates (x, y) can be calculated from the emission spectral data of the phosphor [53]. The coordinates (x, y) marked on a chromaticity diagram specify the color of the light source. The calculated CIE chromaticity coordinates for undoped and $Eu^{3+}$ doped CBV phosphors excited under 342 and 464 nm wavelengths are listed in Table 2. The CIE chromaticity coordinates for the $CaBi_{1-x}VO_5$: $xEu^{3+}$, x=0.0 mol% (undoped) and x=1.0, 2.0, 3.0, 4.0, 5.0 and 6.0 mol% (doped) samples excited under 342 nm wavelength are shown in Fig. 11. As seen from the CIE diagram, chromaticity coordinates shift from yellowish orange to reddish-orange region under 342 nm excitation with increasing $Eu^{3+}$ concentration. Hence, color tunability in $Eu^{3+}$ doped CBV phosphors under 342 nm excitation can be obtained by appropriately adjusting $Eu^{3+}$ ion concentration. The CIE coordinates for all the $Eu^{3+}$ doped CBV phosphors under 464 nm excitation lie in the red region and are in proximity of each other as shown in Fig. 11. All the CIE coordinates calculated for emission spectra recorded under 464 nm excitation wavelength are found close to commercial phosphor $Y_2O_2S$: $Eu^{3+}$ (x= 0.622 and y= 0.351) and the

coordinates designated by National Television Standard Committee (NTSC) (x = 0.670, y = 0.330). The CCT values can be calculated by employing McCamy's polynomial formula [54]:

$$CCT = -449n^3 + 3525n^2 - 6823.3n + 5520.33$$

where, $n = \frac{x-x_e}{y-y_e}$; $x_e$ = 0.332, $y_e$ = 0.186 are the coordinates for epicentre of convergence and *n* represents inverse slope line. The calculated CCT values are listed in Table. 2 for emission spectra recorded at $\lambda_{ex}$ = 342 and 464 nm for 1.0, 2.0, 3.0, 4.0, 5.0 and 6.0 mol% of $Eu^{3+}$ doped CBV phosphor. The purity of the red emission can be calculated by the following formula [55]:

$$color\ purity = \frac{\sqrt{(x-x_{ee})^2+(y-y_{ee})^2}}{\sqrt{(x_d-x_{ee})^2+(y_d-y_{ee})^2}} \quad (4)$$

where, (x, y) represents the color coordinates of the phosphor, $(x_{ee}, y_{ee})$ are the coordinates of standard white point in the CIE diagram and ($x_d$, $y_d$) are the coordinates of dominate wavelength point. The calculated color purity for the optimized CBV phosphors excited under 464 nm wavelength is 99.36 %. The high value of color purity for the optimized phosphor indicates pure red emission from the proposed CBV: $Eu^{3+}$ phosphors under 464 nm excitation. Thus, optimized phosphor can be used as a potential candidate for red component in w-LEDs under the blue light emitting excitation source.

*3.6. Time-resolved PL analysis*

The decay profiles of $CaBi_{1-x}VO_5$: $xEu^{3+}$ (1.0 mol% ≤ x ≤ 6.0 mol%) phosphors by monitoring the emission at 613 nm corresponding to $^5D_0 \rightarrow {}^7F_2$ transition excited under 342 and 464 nm wavelength have been displayed in Fig. 12 (a & b) respectively. The decay curves have been fitted with various exponential equations. The decay curve for 1.0 mol% $Eu^{3+}$ doped CBV phosphor shows best fit with single exponential function. However, for $Eu^{3+}$ ion concentrations higher than 1.0 mol%, decay curves deviate from single exponential and the

best fit is obtained for bi-exponential function. The decay profiles for the CBV phosphors excited under 464 nm also exhibit similar behavior. The single and bi-exponential equations can be given as follow: [56]

$$I = I_0 + A_1 e^{-(t/\tau)} \quad (5)$$

$$I = I_0 + A_1 e^{-(t/\tau 1)} + A_2 e^{-(t/\tau 2)} \quad (6)$$

Where, $I_0$ and I are the luminescence intensities at t=0 and some instant of time 't' respectively. $A_1$ and $A_2$ are curve fitting constants; $\tau_1$ and $\tau_2$ are the decay time for the two exponential terms respectively. In general, single exponential fitting indicates no energy transfer between the luminescent centers as they are separated far apart at lower concentration. While, the bi-exponential fitting indicates the probability of energy transfer via interaction between proximal $Eu^{3+}$ ions in the host lattice. The lifetime value for the single exponential fitted plots is provided by τ, while average lifetime of the phosphors showing bi-exponential decay profiles can be calculated as [57]

$$\tau_{avg} = \left[\frac{A_1\tau_1^2 + A_2\tau_2^2}{A_1\tau_1 + A_2\tau_1}\right] \quad (7)$$

The lifetime ($\tau_{avg}$) values for the $CaBi_{1-x}VO_5$: $xEu^{3+}$ (1.0 mol% ≤ x ≤ 6.0 mol%) excited under 342 and 464 nm wavelengths have been calculated and listed in Table 2. The decreasing trend of average life time along with increasing concentration of $Eu^{3+}$ ions also indicates energy transfer among $Eu^{3+}$ ions. The calculated values of lifetimes for as-synthesized phosphors under 342 nm excitation asserts energy transfer from host to $Eu^{3+}$ ions as there is only slight change in the values with varying doping concentration. Average life time value in micro-range suggests that as-synthesized phosphor can be potentially utilized for luminescent device applications.

**4. Conclusions**

ACCEPTED MANUSCRIPT

Monophasic $Eu^{3+}$ doped $CaBiVO_5$ phosphors have been successfully prepared by solid-state reaction method. The XRD patterns confirm the pure phase formation with an orthorhombic structure $CaBiVO_5$ with the average crystallite size lies in the 28-45 nm range. The SEM micrographs depict irregular and non-uniform agglomerated morphology of microcrystalline structures. The emission spectrum of the host exhibits a weak broad emission band in the visible region from 400 to 680 nm with an approximate maximum at 542 nm, originating from $[VO_4]^{3-}$ groups under 342 nm excitation. The emission spectra of $Eu^{3+}$ doped CBV phosphor exhibits color tunabilty under 342 nm excitation with varying dopant concentration and intense red emission under blue light excitation. The concentration quenching beyond 4.0 mol% of $Eu^{3+}$ ion concentration is observed due to dipole-dipole interaction between proximal $Eu^{3+}$ ions. Average lifetime values for $^5D_0$ energy level of the $Eu^{3+}$ ions with varying concentrations in CBV phosphors were found to be in the range of microseconds under $\lambda_{ex}$ = 342 and 464 nm. The calculated CIE chromaticity coordinates for undoped CBV sample under the excitation of 342 nm are located at yellow region of CIE 1931 diagram. However, the CIE chromaticity coordinates (0.640, 0.354) of the most intense $Eu^{3+}$ doped CBV phosphor excited under 464 nm wavelength are found to be close with commercial red phosphor $Y_2O_2S$: $Eu^{3+}$ (x= 0.622 and y= 0.351) and with the coordinates designated by NTSC. All the above-mentioned results imply that as-synthesized $Eu^{3+}$ activated $CaBiVO_5$ phosphors are potential to use as a red component in w-LEDs and other luminescent devices.

**Acknowledgements**

One of the authors, Ms. Harpreet Kaur is thankful to Delhi Technological University for providing JRF fellowship. The author (Dr. M. Jayasimhadri) would like to thank DST-

SERB, Govt. of India for sanctioning the research project (No. SB/FTP/PS-082/2014, dt. 02/03/2015)

**References**

[1] J. McKittrick, L.E. Shea-Rohwer, Review: Down conversion materials for solid-state lighting, J. Am. Ceram. Soc. 97 (2014) 1327–1352.

[2] T. Takeda, R.J. Xie, T. Suehiro, N. Hirosaki, Nitride and oxynitride phosphors for white LEDs: Synthesis, new phosphor discovery, crystal structure, Prog. Solid State Chem. 51 (2018) 41–51.

[3] D. Zhang, G. Li, W. Yang, W. Huang, Z. Leng, S. Fang, L.Li, Tunable photoluminescence properties and energy transfer of $Ca_5(BO_3)_3F:Tb^{3+}/Eu^{3+}$ phosphors for solid state lighting, J. Alloys Compd. 208 (2019) 155-163.

[4] X. Huang, H. Guo, A novel highly efficient single-composition tunable white-light-emitting $LiCa_3MgV_3O_{12}:Eu^{3+}$ phosphor, Dye. Pigment. 154 (2018) 82–86.

[5] K. Jha, A.K. Vishwakarma, M. Jayasimhadri, D. Haranath, Multicolor and white light emitting $Tb^{3+}/Sm^{3+}$ co-doped zinc phosphate barium titanate glasses via energy transfer for optoelectronic device applications, J. Alloys Compd. 719 (2017) 116–124.

[6] P. Pust, V. Weiler, C. Hecht, A. Tücks, A.S. Wochnik, A.K. Henb, D. Wiechert, C. Scheu, P.J. Schmidt, W. Schnick, Narrow-band red-emitting $Sr[LiAl_3N_4]:Eu^{2+}$ as a next-

generation LED-phosphor material, Nat. Mater. 13 (2014) 891–896.

[7] B.V. Ratnam, M. Jayasimhadri, K. Jang, Luminescent properties of orange emissive $Sm^{3+}$-activated thermally stable phosphate phosphor for optical devices, Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 132 (2014) 563–567.

[8] I. Ahemen, Dilip, K. De, A.N. Amah, A Review of Solid State White Light Emitting Diode and Its Potentials for Replacing Conventional Lighting Technologies in Developing Countries, Appl. Phys. Res. 6 (2014) 95–108.

[9] A. David, L.A. Whitehead, LED-based white light, Comptes Rendus Phys. 19 (2018) 169–181.

[10] N. Thejokalyani, S.J. Dhoble, Novel approaches for energy efficient solid state lighting by RGB organic light emitting diodes - A review, Renew. Sustain. Energy Rev. 32 (2014) 448–467.

[11] S. Nakamura, Blue-green light-emitting diodes and violet laser diodes, Mater. Res. Bull. 22 (1997) 29–35.

[12] G. Blasse, M.G.J. van Leur, Luminescence and energy transfer in the columbite structure, Mater. Res. Bull. 20 (1985) 1037–1045.

[13] G. Blasse, A. Bril, A new phosphor for flying-spot cathode-ray tubes for color television: Yellow-emitting $Y_3Al_5O_{12}$-$Ce^{3+}$, Appl. Phys. Lett. 11 (1967) 53–55.

[14] J.S. Kim, P.E. Jeon, Y.H. Park, J.C. Choi, H.L. Park, G.C. Kim, T.W. Kim, White-light generation through ultraviolet- emitting diode and white-emitting phosphor, Appl. Phys. Lett. 85 (2004) 3696–3698.

[15] R. Saraf, C. Shivakumara, S. Behera, H. Nagabhushana, N. Dhananjaya, Photoluminescence, photocatalysis and Judd-Ofelt analysis of $Eu^{3+}$-activated layered BiOCl phosphors, RSC Adv. 5 (2015) 4109–4120.

[16] A.K. Parchur, R.S. Ningthoujam, Behaviour of electric and magnetic dipole transitions of $Eu^{3+}$, $^5D_0 \rightarrow {}^7F_0$ and Eu–O charge transfer band in $Li^+$ co-doped $YPO_4:Eu^{3+}$, RSC Adv. 2 (2012) 10859-10868.

[17] B.C. Feldmann, T. Jüstel, C.R. Ronda, P.J. Schmidt, Inorganic Luminescent Materials : 100 Years of Research and Application, 13 (2003) 511–516. doi:10.1002/adfm.200301005.

[18] Y. Chen, W. Huang, T. Chen, Enhancing the performance of photovoltaic cells by using down-converting, J. Rare Earths. 29 (2011) 907–910. doi:10.1016/S1002-0721(10)60565-0.

[19] S. Neeraj, N. Kijima, A.. Cheetham, Novel red phosphors for solid-state lighting: the system $NaM(WO_4)_{2-x}(MoO_4)x:Eu^{3+}$ (M=Gd, Y, Bi), Chem. Phys. Lett. 387 (2004) 2–6.

[20] S. Kaur, M. Jayasimhadri, A.S. Rao, A novel red emitting $Eu^{3+}$ doped calcium aluminozincate phosphor for applications in w-LEDs, J. Alloys Compd. 697 (2017) 367–373.

[21] S.K. Mahesh, P.P. Rao, M. Thomas, T.L. Francis, P. Koshy, Influence of cation substitution and activator site exchange on the photoluminescence properties of $Eu^{3+}$-doped quaternary pyrochlore oxides, Inorg. Chem. 52 (2013) 13304–13313.

[22] V.R. Bandi, B.K. Grandhe, K. Jang, H.S. Lee, D.S. Shin, S.S. Yi, J.H. Jeong, Citric based sol-gel synthesis and luminescence characteristics of $CaLa_2ZnO_5:Eu^{3+}$ phosphors for blue LED excited white LEDs, J. Alloys Compd. 512 (2012) 264–269.

[23] R. Srinivasan, N.R. Yogamalar, J. Elanchezhiyan, R.J. Joseyphus, A.C. Bose, Structural and optical properties of europium doped yttrium oxide nanoparticles for phosphor applications, J. Alloys Compd. 496 (2010) 472–477.

[24] X. Zhang, Z. Zhu, Z. Sun, Z. Guo, J. Zhang, Host-sensitized color-tunable luminescence

properties of self-activated and $Eu^{3+}$ -doped $Ca_3LiZnV_3O_{12}$ phosphors, J. Lumin. 203 (2018) 735-740.

[25] Z. Zhu, Z. Guo, Z. Sun, Z. Wu, L. Zhou, X. Zhang, Synthesis, structure and luminescent properties of $Eu^{3+}$ doped $Ca_3LiMgV_3O_{12}$ color-tunable phosphor, Ceram. Int. 44 (2018) 16514–16521.

[26] L. Han, Y. Wang, J. Zhang, Y. Wang, Enhancement of red emission intensity of $Ca_9Y(VO_4)_7$:$Eu^{3+}$ phosphor via Bi co-doping for the application to white LEDs, Mater. Chem. Phys. 139 (2013) 87–91.

[27] G. Gundiah, Y. Shimomura, N. Kijima, A.K. Cheetham, Novel red phosphors based on vanadate garnets for solid state lighting applications, Chem. Phys. Lett. 455 (2008) 279–283.

[28] F. Wang, C. Liu, Z. Zhou, P. Jia, J. Lin, Molten salt synthesis and luminescent properties of $YVO_4$:Eu nanocrystalline phosphors, J. Rare Earths. 30 (2012) 202–204.

[29] Y. Wang, X. Liu, Y. Li, L. Jing, Synthesis and photoluminescence of multicolor tunable $K_3Gd(VO_4)_2$:$Eu^{3+}$ phosphors, J. Alloys Compd. 653 (2015) 315–320.

[30] P. Du, J.S. Yu, Self-activated multicolor emissions in $Ca_2NaZn_2(VO_4)_3$:$Eu^{3+}$ phosphors for simultaneous warm white light-emitting diodes and safety sign, Dye. Pigment. 147 (2017) 16–23.

[31] J. Boje and Hk.Muller-bushbaum, Synthese und Kristallstruktur von $CaBiVO_5$ Zeitschrift Fur Anorg. Und Allg. Chemie. 619 (1993) 521–524,

[32] A.K. Vishwakarma, M. Jayasimhadri, Pure orange color emitting $Sm^{3+}$ doped $BaNb_2O_6$ phosphor for solid - state lighting applications, J. Lumin. 176 (2016) 112–117.

[33] V. Mote, Y. Purushotham, B. Dole, Williamson-Hall analysis in estimation of lattice strain in nanometer-sized ZnO particles, J. Theor. Appl. Phys. 6 (2012) 6-13.

[34] A.N. Mallika, A.R. Reddy, K.S. Babu, C. Sujatha, K.V. Reddy, Structural and photoluminescence properties of Mg substituted ZnO nanoparticles, Opt. Mater. (Amst). 36 (2014) 879–884.

[35] S. Muthukumaran, R. Gopalakrishnan, Structural , FTIR and photoluminescence studies of Cu doped ZnO nanopowders by co-precipitation method, Opt. Mater. (Amst). 34 (2012) 1946–1953.

[36] B.D. Cullity, Elements of X-ray diffraction, second ed., Addison- Wesley Publishing Company, Inc, Philippines, 1978.

[37] L. Qin, D.L. Wei, Y. Huang, C. Qin, P. Cai, S. Il Kim, H.J. Seo, Ortho-vanadates $K_3RE(VO_4)_2$ (RE = La, Pr, Eu, Gd, Dy, Y) for near UV-converted phosphors, Mater. Chem. Phys. 147 (2014) 1195–1203.

[38] H. Ronde and G.Blasse, The nature of the electronic transitions of the vanadate group, Jorurnal Inorg. Nucl. Chem. 40 (1978) 215–219.

[39] X. Huang, H. Guo, $LiCa_3MgV_3O_{12}:Sm^{3+}$: A new high-efficiency white-emitting phosphor, Ceram. Int. 44 (2018) 10340–10344.

[40] Z. Pei, A. van Dijken, A. Vink, G. Blasse, Luminescence of calcium bismuth vanadate ($CaBiVO_5$), J. Alloys Compd. 204 (1994) 243–246.

[41] W.T. Carnall, P.R. Fields, K. Rajnak, Electronic Energy Levels of the Trivalent Lanthanide Aquo Ions. IV. $Eu^{3+}$, J. Chem. Phys. 49 (1968) 4450–4455.

[42] F. Wang, C. Liu, Z. Zhou, P. Jia, J. Lin, Molten salt synthesis and luminescent properties of $YVO_4$:Eu nanocrystalline phosphors, J. Rare Earths. 30 (2012) 202–204.

[43] Y. Li, X. Wei, H. Chen, G. Pang, Y. Pan, L. Gong, L. Zhu, G. Zhu, Y. Ji, A new self-activated phosphor of $Na_2YMg_2(VO_4)_3$ and luminescence properties in $Eu^{3+}$ doped $Na_2YMg(VO_4)_3$, J. Lumin. 168 (2015) 124–129.

[44] A.K. Parchur, R.S. Ningthoujam, Behaviour of electric and magnetic dipole transitions of $Eu^{3+}$,$^5D_0 \rightarrow {}^7F_0$ and Eu-O charge transfer band in $Li^+$ co-doped $YPO_4:Eu^{3+}$, RSC Adv. 2 (2012) 10859–10868.

[45] A.K. Vishwakarma, K. Jha, M. Jayasimhadri, A.S. Rao, K. Jang, B. Sivaiah, D. Haranath, Red light emitting $BaNb_2O_6 : Eu^{3+}$ phosphor for solid state lighting applications, J. Alloys Compd. 622 (2015) 97–101.

[46] G. Blasse, Concentration Quenching of $Eu^{3+}$ Fluorescence, J. Chem. Phys. 46 (1967) 2583–2585.

[47] L. Zhang, H. Zhong, X. Li, L. Cheng, L. Yao, J. Sun, J. Zhang, R. Hua, B. Chen, Solid state reaction synthesis and luminescence properties of $Dy^{3+}$-doped $Gd_2Mo_3O_9$ phosphor, Phys. B Condens. Matter. 407 (2012) 68–72.

[48] G. Blasse, Energy transfer in oxidic phosphors, Philips Res. Repts. 24 (1969) 131–144.

[49] L.G. Van Uitert, Characterization of Energy Transfer Interactions between Rare Earth Ions, J. Electrochem. Soc. 114 (1967) 1048–1053.

[50] D.L. Dexter, A Theory of Sensitized Luminescence in Solids, J. Chem. Phys. 21 (1953) 836–850.

[51] Q. Xu, J. Sun, D. Cui, Q. Di, J. Zeng, Synthesis and luminescence properties of novel $Sr_3Gd(PO_4)_3:Dy^{3+}$ phosphor, J. Lumin. 158 (2015) 301–305.

[52] X. Huang, H. Guo, $LiCa_3MgV_3O_{12}:Sm^{3+}$: A new high-efficiency white-emitting phosphor, Ceram. Int. 44 (2018) 10340–10344.

[53] H.J. Woo, S. Gandhi, M. Jayasimhadri, D.S. Shin, H.S. Lee, K. Jang, Engendering color tunable emission in calcium silicate based phosphors via ageing of silicate source, Sensors Actuators, B Chem. 241 (2017) 1106–1110.

[54] C.S. McCamy, Correlated color temperature as an explicit function of chromaticity

coordinates, Color Res. Appl. 17 (1992) 142–144.

[55] S. Gai, P. Yang, D. Wang, C. Li, N. Niu, F. He, X. Li, Monodisperse $Gd_2O_3$:Ln (Ln = $Eu^{3+}$, $Tb^{3+}$, $Dy^{3+}$, $Sm^{3+}$, $Yb^{3+}$/$Er^{3+}$, $Yb^{3+}$/$Tm^{3+}$, and $Yb^{3+}$/$Ho^{3+}$) nanocrystals with tunable size and multicolor luminescent properties, CrystEngComm. 13 (2011) 5480–5487.

[56] Z.H. Ju, R.P. Wei, J.X. Ma, C.R. Pang, W.S. Liu, A novel orange emissive phosphor $SrWO_4$:$Sm^{3+}$ for white light-emitting diodes, J. Alloys Compd. 507 (2010) 133–136.

[57] S. Kaur, A.S. Rao, M. Jayasimhadri, Enhanced red down-conversion luminescence and high color purity from flux assisted $Eu^{3+}$ doped calcium aluminozincate phosphor, J. Lumin. 202 (2018) 461–468.

ACCEPTED MANUSCRIPT

**Figure Captions**

**Fig. 1.** TGA curve for $CaBiVO_5$ (CBV) sample.

**Fig. 2.** (a) XRD patterns of CBV samples: standard data for $CaBiVO_5$, samples calcined at 750 °C, 850 °C and 950 °C for 3 hrs. (b) Experimental, refined and the difference between XRD pattern of $CaBiVO_5$ sample. (c) XRD patterns of $CaBi_{1-x}VO_5$: $xEu^{3+}$ (x = 0.0, 2.0, 4.0, and 6.0 mol%) sintered at 950° C for 3hr. (d) W-H plot of $CaBi_{1-x}VO_5$: $xEu^{3+}$ (x = 0.0, 2.0, 4.0, and 6.0 mol%).

**Fig. 3.** SEM micrographs of as-synthesized $CaBi_{1-x}VO_5$, x = 4.0 mol% phosphor at different resolutions (a) at 1 µm (b) at 2 µm.

**Fig. 4.** PLE spectrum monitored at 542 nm wavelength (left) and emission spectrum under 342 nm excitation of $CaBiVO_5$ host lattice (right).

**Fig. 5.** PLE spectra of $CaBi_{1-x}VO_5$ :$xEu^{3+}$ (1.0 mol% ≤ x ≤ 6.0 mol%) monitored at 613 nm wavelength.

**Fig. 6.** Comparison of the emission intensity of $Eu^{3+}$ doped CBV phosphor under different excitations.

**Fig. 7.** (a) Emission spectra of $CaBi_{1-x}VO_5$ :$xEu^{3+}$ (1.0 mol% ≤ x ≤ 6.0 mol%) under 342 nm excitation (Inset: emission intensity variation with $Eu^{3+}$ ion concentration). (b) Emission spectra normalized at 613 nm of all $Eu^{3+}$ doped CBV phosphors.

**Fig. 8.** (a) Emission spectra of $CaBi_{1-x}VO_5$ :$xEu^{3+}$ (1.0 mol% ≤ x ≤ 6.0 mol%) under 464 nm excitation; (b) $Eu^{3+}$ ions concentration dependence on emission intensity of CBV phosphors).

**Fig. 9.** Relationship of log I/x and log x for different concentration of $Eu^{3+}$ ions.

**Fig. 10.** Energy level diagram for $Eu^{3+}$ doped CBV phosphors.

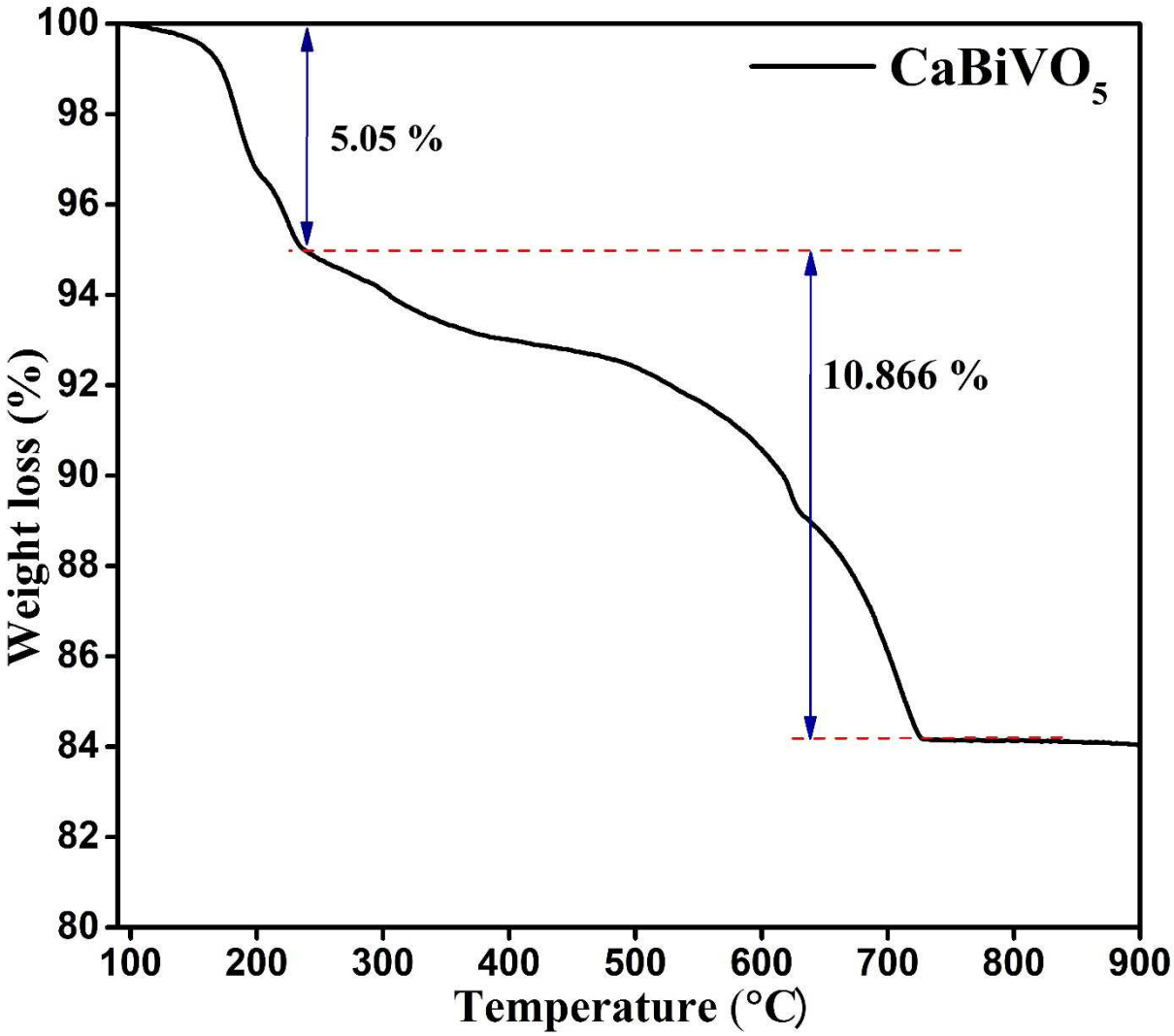
**Fig. 11.** CIE chromaticity coordinates of undoped ($\lambda_{ex}$=342 nm) and $Eu^{3+}$ doped CBV phosphor ($\lambda_{ex}$=342 and 464 nm).

**Fig. 12.** (a) TR-PL decay curve ($\lambda_{em}$ = 613 nm ) of $CaBi_{1-x}VO_5$: $xEu^{3+}$ (1.0 mol% ≤ x ≤ 6.0 mol%) under 342 nm. (b) under 464 nm excitation.

ACCEPTED MANUSCRIPT

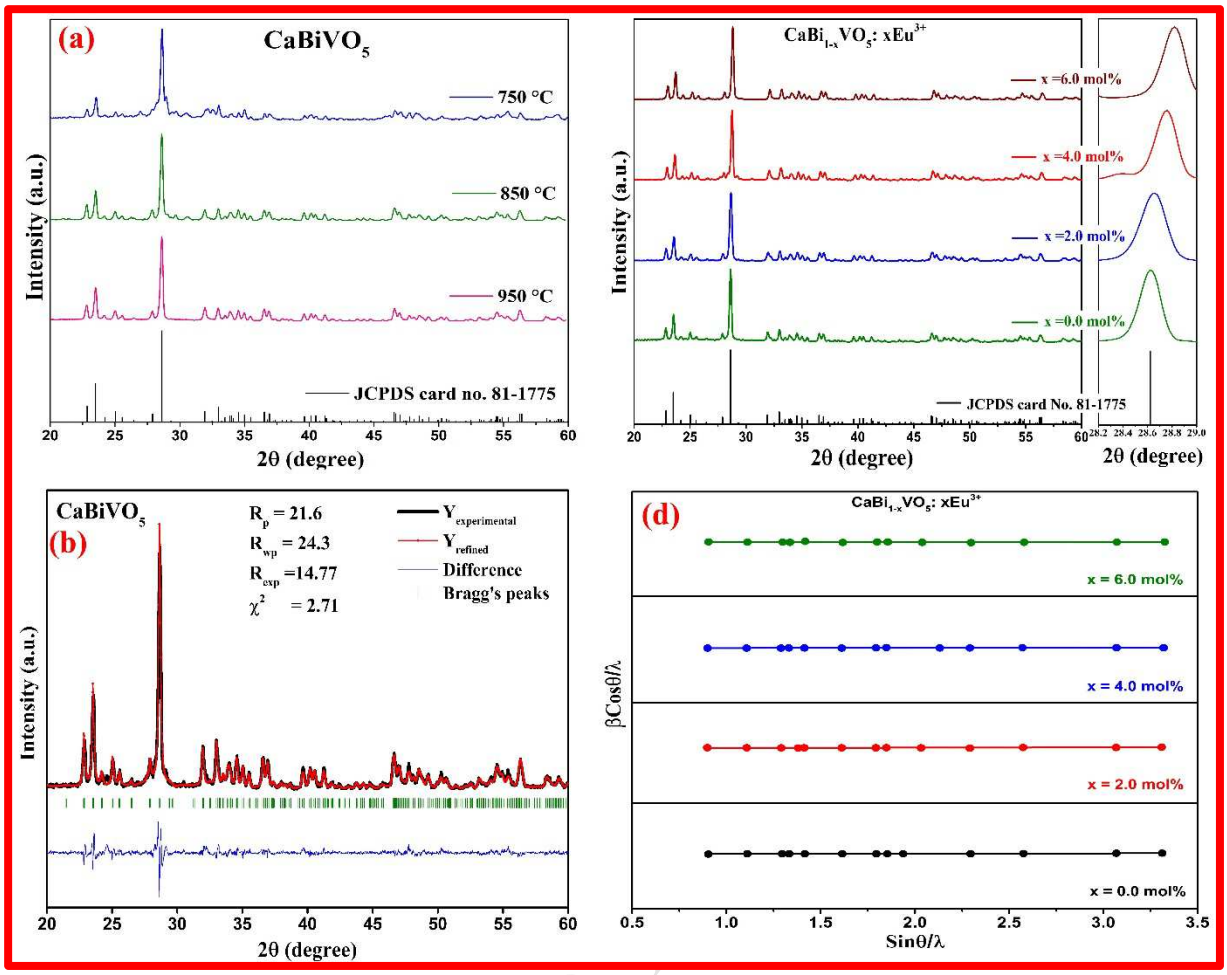
ACCEPTED MANUSCRIPT

**Fig. 1**



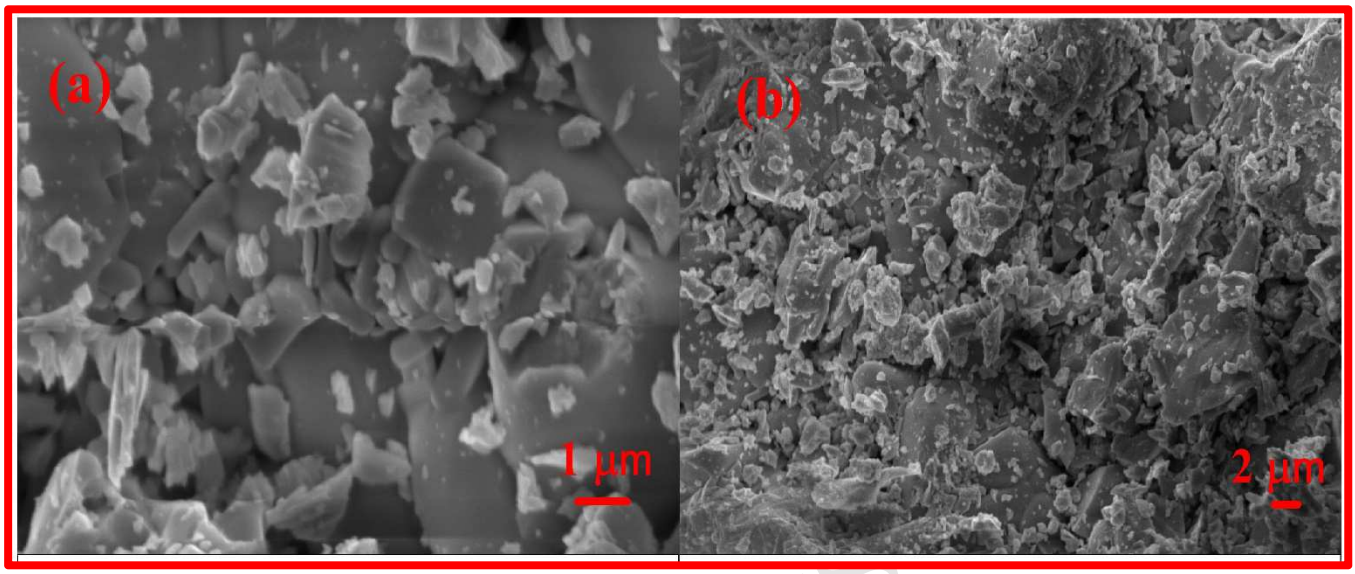

ACCEPTED

**Fig. 2 (a, b, c &d)**



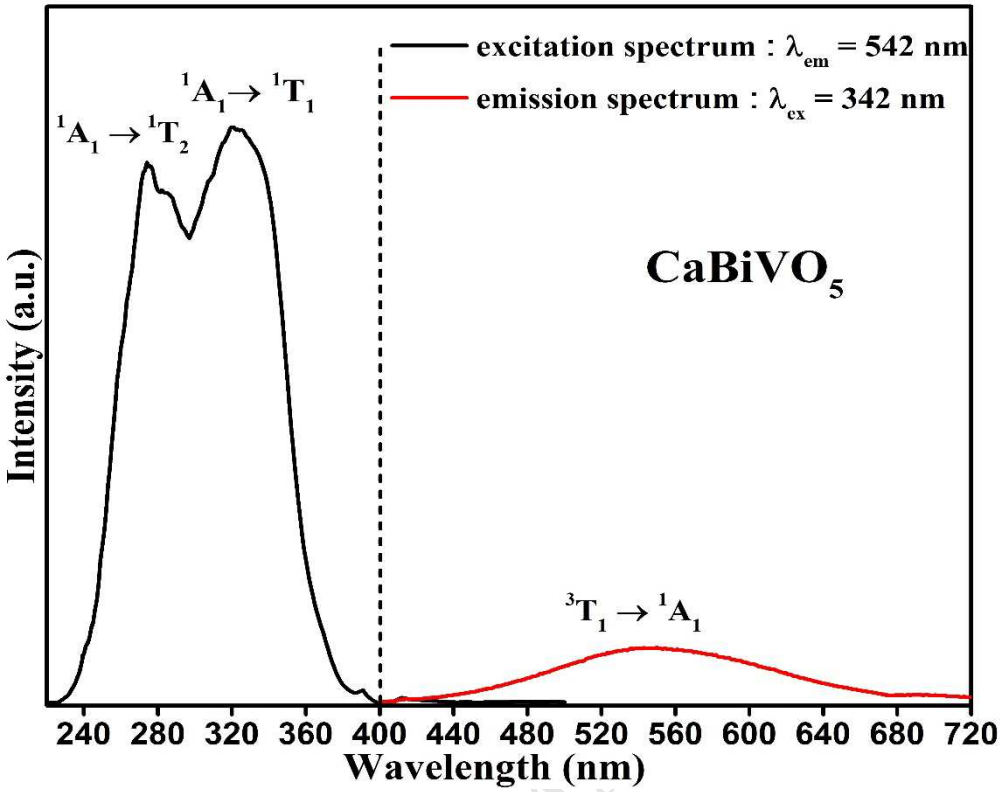

ACCEPTED MANUSCRIPT

**Fig. 3**



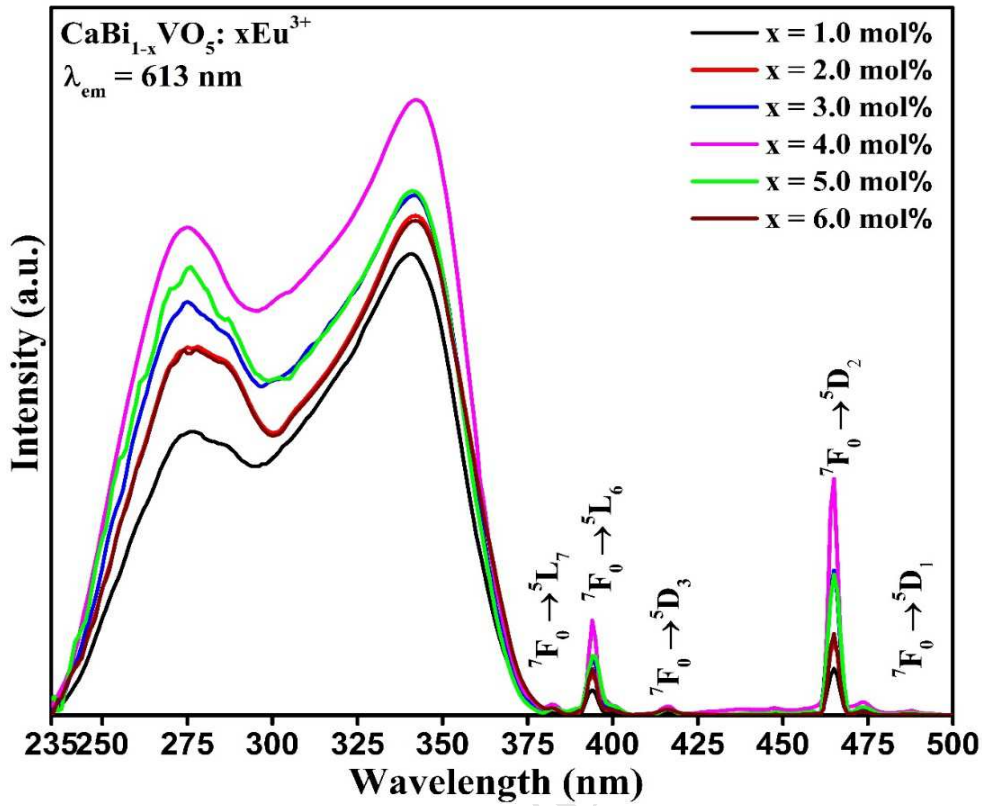

ACCEPTED MANUSCRIPT

**Fig. 4**



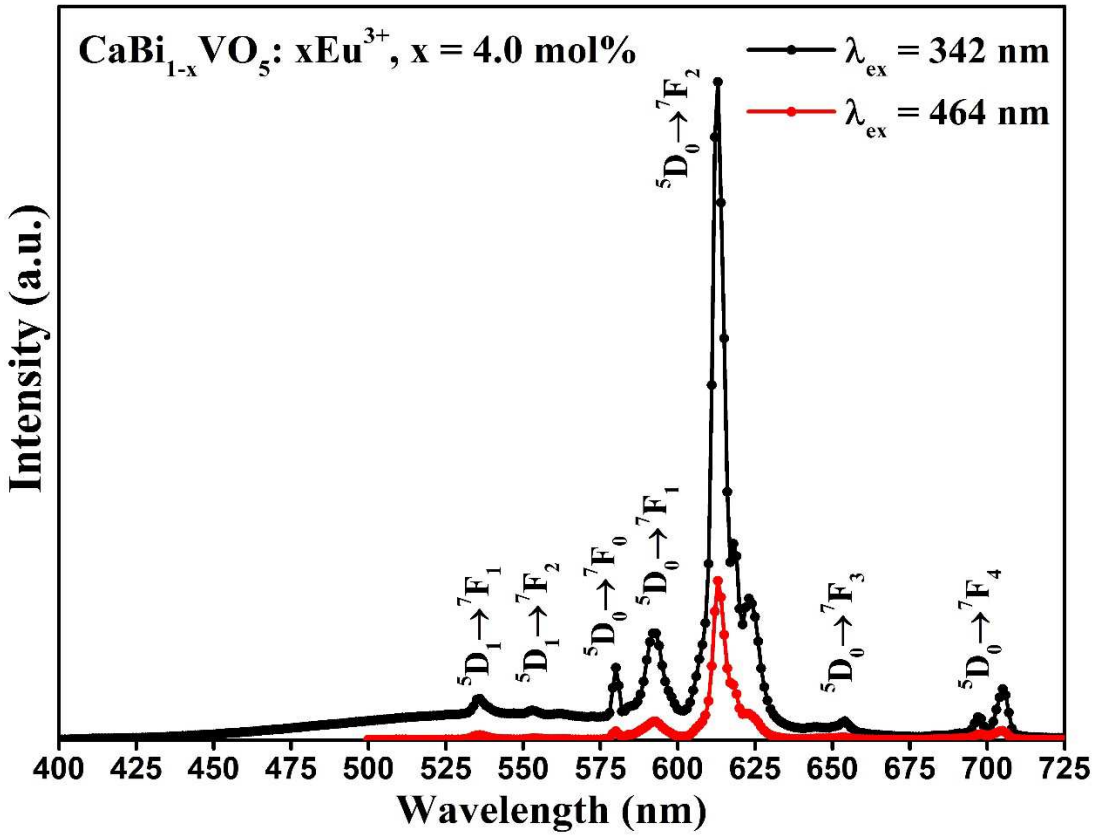

ACCEPTED MANUSCRIPT

**Fig. 5**



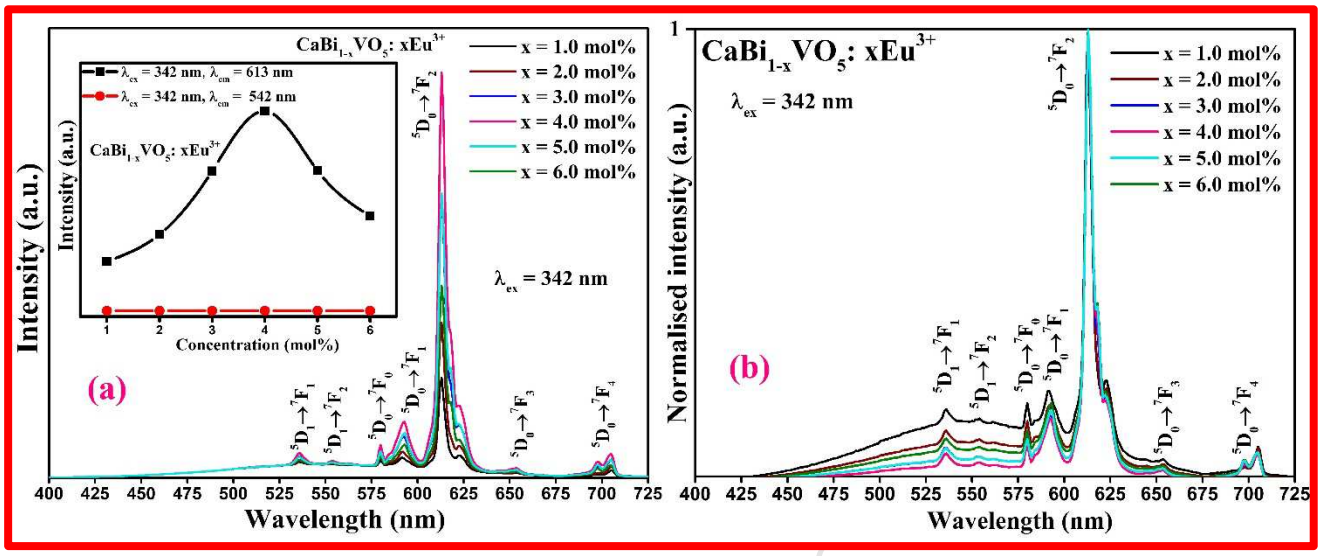

ACCEPTED MANUSCRIPT

**Fig. 6**



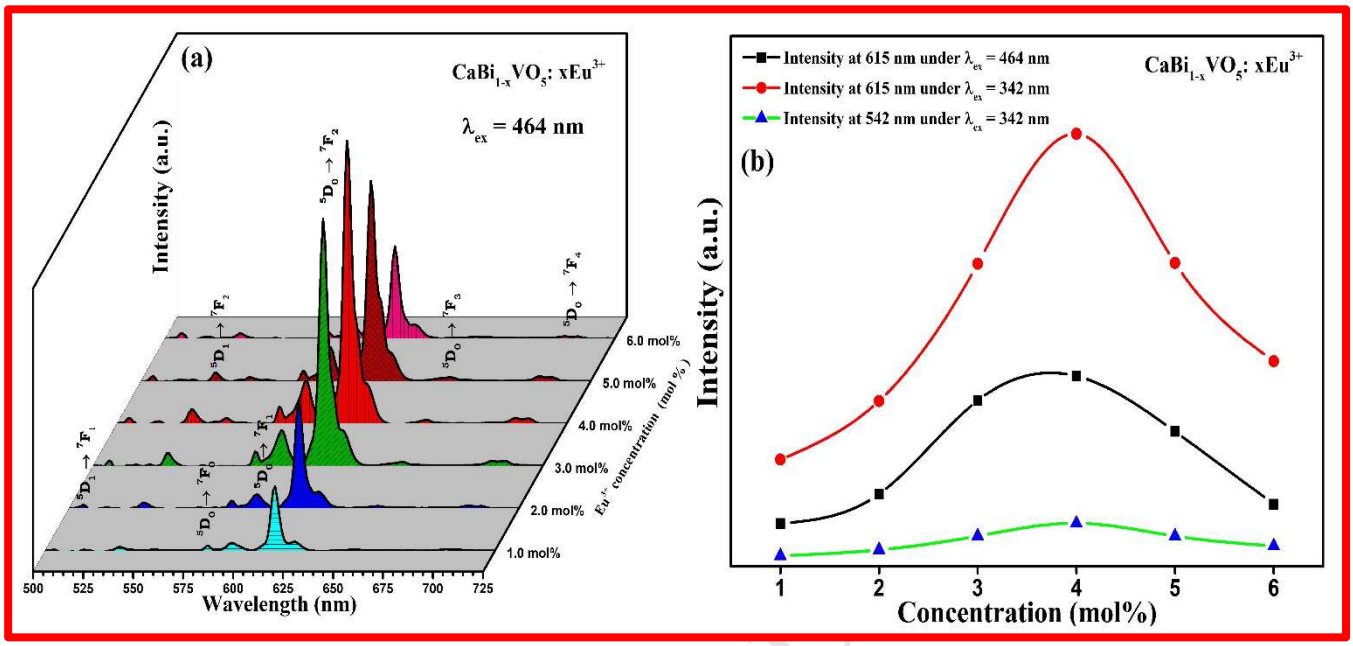

ACCEPTED MANUSCRIPT

**Fig. 7 (a & b)**



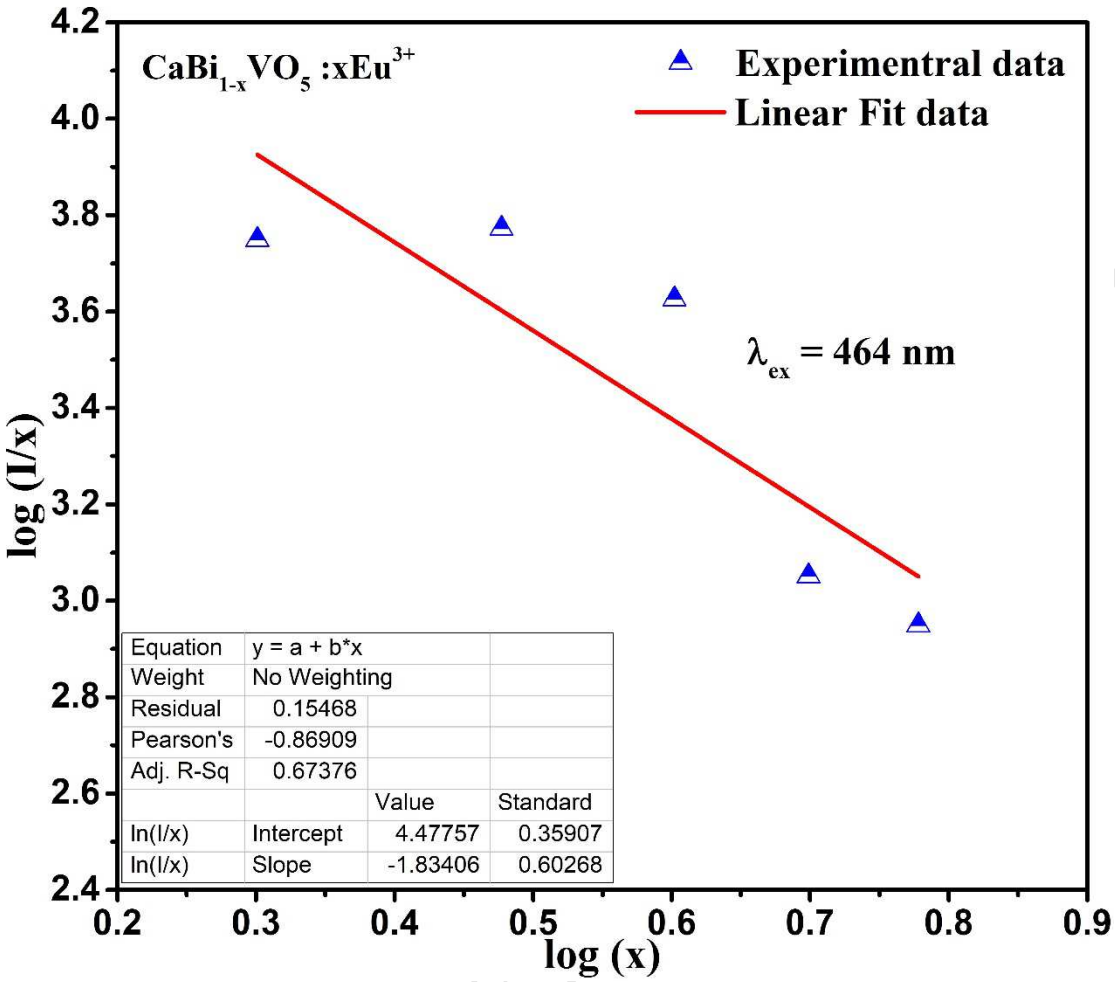

ACCEPTED MANUSCRIPT

**Fig. 8 (a & b)**



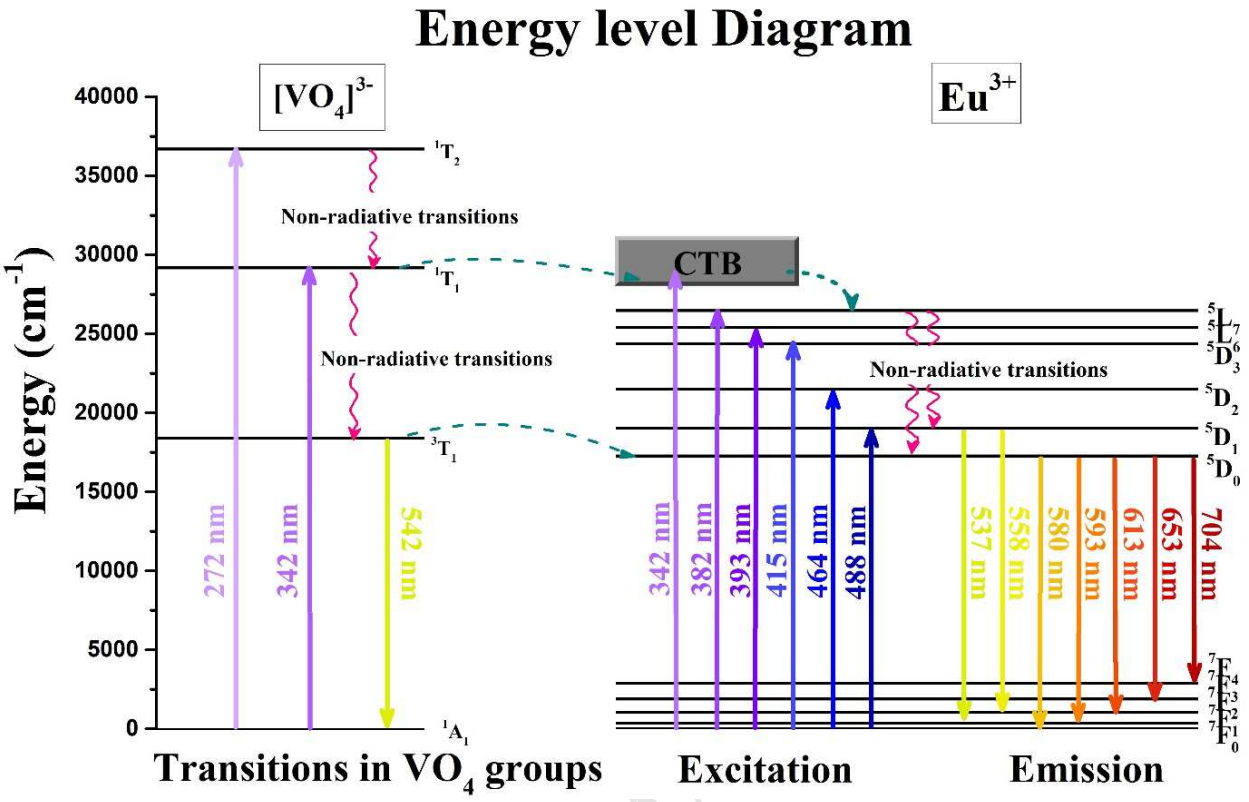

ACCEPTED MANUSCRIPT

**Fig. 9**



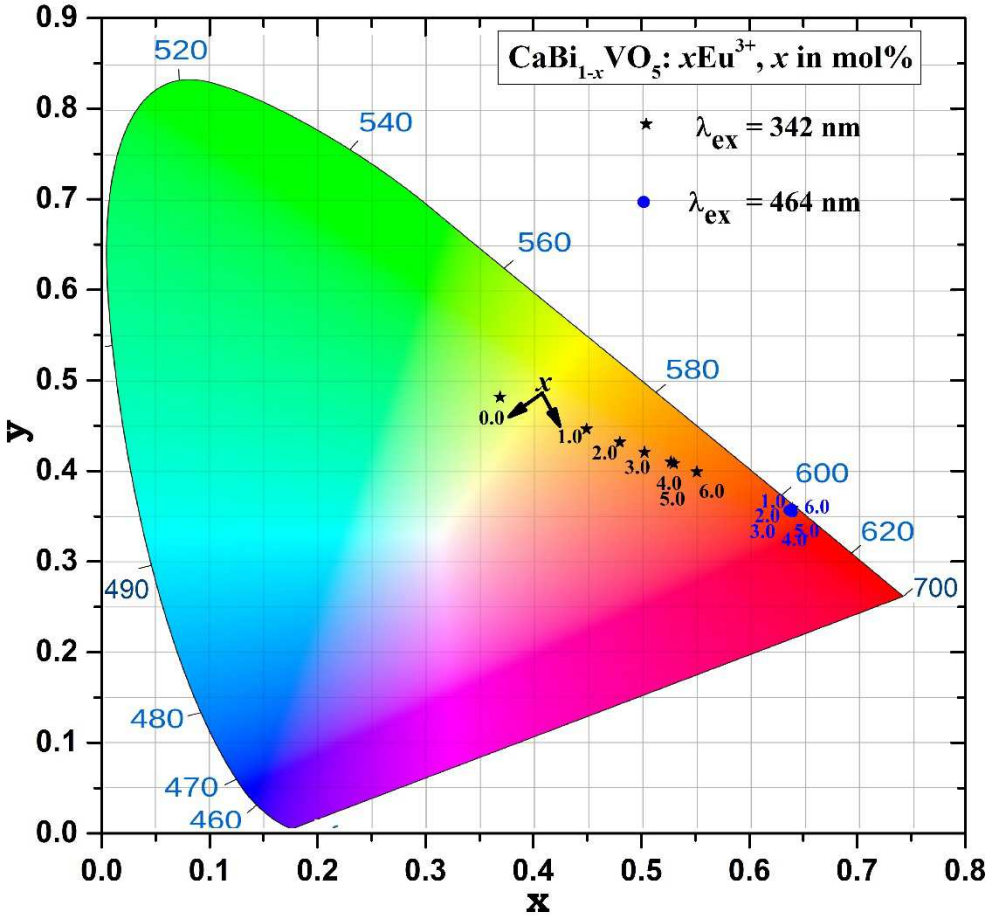

ACCEPTED MANUSCRIPT

**Fig. 10**



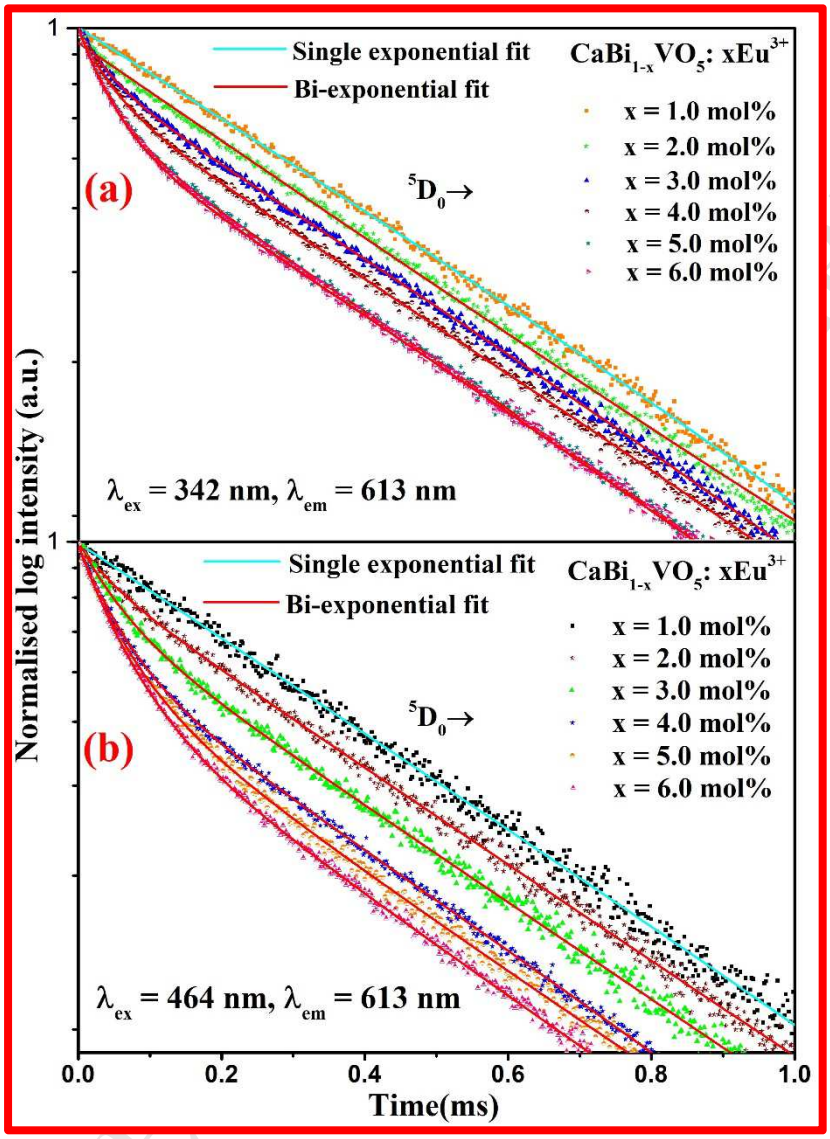

ACCEPTED MANUSCRIPT

**Fig. 11**




ACCEPTED MANUSCRIPT

**Fig. 12(a &b)**


Single exponential fit
Bi-exponential fit
CaBi1-xVO5: xEu3+
x = 1.0 mol%
x = 2.0 mol%
x = 3.0 mol%
x = 4.0 mol%
x = 5.0 mol%
x = 6.0 mol%
(a)
5D0→
λex = 342 nm, λem = 613 nm
(b)
λex = 464 nm, λem = 613 nm
Normalised log intensity (a.u.)
Time(ms)
0.0
0.2
0.4
0.6
0.8
1.0
1

## Tables

**Table 1** Rietveld refined parameters of $CaBiVO_5$ sample.

| **Formula** | **$CaBiVO_5$** |
|---|---|
| Structure | Orthorhombic |
| Space group | Pbca (61) |
| a (Å) | 11.1973 |
| b (Å) | 5.4246 |
| c (Å) | 15.5670 |
| α=β=γ | 90 ° |
| Z | 8 |
| $R_p$ | 21.56 |
| $R_{wp}$ | 24.30 |
| $\chi^2$ | 2.71 |
| V ($Å^3$) | 945.61 |
| Density ($mg/m^3$) | 5.33 |

ACCEPTED MANUSCRIPT

**Table 2** CIE coordinates, CCT and lifetime of undoped and doped CBV samples under different excitations.

| $CaBi_{1-x}VO_5$: $xEu^{3+}$ (x in mol%) | Under 342 nm excitation | | | Under 464 nm excitation | | |
|---|---|---|---|---|---|---|
| | CIE coordinates (x, y) | CCT values (K) | Lifetime (µs) | CIE coordinates (x, y) | CCT values (K) | Lifetime (µs) |
| **0.0** | (0.369, 0.481) | 4719 | - | - | - | - |
| **1.0** | (0.449, 0.446) | 3123 | 498 | (0.631, 0.366) | 2053 | 469 |
| **2.0** | (0.480, 0.431) | 2603 | 478 | (0.636, 0.361) | 2171 | 449 |
| **3.0** | (0.527, 0.409) | 1989 | 463 | (0.639, 0.358) | 2254 | 438 |
| **4.0** | (0.551, 0.398) | 1799 | 450 | (0.639, 0.358) | 2258 | 422 |
| **5.0** | (0.503, 0.420) | 2268 | 447 | (0.640, 0.358) | 2277 | 391 |
| **6.0** | (0.530, 0.408) | 1960 | 441 | (0.635, 0.362) | 2148 | 384 |

ACCEPTED MANUSCRIPT

Multimedia Tools and Applications
https://doi.org/10.1007/s11042-019-7677-2

# Design and analysis of an optimal ECC algorithm with effective access control mechanism for big data

Om Prakash Verma[1] · Nitin Jain[2] · Saibal Kumar Pal[3]

Received: 2 February 2019 / Revised: 3 April 2019 / Accepted: 24 April 2019
Published online: 07 May 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

Big data is a high volume data, as it comprises complex and large volume of information. A successful solution is to redistribute the data to a cloud server that has the capacity of storing and processing big data in an effective manner. The main intention of the research is to secure storage of big data and effective access control mechanism. The main stages of the proposed method are map reduce framework, secure storage process and access control mechanism process. Map Reduce is a distributed programming framework used to process big data. In mapper, the input dataset is grouped using hybrid kernel fuzzy c means (HKFCM) clustering algorithm. Finally, the reduced output is fed to the data owner for secure storage. In secure storage process, the suggested method utilizes optimal elliptic curve cryptography (OECC). Here the fundamental values are optimally selected by Modified grasshopper optimization algorithm (MGOA). In the access control mechanism, the effective policy update is proposed along with data storage construction and data deconstruction stage. The routine of the recommended method is assessed using memory and execution time by differentiating the number data size, number cluster size and the number of mapper. The proposed method attains the minimum time and memory utilization when compared to the existing method. The suggested method is implemented in cloud sim with Hadoop Map-reduce framework.

**Keywords** Big data · Elliptic curve cryptography · Fuzzy c means · Grasshopper optimization · Access control mechanism · Map reduce

✉ Nitin Jain
garg.nitin007@gmail.com

[1] Department of ECE, Delhi Technological University, Delhi, India

[2] Department of Information Technology, Delhi Technological University, Delhi, India

[3] Directorate of Information Technology & Cyber Security, DRDO, New Delhi, India

## 1 Introduction

In recent days, the most discussed and well known technological term is big data. It is defined as a collection of huge data sets and its primary characteristics are velocity, volume, and variety [13]. Nowadays, the usage of big data is inevitable in various fields such as healthcare, education, natural resources, social networking and so on. Some of the benefits of big data towards organization are considered as developing operations, facilitating innovation and adaptability and optimizing resources allocations [7, 17]. In general, the cloud is believed to be promising platform which is widely used for saving and processing big data. Moreover, cloud computing provides a big resource pool and enables virtual unlimited storage space by connecting the network resources together [22]. The most desirable properties of cloud are pay-per-use, scalability, rapid elasticity, fault-tolerance, and resource pooling [1]. In addition, the most essential part of cloud is data storage service. The users are most concerned with confidentiality, integrity, and approach control for the security and privacy of open networks including cloud computing [20]. Yet some users and companies hesitate to use cloud storage for maintaining and storing data due to security and privacy concerns [10]. In cloud computing, some of the issues are considered as the biggest challenges, which degrade the trustworthiness of cloud users. First, users concern about the risk of unauthorized access and data leakage due to storing data at cloud increases. Second, some attacks and intrusions towards data are biggest challenges of cloud service providers (CSP). Third, cloud user may not rely on the cloud service provider for managing their data for data storage, backup, migration, deletion, updation, search, query and data owner may not reliance upon service provider regarding the access. Fourth, data process and computation in the cloud could uncover the privacy of data proprietors [11]. There are distinctive security schemes are followed in cloud computing to reduce the security and privacy risk of big data, for example, data leak prevention systems (DLPSs), intrusion detection system (IDSs), encryption, authentication, access control, and firewalls, address these security issues [23]. Cryptography is the widely used technique which helps to hide sensitive or confidential data from the unauthenticated user. In order to maintain data confidentiality and security, the most trusted way is to outsource the data in the encrypted from [11]. Some of the techniques which are widely used for cryptography such as Data Encryption Standard (DES) algorithm, Advanced Encryption Standard (AES) algorithm, RSA algorithm, Diffie-Hellman Key Exchange and so forth. The two DES and AES algorithm is symmetric-key algorithm which is a sort of cryptographic strategy which utilizes same key for both encryption and decryption technique. RSA and Diffie-Hellman Key Exchange are the most widely recognized asymmetric-key algorithm which is another kind of cryptographic strategy which applies the different key for decryption and encryption in clouds [12]. Besides the availability of security recommendations, it is furthermore required to envision the security of the IT infrastructure itself [9]. Along with watching, separating and picking up information from the stored data keeping an eye upon the security is an earnest point [18].

## 2 Related work

Bruno Guazzelli Batista, et al. [2] have proposed a strategy for deciding QoS-driven methodologies in the cloud environments based on the different security mechanisms

were applied. Those mechanisms consolidated extra overhead on the performace, and to counter that, an endeavor was created to modify the computational assets progressively and on-the-fly.

Laurence T. Yang, et al. [24] have proposed an algorithm which reduces the overhead to solve large linear system and this algorithm is parallel in nature. The novelty is to use Wiedemann algorithm which includes many techniques such as strip partitioning, cyclic partitioning, and improved strip partitioning to improve the speed of algorithm by taking different step size. By doing this way, GNFS outperform the existing techniques.

Muthu Ramachandran [14], has given concise strategies, procedures, and best preliminary designing for requirement engineering. They furthermore rendered rules on software security as an service. He additionally talked about an Integrated-Secure SDLC model (IS-SDLC), which would profit experts, specialists, students, and instructors. That examination outlined that the approach can be extensively used for cloud framework Amazon EC2 benefit.

Dinh-Mao Bui et al. [3] recommend an energy-efficient answer for arranging the resounces in cloud computing. The proposed methodology followed up on the Gaussian procedure regression strategy for foreseeing the resouces usage of things in the coming period. In their examination, they applied convex optimization system for computing a proper amount of physical servers for each controlling window. They needed to guarantee that a low number of servers could be rendered nature of administration with the goal that they determined the amount of intrigue. In their last strides, for accomplishing the goal of energy reserve funds, they killed the inactive physical servers by a equivalent migrating instruction was issued to stack the virtual machines.

Hu Xiong et al. [21] have proposed a scheme which is identity-based and offers end to end security between the analytical system and data collectors. The property of the identity-based (ID-based) signcryption scheme with efficient revocation as well as the feature to outsource unsigncryption to enable secure big data communications between data collectors and analytical data systems. All the major security requirements such as confidentiality, authentication, integrity and non-repudiation were achieved by the scheme.

Jian Shen et al. [16] have proposed a third-party-helped scheme which uses cloud computing as technology that is searchable and verrifies data. For a superior depiction of the protocol, they have presented a client separated framework and a data storage structure that is cube based. Based on the framework model and information structure, the plan enables the clients to survey the trustworthiness of their transferred or downloaded information whenever and look through the online academic information.

Gang Chen et al. [4] have experienced in also improving, keeping up the information security framework for Netease, for instance, to portray how those big data frameworks advance. Specifically, their first form was a Hadoop-based disconnected location framework, which was immediately adjusted by an increasingly adaptable web-based gushing framework. Their progressing work is to manufacture a nonexclusive ongoing expository framework for Netease to deal with various occupations, for example, email spam location, client pattern mining, amusement log examination, and so forth. A built up on their experience, they have additionally proposed some key plan factors and difficulties for future big data frameworks.

Kan Yang et al. [25] have proposed a practical and fine-grained enormous information (big data) access control algorithm with the privacy-preserving arrangement. In particular, they conceal the entire attribute (instead of just its qualities) in the access arrangements. To help

information decoding, they have additionally designed a novel Attribute Bloom Filter to evaluate whether an attribute was in the access arrangement and pinpoint the correct position in the policy arrangement if, it was in the access strategy.

Yang et al. [26] have proposed a novel algorithm for healthcare system that stores big data and it IoT based. The point of their article was to improve the patient's security, for example, healthcare data, acknowledge access control for typical and crises, and bolster smart deduplication and spare the storage space in big data storage system. A safe deduplication strategy was work to expel the copy restorative records with indistinguishable data, which might be scrambled with different access arrangements and furthermore that method was utilized to spare the storage overhead in the big data storage system. The principal objective of those strategies was those the staying restorative records after the deduplication can be accessed by every one of the data clients approved by the distinctive, unique access approaches. Those smart healthcare big data storage system was formally demonstrated secure and broad correlation and reproductions show that it is useful.

Fugkeaw et al. [6] have presented a method for policy update, which is very secure and can be used for light weight applications. Their presented access control mechanism that demonstrates secure policy refreshes in a big data setup. The significant point of their technique was to reduce the re-encryption cost that was a successful policy updating strategy. Their trial result demonstrates that their presented a strategy presents a effective policy updating plan with less computation cost. Amid the few occasions of policy updation, that plan was supporting a high number of re-requesting.

Wei Zhou et al. [27] have exhibited an proposed model to quantitatively weight the capacity execution irregularity. Their technique was utilized amid the capacity execution when data are disseminated on various gadgets, and afterwards appropriates data on capacity gadget whose execution productively coordinates data get to qualities. The point of their article was to enhance the data have their capacity gadgets, and data like to live on gadgets which were most liked to get to attributes subsequently introducing best access execution. Aside from existing works, the model introduced in those articles delicate to the asymmetric execution of capacity gadgets and the asymmetric access of data. The degree was characterized to weight the entrance execution awkwardness among various gadgets. At that point data were conveyed on gadgets which best match the entrance attributes of data as indicated by inclination degree, subsequently the gadgets will be adequately used and the capacity execution of big data can be effectively made strides. Plus, little records with high access recurrence were offered need to disperse on developing stockpiling gadgets, for example, Flash SSDs, and the speeding up on little documents may enhance the proficiency of the entire framework.

Kavitha et al. [15] proposed a secure healthcare private data cloud which fools the attackers by providing the false and concealing the original data. Security of the cloud is achieved by sharing key using bilinear map that is very secure, and users can communicate securely by generating session keys. Decoy technique is used for the implementation process which provides additional security.

## 2.1 Problem definition

Proliferation of Big data leads to certain problems such as data hugeness, data security and many more. But in the today era, there is a urgent need to handle issue because of gaining

popularity. Data generated by various internet devices is so huge that it must be stored on to cloud, so leads to various other related issue of cloud such as security of data over cloud. The limitations of existing big data access control schemes are as follows.

- Existing big data processing methods faces many scalability and efficiency issues, mainly due to the increased volume of data. The increased volume of data occupies more space for storing the data in the cloud and hence they are difficult to process.
- Traditional encryption schemes have many flexibility issues and hence they could not provide complete security to the data.
- Since the volume of data is more it is impossible to identify the sensitive part of the information.
- In existing ABE based methodologies [8] provides the adaptability for a data proprietor to predefined the arrangement of clients who are qualified for accessing the data however they may experience the high complexity of effectively refreshing the access policy and ciphertext.
- The current attribute-based access control schemes may not handle with the attribute revocation issue.
- To maintain a privacy protection in the access policy, an immediate system is to disguise the attributes in the access policy. When the attributes are concealed, the unapproved and the unapproved customers cannot understand which attributes are defined in the access policy, which makes the deciphering a difficult issue.

These are the real disadvantages of various existing works, which inspire us to do this exploration on the access control scheme over big data.

## 3 Proposed method

Big data is regularly portrayed by 3Vs: i.e. volume, variety and velocity. These three categories must be handled anyway. The volume of data is coming from various sources such as business records, share market, healthcare and many more. There is a possibility of many security issues in the big data stored in the cloud. Existing access control technique is weak and hence leads to different attacks like collusion attacks, denial of service attack and insider attack to the data stored in the cloud. Hence, in order to get rid of all these issues and to preserve the data from attacks, the proposed method be used. Here we are considering the input big data as KDD data set. In the very first phase, input is fed to Map-Reduce Framework. In mapper, the big data is clustered into groups using the clustering algorithm. For clustering, the suggested technique utilizes hybrid kernel fuzzy c means algorithm (HKFCM). Then the clustered data of every mapper is fed as input to the reducer. Next, the reducer is utilized to decrease the number of clusters produced by the mapper without any data loss. Finally, the reduced output is sent to the data owner for further process. The proposed method has designed an optimal cryptographic algorithm for secure storage of big data. The main stage of the proposed method is the Setup and encryption stage, Data storage construction stage, Data reconstruction stage and Access policy update stage. The proposed method utilizes the optimal elliptic curve cryptography (OECC) method for security. Here the key values are optimally selected using Modified grasshopper optimization algorithm (MGOA). Finally, the secured data is stored in the cloud

along with the access policy of the data owner. The overall proposed system model is demonstrated in Fig. 1.

### 3.1 Map-reduce framework

Map reduce is a system for effective analysis of of big data on servers having large data size.It is a parallel and large scale data preparation framework which has been widely investigated and extensively received attention as of late. Map Reduce gets considerably more powerful

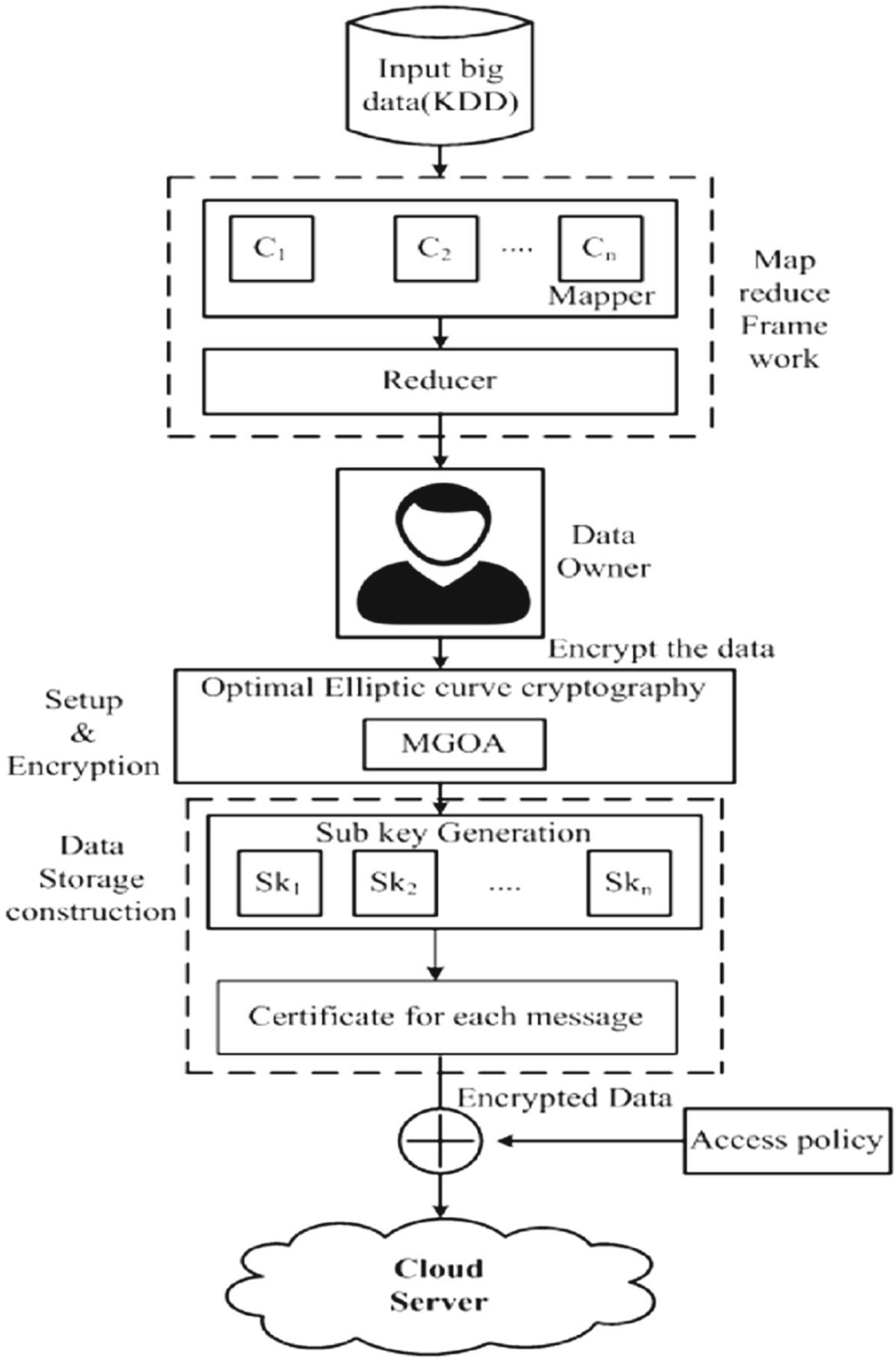



**Fig. 1** The proposed system model

and versatile because of the striking attributes of cloud computing. The map reduces framework consists of two steps namely mapper step and reducer step. The flow diagram of the map-reduce framework is shown in Fig. 2.

*Input Segment*: Here the big data (Data Mining (KDD) dataset and Knowledge Discovery) is considered as the input for the proposed method.

*Map Segment:* In this segment, kernel fuzzy c means clustering algorithm is used to group the available data. Moreover, then partitions it into M Map functions, each Map function run in parallel. Here the proposed clustering process is illustrated in beneath,

### 3.2 Clustering

FCM algorithm is dependent on apportioning, which builds the thought is to be classified into the greatest likeness among objects on a similar cluster, while the low likeness among the various clusters. FCM is an enhancement of the common c-means algorithm for data classification is inflexible, while the FCM is an adaptable fuzzy segment. The fuzzy clusters are controlled by the parcel of preparing tests as per the membership functions. The standard FCM applies the Euclidean distance as an objective function. In our proposed technique, the distance taken is Euclidean distance and this is adjusted using kernel-induced distance, and hence the speaking to the algorithm is inferred and called as the fuzzy kernel C-means (KFCM) algorithm, which is heartier than FCM. In the proposed method, two kernel values are hybridized, and then the hybrid kernel value is used for the further process. The detailed explanation of HKFCM algorithm is illustrated in beneath,

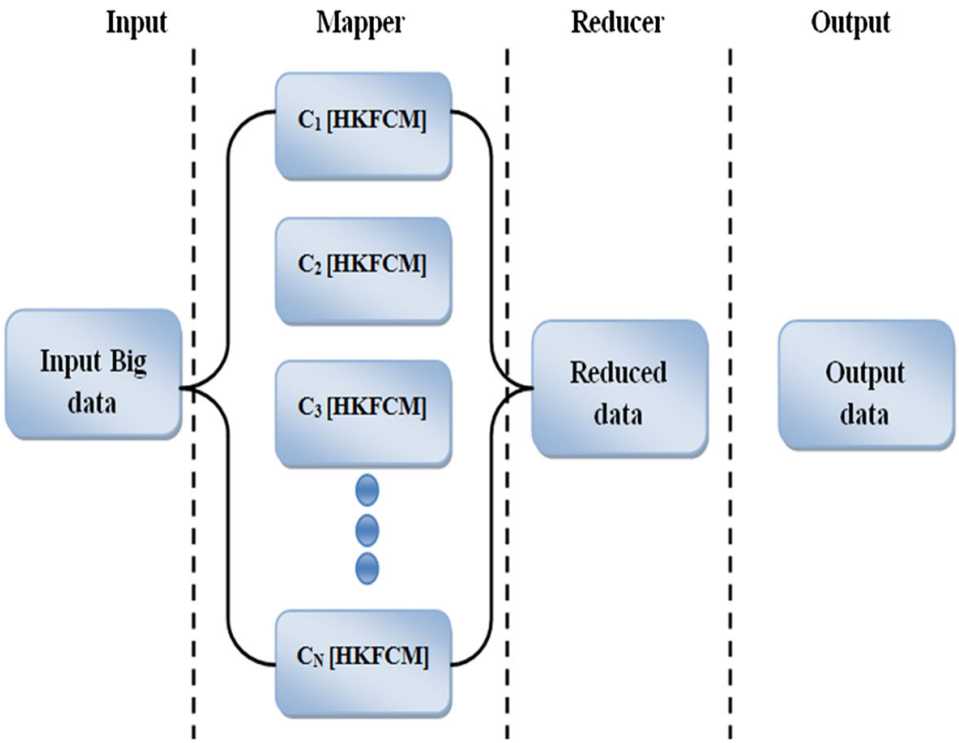



**Fig. 2** The proposed Map-Reduce Framework. The general arrangement of the proposed map reduces structure; in Map Reduce there are four segments such as Input data, Map Reduce and Output (collect all the reduce output). Map Reduce framework are normally exploited to practice big files

### 3.3 Hybrid kernel fuzzy C means clustering algorithm (HKFCM)

Hybrid kernel based fuzzy c means clustering algorithm is used for grouping the input data so that information loss can be minimized. The objective function (OF) of the proposed approach is described below,

$$OF(A,C) = \sum_{i=1}^{n}\sum_{j=1}^{c} A_{ij}^{m}\left(1-K_H\left(B_i,C_j\right)\right) \tag{1}$$

where $m$ represents the R (real number) >1, $A_{ij}^{m}$ represents the membership function, $C_j$ represents the cluster centre. This iteration will stop when max*imum* $\{|A_{ij}^{(k+1)}-A_{ij}^{k}|\}<t$, where $t$ is a termination criteria [0,1], whereas $k$ represents the iteration step. The kernel function is furnished as follows,

$$K(p,q) = \varphi(p)^{T}\varphi(q) \tag{2}$$

The linear and quadratic kernels are mostly used kernel function. The expressions are given below.

$$\begin{aligned} &\text{Linear Kernel :} && linear_k(p,q) = p^{T}q + r \\ &\text{Quadratic Kernel :} && quad_k(p,q) = 1-\frac{\|p-q\|^{2}}{\|p-q\|^{2}+r} \end{aligned} \tag{3}$$

Here

$p$, $q$ refer the inner products.
r represents constant value;

The proposed hybrid kernel represents the $K_{H,}$

$$K_H(p,q) = \frac{linera_k(p,q) + quad_k(p,q)}{2} \tag{4}$$

The fuzzy centres of the proposed kernel based FCM is adequately explained as follows.

$$C_j = \frac{\sum_{i=1}^{n} A_{ij}^{m}K_H\left(B_i,C_j\right)B_i}{\sum_{i=1}^{n} A_{ij}^{m}K_H\left(B_i,C_j\right)}; j = 1,2....C \tag{5}$$

The fuzzy membership function of the proposed kernel based FCM is adequately explained as follows.

$$A_{ij} = \frac{\left(1-K_H\left(B_i,C_j\right)\right)^{-1/(m-1)}}{\sum_{i=1}^{n}\left(1-K_H\left(B_i,C_j\right)\right)^{-1/(m-1)}}; j = 1,2....C \tag{6}$$

The steps are repeated until the maximum value of the objective function is achieved. From the above process, the input big data is clustered. Then the clustered output is fed to the input for the reducer.

*Reduce Segment:* In this segment, every one obtains the condensed data from the mapper and constructs the productivity of the Map-Reduce scheme. Map-reduce framework does not allow the duplicate data; it contains only the consistent data without duplication.
*Output segment:* Here, the reduced data is collected from the map-reduce framework. Then the reduced output is sent to the data owner.

In the data owner side, the main stage of the proposed method are setup and encryption stage, data storage construction stage, data reconstruction stage and access policy update stage. The detailed process of every stage is illustrated in beneath,

### 3.4 Setup and encryption stage

In the setup and encryption stage, the data owner is initializing the system parameter to generate the public and private keys for optimal encryption. The proposed method utilizes optimal elliptic curve cryptography (OECC) method for encrypting the data. Here the key values are optimally selected using Modified grasshopper optimization algorithm (MGOA). The step by step procedure of the OECC is described in the following section.

#### 3.4.1 Optimal elliptic curve cryptography (OECC)

In the proposed work, for the ECC part, key values are taken optimally using the modified grasshopper optimization algorithm (MGOA). The public and private keys construct encrypted data safer. The normal elliptic curve equation is written below,

$$y^2 = x^3 + ax + b \tag{7}$$

To develop both private and public key, the suggested method needs a key generation part. The message is encrypted by the receiver's public key and finally, the receiver will decrypt the encrypted data by using its private key. The key generation part is clarified underneath.

#### 3.4.2 Key generation by ECC

The process of elliptic curve cryptography are is explained for two pre-evaluated fields. For cryptographic tasks the proper field is picked with a massive quantity of points. The prime field selects a prime number, and finite number ofl points is produced on the elliptic curve.

$$PU_{key} = a*P_C \tag{8}$$

where,

a Random values
$P_C$ Point of the curve
$PU_{key}$ Public key

Here the proposed method optimally selects the "a" values based on the optimization technique. For optimization, the MGOA is used.

### 3.4.3 Modified grasshopper optimization algorithm (MGOA)

In the suggested work, the random values are selected by a MGOA. Grasshopper Optimization algorithm (GOA) is s swarm based algorithm which follows the model grasshopper swarm in nature. The grasshopper swarm has a unique featurethat similar behavior is shown by both the nymph and adulthood in grasshopper. The nymph grasshopper moves like moving barrels in a great numbers. The nymph grasshopper nearly eats all the vegetation which comes in their way during their development. When they get grown-up from nymph, they frame a swarm in the air and afterwards they move over a substantial separation. In the suggested work, GOA is modified by oppositional based learning strategy (OBL) which increases the searching ability. In the proposed method, a grasshopper is used to represent the random values represented by prime numbers. The objective function is defined as the key breaking time. The step by step procedure of MGOA is described below. The accompanying procedures are utilized to choose the optimal key value. The numerical model utilized to reproduce the swarming conduct of MGOA is introduced as pursues,

Step 1: *Solution representation*

For optimization of key value, MGOA is primarily making the arbitrary solution. Solution creation is a vital step of optimization calculation that generates the optimal arrangement quickly. In MGOA, key values are randomly initializing the position of all grasshoppers. Here, each in the swarm is treated as a grasshopper in a D-dimensional search space.

Step 2: *Opposite solution generation*

After the initial solution generation, the method creates an opposite solution using (9). Here, every solution $P_{ij}$ has a unique opposite $P'_{ij}$ solution. The opposite solution $OP\,(P'_{11}, P'_{12}, \ldots, P'_{1n})$ is calculated as.

$$P'_{ij} = L_{W_i} + U_{Pi} - P_{ij} \tag{9}$$

where $L_{W_i}$ refers to the lower bound coefficient, $U_{Pi}$ refers to the upper bound coefficient, $P_{ij}$ refers to the old solution. Then, the method combines the initial and opposite solution for further processing.

Step 3: *Fitness calculation*

The evaluation of fitness is accomplished for each solution using (10),

$$\text{Fitness} = \textit{Max key breaking time} \tag{10}$$

Step 4: *Updation using grasshopper optimization*

After computing the fitness value, the method updates the solution by grasshopper optimization algorithm. The updated equation is given as

$$P_i = S_i + G_i + W_i \tag{11}$$

where $P_i$ refers to the position, $S_i$ refers the social interaction, $G_i$ refers to the gravity force and $W_i$ refers to the wind advection.

$$S_i = \sum_{\substack{j=1 \\ j \neq i}}^{N} s(d_{ij})\hat{d}_{ij} \tag{12}$$

$$d_{ij} = |P_j - P_i| \tag{13}$$

$$\hat{d}_{ij} = \frac{P_j - P_i}{d_{ij}} \tag{14}$$

where $d_{ij}$ refers to the distance and s is the strength of social force. The gravity force ($G_i$) is calculated using (15).

$$G_i = -g\hat{e}_g \tag{15}$$

Here, g refers to gravitational constant and $\hat{e}_g$ refers a unity vector. The wind advection $W_i$ is estimated using Eq. (16).

$$W_i = u\hat{e}_w \tag{16}$$

Here, u refers to the constant drift and $\hat{e}_w$ refers unity vector. The Eq. 11 is altered by substituting S, G and W in Eq. (11).

$$P_i = \sum_{\substack{j=1 \\ j \neq i}}^{N} s(|P_j - P_i|)\frac{P_j - P_i}{d_{ij}} - g\hat{e}_g + u\hat{e}_w \tag{17}$$

where N is the number of grasshoppers. Using (17), the proposed method can update the solution.

By using the above equation, grasshoppers rapidly achieve the safe place and the swarm does not reach to point of convergence. In order to avoid this issue, an proposed version of this equation is described below,

$$P_i = C\left(\sum_{\substack{j=1 \\ j \neq i}}^{N} C\frac{U_P - L_W}{2} s(|P_j - P_i|)\frac{P_j - P_i}{d_{ij}}\right) + T \tag{18}$$

where T represents the target or best solution. C is a coefficient

$$C = C_{\max} - t\frac{C_{\max} - C_{\min}}{t_{\max}} \tag{19}$$

where $C_{max}$ represents higher value and $C_{min}$ represents a lower value. Moreover, t represents the current iteration and $t_{max}$ represents the maximum iteration.

Step 5: *Termination criteria*

The process stops its execution just if the most extreme number of iterations is accomplished and the arrangement which comprises the best fitness value is chosen as the optimal one.

Thus the optimized value will be examined for the prime value. Based on the optimal prime value the method selects the public $pu_{ky}$ and private key $pv_{ky}$. The encryption and decryption process of optimal ECC algorithm is explained beneath,

### 3.4.4 Encryption and decryption

In optimal ECC, the input data is encrypted and the output is split into two cipher text as $T_1$ and $T_2$.

$$T_1 = k*P_C \tag{20}$$

$$T_2 = Data + k*P_C \tag{21}$$

Where $k$ symbolizes the random value [1 to –(n-1)]. This encrypted message $T_1$ and $T_2$ is sent to the receiver. After receiving the ciphertext, the receiver decrypts the data using the following equations,

$$Data = T_2 - a*T_1 \tag{22}$$

Based on the above process, the input data is encrypted in the data owner side. After encrypting the input data, the resultant output is given to the data storage construction stage. Data storage construction stage is explained in beneath.

### 3.4.5 Data storage construction stage

In data storage construction stage, the data owner controls the subkey for each genuine client, who requires utilizing the information. At that point a certificate is generated for each message in a set of the message, and supplies the encrypted information into the cloud server along with the access policy. Every process of subkey generation and certificate generation is described beneath,

- Subkey generation
- Certificate generation

*Subkey generation*

Here, each legitimate user selects the random number and encrypt the random number with the data owner public key and generate the ciphertext. Then, he sends the user id ($Id_i$), a hash function (H) and ciphertext to the data owner. After receiving the ciphertext, the data owner checks the two conditions: the first condition is the new hash value is equal to the old hash

value for the corresponding user and the second one is the random value of the corresponding user is not equal to the other user value. If the two conditions are verified, the data owner controls the subkey for the corresponding user. Then securely broadcasts the $Id_i$ and subkey to the entire user.

*Certificate generation*

To validate the user legitimacy to access the data and to avoid the cheating behaviours of the user, the proposed method generates the certificate for each message. Finally, along with encrypted data, certificate message and access policy are sent to the cloud server (Fig. 3).

### 3.4.6 Data reconstruction stage

In the data reconstruction stage, if any user requires the data from the cloud server, first the user download the data and asks help from the other legitimate user to decrypt the data. The user sends the demand to the data owner to get the message certificate. The data owner encrypts certificate by usingthe user secret number after getting the request and the ciphertext sends to the user. Next, the user applies the subkey to compute the exchange certificate and send to the other legitimate user. After that, the other user verifies the certificate and if the verification is a success then the corresponding user is a valid user of the message. Once the verification gets completed the other user sends the $Id_i$ and the random secret number to the corresponding user. After receiving the idea and secret number from all the users, the corresponding user verifies the $Id_i$ is correct. In case, if the secret number does not pass the verification, the subkey value cannot be applied by the corresponding user. When the users obtain the authorization from the legal user, the corresponding user recovers the original data.

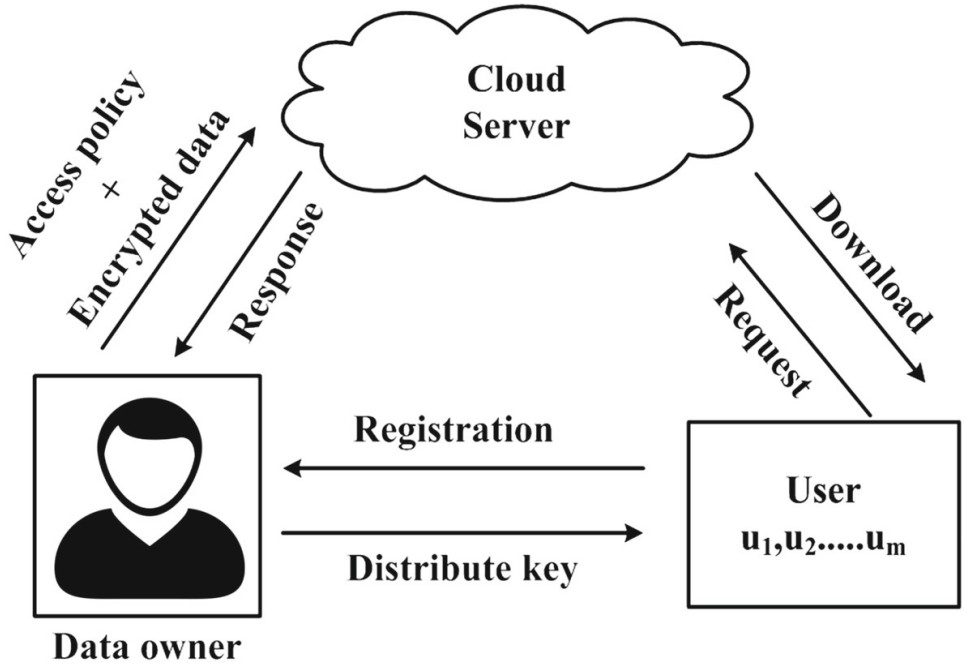



**Fig. 3** The data storage construction model

**Table 1** Experimental analysis of time and memory by varying data size

| Data size (MB) | Execution time (ms) | Memory (bits) |
|---|---|---|
| 10,000 | 113,485 | 1,111,745 |
| 20,000 | 156,984 | 1,245,787 |
| 30,000 | 214,785 | 1,364,575 |
| 40,000 | 286,545 | 1,457,745 |

### 3.4.7 Policy update stage

The final stage is access policy update stage, the cloud server can upgrade the cipher content when a different access policy is specified by the data owner, and the updated value is to be verified by the data owner. In the meantime, every client can quickly upgrade its secret number and the sub-key. The performance of the suggested method is tested and its effectiveness is equated with another methods in the result and discussion section.

## 4 Result and discussion

We have explained the results obtained from the suggested secure storage over Big Data through efficient access control mechanism and Map Reduce Framework in a Cloud environment is experimented with the KDD dataset in this section. The experimentation is led in a cloud situation called as CloudSim. We have implemented our proposed method in Java (JDK 1.8) and a number of tests were performed on a 2 GHz machine with 4GB of RAM.

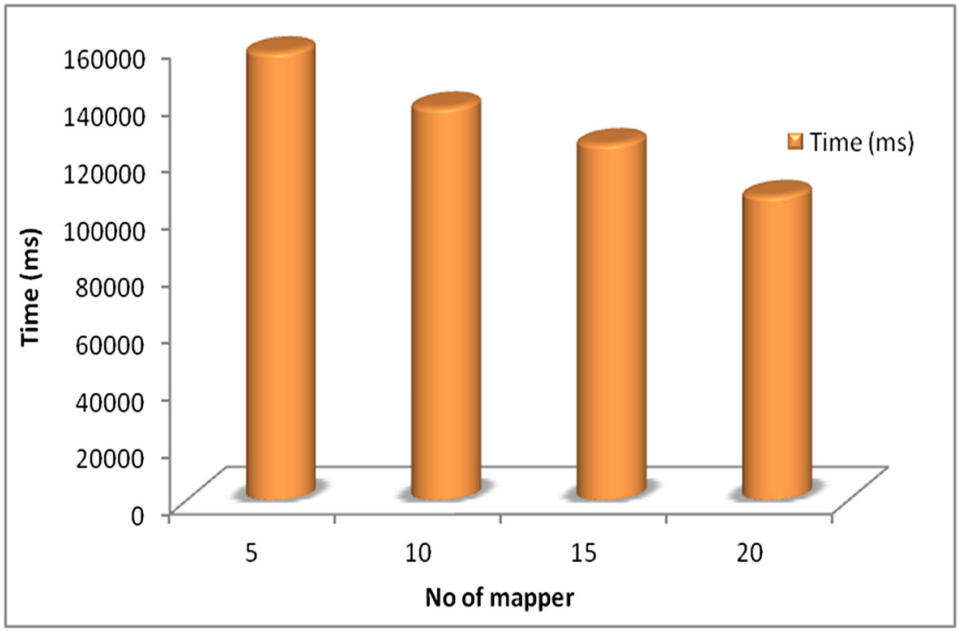



**Fig. 4** Performance analysis of execution time

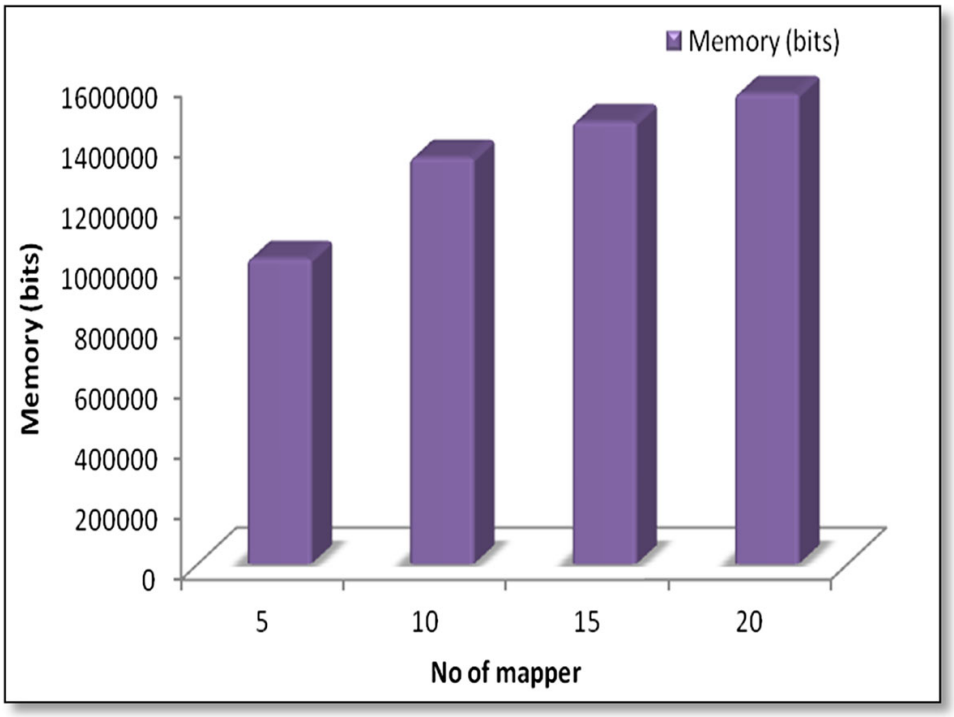



**Fig. 5** Performance analysis of memory

## 4.1 Dataset description

In our experimentation, the proposed technique uses the KDD data set [5]. This data set is used cryptography and its application. This database is very exhaustive and contains a standard data which is used for inspection and research. KDD preparing dataset involves around 4,900,000 single affiliation vectors all of which comprises 41 incorporates and is named as either conventional or a strike, with correctly one express attack type.

## 4.2 Experimental results

Experimental analysis of our suggested technique is shown in the below section. Here Table 1 shows the performance measure of the recommended technique. Here we are considering the performance measure as memory and time by varying the data size. The experimental analysis of the proposed technique is shown by changing the number of data size.

**Table 2** Experimental analysis of execution time and memory by differentiating cluster size

| Cluster size | Execution time (ms) | Memory (bits) |
|---|---|---|
| 4 | 106,545 | 1,111,745 |
| 6 | 124,785 | 1,317,383 |
| 8 | 137,573 | 1,415,869 |
| 10 | 156,984 | 1,531,727 |

**Table 3** Comparison of the proposed method and existing methods with time-based on data size

| Data size | Time (ms) MGOA+ECC | Time (ms) GOA + ECC | Time (ms) GA + ECC | Time (ms) ECC |
|---|---|---|---|---|
| 10,000 | 113,485 | 136,537 | 158,648 | 175,754 |
| 20,000 | 156,984 | 185,436 | 194,723 | 226,648 |
| 30,000 | 214,785 | 248,765 | 259,648 | 278,538 |
| 40,000 | 286,545 | 359,473 | 378,386 | 407,285 |

Table 1 shows the execution time and memory by altering the information measurement. For the information measurement of 10,000, the execution time found by the procedure is 113,485 ms and the memory usage of the method is 1,111,745 bits. For the information measurement of 20,000, the execution time was 156,984 ms and the memory use is 1,245,787 bits. For the information measurement of 30,000, the usage time got is 214,785 ms the memory utilization is 1,364,575 bits. The implementation time obtained is 286,545 ms and the memory utilization is 1,457,745 bits for the data dimension 40,000. Figures 4 and 5 represents the time and memory value of the proposed method by varying the number of the mapper.

Figure 4 represents the investigation of usage time by varying the number of mappers. For 5 mappers, the execution time acquired is too far-reaching the technique for the proposed strategy is 156,984 ms, for 10 mappers the usage time got is 137,573 ms, for 15 mappers the execution time got is 124,785 ms and for 20 mappers the execution time got is 106,545 ms by the procedure.

Figure 5 determines the investigation of memory by a variable measure of mappers. For 5 mappers the memory use is 1,011,745 bits, for 10 mappers the memory use is 1,345,787 bits,

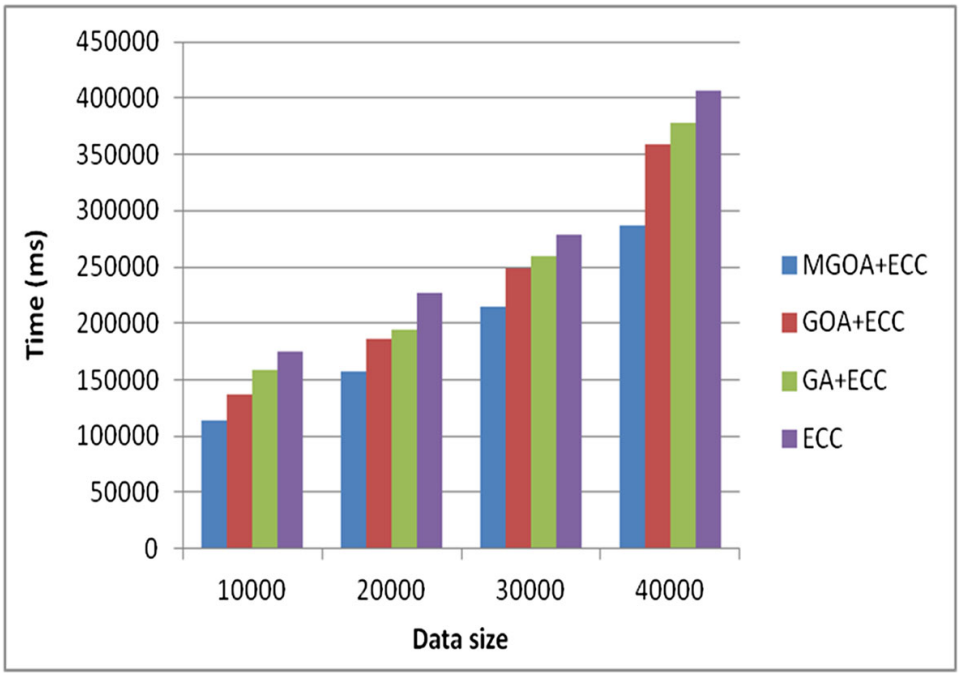



**Fig. 6** Performance comparison of proposed and existing methods with time-based on data size

**Table 4** Comparison of the proposed method and existing methods with memory based on data size

| Data size (MB) | Memory (bits) MGOA+ECC | Memory (bits) GOA + ECC | Memory (bits) GA + ECC | Memory (bits) ECC |
|---|---|---|---|---|
| 10,000 | 1,111,745 | 1,365,743 | 1,574,362 | 1,754,382 |
| 20,000 | 1,245,787 | 1,574,362 | 1,765,285 | 1,976,431 |
| 30,000 | 1,364,575 | 1,795,421 | 1,965,473 | 2,365,471 |
| 40,000 | 1,457,745 | 1,896,542 | 2,287,452 | 2,576,352 |

for 15 mappers the memory usage is 1,464,575 bits and for 20 mappers the memory use is 1,557,745 bits by the proposed procedure.

Table 2 represents the proposed time and memory by changeable cluster size. For the cluster size of 4, the implementation time obtained by the proposed method is 106,545 ms, and the memory utilization of the projected technique is 1,111,745 bits. For the cluster size of 6, the time obtained is 124,785 ms and the memory utilization is 1,317,383 bits. For the cluster size of 8, the time obtained is 137,573 ms the memory utilization is 1,415,869 bits. The time obtained is 156,984 ms and the memory utilization is 1,531,727 bits for the cluster size 8.

### 4.2.1 Comparison analysis of recommended method based on data size

Table 3 shows the Performance comparison of MGOA+ECC, GOA + ECC, GA + ECC and ECC with time. The time under the mentioned techniques for various data sizes is depicted in the table.

The comparison for the time under different techniques such as MGOA+ECC, GOA + ECC, GA + ECC and ECC is shown below. When the data size is 10,000, the time under MGOA+ECC is 113,485 ms, under GOA + ECC is 136,537 ms, under GA + ECC is 158,648 ms and the time under ECC is 175,754 ms. When data size is 20,000, the time under

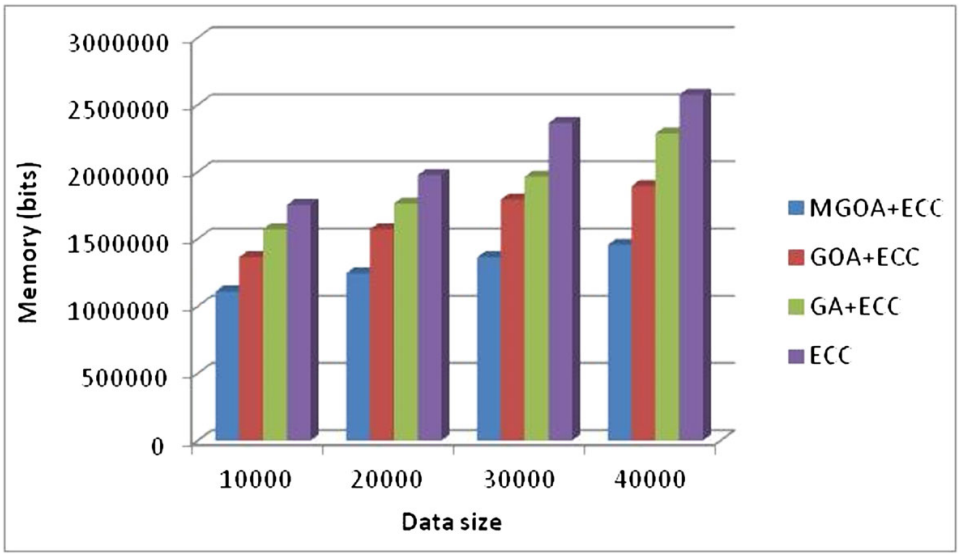



**Fig. 7** Performance comparison of proposed and existing methods with memory based on data size

**Table 5** Comparison of the proposed method and existing methods with time for based on mapper

| Number of mappers | Time (ms) MGOA+ECC | Time (ms) GOA+ECC | Time (ms) GA+ECC | Time (ms) ECC |
|---|---|---|---|---|
| 5 | 156,984 | 185,436 | 194,723 | 226,648 |
| 10 | 137,573 | 174,636 | 184,747 | 205,815 |
| 15 | 124,785 | 148,765 | 169,648 | 198,538 |
| 20 | 106,545 | 139,473 | 158,386 | 177,285 |

MGOA+ECC is 156,984 ms, under GOA + ECC is 185,436 ms, under GA + ECC is 194,723 ms and the time under ECC is 226,648 ms. When the data size is 30,000, the time under MGOA+ECC is 214,785 ms, under GOA + ECC is 248,765 ms, under GA + ECC is 259,648 ms and the time under ECC is 278,538 ms. When the data size is 40,000, the time under MGOA+ECC is 286,545 ms, under GOA + ECC is 359,473 ms, under GA + ECC is 378,386 ms and the time under ECC is 407,285 ms. The proposed MGOA+ECC method has lower time than all the other compared methods in all the data sizes considered for experimentation. The time consumption has to be low for a method to be efficient. Therefore, from the following observation, it is shown that the proposed MGOA+ECC is more efficient than the compared existing methods. The graphical representation of this performance analysis based on time consumption is shown in Fig. 6.

Table 4 shows the Performance comparison of MGOA+ECC, GOA + ECC, GA + ECC and ECC with memory. The memory under the mentioned techniques for various data sizes is tabulated in below.

The comparison for the memory under different techniques such as MGOA+ECC, GOA + ECC, GA + ECC and ECC is shown in Table 4. When the data size is 10,000, the memory

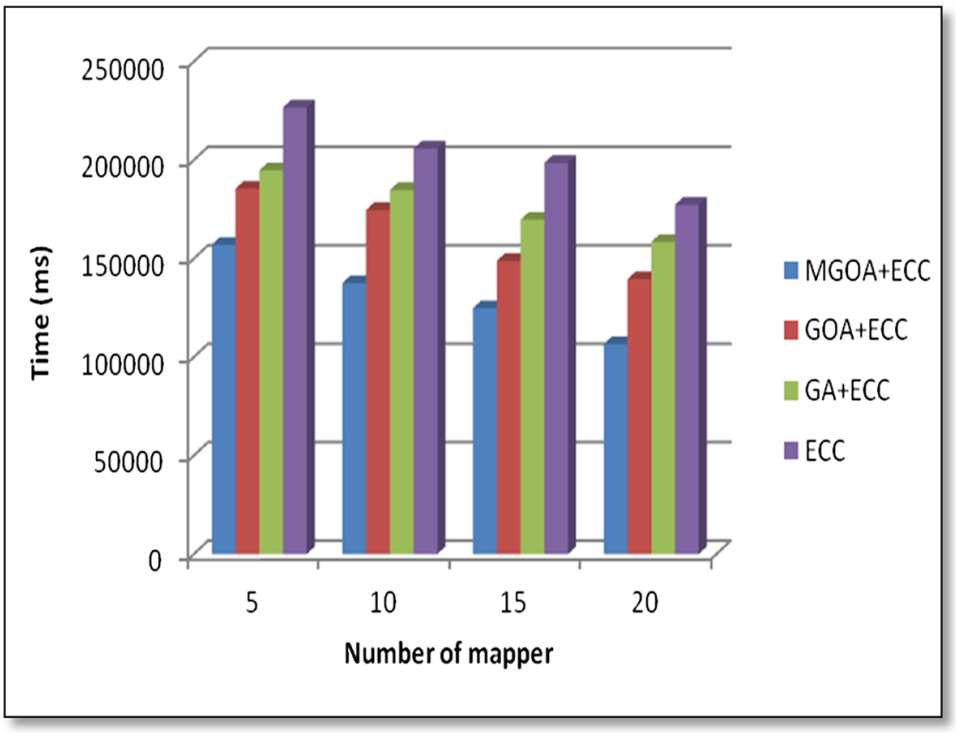



**Fig. 8** Performance comparison of proposed and existing methods with time-based on numbers of mapper

**Table 6** Comparison of the proposed method and existing methods with memory based on mapper

| Number of mappers | Memory (bits) MGOA+ECC | Memory (bits) GOA + ECC | Memory (bits) GA + ECC | Memory (bits) ECC |
|---|---|---|---|---|
| 5 | 1,011,745 | 1,465,743 | 1,774,362 | 1,954,382 |
| 10 | 1,345,787 | 1,674,362 | 1,865,285 | 2,076,431 |
| 15 | 1,464,575 | 1,895,421 | 2,065,473 | 2,465,471 |
| 20 | 1,557,745 | 1,996,542 | 2,387,452 | 2,676,352 |

under MGOA+ECC is 1,111,745 bits, under GOA + ECC is 1,365,743 bits, under GA + ECC is 1,574,362 bits and the memory under ECC is 1,754,382 bits. When data size is 20,000, the memory under MGOA+ECC is 1,245,787 bits, under GOA + ECC is 1,574,362 bits, under GA + ECC is 1,765,285 bits and the memory under ECC is 1,976,431 bits. When the data size is 30,000, the memory under MGOA+ECC is 1,364,575 bits, under GOA + ECC is 1,795,421 bits, under GA + ECC is 1,965,473 bits and the memory under ECC is 2,365,471 bits. When the data size is 40,000, the memory under MGOA+ECC is 1,457,745 bits, under GOA + ECC is 1,896,542 bits, under GA + ECC is 2,287,452 bits and the memory under ECC is 2,576,352 bits. The proposed MGOA+ECC method utilizes lower memory than all the other compared methods in all the data sizes taken into account. The memory utilization has to be low for a method to be efficient. Therefore, from the observation, it is established that the suggested MGOA+ECC method is more efficient than the compared existing methods. The graphical representation of this performance analysis based on memory consumption is shown in Fig. 7.

### 4.2.2 Comparison analysis of a recommended method based on mapper

The comparison of time consumption under various techniques such as MGOA+ECC, GOA + ECC, GA + ECC and ECC for a varying number of mappers is shown in Table 5. When the number of the mapper is 5, the time under MGOA+ECC is 156,984 ms, under GOA + ECC is

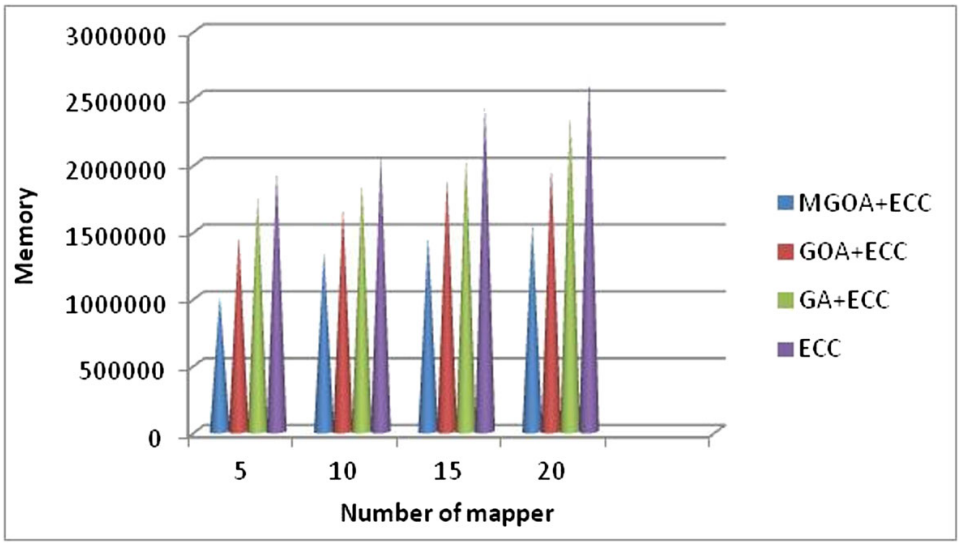



**Fig. 9** Performance comparison of proposed and existing methods with memory based on the number of mappers

**Table 7** Comparison of the proposed method and existing methods with time-based on cluster sizes

| Cluster size | Time (ms) MGOA+ECC | Time (ms) GOA + ECC | Time (ms) GA + ECC | Time (ms) ECC |
|---|---|---|---|---|
| 4 | 106,545 | 139,473 | 158,386 | 177,285 |
| 6 | 124,785 | 148,765 | 169,648 | 198,538 |
| 8 | 137,573 | 174,636 | 184,747 | 205,815 |
| 10 | 156,984 | 185,436 | 194,723 | 226,648 |

185,436 ms, under GA + ECC is 194,723 ms and the encryption time under ECC is 226,648 ms. This order of time is the same for all the number of mappers.

The graphical representation of this performance analysis based on time for the variable number of mappers is shown in Fig. 8.

The Performance comparison of MGOA+ECC, GOA + ECC, GA + ECC and ECC methods with memory is shown in Table 6. The memory under the mentioned techniques for various quantities of mappers is depicted in the table.

The proposed MGOA+ECC method takes lesser time than all the other compared methods in all the number of mappers taken into account. The time consumption has to be low for a method to call it as an efficient one. Therefore, from the observation, it is complete that the suggested MGOA+ECC method has performed much better than when compared with the existing methods.

The comparison of memory under various techniques such as MGOA+ECC, GOA + ECC, GA + ECC and ECC for a various number of mappers is shown in Table 6. When the number of the mapper is 10, the memory under MGOA+ECC is 1,345,787 bits, under GOA + ECC is 1,674,362 bits, under GA + ECC is 1,865,285 bits and the memory under ECC is 2,076,431 bits. This order of memory utilization value is same for all the number of the mapper.

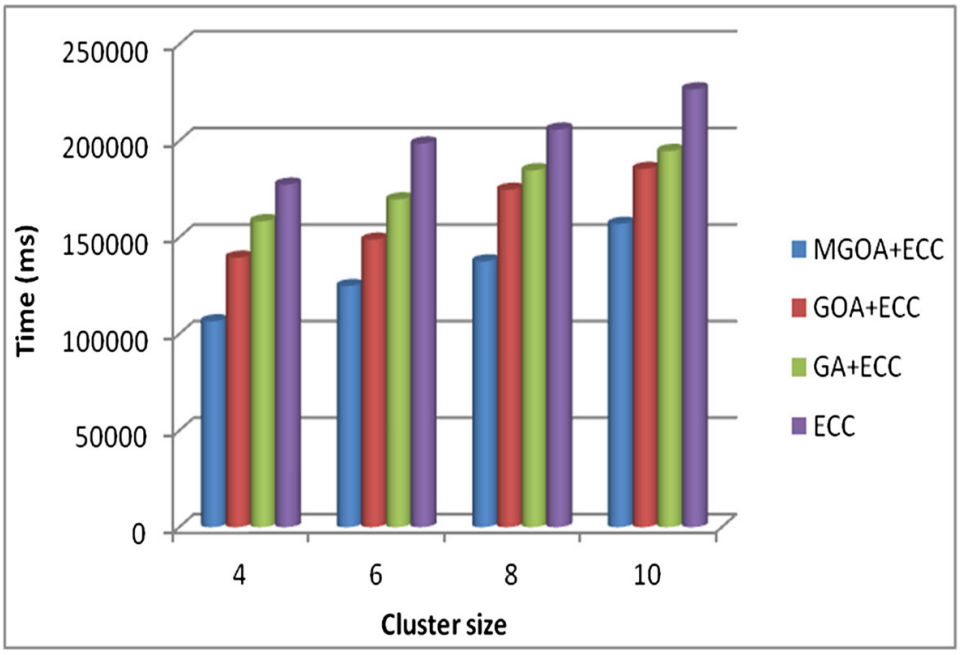



**Fig. 10** Performance comparison of proposed and existing methods with time-based on cluster size

**Table 8** Comparison of the proposed method and existing methods with memory based on cluster sizes

| Cluster size | Memory (bits) MGOA+ECC | Memory (bits) GOA + ECC | Memory (bits) GA + ECC | Memory (bits) ECC |
|---|---|---|---|---|
| 4 | 1,111,745 | 1,446,786 | 1,874,362 | 1,908,643 |
| 6 | 1,317,383 | 1,777,362 | 1,965,285 | 2,076,431 |
| 8 | 1,415,869 | 1,895,421 | 2,065,473 | 2,465,471 |
| 10 | 1,531,727 | 2,096,542 | 2,387,452 | 2,676,352 |

Figure 9 shows the graphical representation of this performance analysis based on memory. The proposed MGOA+ECC method utilizes lower memory than all the other compared methods in all the number of mapper taken into account. For a method to be efficient, the memory utilization has to be low. Therefore, from the observation, it is demonstrated that the suggested MGOA+ECC method is more efficient than the compared existing methods.

### 4.2.3 Comparison analysis of a recommended method based on cluster size

Table 7 shows the Performance comparison of MGOA+ECC, GOA + ECC, GA + ECC and ECC methods with time. The time under the mentioned techniques for various cluster sizes is depicted in the table.

The comparison of time consumption under various techniques such as MGOA+ECC, GOA + ECC, GA + ECC and ECC for various sizes of the cluster is shown in Table 7. When the cluster size is 10, the time under MGOA+ECC is 156,984 ms, under GOA + ECC is 185,436 ms, under GA + ECC is 194,723 ms and the time under ECC is 226,648 ms. This order of time is the same for all the cluster sizes. The proposed MGOA+ECC method takes lesser time than all the other compared methods in all the sizes of cluster taken into account. The time consumption has to be low for a method to call it as an efficient one. Therefore, from the observation, it is shown that the proposed MGOA+ECC method has performed much

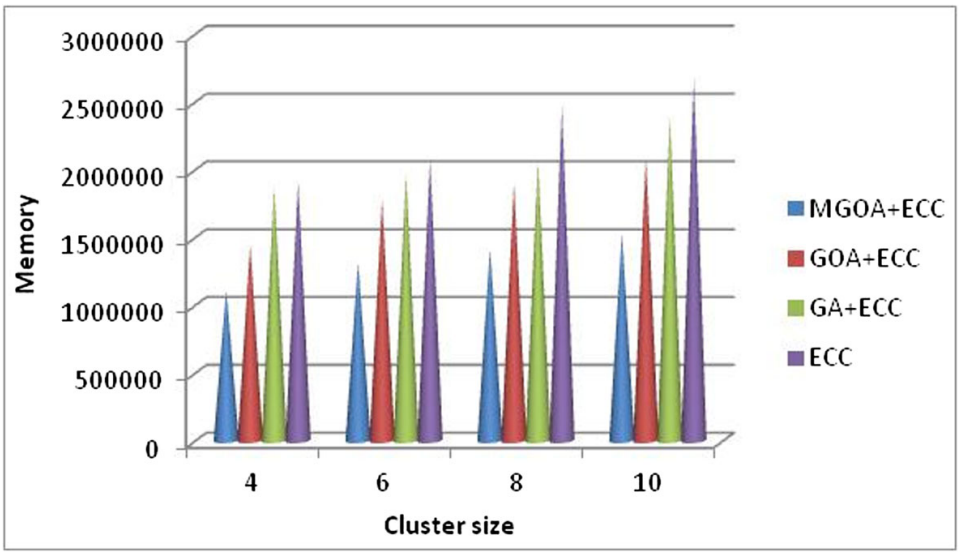



**Fig. 11** performance comparison of proposed and existing methods with decryption time based on cluster size

**Table 9** Comparison of the proposed method and existing methods

| Data size (MB) | Time (ms) | | Memory (bits) | |
|---|---|---|---|---|
| | Proposed (MGOA+ECC) | Existing (VSSFA) | Proposed (MGOA+ECC) | Existing (VSSFA) |
| 10,000 | 113,485 | 254,684 | 1,111,745 | 1,258,462 |

better than the existing methods. The graphical representation of this performance analysis based on time for different cluster sizes is shown in Fig. 10.

The Performance comparison of MGOA+ECC, GOA + ECC, GA + ECC and ECC methods with memory is shown in Table 8. The memory under the mentioned techniques for various sizes of the cluster is depicted in the table.

The comparison of memory under various techniques such as MGOA+ECC, GOA + ECC, GA + ECC and ECC for various cluster sizes is shown in Table 8. When the cluster size is 8, the memory under MGOA+ECC is 1,415,869 bits, under GOA + ECC is 1,895,421 bits, under GA + ECC is 2,065,473 bits and the memory under ECC is 2,465,471 bits. This order of memory utilization value is same for all the sizes of the clusters.

Figure 11 shows the graphical representation of this performance analysis based on memory. The proposed MGOA+ECC method utilizes memory very much lower than all the other compared methods in all the sizes of clusters taken into account. For a method to be efficient, the memory utilization has to be low. From the observation, it is clear that the suggested MGOA+ECC method is more efficient than the compared existing methods.

All the comparative analysis from Tables 1, 2, 3, 4, 5, 6, 7 and 8 and from Figs. 4, 5, 6, 7, 8, 9, 10 and 11 have shown that the proposed MGOA+ECC method has performed much better than the existing methods such as GOA + ECC, GA + ECC and ECC which were taken for comparison.

Here also, the suggested method is compared with the existing research [19]. In existing research (VSSFA) [19], secure storage of big data is done by varying step size firefly algorithm; the comparison results are tabulated beneath (Table 9),

When analyzing the above table, it is clear that the suggested method takes less time and memory utilization for 10,000 data size when compared to the existing research [19], here the proposed method takes 113,485 ms for 10,000 data, but the existing method takes 254,684 ms which is higher when compared to the proposed method. The memory utilization of the proposed method is 1,111,745 bits for 10,000 data, but the existing research [19] utilizes 1,258,462 bits which are a minimum value when equated to the existing method. Table 10 compares the time and memory usage against the mapper size of 20. From the Table, it is observed that the proposed method is efficient.

**Table 10** Comparison of the proposed method and existing methods by varying mapper

| Mapper size | Time (ms) | | Memory (bits) | |
|---|---|---|---|---|
| | Proposed (MGOA+ECC) | Existing (VSSFA) | Proposed (MGOA+ECC) | Existing (VSSFA) |
| 20 | 106,545 | 218,652 | 1,557,745 | 2,500,000 |

**Table 11** Comparison of the proposed method and existing methods with encryption time based on data size

| Data size | Encryption time (ms) MGOA+ECC | Encryption time (ms) GOA + ECC | Encryption time (ms) GA + ECC | Encryption time (ms) ECC |
|---|---|---|---|---|
| 10,000 | 11,245 | 13,764 | 15,835 | 17,363 |
| 20,000 | 16,547 | 18,544 | 20,748 | 22,857 |
| 30,000 | 21,444 | 23,765 | 25,745 | 27,369 |
| 40,000 | 26,574 | 27,564 | 18,574 | 21,463 |

The proposed method takes minimum time and less memory utilization because the proposed method uses an efficient method for secure storage.

### 4.2.4 Comparison analysis of a recommended method based on encryption and decryption time

Table 11 shows the Performance comparison of MGOA+ECC, GOA + ECC, GA + ECC and ECC with encryption time. The encryption time under the mentioned techniques for various data sizes is depicted in the table.

The comparison for encryption time under various techniques such as MGOA+ECC, GOA + ECC, GA + ECC and ECC is shown in Table 11. When the data size is 10,000, the encryption time under MGOA+ECC is 11,245 ms, under GOA + ECC is 13,764 ms, under GA + ECC is 15,835 ms and the encryption time under ECC is 17,363 ms. This order of encryption time is the same for most of the data sizes. The proposed MGOA+ECC method takes lesser time to encrypt than all the other compared methods in most of the data sizes taken into account. The encryption time has to be low for a method to be efficient. Therefore from the observation, it is clear that the proposed MGOA+ECC method is more efficient than the compared existing methods. The graphical representation of this performance analysis based on encryption time is shown in Fig. 12.

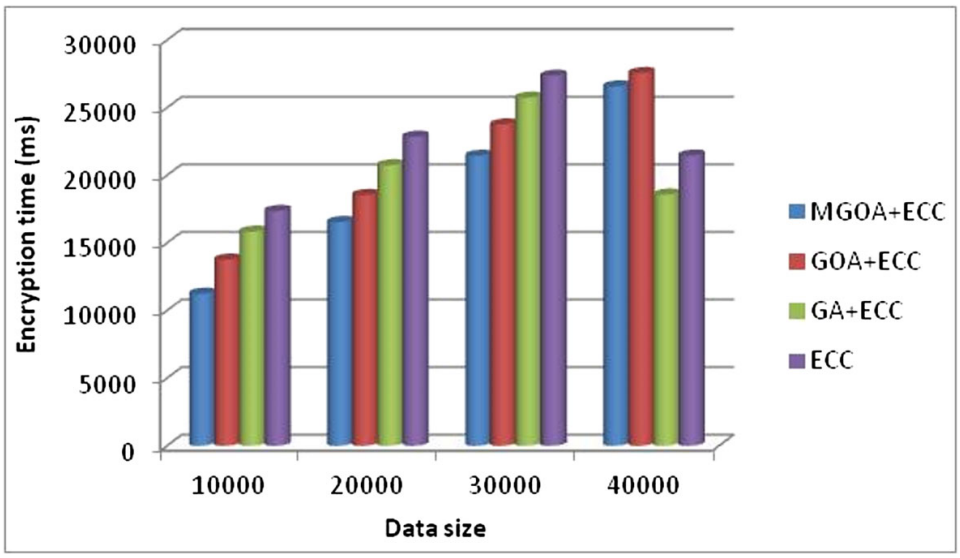



**Fig. 12** Performance comparison of proposed and existing methods with encryption time based on data size

**Table 12** Comparison of the proposed method and existing methods with decryption time based on data size

| Data size | Decryption time (ms) MGOA+ECC | Decryption time (ms) GOA + ECC | Decryption time (ms) GA + ECC | Decryption time (ms) ECC |
|---|---|---|---|---|
| 10,000 | 8954 | 10,647 | 12,574 | 14,257 |
| 20,000 | 9654 | 12,464 | 14,734 | 16,478 |
| 30,000 | 10,245 | 14,647 | 16,346 | 18,568 |
| 40,000 | 11,254 | 15,747 | 17,357 | 20,575 |

The Performance comparison of MGOA+ECC, GOA + ECC, GA + ECC and ECC with decryption time is shown in Table 12. The decryption time under the mentioned techniques for various data sizes is depicted in the table.

The comparison for decryption time under various techniques such as MGOA+ECC, GOA + ECC, GA + ECC and ECC is shown in Table 12. When the data size is 40,000, the decryption time under MGOA+ECC is 11,254 ms, under GOA + ECC is 15,747 ms, under GA + ECC is 17,357 ms and the decryption time under ECC is 20,575 ms. This order of decryption time is the same for all the data sizes. The proposed MGOA+ECC method takes lesser time to decrypt than all the other compared methods in all the data sizes taken into account. For a method to be efficient the decryption time has to be low. Therefore, from the observation, it is clear that the proposed MGOA+ECC method is more efficient than the compared existing methods. The graphical representation of this performance analysis based on decryption time is shown in Fig. 13.

## 5 Conclusion

In this work, secure storage of big data and effective access control mechanism is proposed. The implementation is done using cloudsim with Hadoop Map-reduce framework. It is analyzed that the proposed technique outperforms in terms of time and memory usage when

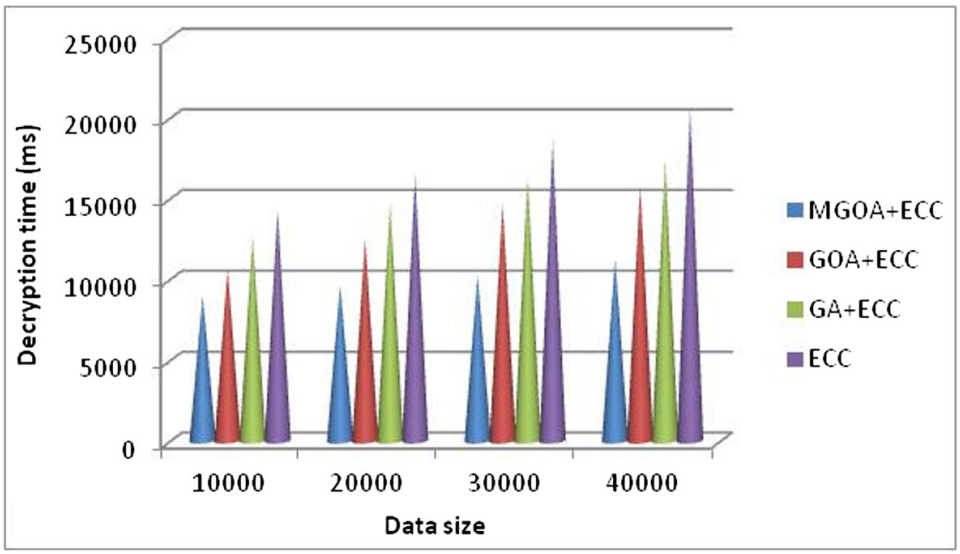



**Fig. 13** Performance comparison of proposed and existing methods with decryption time based on data size

compared to the current techniques. This is due to the fact that proposed technique uses the optimal ECC with the assistance of MGOA calculation for secure storage. In our future research, we can enhance our plan by merging the secret sharing with chosen attribute based access control systems. We will also explore the security issues when a data owner distributes its data to multi-cloud servers in big data.

## Compliance with ethical standards

**Conflict of interest** we as authors declares that there is no conflict of interest.

## References


1. Assunção MD, Calheiros RN, Bianchi S, Netto MA, Buyya R (2015) Big data computing and clouds: trends and future directions. J Parallel Distrib Comput 79:3–15
2. Batista BG, Ferreira CH, Segura DC, Leite Filho DM, Peixoto ML (2017) A QoS-driven approach for cloud computing addressing attributes of performance and security. Futur Gener Comput Syst 68:260–274
3. Bui DM, Yoon Y, Huh EN, Jun S, Lee S (2017) Energy efficiency for cloud computing system based on predictive optimization. J Parallel Distrib Comput 102:103–114
4. Chen G, Wu S, Wang Y (2015) The evolvement of big data systems: from the perspective of an information security application. Big Data Research 2(2):65–73
5. UCI Machine Learning Repository (1999), from http://archive.ics.uci.edu/ml/datasets/kdd+cup+1999+data.
6. Fugkeaw S, Sato H (2018) Scalable and secure access control policy update for outsourced big data. Futur Gener Comput Syst 79:364–373
7. Hashem IA, Yaqoob I, Anuar NB, Mokhtar S, Gani A, Khan SU (2015) The rise of "big data" on cloud computing: review and open research issues. Inf Syst 47:98–115
8. Hu C, Cheng X, Tian Z, Yu J, Akkaya K, Sun L (2015) An attribute-based sign cryption scheme to secure attribute-defined multicast communications. In: Secure comm, pp 418–437
9. Kindervag J, Balaouras S, Hill B, Mak K (2012) Control and protect sensitive information in the era of big data, pp 1–7
10. Kshetri N (2014) Big data′ s impact on privacy, security and consumer welfare. Telecommun Policy 38: 1134–1145
11. Li P, Guo S, Miyazaki T, Xie M, Hu J, Zhuang W (2016) Privacy-preserving access to big data in the cloud. IEEE Cloud Comput 5:34–42
12. Nigoti R, Jhuria M, Singh S (2013) A survey of cryptographic algorithms for cloud computing. J Emerging Tech Comp App Sci 4:141–146
13. Perera C, Ranjan R, Wang L, Khan SU, Zomaya AY (2015) Big data privacy in the internet of things era. IT Prof 17:32–39
14. Ramachandran M (2016) Software security requirements management as an emerging cloud computing service. Int J Inf Manage 4:580–590
15. Shanmugapriya E, Kavitha R (2019) Medical big data analysis: preserving security and privacy with hybrid cloud technology. Soft Comput. https://doi.org/10.1007/s00500-019-03857-z
16. Shen J, Wang C, Wang A, Ji S, Zhang Y (2018) A searchable and verifiable data protection scheme for scholarly big data. IEEE Emerg Top Com 14:8
17. Singh S, Jeong YS, Park JH (2016) A survey on cloud computing security: issues, threats, and solutions. J Netw Comput Appl 75:200–222
18. Tan Z, Nagar UT, He X, Nanda P, Liu RP, Wang S, Hu J (2014) Enhancing big data security with collaborative intrusion detection. IEEE Cloud Comput 3:27–33
19. Thiyagarajan VS, Ayyasamy A (2017) Privacy preserving over big data through VSSFA and MapReduce framework in cloud environment. Wirel Pers Commun 4:6239–6263
20. Wang Z, Cao C, Yang N, Chang V (2017) ABE with improved auxiliary input for big data security. J Comput Syst Sci 89:41–50

21. Xiong H, Choo KK, Vasilakos AV (2017) Revocable identity-based access control for big data with verifiable outsourced computing. IEEE Trans Big Data 6:1
22. Yan Z, Ding W, Yu X, Zhu H, Deng RH (2016) Deduplication on encrypted big data in cloud. IEEE Trans Big Data 2:138–150
23. Yan Z, Deng RH, Varadharajan V (2017) Cryptography and data security in cloud computing. Inf Syst 387: 53–55
24. Yang LT, Huang G, Feng J, Xu L (2017) Parallel GNFS algorithm integrated with parallel block Wiedemann algorithm for RSA security in cloud computing. Inf Sci 387:254–265
25. Yang K, Han Q, Li H, Zheng K, Su Z, Shen X (2017) An efficient and fine-grained big data access control scheme with privacy-preserving policy. IEEE Internet Things J 2:563–571
26. Yang Y, Zheng X, Guo W, Liu X, Chang V (2018) Privacy-preserving smart IoT-based healthcare big data storage and self-adaptive access control system. Inf Sci:1–26
27. Zhou W, Feng D, Tan Z, Zheng Y (2018) Improving big data storage performance in hybrid environment. J Comput Sci 26:409–418

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Om Prakash Verma** received M.Tech (Communication and Radar Engineering) from Indian Institute of Technology (IIT), Delhi, in 1996 and Ph.D from University of Delhi, Delhi, in 2011. He is Professor in Department of Electronics & Communication Engineering, Delhi Technological University, Delhi. Prior to this, he was the head of Department of Computer Science and Engineering, DTU, Delhi. He has recently published two papers in IEEE Transactions. He is also the author of more than 40 publications in both conference proceeding and journal. He has authored a book on Digital Signal Processing. His research interests include: Image processing, application of fuzzy logic in image processing, application of evolutionary algorithm in image processing, artificial intelligent and digital signal processing.



**Nitin Jain** received his Masters Degree in Computer Science and Engineering from University School of Information Technology (USIT), Delhi in 2012 and now pursuing Ph.D from Delhi Technological University (DTU). He has worked as a research intern at Scientific Analysis Group DRDO, Delhi. His area of interest lies in the field of public key cryptography, homomorphic encryption and cloud computing.



**Saibal Kumar Pal** completed his Post-Graduation in Computer Science from University of Allahabad & Ph.D from University of Delhi. Dr. Pal is working as Director, Directorate of Information Technology & Cyber Security and CISO. Prior to this, he worked as Scientist 'G' at Scientific Analysis Group DRDO, Delhi. His area of interest includes cryptography, multimedia & network security, computational intelligence & data mining. He is in the editorial board & review / program committee of a number of journals & international conferences. He has over 200 publications in science, technology & management studies. He was conferred the DRDO Scientist of the Year award in 2012.

2019 6th International Conference on Signal Processing and Integrated Networks (SPIN)

# Design of Low Power ECRL based Power Gated 4:2 Compressor

Naman Malhotra
*Department of Electronics and Communication*
*Delhi Technological University*
Delhi, India
namanmalh_bt2k15@dtu.ac.in

Aarushi Mann
*Department of Electronics and Communication*
*Delhi Technological University*
*Delhi, India*
aarushi_bt2k15@dtu.ac.in

Neeta Pandey
*Department of Electronics and Communication*
*Delhi Technological University*
*Delhi, India*
n66pandey@rediffmail.com

***Abstract*— Advent of new age technologies like augmented reality (AR) has made techniques like digital signal processing (DSP), image and speech processing one of the most frequently used and important techniques. Multipliers are the fundamental components of these techniques which depend on extensive multiplications and numerical manipulations. Compressors are a crucial and important unit of multipliers that largely influence the speed and power dissipation of multipliers. This paper proposes a low power ECRL based 4:2 compressors by employing power gating. Two different power gating methodologies have been evaluated and compared, at varying load capacitances, supply voltage and frequency of the power clock signal, using 32nm PTM CMOS technology parameters. The small overhead observed in active state power dissipations is well compensated for by the significant power savings in idle states.**

***Keywords*—Compressor, Low Power VLSI, ECRL, Power Gating, Adiabatic Logic**

## I. INTRODUCTION

In today's world each person carries a smart phone, that is, one of the most advanced electronic devices ever; right in their pockets at all times [1]. This has increased accessibility to advanced technology, such as 3D graphics, augmented reality(AR) and Bokeh effect photography, for a layman. These technologies make use of various techniques like digital signal processing, image processing, filtering and speech processing, which in turn are based upon various numerical manipulations especially multiplications, hence making low power multipliers essential for development in the future. Compressors are the most important part of multipliers that significantly influence its speed and power dissipation [2]. Hence low power compressors are the need of the hour.

Multipliers are used in real time signal processing, hence along with low power capabilities; they are designed to work at high frequencies so as to perform fast computations. CMOS logic family, one of the most popular logic families for integrated circuits, suffers from large power dissipation at high frequencies, because of its dynamic (switching) power dissipation being directly proportional to the operating frequency [3]. Adiabatic logic family, working on the principle of charge recovery from load capacitance, is a power efficient approach to design compressors for high frequency operation [1]. Amongst various adiabatic logic families like Positive Feedback Adiabatic Logic (PFAL) [4], Diode Free Adiabatic Logic (DFAL) [5], Modified Positive Feedback Adiabatic Logic (MPFAL) [6] and Split- Rail Charge Recovery Logic (SCRL) [7]; Efficient Charge Recovery Logic (ECRL)[8] has been employed in this work for its significant improvement in power dissipation and relative simplicity.

Power efficiency of adiabatic compressors can be increased still, by switching them off when they are not being actively used this is achieved by employing power gating i.e. isolating the circuit from one of the power rails [9]. In this paper, two power gating methodologies, namely, header and footer [10,11], have been used with ECRL based 4:2 compressors to further minimize its power dissipation in order to achieve low power compressors which would result in low power multipliers, enabling increased amounts mathematical manipulations while dissipating less energy which is required for modern day technology.

This paper is organized in six sections including the first introductory one. Section II briefly describes the adiabatic logic style and further reduction in power dissipation using power gating. The compressor design is briefly explained in section III. The functional verification of the design and the performance comparison is presented section IV. The conclusions are drawn in the next section followed by references in the last section.

## II. ADIABATIC LOGIC AND POWER GATING

CMOS circuits have been an integral part of the modern day VLSI devices, but still suffer from various power losses. CMOS logic suffers from primarily three power losses, static (leakage), direct path and dynamic(switching) losses, with the latter dominating the overall power dissipation in CMOS devices. Dynamic power is dissipated while charging and discharging of the load capacitance through pull up and pulls down networks from $V_{DD}$ and to ground nodes respectively, it is also known as active loss [3]. Active losses are minimised by adiabatic circuits. The principle of charge recovery is not possible in the CMOS circuits which make it difficult to save power lost during active switching [1].

Adiabatic logic is an efficient way to minimise power losses using charge recovery from the load capacitance by making use of a four phase power supply. The four phase power supply ensures that there is no abrupt switching between logic levels instead the switching takes place with a finite rise and fall time, hence minimizing dynamic power dissipation [12]. The four phases of the four phase power clock supply are evaluate, hold, recovery and wait.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE
978-1-7281-1380-7/19/$31.00 ©2019 IEEE

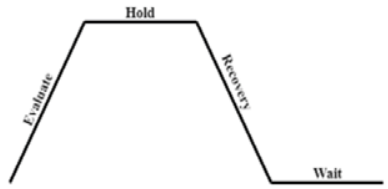



**Fig.1. Four phase power supply**

Further reduction in power losses in adiabatic logic can be achieved by various methods such as transistor sizing, decreasing the feature size or by employing power gating methodologies [13]. In this work, power gating has been applied to ECRL adiabatic logic family due to its coherent nature. Power gating is established by disconnecting the network from either of the two power rails ground (0V) or power supply ($V_{CLK}$) using sleep state control transistors [10]. The sleep state control transistors are responsible for switching the circuit between awake(functioning) and sleep(low power) states.

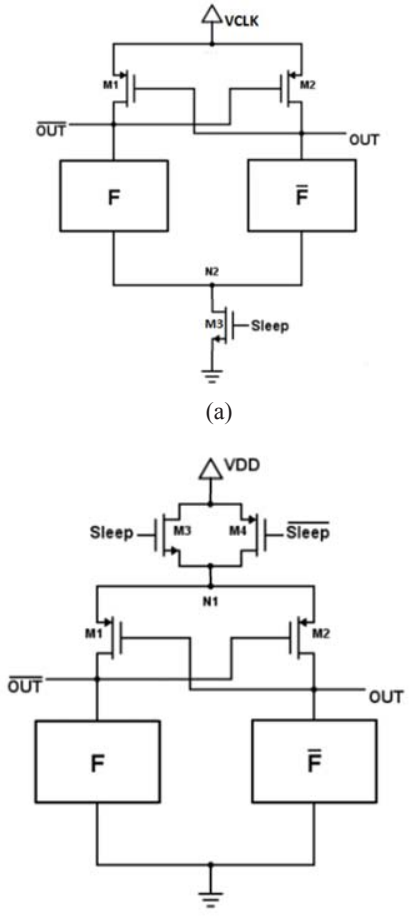



**Fig.2. General Cell of(a)Footer Gated and (b)Header Gated ECRL family**

Figure 2(a) shows the general cell of the ECRL family with footer power gating, as seen in the figure the two logic block F and its complement F̄act as the pull down networks for complementary output nodes $\overline{OUT}$and OUT respectively, also the sleep state control unit employed in footer gating is only a NMOS M3 because it needs to pass only a strong low (0V) to node N2 in awake state(sleep signal is high) [10]. Absence of discharging path, in sleep state (sleep signal is low), for the load capacitances leads to both output nodes remaining in a floating state.

Similarly, as seen in Fig.2(b)header gating is performed by isolating the cross coupled PMOSs M1 and M2 from the power supply ($V_{CLK}$) by using a transmission gate (TG) M3 and M4 as a sleep state control unit, as it required to pass both the strong low and strong high components of the four phase power clock. Absence of charging path for the load capacitances leads to both outputs remaining low in sleep state.

## III. 4:2 COMPRESSOR

Compressor is a crucial unit used in optimizing partial product addition aiding the complex design of multiplier circuits. The data computation units embedded in modern day electronic circuits require multipliers and multiplication process takes place in three steps, namely, partial product generation, partial product accumulation, and final addition [14,15].The path of addition is the most time consuming thereby requiring more power to compute and evaluate data. Thus compressors are utilised to minimise this power and delay requirements [16]. Compressors are needful in minimising data i.e. compressing data. A 4:2 Compressor has 5 inputs, including 4 input bits and one carry in ($C_{in}$) and has 3 outputs i.e. sum, carry out ($C_{out}$) and final carry (Carry) [17].

The basic sub cells of a compressor are exclusive-OR (XOR) gates [18]and 2:1 multiplexer (MUX) [19, 20]. All other higher blocks of compressors are easily realisable using cascaded units of 4:2 compressors. The functionality of this Compressor can be expressed as follows:

$$Sum = I1 \oplus I2 \oplus I3 \oplus I4 \oplus Cin \quad (1)$$

$$Carry = (I1 \oplus I2 \oplus I3 \oplus I4).Cin + \overline{(I1 \oplus I2 \oplus I3 \oplus I4)}.I4 \quad (2)$$

$$Cout = (I1 \oplus I2).I3 + \overline{(I1 \oplus I2)}.I1 \quad (3)$$

Where, I1, I2, I3, I4 and Cin are the inputs to the 4:2 compressors.

Figure 3 depicts the block diagram of a 4:2 compressor[21] that uses, as shown it uses four XOR gates (X1 block)and two 2:1 MUX (X2 block) as fundamental blocks. The functional equation of a 2:1 MUX is expressed as:

$$OUT = A.\bar{S} + B.S \quad (4)$$

Where, A and B are the inputs to the 2:1 MUX and S is the select input.

The functional equation of XOR and XNOR gates are expressed as:

$$XOR = A \oplus B = A.\bar{B} + \bar{A}.B \quad (5)$$

$$XNOR = A \odot B = A.B + \bar{A}.\bar{B} \quad (6)$$

Where, A and B are the inputs to the XOR and XNOR gates

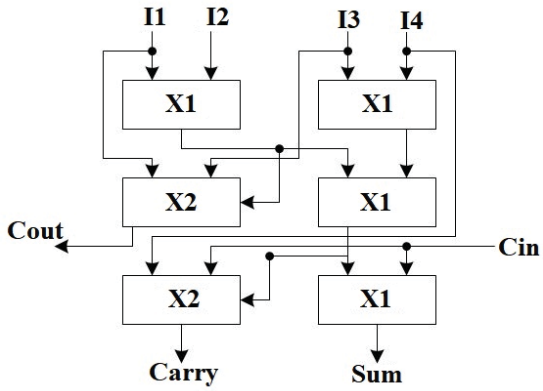



**Fig.3. Block Diagram of 4:2 Compressor**

The truth table of the 4:2 Compressor has been shown in Table I [17]. To further decrease the power dissipation in the designed compressor, power gating has been employed. The compressor is designed using both header and footer gating methodology.

Table I: Truth Table of 4:2 Compressor [17]

| $C_{in}$ | $X_4$ | $X_3$ | $X_2$ | $X_1$ | $C_{out}$ | Carry | Sum |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 |
| 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 1 | 0 | 1 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 |
| 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 |
| 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 |
| 0 | 1 | 1 | 1 | 0 | 1 | 0 | 1 |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 |
| 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 |
| 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 |
| 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 |
| 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 |
| 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 |
| 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

## IV. SIMULATION RESULTS

This section consists of the functional verification and the power evaluation results of ECRL based 4:2 compressor with different gating techniques. The 32nm technology node has been considered for all the simulations.

### *A. Functional Verification*

Figure 4 shows the time domain simulation for an ECRL based 4:2 compressor with a load capacitance of 10fF, supply voltage of 1V and a power clock frequency of 200MHz.The functionality of the 4:2 Compressor has been verifiedusing the truth table mentioned in Table I. The five available inputs (I1, I2, I3, I4 and Cin) have been successfully compressed into the three outputs (Cout, Sum and Carry).

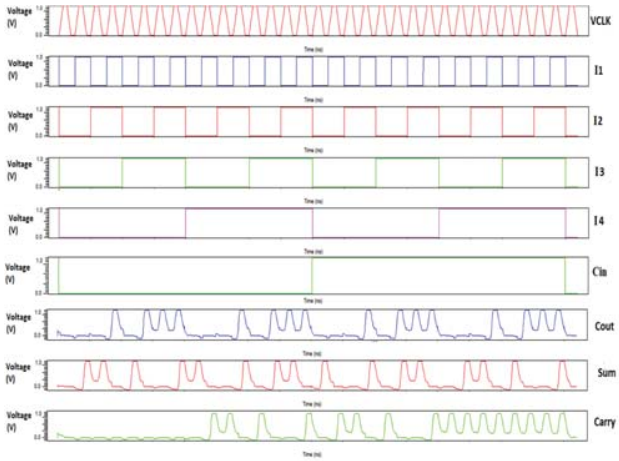



**Fig.4. Time Domain Simulation of ECRL based 4:2 Compressor (a) Power clock (b) I1, (c)I2, (d) I3, (e) I4, (f) Cout, (g) Sum, (h)Carry waveform**

The time domain simulation of footer gated ECRL based 4:2 compressor performed with a load capacitance of 10fF, supply voltage of 1V and a power clock frequency of 200MHz is shown in Fig.5. As seen in the simulation the correct functioning of the compressor is obsereved in the awake state (when sleep signal is high) and the outputs are floating in sleep state (when sleep signal is low). Due to lack of discharging path the floating outputs eventually attain a value of $V_{DD}$ - $|Vtp|$, where $V_{tp}$ is the threshold voltage of the PMOS.

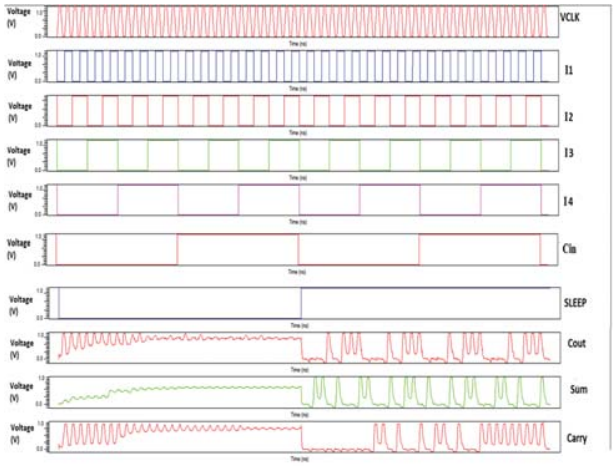



**Fig.5. Time Domain Simulation of ECRL based footer gated4:2 Compressor(a) Power b) I1, (c) I2, (d) I3, (e) I4, (f) Sleep signal, (g) Cout, (h)Sum, (i) Carry waveform**

The time domain simulation of header gated 4:2 compressor with a load capacitance of 10fF, supply voltage of 1V and a power clock frequency of 200MHz is shown in Fig.6. As seen from the simulation the correct functioning of the compressor is obsereved in the awake state (when sleep signal is high) and the outputs are discharged in sleep state (when sleep signal is low) due to absence of a charging path.

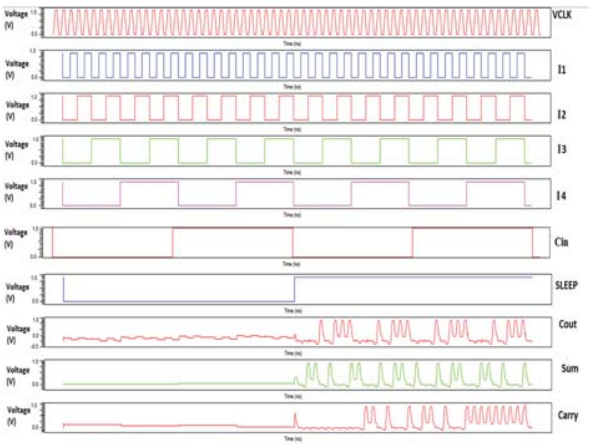



**Fig.6. Time Domain Simulation of ECRL based header gated4:2 Compressor(a) Power clock (b) I1, (c) I2, (d) I3, (e) I4, (f) Sleep signal, (g) Cout, (h)Sum, (i) Carry waveform**

### *B. Comparison of Power Dissipation at varying Load Capacitances*

The power evaluations have been performed at constant power clock frequency of 200MHz and supply voltage of 1V for a range of load capcitance from 10fF to 50fF.The reults have been compiled in a graph shown in Fig.7.

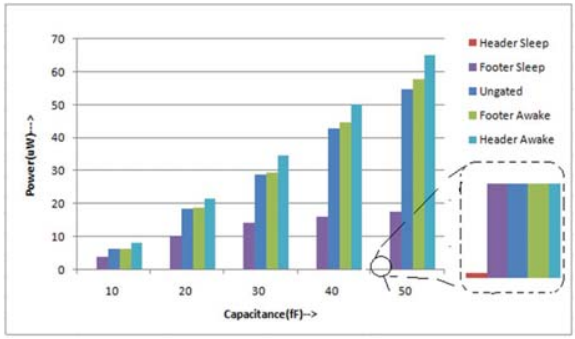



**Fig.7. Power Dissipation vs Load Capacitance**

The power dissipation of all circuits increases with increase in load capacitance. Small overhead power dissipation is observed in gated awake circuits due to more number transistors in the circuits. The header sleep circuit maintains a negligible power dissipation throughout the range of load capacitance as shown in the magnified portion (dotted) of the graph. The overhead observed in footer active circuits is less than that of the header gated however there is notable power dissipation in sleep state of footer gated circuits.

.

### *C. Comparison of Power Dissipation at varying Supply Voltages*

The power evaluations have performed at a constant load capacitance of 10fF and power clock frequency of 200MHz for a range of supply voltage from 0.7V to 1.2V. The results have been compiled in a graph as shown in Fig.8

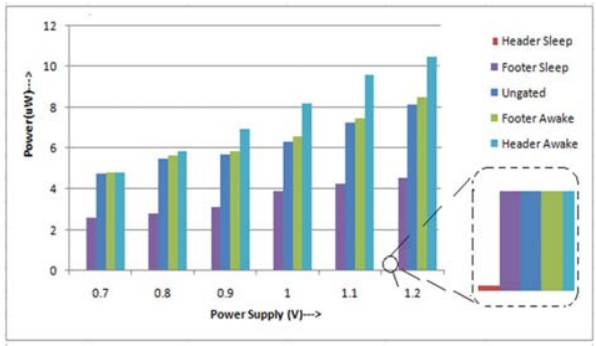



**Fig.8. Power Dissipation VS Supply Voltage**

The graph clearly shows an increase in power dissipation with increase in supply voltage. The power dissipation for gated awake circuits is more than their ungated counter parts throughout the range of supply voltage. A larger overhead power dissipation is observed in gated awake circuits at higher supply voltages as compared to lower supply voltages. The overhead power dissipation in header active circuits is observed to be more than that of footer gated circuits, however, significant power dissipation is observed in footer sleep circuit while header sleep circuit has negligible power dissipation, as seen from the enlarged portion of graph, due to no switching activity throughout the supply voltage range.

### *D. Comparison of Power Dissipation at varying Frequencies of operation*

The analysis has been carried out at constant load capacitance of 10fF and a supply voltage of 1V for the frequency range from 10MHz to 250MHz. The results have been in compiled in graph in Fig. 9. As seen in the graph all the circuits have similar power dissipation at lower frequencies (10-25MHz) and a sharp increase is observed after 125MHz. Header sleep circuit having negligible power dissipation (enlarged portion) due to absence of switching activity whereas footer sleep circuits have a notable amount of power dissipation, especially at higher frequencies. Clearly the power dissipation increases with increase in frequency but the gated awake circuits have a larger power dissipation than their ungated counterpart throughout the frequency range, with the overhead power dissipation also increasing with increase in frequency of operation.

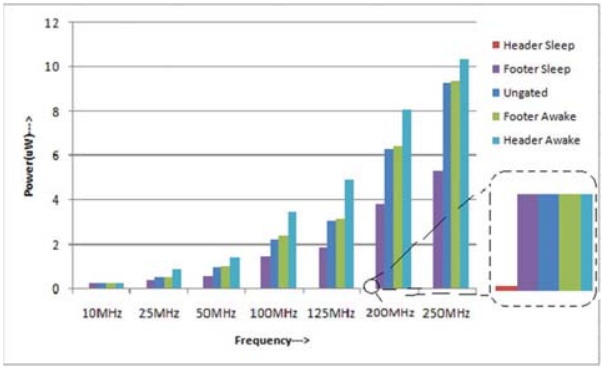



**Fig.9. Power Dissipation vs Frequency**

## V. CONCLUSION

In this paper, 4:2 compressor, a crucial part of multipliers, has been implemented with both gated and ungated ECRL, by employing two power gating methodologies, header and footer. The gated counterparts result in better power analysis especially at higher frequencies and load capacitances. A small overhead power dissipation is observed in the gated circuits which is compensated for by the significant reduction in dissipated power in idle states. Header gated awake state has a larger overhead but its idle state having negligible power dissipation makes it a more suitable gating methodology for adiabatic 4:2 compressors.

## REFERENCES


[1] P. Teichmann, "Adiabatic Logic" in Future Trend and System Level Perspective, Springer Series in Advance Microelectronics, 2012.

[2] C. H. Chang, J. G and M. Zhang, "Ultra low-voltage low-power CMOS 4-2 and 5-2 compressors for fast arithmetic circuits", in IEEE Transactions on Circuits and Systems I: Regular Papers, Volume 51, Issue: 10, pp. 1985-1997, October 2004.

[3] J. M. Rabaey, A. Chandrakasan and B. Nikolic, "Digital Integrated Circuits", A Design Perspective, Second Edition, 2003.

[4] A. Kumar and M. Sharma, "Design and analysis of MUX using adiabatic techniques ECRL and PFAL" International Conference on Advances in Computing, Communications and Informatics, 2013, pp. 1341-1345.

[5] N. Nitish, N. Pandey, R. Pandey, and K. Gupta, "DFAL based implementation of frequency divider-by-3", Annual IEEE India Conference, 2015, pp. 1-6.

[6] S. D. Kumar and S. K. N. Mahammad, "A novel adiabatic SRAM cell implementation using split level charge recovery logic", 19th International Symposium on VLSI Design and Test, 2015

[7] S. P. S. Kushawaha and T. N. Sasamal, "Modified positive feedback adiabatic logic for ultra-low power VLSI", International Conference on Computer, Communication and Control, 2015.

[8] M. L. Keote and P. T. Karule, "Design and implementation of energy efficient Adiabatic ECRL and basic gates", International Conference on Soft Computing Techniques and Implementations, 2015, pp. 87-91.

[9] P. Teichmann, J. Fischer and S. Henzler, "Power clock Gating in Adiabatic Logic Circuits", in Proc. PATMOS, pp. 638– 646, 2005.

[10] W. Morell and A. Srivastava, "Evaluation of Four Power Gating Schemes Applied to ECRL Adiabatic Logic", in 61st International Midwest Symposium on Circuits and Systems (MWSCAS), 2018, pp. 1012-1015.

[11] N. Deo, R. K. Mangang and K. Murugan, "Power gating in FinFET Adiabatic circuits," in 2014 International Conference on Green Computing Communication and Electrical Engineering (ICGCCEE), Coimbatore, 2014, pp. 1-5.

[12] A. K. Bakshi and M. Sharma, "Design of basic gates using ECRL and PFAL", in International Conference on Advances in Computing, Communications and Informatics, 2013, pp. 580-585.

[13] K. K. Sahu and M. N. Islam, "Adiabatic Logic Design for Low Power VLSI Applications", 2015.

[14] N. Ravi, T. J. Prasad, M. Umamahesh and T. S. Rao, "Performance evaluation of high speed compressors for high speed multipliers using 90nm technology", Recent Advances in Space Technology Services and Climate Change, 2010, pp. 189-193.

[15] B. Lamba, and A. Sharma, "A review paper on different multipliers based on their different performance parameters", in 2nd International Conference on Inventive Systems and Control (ICISC), 2018, pp. 324-327.

[16] M. Ha and S. Lee, "Multipliers with Approximate 4-2 Compressors and Error Recovery Modules", in IEEE Embedded Systems Letters, Vol.10, Issue I, March-2018, pp. 6-9.

[17] S. Sowmiya, K Stella and V. M. Senthilkumar, "Design and Analysis of 4-2 Compressor for Arithmetic Application", in Asian Journal of Applied Science and Technology, February 2017, 984-994.

[18] S. Kumar and M. Kumar, "4-2 Compressor Design with New XOR-XNOR Module", in Fourth International Conference on Advanced Computing & Communication Technologies, 2014, pp. 106–111.

[19] M. Mano and M. D. Ciletti, "Digital Design, With an Introduction to the Verilog HDL", FIFTH EDITION, 2013.

[20] R. V. Anugraha, D. S. Durga and R. Avudaiammam, "Design and performance analysis of 2:1 multiplexer using multiple logic families at 180 nm technology", 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology, 2017, pp. 1849-1853.

[21] J. Tonfat and R. Reis, "Low power 3–2 and 4–2 adder compressors implemented using ASTRAN", in IEEE 3rd Latin American Symposium on Circuits and Systems (LASCAS), 2012

Materials Research Express

**ACCEPTED MANUSCRIPT**

# Effects of flyash addition on the dry sliding tribological behavior of aluminum composites

To cite this article before publication: Vipin Kumar Sharma *et al* 2019 *Mater. Res. Express* in press https://doi.org/10.1088/2053-1591/ab24fd

Manuscript version: Accepted Manuscript

Accepted Manuscript is "the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an 'Accepted Manuscript' watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors"

This Accepted Manuscript is **© 2019 IOP Publishing Ltd**.

During the embargo period (the 12 month period from the publication of the Version of Record of this article), the Accepted Manuscript is fully protected by copyright and cannot be reused or reposted elsewhere.
As the Version of Record of this article is going to be / has been published on a subscription basis, this Accepted Manuscript is available for reuse under a CC BY-NC-ND 3.0 licence after the 12 month embargo period.

After the embargo period, everyone is permitted to use copy and redistribute this article for non-commercial purposes only, provided that they adhere to all the terms of the licence https://creativecommons.org/licences/by-nc-nd/3.0

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions will likely be required. All third party content is fully copyright protected, unless specifically stated otherwise in the figure caption in the Version of Record.

View the article online for updates and enhancements.

This content was downloaded from IP address 14.139.251.100 on 02/06/2019 at 06:21

# Effects of Flyash Addition on the Dry Sliding Tribological Behavior of Aluminium Composites

Vipin Kumar Sharma[1,2,*], Ramesh Chandra Singh[1], Rajiv Chaudhary[1], Mayank Saxena[2], Mukul Anand[2]

[1]Department of Mechanical Engineering, Delhi Technological University, Bawana Road, Delhi-110042 India.

[2]Department of Mechanical and Automation Engineering, Maharaja Agrasen Institute of Technology, Rohini Sector-22, Delhi-110086, India

**Corresponding Author* -** vipin.dtu@gmail.com

ORCID ID: https://orcid.org/0000-0002-7356-4590

## Abstract

Flyash is an industrial waste obtained from thermal power plants. The mixing of flyash particles in the air may cause damage to the environment. In this paper, attempts have been made to utilize the waste flyash particle to fabricate an aluminium composite. Three aluminium composites with fly-ash (5wt%-10wt%) were fabricated using a stir casting with vortex technique. Wear and friction behavior of the aluminium flyash composites was evaluated using a linear reciprocating tribometer at 0.38 Mpa pressure and 10Hz frequency against a mild steel counter surface It is established that flyash contents enhanced the wear and friction behavior of the aluminium metal composite. The aluminium-flyash composite with 10wt% flyash particles yields the lowest amount of wear,friction coefficient value and maximum hardness among the fabricated materials. The scanning electron microscope (SEM) images also indicated the lower microcutting and wear wedges on the 10wt% flyash composite.

Keyword: aluminium; stir casting; flyash; composite;

## Introduction

**Aluminium metal** composites (AMC) are the important developed advanced materials which are used in the current engineering applications. The better strength to weight ratio of these materials help in improving the overall performance of the engineering applications. Aluminium silicon alloys are famously used as the potential material for different automobile parts. Low density and high tribological properties of the aluminium silicon alloys provide necessary characteristics to be used as automobile parts. In the automobile field, AMC has been used as pistons, piston rings, and cylinder heads. Aluminium-based composites are substituting many copper alloys and even ferrous alloys in view of the new standard regulations to diminish fuel consumption and preserve the environment [Kumar 2011, Kathiresan 2010]. Many Authors have used different reinforcement agents to the aluminium matrix to obtain specific properties. It is observed that both conventional and non-conventional reinforcements can be used in aluminium for the successful fabrication of the AMC.

### 1.1 Aluminium metal composite with ceramic materials

Accepted Manuscript

Charles and Arunachalam (2008) reported that SiC, quartz, magnesium have proved to improve the hardness, wear and friction properties, and strength of the metals. Ahlatci et al (2004) used the SiC contents in particulate form in the aluminium matrix for the fabrication of Al-SiC material. Improvement in strength was observed for 1% addition of SiC in an aluminium matrix, however, with further increase in SiC contents the strength of the composite decreased. Rao et al. (2009) also proposed an Al-SiC composite material which had better strength and wear resistance as compared to the aluminium. Basavarajappa et al. (2007) used the SiC and graphite powder in the aluminium Al 2024 for the fabrication of AMCs. The mechanical properties improved with the addition of these reinforcing agents. Similarly, Sharma et al. (2018) casted an aluminium composite using graphite particles. Graphite particles in ratio of 2wt%, 4wt%, and 6wt% were used as reinforcement into the aluminium matrix. The presence of graphite particles improved the tribological properties. The graphite particles transfer a graphite rich layer on the testing parts which improved the tribological properties. Omrani et al. (2016) explored the effects of graphite on the wear nature of Al-Si-Ni-Graphite material. Authors performed the experiments in fully flooded as well as starved lubrication conditions. The graphite in the aluminium improved the machine-ability as well as the wear resistance. This additional property of machinability enables the use of aluminium graphite composites in structural engineering applications. Composite materials having higher amounts of carbides, oxides or nitrides are the materials which resulted in maximum wear resistance (Natarajan et al. (2009)).

Zhang and Alpas (1993) conducted a detailed study about the wear resistance of the $Al_2O_3$ reinforced Al 6061 composite. The composite behaved differently at different loading conditions. At low loads (<10N), the composite materials showed excellent wear resistance, this wear resistance was attributed due to the development of translational coat, and as the load was increased, the fracture of composite particles and third body wear decreased the wear resistance significantly.

Deravaju et al. (2013) fabricated the aluminium composite materials by using SiC/Graphite and SiC/$Al_2O_3$ particles. The hardness and tribological properties were better as compared to the pure aluminium material. Chen et al. (2007) also reported the improvement in wear properties of aluminium with the addition of SiC particles. Numerous amount of literature is available where a synthetic reinforcement has been successfully used by researchers to improve the wear properties of base metals. Table 1 presents some recent work in the field of aluminium metal composites.

Shamsipour et al. (2016) utilized the titanium carbide particles and reinforced it into the Al-Si alloy. Authors reported enhancement in hardness of the TiC –Al-Si composite. The wear reisitance of the fabricated composite also found to be improved. $TiB_2$ and $B_4C$ particles were reinforced by Mazahery and Shabani (2012) and Shabani et al. (2012) in the Al-Si matrix.. The quantity of $B_4C$ directly changed the wear resistance of the preparedcomposites.3 wt%-12 wt% of silicon nitride were used as reinforcements by Sharma et al. (2015) in Al6082-T6. The silicon nitride reduced the ductility of the prepared composite material, however, strength and hardness get increased..

### 1.2 Aluminium metal composite with industrial waste combinations

Krishnan et al. (2019) used the vehicle waste scrap aluminium alloy wheels (SAAW), spent alumina catalyst (SAC), AlSi7Mg and alumina particles to fabricate a composite material. Four

Accepted Manuscript

composite materials (i) SAAW with SAE (ii) SAAW with alumina (iii) AlSi7Mg with SAE and (iv) AlSi7Mg with alumina particles were prepared. The prepared composites were tested for evaluation of different properties. It is reported by authors that, alumina particle mixed composites produced low porosity as compared to the SAE mixed composite. The SAAW with alumina particles resulted in the lowest porosity and abrasive wear.

Flyash and Rice husk ash (RHA) were reinforced into the matrix of aluminium A356 for the fabrication of composites by Vinod et al. (2018). Composites with RHA and flyash (0, 5, 7.5, 10 and 12.5wt%) were prepared using the stir casting method. The mechanical properties significantly improved with the addition of RHA and flyash particles until 10wt%. After a further increase in RHA and Flyash particle, the properties tend to reduce. Suleiman et al. (2018) used the Melon Shell Ash (MSA) in 5, 10, 15 and 20wt% were reinforced in the Al-12%Si alloys and examined the hardness, strength and wear properties. It is reported that strength (tensile) improved with nearly 66 % in the prepared composites, the Rockwell hardness number increased to 103.1 HRC and wear rate decreased from to 0.0091 g/min from 0.022g/min. Selvi and Rajasekar (2015) prepared aluminium-flyash (0, 5, 10, 15, 20 wt%) composite with zinc and magnesium. With the addition of more flyash particles the hardness of the fabricated material increased that further improved the wear and friction properties as well. The $SiO_2$ present in the flyash helped in reducing the wear and friction between the tribo-pairs. Prabhu et al. (2019) developed the aluminium composite by using flyash along with $Al_2O_3$. Composites with (5,10,15 and 20wt%) volume fractions were prepared and strength, density and tribological properties of prepared materials were evaluated. It is reported that density of the prepared AMC decreasedwith the addition of flyash and $Al_2O_3$ contents. The $Al_2O_3$ also helped in improving the hardness and tensile properties of the composites. The $Al_2O_3$ and flyash in aluminium matrix also improved the wear resistance.

Anilkumar et al. (2011) used the stir casting method to prepare the aluminium flyash composite. It is concluded by authors that, flyash particles resisted the dislocation phenomenon between the particles and act as a obstraction to the applied load. This cause some enhancement in the strength of the aluminium flyash composites. The size of the flyash particles also affected the tensile strength, the larger particles tends to decrease the tensile strength. Shaikh et al. (2018) fabricated aluminium hybrid composites with (5, 10, 15 and 20 wt%) flyash and 10wt% SiC particles. Pin on disc experiments were performed at (10-30N) of load, (1.5-2.5 m/s) of speed and at (300-900 m) sliding distance. It was observed by authors that wear rate of the developed composite reduced by nearly 40% and hardness improved by nearly 25%. It was also concluded that sliding speed was the prime parameter that affected the wear behavior. Dinaharan et al. (2016) prepared aluminium flyash composite with (6, 12 and 18wt%) flyash contents. The composites were prepared using the friction stir processing technique. The flyash induced composites helped in increasing the surface hardness and wear behavior of the aluminium alloys.

Powder metallurgy, hot pressing, pressure infiltration, squeeze casting, ultrasonic cavitation, compocasting, stir casting are some of the methods which have been used by several researchers to fabricate the materials. Out of all these, stir casting is the most used method owing to its ease of use and economics. The continuous stirring during the casting yields in a uniform mixture of the composite constituents. In the this work, stir casting methods have been used to prepare the aluminium composite. There are several methods for the determination of tribological properties of the material. Rotary pin/ball on disc and reciprocating pin/ball on disc are the famously used

Accepted Manuscript

methods. The tribological properties also depends upon the testing parameters. A brief review about the use and testing parameters of these machines is presented in Table 2.

Table 1: Aluminium composites with some synthetic and industrial waste as reinfocements.

| Reference | Base Metal | Reinforcement | | Reinforcement Percentage (wt%) | Properties |
|---|---|---|---|---|---|
| | | Type | Material | | |
| **Uvaraja and Natrajan (2012)** | AL-7075 | Synthetic | SiC , $B_4C$ | 5%Sic3%$B_4C$<br>10%Sic3%$B_4C$<br>15%Sic3%$B_4C$ | • Composite and unreinforced metal possessed similar properties.<br>• Slightly higher hardness and toughness than 10% fabricated composite. |
| **Suryakumari et al. (2015)** | Al-7075 | Synthetic | $Al_2O_3$, SiC | 2.5%$Al_2O_3$, 2.5%SiC- 7.5%SiC<br>5%$Al_2O_3$, 2.5%SiC-7.5%SiC<br>7.5%$Al_2O_3$, 2.5%Si-7.5%SiC | • Low amounts of abrasive and adhesive wear were reported.<br>• Resulted in high microhardness<br>• The wear particles get cold welded on the wear track by the action of applied load. |
| **Sharma et al. (2017)** | Al | Industrial waste | flyash | 2% flyash<br>4% flyash<br>6% flyash | • The 6% flyash composite resulted in lowest amount of wear. The stir casting setup yields uniform mixing of flyash particles. |
| **Vedrtnam and Kumar (2017)** | Al | Synthetic and Metal | Si,Cu | 3%SiC,0.5%Cu<br>6%SiC,7.5%Cu<br>9%SiC,1%Cu | • Average grain size 31.26µm<br>• 76% higher COF was observed in 9%SiC,1%Cu than 3%SiC,0.5%Cu |
| **Das et al. (2018)** | Al 7075 | | Silicon carbide particulates $SiC_p$ | 3%$SiC_p$ 6.18µm<br>10%$SiC_p$ 8.18µm<br>17%$SiC_p$ 30.65µm | • Small particle size reinforcement increased the density.<br>• With the increase in $SiC_p$ particles, the porosity in the composite increased. The $SiC_p$ particles also helped in improving the hardness. |

Accepted Manuscript

| | | | | | |
|---|---|---|---|---|---|
| **Agbeleye et al. (2017)** | Al6063 | | Clay particle | 5-30% clay particles Size 250micron | • With 15-20% clay, the least amount of wear wasobserved. The hardness also get improved and hard intermetallic in the AA6063 matrix acts as a load supporting parts. |
| **Kumar et al. (2011)** | Al 6061 | Industrial waste | flyash | 10% flyash<br>15% flyash<br>20% flyash<br><br>Particle Size<br>4-25µm<br>45-50µm<br>75-100µm | • Tensile,compressive strength, and Hardness increased with 10% flyash contents, however, these properties decreases with larger flyash particle sizes. |
| **Gladston et al. (2017)** | Al 6061 | Agro Waste | Rice Husk Ash (RHA) | 0%, 2% ,4%, 6%, 8% | • Wear rate of the prepared aluminium RHS composite decreased with increase in% of RHA.<br>• Increased applied load lead to increase in plastic deformation. |
| **Kulkarni (2016) et al.** | Al356 | Industrial waste | (a) Flyash<br>(b) Alumina<br>(c) Hybrid reinforcement | 4%, 8%, 12% | • The density, microhardness, increased with the increase in flyash percentage.<br>• The density was reduced but microhardness, porosity increased with alumina percentage.<br>• For hybrid composite, the density increased as with the percentage of reinforcing agents. |
| **Razzaq et al. (2017)** | AA6063 | Industrial Waste | Flyash | 0%, 2%, 4%,6%, 8%, 10%, 12% | • With 4-12% flyash contents the brittleness was increased.<br>• Porosity was increased by an increase in % of the reinforcement |
| **Prasad et al. (2016)** | Al-Mg Alloy | Industrialwaste | Flyash | 5%, 10%, 15% | • 10% reinforcement sample possessed good resistance. |
| **Escalera-Lozano et al. (2008)** | Aluminum Cans | Agro waste and synthetic waste | SiCp and Rice Husk Ash (RHA) | 30:20 (SiC:RHA) with 8 wt.% reinforcement | The corrosion mechanism in the fabricated composite was attributed to micro-galvanic coupling with the intermetallic compound $Mg_2Si$ and the matrix material. |

Accepted Manuscript

| | | | | | |
|---|---|---|---|---|---|
| **Alaneme et al. (2013)** | Al-Mg-Si alloy | Agro waste and synthetic waste | Bamboo leaf ash (BLA) and Silicon Carbide (SiC) | 10 wt% reinforcements were used in specific ratio 0:10, 2:8, 3:7, 4:6 (BLA:SiC) | • Fracture toughness for hybrid composites was much better than single reinforced Aluminium composite.<br>• In 5wt% NaCl solution testing, the 2-3 wt % bamboo leaf ash showed better corrosion resistance than single reinforced Al - SiC.<br>• In 0.3M $H_2SO_4$ solution Al - SiC composite had better corrosion resistance than hybrid composite |

Table 2: Measurement methods for tribological properties.

| Reference | Test Performance | Counter Body | Material | Testing Parameter | Result |
|---|---|---|---|---|---|
| **Saeidi et al. (2016)** | Pin-on-Disc (Sliding) | Grey cast iron hardness 0f 200HV | Hardened 42 Cr Mob Steel with a hardness of 600HV | Const. Pressure=24Mpa (600N)<br>Temperature=35°C<br>Humidity=30%RH | • High Local Pressure<br>• Lower COF in Steady state<br>• High COF in running. |
| **Pattanaik et al. (2016)** | Pin-on-Disc (Sliding) | Stainless Steel Disc | Epoxy Resin (LY556) Flyash pin | Dry sliding<br>Time=5min<br>% flyash=40<br>Speed=400rpm<br>Load=10Kg<br>Track dia. =40mm | • Change of secondary wear due to the eroded particle from the surface which gets trapped in between plate and sample.<br>• Agglomeration of eroded material due to the heat generated during sliding. |
| **Kumar et al. (2015)** | Pin-on-Disc (Sliding) | EN19 Disc | EN31 Steel pin | Dry sliding<br>Time=15min<br>Speed=200-2000rpm (200 difference)<br>Load=10-70N<br>Track dia. =90mm | • COF increased slowly with sliding speed at lower load and become stable at a higher speed at higher load.<br>• Showed stable behavior for initial running and then decreases considerably at a higher speed. |
| **Liu et al. (2015)** | Ball-on-Disk (Sliding) | Ni-Cr Sintered Disc | Ball-$Al_2O_3$ Ceramic | Time=60min<br>Speed=0.1m/s<br>Load=10Kg<br>Diameter=10mm<br>Temperature=200-1000ºC | • NiCr 40wt%.$Al_2O_3$ exhibited the satisfactory wear resistance over entire range of temperature. |
| **Zhang et al. (2016)** | Pin-on-Disc (Sliding) | Textured Babbitt alloy | PIN-C45 steel | Lubricant-Antirust anti-oxygen steam Turbine oil | • The COF first rapidly decreased with increase in lubricant parameter. |

Accepted Manuscript

| | | | | | |
|---|---|---|---|---|---|
| | | | | Speed=0.3-0.8m/s | |
| **Ibrahim (2015)** | Pin-on-Disc (Sliding) | Steel Disc | PIN-Polymer composite | Dry Sliding<br>Surface finish=0.13µm<br>Time=3min<br>Speed=0.8m/s<br>Load=2-6N<br>Diameter=90mm | • The increase of hard (coarse) particle consists of palm fronds in polyester composite increased COF.<br>• Polyester composites reinforced with palm fronds were recommended at high friction and low wear rate material for industrial application |
| **Omrani et al. (2015)** | Pin-on-Disc (Sliding) | Hardened 440C steel disc | Carbon reinforced PIN | Dry Sliding<br>Room Temprature<br>Speed=0.15-0.35m/s<br>Load=10-20N<br>Distance=1Km | • Carbon fiber serve as solid Lubricant.<br>• At higher normal load COF decreases due to higher number of Carbon fiber act as a solid lubricant between the surfaces while the wear loss increases showing more roughness at the contact surface. |
| **Carrera-Espinoza et al. (2016)** | Ball-on-Disk (Reciprocating motion) | Borided and Unborided AISI 1018 low-carbon steel | $Al_2O_3$ ball | Stroke length-5mm and Sliding speed-5mm/s<br>dry sliding and lubricated condition<br>Loads - 5 N<br>Lubricant used- engine oil HX710W-40<br>Sliding distances - 81 m, 108 m and 135 m. | • Surface roughness value for the unborided was 0.02 µm and 0.23–0.33 µm for the boriaded.<br>• The friction coefficient and wear rates in lubricated condition was approximately 85.7 % and 89.6 % respectively, less than dry sliding conditions. |
| **Sharma et al. (2017)** | Ball-on-Disk (Reciprocating motion) | Hot pressed silicon carbide (SiC) with 10, 30, 50 wt% tungsten carbide (WC) | SiC ball diameter 6.35 mm | Frequency - 8.3 Hz<br>Stroke- 4.71 mm (speed ≈ 78 mm/s)<br>Cycles-1000<br>Load - 6, 9 or 19 N<br>Ambient conditions (28±2 °C and 55±5% RH), | • The COF values for the SiC ceramics were in range of 0.2–0.7.<br>• Mechanical wear included abrasion and chipping. |
| **Mishra et al. (2018)** | Ball-on-Disk (Reciprocating motion) | Zircaloy-4 | D2 steel half Cylinder | Frequency - 2 Hz<br>Sliding amplitude-6mm<br>Load - 8kN<br>Lubricants<br>Mineral oil having chlorinated EP additives and solid lubricants | • COF was not much affected by Lubrication. |

Accepted Manuscript

| | | | | | |
|---|---|---|---|---|---|
| **Peruzzo et al. (2019)** | Ball-on-Disk (Reciprocating motion) | Boron-free and Boron-containing sintered AISI 316L | Alumina (Al2O3116 ) ball | Load -8 N<br>Frequency -3 Hz.<br>Stroke length - 4 mm,<br>Sliding distance -172.8 m,<br>Sliding time - 7200 s.<br>Dry sliding at room temperature | • No specific change in COF value was obtained with the addition of boron addition.<br>• The wear mechanism was mostly Adhesive.<br>• COF of Boron-containing stainless steels was 0.85<br>• Specific wear decreased significantly with additions of Boron |
| | | | | | |

From the literature review, it is found that ceramic reinforcements have been largely used for the fabrication of aluminium composites. Limited work is reported by using industrial and agro wastes as reinforcements. The available literature lacks the usage of flyash particles in the aluminium matrix as a piston rings material. Therefore, in this work, 5wt%, 7.5wt% and 10wt% flyash particles are used to fabricate the aluminium flyash composite for the piston rings material. To simulate the motion of piston ring and cylinder body, a linear reciprocating tribometer has been used and wear and friction properties were evaluated in unlubricating condition. Optical micrographs, scanning electron microscope images and vicker hardness tests were performed to analyze the tribological behavior of the aluminium flyash composite.

## 2. Materials and Methods

### 2.1 Matrix material and Flyash

Aluminium have been used as the matrix material. The elemental composition of the matrix material is presented in Table 3. It contains magnesium (Mg), silicon (Si) and zinc (Zn) as the main alloying elements. The Si and Mg provide high hardness and strength and has major applications in the field of automobile, marines, and aircraft.

Table 3: Elemental composition of aluminium alloy

| Element | wt% |
|---|---|
| Si | 0.1697 |
| Mg | 0.6492 |
| Zn | 1.0259 |
| Ni | 0.1207 |
| Cu | 1.2 |
| Fe | 0.4102 |
| Al | Balance |

Flyash is a fine waste product which is produced in the coal-based thermal power plants. The major elements in the flyash are alumina, silica, calcium sulphate and unburnt carbon. The flyash particles were converted to fine powder by using the ball milling method. The ball milling setup consists of a steel container which is rotated with the help of a motor. The motor rotates the chamber for a predefined time. In the steel chamber, the flyash particles were poured along with hardened steel balls of

Accepted Manuscript

different diameters. Three balls of 40mm diameter and 6 balls of 20 mm diameter were used for the ball milling process. The mill was rotated at a speed which is not to fast or slow. During the fast-milling operation, the steel balls strike with the wall of the cylindrical chamber and do not play much role in the powder production. The slow speed does create much energy for breaking the power to finer diameters. The ball milling setup was operated for 24 hours in successive steps of 4 hours. After the running of the ball mill, the sample powder was inspected to check the fineness. For that, the milled flyash powder was filtered by using an Ultra-filtration membrane (>0.001nm).

## 2.2 Stir Casting Method

For the fabrication of the aluminium composites, powder metallurgy processing, plasma spray, spray atomization, and co-deposition, plasma spraying, conventional casting, and stir casting are some of the methods which are used by researchers. Out of all the methods, the stir casting process is the most economical method. In the stir casting method, the reinforcement particles are added in the melted aluminium material and after that, the mixture is allowed to solidify it. To create the desired wettability between the reinforcement and material, the vortex technique is used by the authors as explained in the research article by Surappa, (2003). In the vortex method, the preheated reinforcement particles are added to the vortex on the aluminium melt. This vortex on the aluminium melt is created by the rotating stirrer. In this work, aluminium alloys with 5 wt%, 7.5 wt% and 10wt% flyash particles are used to fabricate the aluminium flyash composite. A stir casting furnace having an ultrasonic stirrer has been used (Figure 1). The technical details of the stir casting furnace are presented in Table 3. In the casting process, the aluminium material was heated in a graphite crucible till 650ºC and the flyash particles were preheated at a temperature of 200ºC to release any kind of dampness. The flyash particles were poured into the aluminum melt and the mixture was first manually stirred for 15 minutes at a constant temperature of 650ºC. After that, a mechanical stirrer having cross blades was used to stirrer the mixture for the next 15 minutes. The melt was poured into a preformed mold and different specimens were machined as per the requirement. For observing the structure of the prepared samples, the samples were polished with the help of abrasive paper of 250, 600, 800 grade and diamond polished with 0.5 µm alumina powder. The finished surfaces of the prepared samples were observed by using an Olympus made optical microscope. The images were obtained at 20X magnification and are presented in Figure 3.

Table 4: Technical specifications of stir casting setup.

| Technical Detail | Value |
|---|---|
| Maximum Capacity (Kg) | 1-5 |
| Maximum Temperature (ºC) | 1250 |
| Maximum rotational speed of Stirrer | 1500 (DC Drive for control) |
| Door opening | Split Type |
| Furnace Outer Shell | Corrosion resistant Stainless steel |
| Stirrer Blade | Cross shape (90º) |
| Stirrer length (cm) | 95 cm |
| Blade length | 9 cm each |

Accepted Manuscript

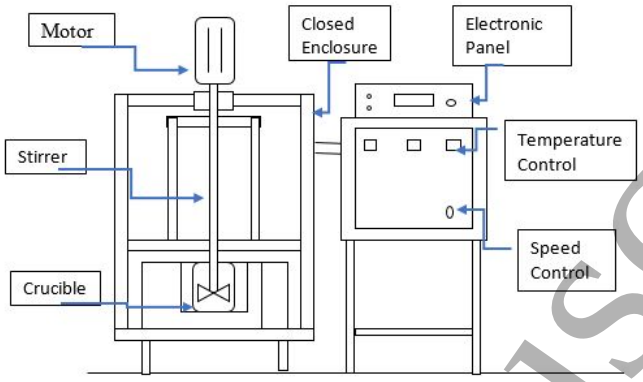



Figure 1: Stir casting setup used to fabricate the composites.

### 2.3 Measurement of Hardness, Wear and Friction

Pin on disc [Sharma et al. (2017), Singh et al. (2019)] is the most used wear and friction testing method In this a pin is pressed against a rotating disc. For applications where the tribo-pairs are in a reciprocating motion (piston-cylinder) a linear reciprocating tribometer (LRT) has been used by several researchers (Nazir et al. (2018), Peruzzo(2019)). In the current research, the composite fabrication is based on the application of piston rings material (Prasad and Asthana (2004)), and piston rings in an engine and cylinder assembly undergo a reciprocating motion. So, a linear reciprocating tribometer has been used to evaluate the wear and friction properties of Al-F composite. The technical specifications of the LRT are given in Table 5. Tribometer consists of a fixed steel specimen at the bottom on which the prepared composite pin reciprocates. The pin reciprocates under different input values of applied load, stroke length, temperature, and frequency. For the experiments, the frequency, test time, and stroke length are provided to the LRT with the help of software and load is applied on the pin specimens with the help of loading arm as shown in Figure 2. The friction force obtained by the reciprocating motion between the pin and plate specimens was measured with the help of friction sensor installed on the LRT. The permissible error and specifications of the senses installed on LRT are given in Table 6. The friction force was continuously measured by the sensor for the complete duration of the experiments. This friction force is later converted to friction coefficient and plotted with sliding time for the analysis. The wear from the pin specimens was measured before and after each experiment by using a weighing balance with 0.0001 g accuracy. The various testing conditions used in the present research are given in Table 7.

Table 5: Specifications of the linear reciprocating tribometer.

| Input Parameter | Values |
|---|---|

Accepted Manuscript

| | |
|---|---|
| Stroke Length (mm) | 1 - 20 |
| Frequency (Hz) | 50 (Max) |
| Normal Load (N) | 5 – 50 |
| Frictional Force (N) | 0 – 50 |
| Temperature (ºC) | 550 (Max) |
| Pin Diameter, Length (mm) | 10, 15 |

Table 6: Details of sensor installed on linear reciprocating tribometer

| **Sensor** | **Specification** | **Value** |
|---|---|---|
| Friction Sensor | Sensor | Piezo Sensor |
| | Least Count (N) | 0.01 |
| | Accuracy | 0.01±1% observed value |
| Frequency Measurement Sensor | Sensor | K-Type Thermocouple |
| | Least Count (ºC) | 1 |
| | Accuracy | 1±1% observed value |
| Temperature Sensor | Sensor | Encoder Output |
| | Least Count (Hz) | 0.01 |
| | Accuracy | 0.01±1% observed value |

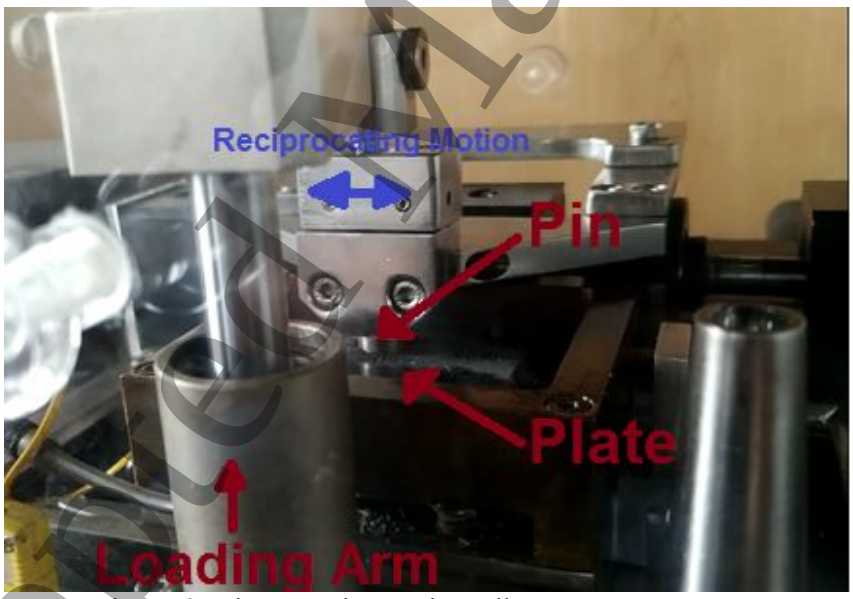



Figure 2: Linear reciprocating tribometer components

Table 7: Experimental conditions

| **Property** | **Value** |
|---|---|
| Pin diameter, height (mm) | 10, 15 |
| Plate height, length, breath (mm) | 5, 50, 50 |
| Frequency (Hz) | 10 |

Accepted Manuscript

| | |
|---|---|
| Stroke Length (mm) | 10 |
| Load (MPa) | 0.38 |
| Time (s) | 900 |
| Lubricating Condition | Dry |
| Temperature (ºC) | 28 |
| Humidity (%) | 50-55 |

It is well known that, volume loss for the material during the sliding action directly depends on its hardness. A Fischer made microhardness tester was used in the present investigation for the measurement of hardness of the fabricated materials. The hardness measurement was done as per ASTM E-384 standards (ASTM E-384 standards, (2017)). In this test, the test surfaces were cleaned and placed perpendicularly to the indenter. For obtaining the indentation load is applied very smoothly on the surface. A diamond indenter of square base pyramid shape was used to indent on a surface of the specimens at a load of 3000 mN for 20 seconds. To obtain the quality of tests, the load should be accurately applied on each measurement. The hardness measurement instrument measures the diagonals of the indentation and calculates the Vicker hardness number. The Vicker hardness number is defined as the ratio of the load applied by the indenter to the pyramidal contact area of the indentation. In this work for each specimen, three hardness values were measured and an average of three was taken for the analysis.

All experiments were performed three times and an average of three observations was used for evaluating the results. Before starting the experiments, the pin and disc samples were cleaned properly and a surface roughness value in the range of 0.2-0.3µm was achieved. For measuring the roughness value, a Taylor Hobson, pocket surf instrument was used. The experiments were performed between the prepared Al-F composites and a mild steel plate.

## 3. Results and Discussion

### 3.1 Hardness and Optical Micrographs

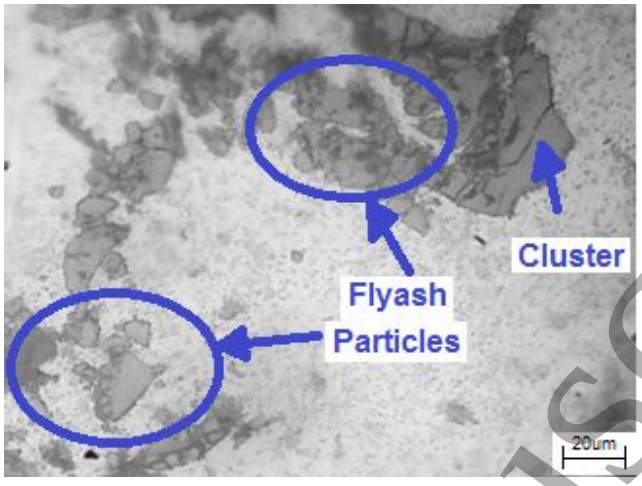
The uniform distribution of the flyash particles in the aliminium matrix helps in improving the mechanical properties of the composite. To obtain the optical micrographs of the prepared composites, the composite surfaces were cleaned with the help of several grades of emery papers and diamond polishing. The optical micrographs (Figure 3 (a-c)) were obtained after performing the etching operation. The Images were captured at 20X magnification and are shown in Figure 3.

The optical micrographic images indicate the good interfacial bonding between the aluminium matrix and flyash particles. This interfacial bonding helps in transferring the applied load effectively during the sliding wear tests (Dinaharan et al. 2016)).

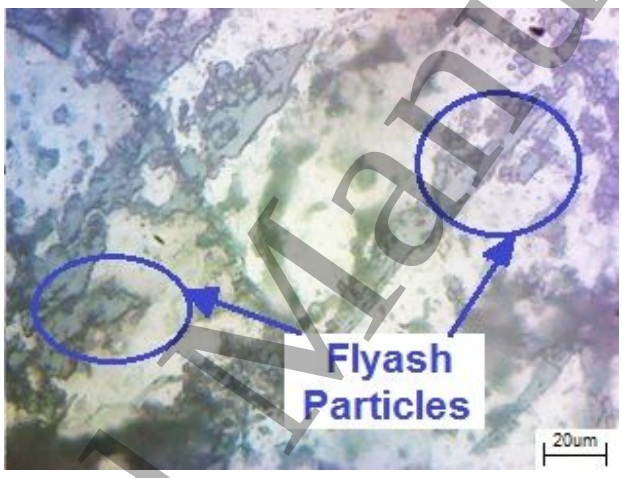
A near homogeneous dispersion of the flyash is obtained for the fabricated materials. However, the composite Al-5wt% flyash has some agglomeration of the flyash particles. The uniform distribution enables isotropic properties for the materials.

(a)

Accepted Manuscript




(b)




(c)

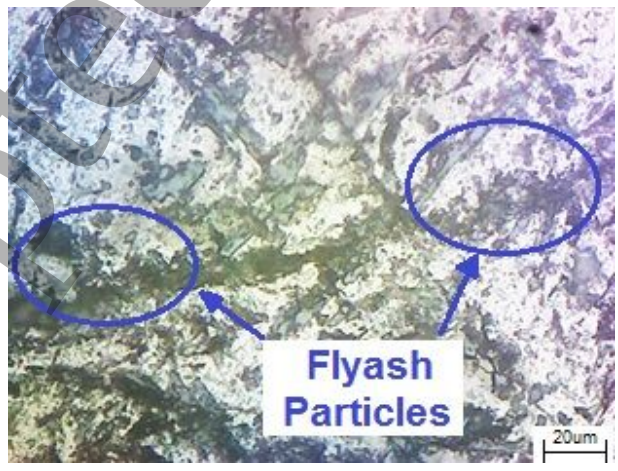



Figure 3: Optical microscopic image for (a) 5 wt% flyash (b) 7.5 wt% flyash, (c) 10wt% flyash

Accepted Manuscript

The hardness of the material directly affects its wear behavior. It is reported in previous studies that flyash particles improved the hardness of the aluminium composites (Guo et al. (1997), Arif et al. (2017)). The three observations for the calculation of vicker hardness were obtained and the average of three was taken for evaluation of surface hardness of the prepared materials. The observations obtained are presented in Table The 10wt% and 7.5wt% flyash composite have 109.88HV and 99.10HV vicker hardness. The 5wt% flyash composite have 98.55 HV hardness.. The hard flyash particles provided a higher resistance to the plastic deformation. The flyash particles also acts as obstruction to the movement of dislocations (Basak et al. (2015)) which further improved the hardness. These effects get multiply with the increase in flyash contents in aluminium matrix (Dinaharan et al. (2016)).

The distribution of the flyash particles in the aluminium matrix directly affects the surface hardness. It is seen from Figure 3 (b) and (c) that flyash particles are uniformly distributed in the aluminium matrix. However, some clustering of flyash particles was seen in the case of 5wt% flyash composite (Figure 3 (a)). This clustering of flyash particles shall increase the hardness. Due to this, one of the hardness value gets a sharp increase which is shown in Table 8. Deng and Chawla (2006) performed a study about clustering of reinforcements and it is concluded that higher clustering incareses the work hardening of the composite which in result reduces the ductility. However, the average hardness for the 7.5wt% flyash composite is still higher than the 5wt% flyash composite. The vicker hardness for the fabricated materials is presented in Figure 4 and error bars have also been drawn to indicate the variation in measurements.

Table 8: Vicker Hardness observations

| **Al-flyash Composite** | **Vicker Hardness$_1$ (HV)** | **Vicker Hardness$_2$ (HV)** | **Vicker Hardness$_3$ (HV)** | **Vicker Hardness$_{Avg}$ (HV)** |
|---|---|---|---|---|
| **Al-5wt% flyash** | 96.10 | 96.20 | 103.35 | 98.55 |
| **Al-7.5wt% flyash** | 99.20 | 98.80 | 98.90 | 99.10 |
| **Al- 10wt% flyash** | 109.81 | 109.9 | 110.23 | 109.88 |

Accepted Manuscript

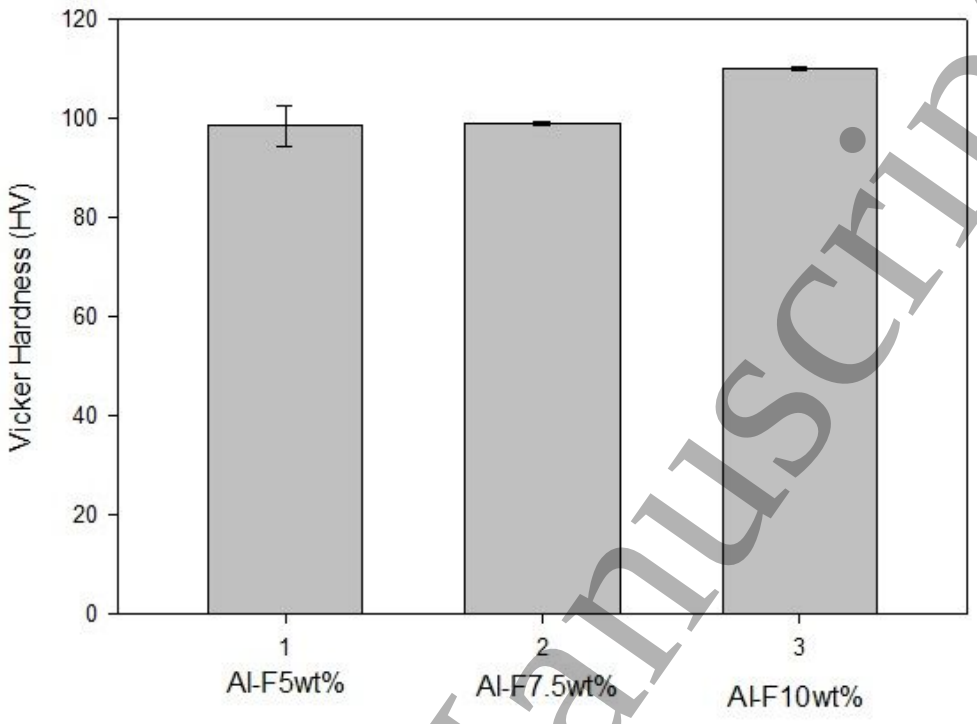



Figure 4: Variation of Vicker Hardness number for the prepared samples.

### 3.2 Wear and Friction

Experiments were conducted for the prepared aluminium flyash composites using a linear reciprocating tribometer. Mild steel plates was used as the counter surface for all the experiments.

Figure 5 presents the variation of wear as a function of wt% of flyash particles. . The wear in gram is assessed to be 0.007 g at 5wt% flyash, 0.0025 g at 7.5 wt% flyash and 0.0022 g at 10wt% of flyash contents. The presence of flyash particles significantly improves the wear resistance of the AMC. As discussed in section 3.1, the hardness of the AMC increases with the increase in flyash particles, this increase in hardness provides the wear resistance to the AMC. The flyash particles have high hardness and it retains their hardness at elevated temperatures. The uniform and smooth distribution of the flyash particles in the aluminium matrix also helps in increasing the wear resistance. This is clearly indicated by the optical micrograph of Al-F7.5wt% and Al-F10wt% (Figure 6). It may also be concluded that the improvement in wear is not linear as the amount of flyash contents increases. This is due to the complex wear mechanism and distribution of flyash particles.For inspecting the worn surfaces, the pin samples were observed under an optical microscope and a scanning electron microscope (SEM). Figure 6 (a-c) presents the optical microscope (OM) images of the pin and plate surfaces. The OM images were obtained at 10X

Accepted Manuscript

magnification. The worn surface of the 5wt% flyash composite reveals a broader worn area as compared to the 7.5wt% and 10wt% flyash composite. There is a large amount of material pull out and plastic flow for the 5wt% flyash composite as well.

A similar pattern of wear area was obtained for the plate samples as well. This decrease in wear is in a similar term as obtained by Sharma et al. (2017b). Figure 7 (a-c) presents the SEM images for the worn surface areas for Al-F composites. The SEM images were obtained at a magnification of 100X with 12500 µm working distance. 115000 nA current and 15kV voltage were used to generate the image of the worn areas. Figure 7 (a) shows the heavy micro-cutting and grain pull outs in case of 5wt% flyash composite. These micro cuttings indicate the heavy wear on the pin surface. There is relatively less particle micro cutting on Al-F7.5wt% and Al-F10wt% composite. The frictional heat produced during the sliding of the pin forces the material to plasticize. The 5wt% flyash composite is seen to have a more plastic flow of the aluminium. With further increase in the flyash particles in aluminium matrix, the plastic flow is very less due to the presence of hard flyash particles. It is also seen that, wear debris neither stick with the pin nor with plate surfaces. This happened due to the hard nature of the flyash particles. The SEM images also indicated that the wear mode is mostly abrasive for the complete duration of the sliding. The similar wear mechanism was also obtained by Dinaharan et al. (2016).

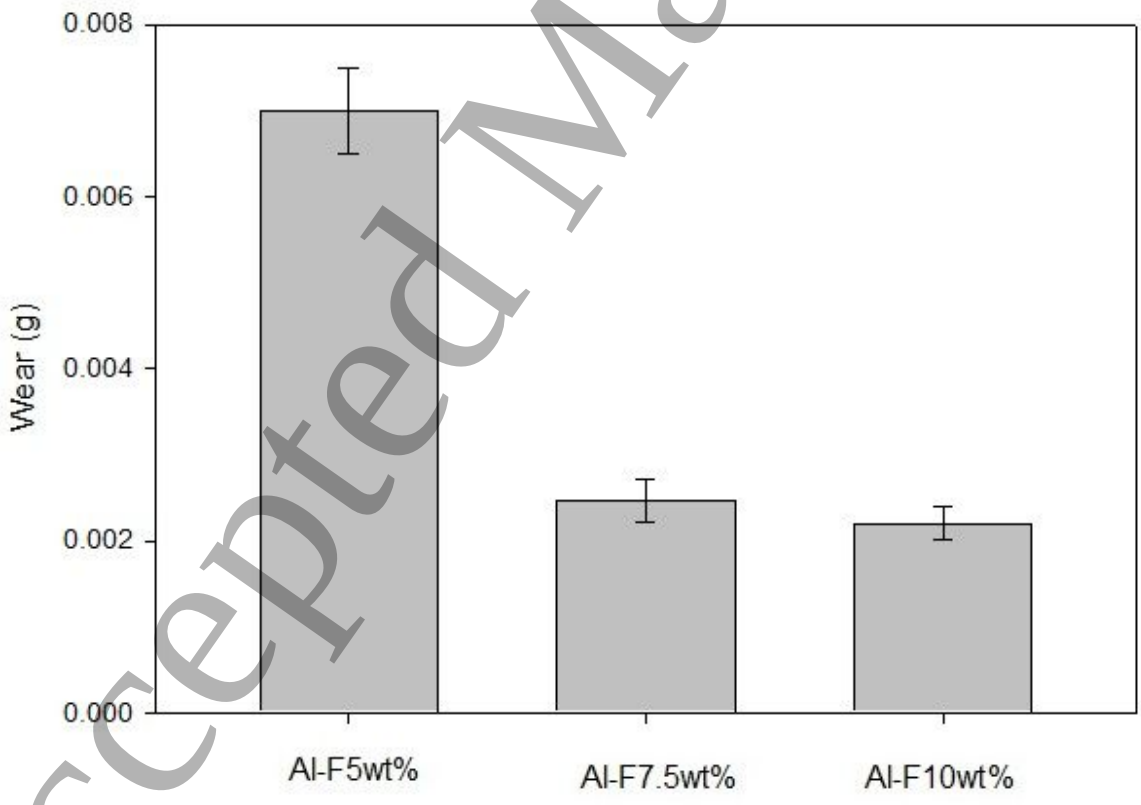



Figure 5: Wear loss for the prepared samples

Accepted Manuscript

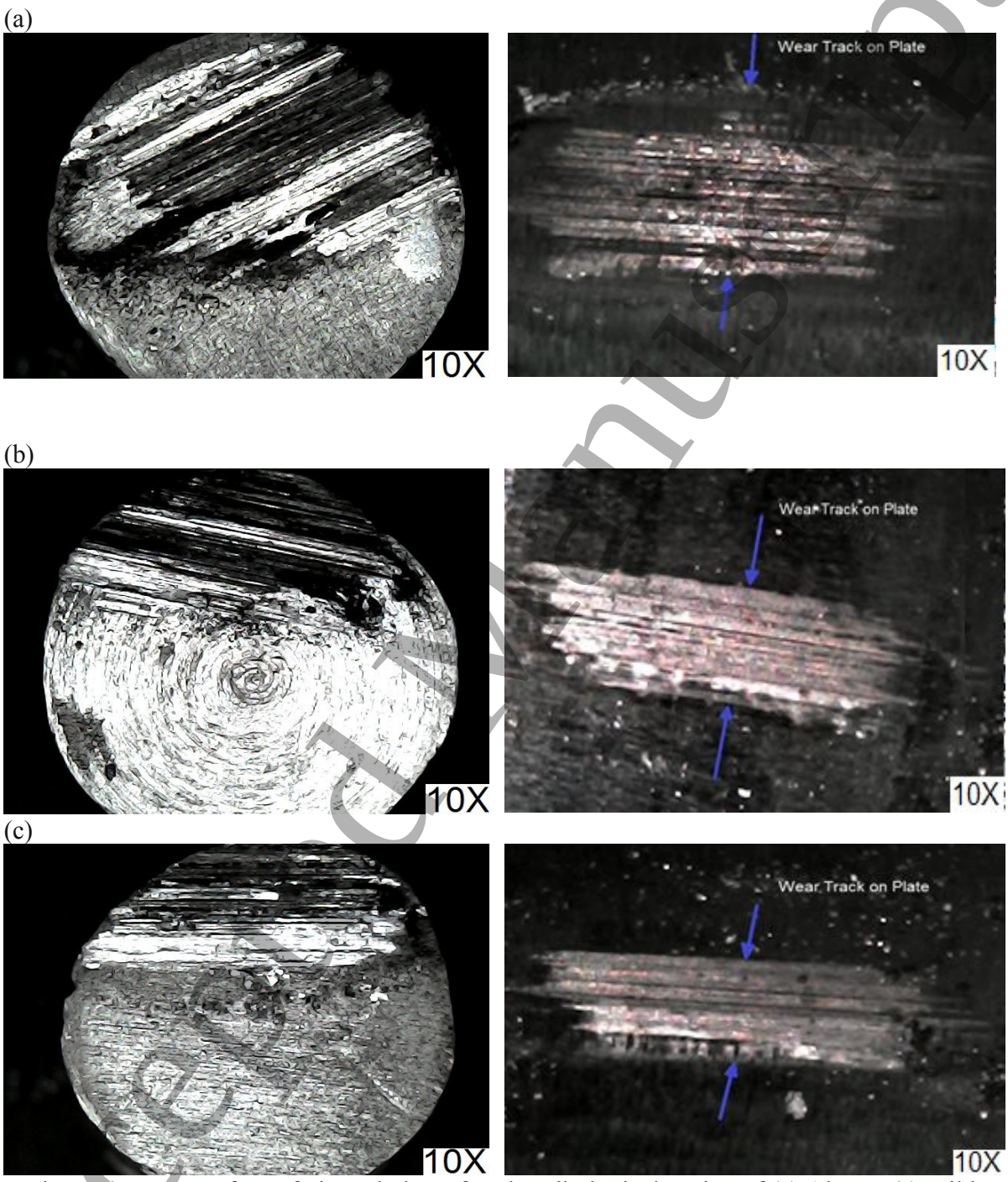
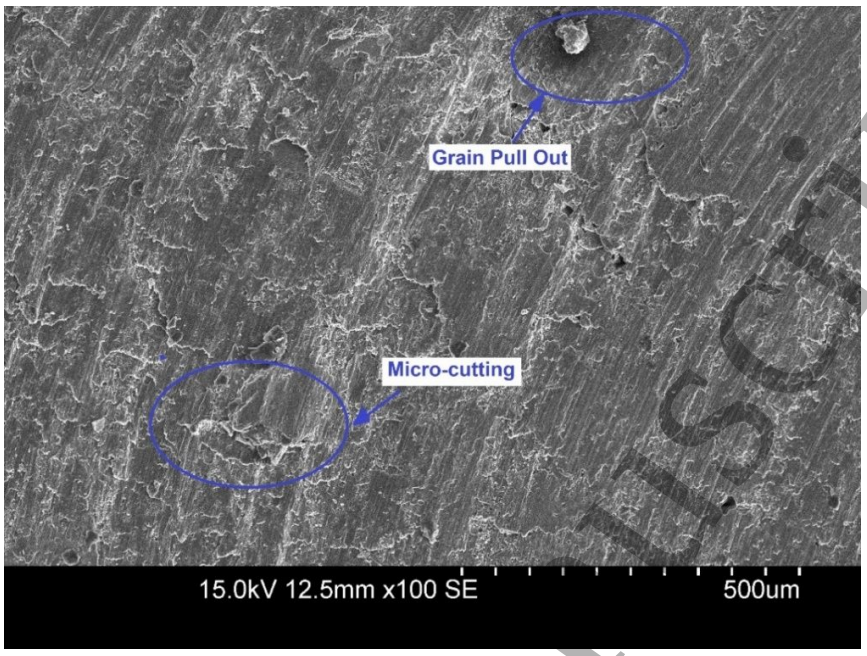


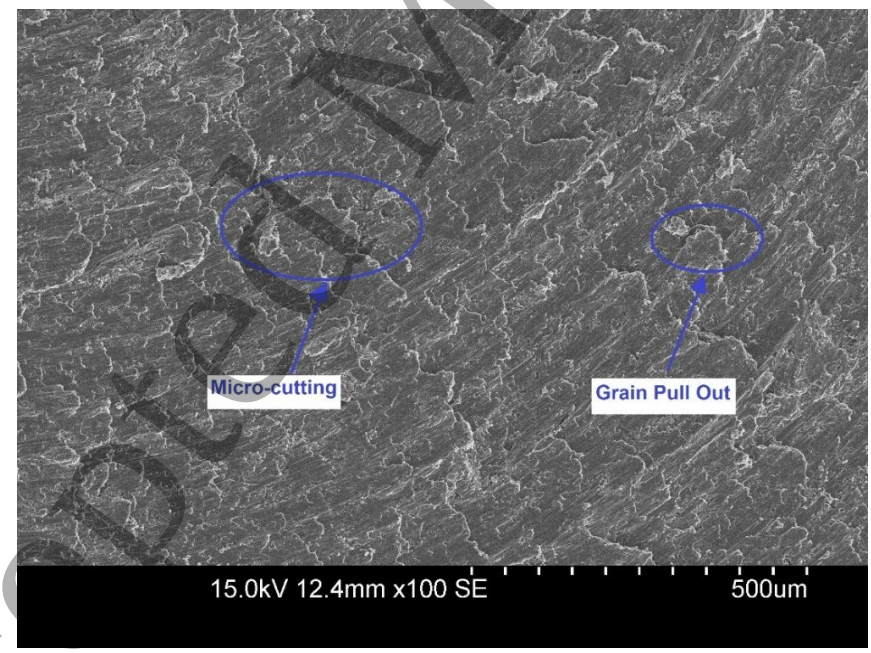

Figure 6: Worn surface of pin and plate after the tribological testing of (a) Al-F5wt%-Mild steel (b) Al-F5wt%-Mild steel (c) Al-F5wt%-Mild steel

(a)

Accepted Manuscript




(b)




Accepted Manuscript

(c)

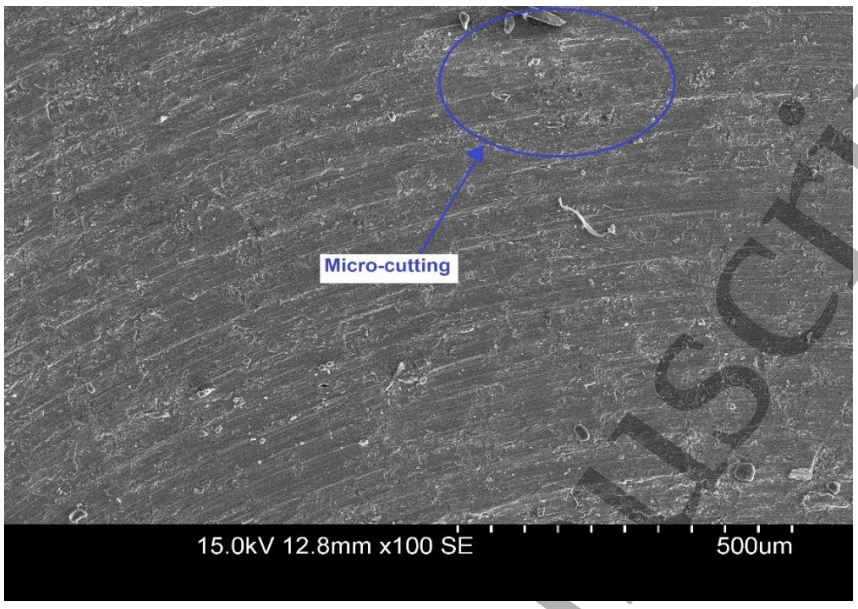



Figure 7: Scanning electron microscope images for (a) worn Al-F5wt% pin (b) worn Al-F7.5wt% pin (c) worn Al-F10wt% pin

Friction between the tribopairs is the results of asperity formation, ploughing, and adhesion (Pirso et al .(2006)). The variation of coefficient of friction (COF) with time is presented in Figure 8. The COF curve has been divided into four parts, namely (i) running in (ii) mixed zone (iii) steady state (iv) wear out. The running in region presents the variation of coefficient of friction during the initial sliding of the pin on the plate surface. During this region, the surface asperities of the pin samples make the first contact with the plate surface. The COF for Al-F5wt%, Al-F7.5wt%, and Al-F10wt% composites rises to 0.59, 0.50 and 0.40 respectively. After achieving the maximum COF value, it tends to decrease in the mixed zone and attains a near constant value. The part of the COF curve, where the COF value attains a nearly constant value is termed a steady state. Due to the thermal heating and excessive wearing the COF value starts to rise again. This part of the curve where the COF starts to rise again is called as wear out region.

Accepted Manuscript

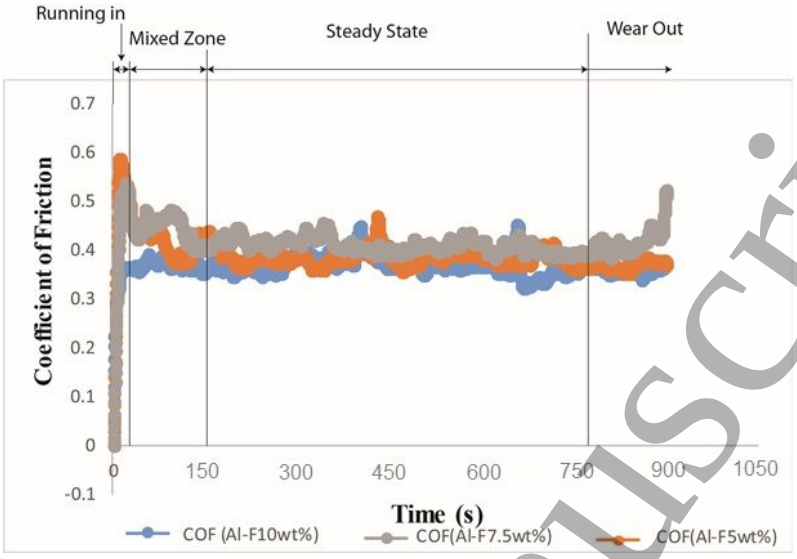



Figure 8: Variation of the coefficient of friction with time

After, attaining the maximum 0.59, 0.50 and 0.40 COF values for Al-F5wt%, Al-F7.5wt%, and Al-F10wt% composites respectively, the COF tents to decreases and attained a near constant value. It is observed that the COF values during the steady-state region are overlapping with each other in the range of 0.35-0.45. The overall average COF values for Al-F5wt%, Al-F7.5wt%, and Al-F10wt% composites are 0.47, 0.45 and 0.38. The aluminium composite with 10 wt% flyash particles yields the minimum average coefficient of friction values.

## 4. Conclusions

- Aluminium-Flyash (5wt%, 7.5wt%, 10wt%) composites were fabricated using the conventional stir casting method. The microhardness, microstructure and reciprocating sliding wear behavior were analyzed. From the current study, the following conclusions were drawn. The flyash particles improved the microhardness of the aluminium composite. The measured microhardness for the 5wt% flyash composite was 98.55 HV and 109.88HV for the 10wt% flyash composite. There was an increase of 11.49 % in the hardness value for the 10wt% flyash composite as compared to the 5wt% flyash composite.
- The flyash particles enhanced the wear and friction properties of the aluminium-flyash composite. The wear resistance improved with the increase in flyash particles in the aluminium matrix. The wear in gram was observed to be 0.007 g at 5wt% flyash, 0.0025 g at 7.5 wt% flyash and 0.0022 g at 10wt% flyash particles in aluminium matrix. The scanning electon microscope image also indicated fewer microcutting on the 10%flyash composite.
- The coefficient of friction value for the fabricated composite was in the range of 0.38-0.47. The 5wt% flyash composite resulted with 0.47 coefficient of friction value and minimum COF was achieved for 10wt% flyash composite.
- The improvement in the tribological properties of the aluminium composite with the addition of higher amounts of flyash particles validated the used of flyash in aluminium as a possible reinforcement. With the use of flyash in aluminium, the flyash contents from the environment may get reduced which would reduce the pollution.

Accepted Manuscript

.

## 5. Future Scope

- There are many challenges which are related with the development of the aluminium-flyash composites. The microstructure of the aluminium-flyash composites fabricated by different methods may be compared for a better understanding of the mechanical behavior.
- The performance study of the aluminium-flyash composite with different size of flyash particle reinforcement may be performed to inspect the effect of reinforcement size.
- Experiments in fully flooded and starved lubrication condition may be performed for the complete understanding of the tribological behavior of the aluminium-flyash composites.

## Refernces


1. Agbeleye, A.A., D.E. Esezobor, S.A. Balogun, J.O. Agunsoye, J. Solis, and A. Neville. "Tribological Properties of Aluminium-Clay Composites for Brake Disc Rotor Applications." *Journal of King Saud University - Science*, September 2017. https://doi.org/10.1016/j.jksus.2017.09.002.
2. Ahlatci, Hayrettin, Ercan Candan, and Hüseyin Çimenoğlu. "Abrasive Wear Behavior and Mechanical Properties of Al–Si/SiC Composites." *Wear* 257, no. 5–6 (September 2004): 625–32. https://doi.org/10.1016/j.wear.2004.03.006.
3. Arif, Sajjad, Md Tanwir Alam, Akhter H Ansari, M Arif Siddiqui, and Mohammad Mohsin. "Study of Mechanical and Tribological Behaviour of Al/SiC/ZrO 2 Hybrid Composites Fabricated through Powder Metallurgy Technique." *Materials Research Express* 4, no. 7 (July 27, 2017): 076511. https://doi.org/10.1088/2053-1591/aa7b5f.
4. ASTM Standard E387. "Standard Test Methods for Microindentation Hardness of Materials." ASTM International, West Conshohocken, PA, United, 2017.
5. Basavarajappa, S., G. Chandramohan, Arjun Mahadevan, Mukundan Thangavelu, R. Subramanian, and P. Gopalakrishnan. "Influence of Sliding Speed on the Dry Sliding Wear Behaviour and the Subsurface Deformation on Hybrid Metal Matrix Composite." *Wear* 262, no. 7–8 (March 2007): 1007–12. https://doi.org/10.1016/j.wear.2006.10.016.
6. Carrera-Espinoza, R., U. Figueroa-López, J. Martínez-Trinidad, I. Campos-Silva, E. Hernández-Sánchez, and Amir Motallebzadeh. "Tribological Behavior of Borided AISI 1018 Steel under Linear Reciprocating Sliding Conditions." *Wear* 362–363 (September 2016): 1–7. https://doi.org/10.1016/j.wear.2016.05.003.
7. Chen, Rong, Akira Iwabuchi, Tomoharu Shimizu, Hyung Seop Shin, and Hidenobu Mifune. "The Sliding Wear Resistance Behavior of NiAl and SiC Particles Reinforced Aluminium Alloy Matrix Composites." *Wear* 213, no. 1–2 (December 1997): 175–84. https://doi.org/10.1016/S0043-1648(97)00123-3.
8. Das, Diptikanta, Vivek Chakraborty, Basanta Kumar Nanda, and Bharat Chandra Routara. "Turning Performance of Al 7075/SiCp MMC and Multi-Response Optimization Using


Accepted Manuscript

WPCA and Taguchi Approach." *Materials Today: Proceedings* 5, no. 2 (2018): 6030–37. https://doi.org/10.1016/j.matpr.2017.12.207.

9. Deng, X., and N. Chawla. "Modeling the Effect of Particle Clustering on the Mechanical Behavior of SiC Particle Reinforced Al Matrix Composites." *Journal of Materials Science* 41, no. 17 (September 1, 2006): 5731–34. https://doi.org/10.1007/s10853-006-0100-1.
10. Devaraju A, Kumar A, Kotiveerachari B. "Influence of Additionof Grp/Al2O3p with SiCp on Wear Properties of Aluminumalloy 6061-T6 Hybrid Composites via Friction Stir Processing." *Trans Nonferrous Met Soc China* 23, no. 5 (2013): 1275–80.
11. Dinaharan, I., R. Nelson, S.J. Vijay, and E.T. Akinlabi. "Microstructure and Wear Characterization of Aluminum Matrix Composites Reinforced with Industrial Waste Fly Ash Particulates Synthesized by Friction Stir Processing." *Materials Characterization* 118 (August 2016): 149–58. https://doi.org/10.1016/j.matchar.2016.05.017.
12. Escalera-Lozano, R., C.A. Gutiérrez, M.A. Pech-Canul, and M.I. Pech-Canul. "Degradation of Al/SiCp Composites Produced with Rice-Hull Ash and Aluminum Cans." *Waste Management* 28, no. 2 (January 2008): 389–95. https://doi.org/10.1016/j.wasman.2006.12.005.
13. Gladston, J. Allwyn Kingsly, I. Dinaharan, N. Mohamed Sheriff, and J. David Raja Selvam. "Dry Sliding Wear Behavior of AA6061 Aluminum Alloy Composites Reinforced Rice Husk Ash Particulates Produced Using Compocasting." *Journal of Asian Ceramic Societies* 5, no. 2 (June 20, 2017): 127–35. https://doi.org/10.1016/j.jascer.2017.03.005.
14. Govindarao, R., R. I. Ganguly, R. R. Dash, P. Surya Prakash Rao, G. S. Reddy, and S. K. Singh. "Development of a Novel Aluminium Based Metal Matrix Composite Using Insitu Ternary Mixture (Al2O3-SiC-C) Prepared by Thermal Treatment of Fly-Ash." *Transactions of the Indian Institute of Metals* 68, no. 5 (October 23, 2015): 951–58. https://doi.org/10.1007/s12666-015-0532-x.
15. Guo, R.Q., Rohatgi, P.K. & Nath, D. "Preparation of Aluminium-Fly Ash Particulate Composite by Powder Metallurgy Technique." *Journal of Materials Science* 32, no. 15 (1997): 3971–74. https://doi.org/https://doi.org/10.1023/A:1018625118090.
16. H.C. Anilkumar, H.S. Hebbar and K.S. Ravishankar. "Mechanical Properties Of Fly Ash Reinforced Aluminium Alloy (Al6061) Composites." *International Journal of Mechanical and Materials Engineering* 6, no. 1 (2011): 41–45.
17. Hemanth, Joel. "Tribological Behavior of Cryogenically Treated B4Cp/Al–12% Si Composites." *Wear* 258, no. 11–12 (June 2005): 1732–44. https://doi.org/10.1016/j.wear.2004.12.009.
18. Ibrahim, R.A. "Tribological Performance of Polyester Composites Reinforced by Agricultural Wastes." *Tribology International* 90 (October 2015): 463–66. https://doi.org/10.1016/j.triboint.2015.04.042.
19. K.K.Alaneme, B.O. Ademilua, M.O. Bodunrin. "Mechanical Properties and Corrosion Behaviour of Aluminium Hybrid Composites Reinforced with Silicon Carbide and Bamboo Leaf Ash." *Tribology in Industry* 35, no. 01 (2013): 25-35.
20. Kathiresan, M., and T. Sornakumar. "Friction and Wear Studies of Die Cast Aluminum Alloy-aluminum Oxide-reinforced Composites." *Industrial Lubrication and Tribology* 62, no. 6 (September 28, 2010): 361–71. https://doi.org/10.1108/00368791011076263.
21. Krishnan, Pradeep Kumar, John Victor Christy, Ramanathan Arunachalam, Abdel-Hamid I. Mourad, Rajaraman Muraliraja, Majid Al-Maharbi, Venkatraman Murali, and Majumder

Accepted Manuscript

Manik Chandra. "Production of Aluminum Alloy-Based Metal Matrix Composites Using Scrap Aluminum Alloy and Waste Materials: Influence on Microstructure and Mechanical Properties." *Journal of Alloys and Compounds* 784 (May 2019): 1047–61. https://doi.org/10.1016/j.jallcom.2019.01.115.

22. Kumar, G. B. Veeresh, C. S. P. Rao, and N. Selvaraj. "Mechanical and Tribological Behavior of Particulate Reinforced Aluminum Metal Matrix Composites – a Review." *Journal of Minerals and Materials Characterization and Engineering* 10, no. 01 (2011): 59–91. https://doi.org/10.4236/jmmce.2011.101005.
23. Kumar, Paras, Harish Hirani, and Atul Agrawal. "Scuffing Behaviour of EN31 Steel under Dry Sliding Condition Using Pin-on-Disc Machine." *Materials Today: Proceedings* 2, no. 4–5 (2015): 3446–52. https://doi.org/10.1016/j.matpr.2015.07.320.
24. Liu, Feng, Junhong Jia, Gewen Yi, Wenzhen Wang, and Yu Shan. "Mechanical and Tribological Properties of NiCr–Al2O3 Composites at Elevated Temperatures." *Tribology International* 84 (April 2015): 1–8. https://doi.org/10.1016/j.triboint.2014.11.023.
25. M. Shamsipour, Z. Pahlevani, M.O. Shabani, A. Mazahery. "Mazahery, Squeeze Casting of Electromagnetically Stirred Aluminum Matrix Nanocomposites in Semi-Solid Condition Using Hybrid Algorithm Optimized Parameters." *Kovove Mater.*, 2017, 33–43. https://doi.org/10.4149/km 2017 1 33.
26. Mazahery, Ali, and Mohsen Ostad Shabani. "Tribological Behaviour of Semisolid–semisolid Compocast Al–Si Matrix Composites Reinforced with TiB2 Coated B4C Particulates." *Ceramics International* 38, no. 3 (April 2012): 1887–95. https://doi.org/10.1016/j.ceramint.2011.10.016.
27. Mishra, Dewika, Farjana J. Sonia, Dinesh Srivastava, G.N. Ganesha, Utpal Singha, and Amartya Mukhopadhyay. "Wear Damage and Effects of Graphene-Based Lubricants/Coatings during Linear Reciprocating Sliding Wear at High Contact Pressure." *Wear* 400–401 (April 2018): 144–55. https://doi.org/10.1016/j.wear.2017.12.024.
28. Mohammed Razzaq, Alaa, Dayang Majid, Mohamad Ishak, and Uday Basheer. "Effect of Fly Ash Addition on the Physical and Mechanical Properties of AA6063 Alloy Reinforcement." *Metals* 7, no. 11 (November 4, 2017): 477. https://doi.org/10.3390/met7110477.
29. Mohammed Razzaq, Alaa, Dayang Majid, Mohamad Ishak, and Uday Basheer. "Effect of Fly Ash Addition on the Physical and Mechanical Properties of AA6063 Alloy Reinforcement." *Metals* 7, no. 11 (November 4, 2017): 477. https://doi.org/10.3390/met7110477.
30. Natarajan, S., R. Narayanasamy, S.P. Kumaresh Babu, G. Dinesh, B. Anil Kumar, and K. Sivaprasad. "Sliding Wear Behaviour of Al 6063/TiB2 in Situ Composites at Elevated Temperatures." *Materials & Design* 30, no. 7 (August 2009): 2521–31. https://doi.org/10.1016/j.matdes.2008.09.037.
31. Nazir, Mian Hammad, Zulfiqar Ahmad Khan, Adil Saeed, Vasilios Bakolas, Wolfgang Braun, and Rizwan Bajwa. "Experimental Analysis and Modelling for Reciprocating Wear Behaviour of Nanocomposite Coatings." *Wear* 416–417 (December 2018): 89–102. https://doi.org/10.1016/j.wear.2018.09.011.
32. Omrani, Emad, Bamdad Barari, Afsaneh Dorri Moghadam, Pradeep K. Rohatgi, and Krishna M. Pillai. "Mechanical and Tribological Properties of Self-Lubricating Bio-Based Carbon-Fabric Epoxy Composites Made Using Liquid Composite Molding." *Tribology International* 92 (December 2015): 222–32. https://doi.org/10.1016/j.triboint.2015.06.007.

Accepted Manuscript

33. Omrani, Emad, Afsaneh Dorri Moghadam, Mahmoud Algazzar, Pradeep L. Menezes, and Pradeep K. Rohatgi. "Effect of Graphite Particles on Improving Tribological Properties Al-16Si-5Ni-5Graphite Self-Lubricating Composite under Fully Flooded and Starved Lubrication Conditions for Transportation Applications." *The International Journal of Advanced Manufacturing Technology* 87, no. 1–4 (October 2, 2016): 929–39. https://doi.org/10.1007/s00170-016-8531-6.
34. Pattanaik, Ashutosh, Mantra Prasad Satpathy, and Subash Chandra Mishra. "Dry Sliding Wear Behavior of Epoxy Fly Ash Composite with Taguchi Optimization." *Engineering Science and Technology, an International Journal* 19, no. 2 (June 2016): 710–16. https://doi.org/10.1016/j.jestch.2015.11.010.
35. Peruzzo, M., F.L. Serafini, M.F.C. Ordoñez, R.M. Souza, and M.C.M. Farias. "Reciprocating Sliding Wear of the Sintered 316L Stainless Steel with Boron Additions." *Wear* 422–423 (March 2019): 108–18. https://doi.org/10.1016/j.wear.2019.01.027.
36. Pirso, Jüri, Mart Viljus, and Sergei Letunovitš. "Friction and Dry Sliding Wear Behaviour of Cermets." *Wear* 260, no. 7–8 (April 2006): 815–24. https://doi.org/10.1016/j.wear.2005.04.006.
37. Prasad, K. Srinivasa, and G. Swami Naidu. "WEAR Behaviour of Nano Fly Ash Reinforced Aluminium-10 Magnesium Alloy Matrix Composites." In *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, 3839–45. IEEE, 2016. https://doi.org/10.1109/ICEEOT.2016.7755432.
38. Prasad, S.V., and R. Asthana. "Aluminum Metal–Matrix Composites for Automotive Applications: Tribological Considerations." *Tribology Letters* 17, no. 3 (October 2004): 445–53. https://doi.org/10.1023/B:TRIL.0000044492.91991.f3.
39. Ram Prabhu, T., M. Murugan, B. P. Chiranth, R. K. Mishra, N. Rajini, P. Marimuthu, P. Dinesh Babu, and G. Suganya. "Effects of Dual-Phase Reinforcement Particles (Fly Ash + Al2O3) on the Wear and Tensile Properties of the AA 7075 Al Alloy Based Composites." *Journal of The Institution of Engineers (India): Series D*, March 6, 2019. https://doi.org/10.1007/s40033-019-00172-7.
40. Rao, R.N., S. Das, D.P. Mondal, and G. Dixit. "Dry Sliding Wear Behaviour of Cast High Strength Aluminium Alloy (Al–Zn–Mg) and Hard Particle Composites." *Wear* 267, no. 9–10 (September 2009): 1688–95. https://doi.org/10.1016/j.wear.2009.06.034.
41. S Charles and V P Arunachalam. "Property Analysis and Mathematical Modeling of Machining Properties of Aluminium Alloy Hybrid (Al-Alloy/SiC/Fly Ash) Composites Produced by Liquid Metallurgy and Powder Metallurgy Techniques." *Indian Journal of Engg. & Material Science* 11 (2008): 473–80.
42. S G Kulkarni, J V Menghani, Achchhe Lal. "Investigation of Mechanical Properties of Fly Ash and Al2O3 Reinforced A356 Alloy Matrix Hybrid Composite Manufactured by Improved Stir Casting." *Indian Journal of Engg. & Material Science* 23, no. 1 (2016): 27–36.
43. Saeidi, F., B. Meylan, P. Hoffmann, and K. Wasmer. "Effect of Surface Texturing on Cast Iron Reciprocating against Steel under Starved Lubrication Conditions: A Parametric Study." *Wear* 348–349 (February 2016): 17–26. https://doi.org/10.1016/j.wear.2015.10.020.
44. Selvi, S., and E. Rajasekar. "Theoretical and Experimental Investigation of Wear Characteristics of Aluminum Based Metal Matrix Composites Using RSM." *Journal of*

Accepted Manuscript

*Mechanical Science and Technology* 29, no. 2 (February 15, 2015): 785–92. https://doi.org/10.1007/s12206-015-0140-z.

45. Shabani, M O, A Mazahery, P Davami, and M Razavi. "Silicon Morphology Modelling during Solidification Process of A356 Al Alloy." *International Journal of Cast Metals Research* 25, no. 1 (January 12, 2012): 53–58. https://doi.org/10.1179/1743133611Y.0000000018.
46. Shaikh, Mohd Bilal Naim, Sajjad Arif, and M Arif Siddiqui. "Fabrication and Characterization of Aluminium Hybrid Composites Reinforced with Fly Ash and Silicon Carbide through Powder Metallurgy." *Materials Research Express* 5, no. 4 (April 6, 2018): 046506. https://doi.org/10.1088/2053-1591/aab829.
47. Shamsipour, Majid, Zahra Pahlevani, Mohsen Ostad Shabani, and Ali Mazahery. "Optimization of the EMS Process Parameters in Compocasting of High-Wear-Resistant Al-Nano-TiC Composites." *Applied Physics A* 122, no. 4 (April 29, 2016): 457. https://doi.org/10.1007/s00339-016-9840-1.
48. Sharma, Pardeep, Satpal Sharma, and Dinesh Khanduja. "Production and Some Properties of Si 3 N 4 Reinforced Aluminium Alloy Composites." *Journal of Asian Ceramic Societies* 3, no. 3 (September 20, 2015): 352–59. https://doi.org/10.1016/j.jascer.2015.07.002.
49. Sharma, Sandan Kumar, B. Venkata Manoj Kumar, Blaz Brodnik Zugelj, Mitjan Kalin, and Young-Wook Kim. "Room and High Temperature Reciprocated Sliding Wear Behavior of SiC-WC Composites." *Ceramics International* 43, no. 18 (December 2017): 16827–34. https://doi.org/10.1016/j.ceramint.2017.09.080.
50. Sharma, Vipin K., R.C. Singh, and Rajiv Chaudhary. "Effect of Flyash Particles with Aluminium Melt on the Wear of Aluminium Metal Matrix Composites." *Engineering Science and Technology, an International Journal* 20, no. 4 (August 2017): 1318–23. https://doi.org/10.1016/j.jestch.2017.08.004.
51. Sharma, Vipin Kumar, Ramesh Chandra Singh, and Rajiv Chaudhary. "Wear and Friction Behaviour of Aluminium Metal Composite Reinforced with Graphite Particles." *International Journal of Surface Science and Engineering* 12, no. 5/6 (2018): 419. https://doi.org/10.1504/IJSURFSE.2018.096753.
52. Singh, Ramesh Chandra, Rajiv Chaudhary, and Vipin Kumar Sharma. "Fabrication and Sliding Wear Behavior of Some Lead-Free Bearing Materials." *Materials Research Express* 6, no. 6 (March 20, 2019): 066533. https://doi.org/10.1088/2053-1591/ab0be4.
53. Suleiman, I. Y., Sani A. Salihu, and T. A. Mohammed. "Investigation of Mechanical, Microstructure, and Wear Behaviors of Al-12%Si/Reinforced with Melon Shell Ash Particulates." *The International Journal of Advanced Manufacturing Technology* 97, no. 9–12 (August 7, 2018): 4137–44. https://doi.org/10.1007/s00170-018-2157-9.
54. Surappa, M. K. "Aluminium Matrix Composites: Challenges and Opportunities." *Sadhana* 28, no. 1–2 (February 2003): 319–34. https://doi.org/10.1007/BF02717141.
55. Suryakumari, T.S.A., S. Ranganathan, and P. Shankar. "Study on Mechanical Properties of Al 7075 Hybrid Metal Matrix Composites." *Applied Mechanics and Materials* 813–814 (November 2015): 230–34. https://doi.org/10.4028/www.scientific.net/AMM.813-814.230.
56. Uvaraja VC, Natrajan N. "Optimization of Friction and Wearbehavior in Hybrid Metal Matrix Composites Using Taguchi Technique." *J Miner Mater Charact Eng* 11 (2012): 757–68.

Accepted Manuscript

57. Vedrtnam, Ajitanshu, and Anuj Kumar. "Fabrication and Wear Characterization of Silicon Carbide and Copper Reinforced Aluminium Matrix Composite." *Materials Discovery* 9 (September 2017): 16–22. https://doi.org/10.1016/j.md.2018.01.002.
58. Vinod, B., S. Ramanathan, V. Ananthi, and N. Selvakumar. "Fabrication and Characterization of Organic and In-Organic Reinforced A356 Aluminium Matrix Hybrid Composite by Improved Double-Stir Casting." *Silicon*, May 5, 2018. https://doi.org/10.1007/s12633-018-9881-5.
59. Zhang, Hui, Meng Hua, Guang-neng Dong, Dong-ya Zhang, and Kwai-Sang Chin. "A Mixed Lubrication Model for Studying Tribological Behaviors of Surface Texturing." *Tribology International* 93 (January 2016): 583–92. https://doi.org/10.1016/j.triboint.2015.03.027.
60. Zhang, J., and A.T. Alpas. "Wear Regimes and Transitions in Al2O3 Particulate-Reinforced Aluminum Alloys." *Materials Science and Engineering: A* 161, no. 2 (April 1993): 273–84. https://doi.org/10.1016/0921-5093(93)90522-G.

Accepted Manuscript

arXiv:1905.05249v1 [cond-mat.mtrl-sci] 13 May 2019

# Electrochemical properties of $Na_{0.66}V_4O_{10}$ nanostructures as cathode material in rechargeable batteries for energy storage applications

Rakesh Saroha,[1] Tuhin S. Khan,[2] Mahesh Chandra,[1] Rishabh Shukla,[1] Amrish K. Panwar,[3] Amit Gupta,[4] M. Ali Haider,[2] Suddhasatwa Basu,[2] and Rajendra S. Dhaka[1,*]

[1]*Department of Physics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi-110016, India*
[2]*Department of Chemical Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi-110016, India*
[3]*Department of Applied Physics, Delhi Technological University, Delhi-110042, India*
[4]*Department of Mechanical Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi-110016, India*

(Dated: May 15, 2019)

**ABSTRACT:** We report the electrochemical performance of nanostructures of $Na_{0.66}V_4O_{10}$ as cathode material for rechargeable batteries. The Rietveld refinement of room temperature x-ray diffraction pattern shows the monoclinic phase with C2/m space group. The cyclic voltammetry curves of prepared half-cells exhibit redox peaks at 3.1 and 2.6 V, which are due to two-phase transition reaction between $V^{5+/4+}$ and can be assigned to the single step deintercalation/intercalation of Na-ion. We observe a good cycling stability with specific discharge capacity (measured vs. $Na^+/Na$) between 80 (±2) and 30 (±2) mAh $g^{-1}$ at a current density 3 and 50 mA $g^{-1}$, respectively. The electrochemical performance of $Na_{0.66}V_4O_{10}$ electrode was also tested with Li anode, which showed higher capacity, but decay faster than Na. Using density functional theory, we calculate the Na vacancy formation energies; 3.37 eV in the bulk of the material and 2.52 eV on the (100) surface, which underlines the importance of nanostructures.

## INTRODUCTION

After commercialization in early 1990s, rechargeable Li-ion batteries (LIBs) are being widely investigated and used as energy storage to power portable electronic devices and hybrid electric vehicles [1, 2]. This is mainly because LIBs show the lowest redox potential of the Li ($E_{Li/Li^+}$ = -3.04 V vs. SHE), which allows possessing high voltage and energy density. Also, small ionic radius of $Li^+$(∼0.76 Å in six coordination state) allow its smooth diffusion during charging/discharging, which makes long cycle life [3, 4]. However, due to huge demand, high cost, safety concerns, limited and non-uniform Li resources [5], the Na-ion batteries (NIBs) have generated considerable interest as the most promising alternative (due to the uniform distribution and abundance of Na resources in the earth crust) to the LIBs for large-scale energy storage systems [6–12]. The elements Na and Li are from the same group in the periodic table and have similar properties such as reactivity, physical strength, etc. The global abundance, low cost and appropriate redox potential ($E_{Na/Na^+}$ = -2.71 V vs. SHE) prove its suitability as a good substitute for Li. However, large ionic radius of Na-ion (∼1.02 Å in the six coordination environment) as compared to the Li leads to slow ionic diffusion and lower energy density [13, 14]. Owing to these facts, materials with open framework or materials which can exchange more than one Na per formula unit, or Na-rich materials are crucial for the possibility of having high power and energy density [15, 16].

In this context, layered transition metal oxides having formula $NaTO_2$ (T= transition metal) and polyanionic compounds have been extensively explored and used as a cathode material for Na-ion battery [17–24]. For example, vanadium oxychloride [25], $Na_{0.44}Mn_{1-x}Zn_xO_2$ [26, 27], $NaNi_{1/3}Mn_{1/3}Co_{1/3}O_2$ [28], $Na_3V_2(PO_4)_3$ [29], olivine-type phospahte and sulphate based cathodes like $NaFePO_4$ [30], $NaFeSO_4F$ [31] and even prussian blue derivatives which contain a suitable transition metal such as $KTFe(CN)_6$ [32], which acts as an open host framework with a large interstitial space to absorb bigger $Na^+$. However, due to their complex reaction mechanism and in order to improve thermal stability, energy density and cycling performance, it is vital to search new cathode materials and investigate their physical/electro-chemical properties for Na-ion batteries [33, 34].

Interestingly, in recent years, vanadium based oxide materials have attracted worldwide attention as a possible alternate cathode material for both Li-ion and Na-ion batteries [35–41]. Because of the variable oxidation state of vanadium (+5 in $V_2O_5$ to +2 in VO), it acts as electron donor and acceptor during the process of Na extraction and insertion [42]. The crystal structure of $V_2O_5$ consists of three different vanadium sites labeled as V(1), V(2) and V(3) [39]. The edge sharing vanadium octahedra [V(1)O6] and corner sharing vanadium octahedral [V(2)O6] combine to form zigzag and double chains along the *b* axis, respectively. The edge sharing [V(3)O5] pyramid forms connection to the $[V_4O_{11}]$ in layers, which are formed from the oxygen linkage to the V(1) and V(2) octehedra, thus yielding a 3D tunnelled structure [43]. The formed 3D tunnelled structure is not only more stable as compared to the layered structures, but also provide fast ion reaction kinetics. It has been reported that depending upon the $Li^+/Na^+$ intercalation amount, $V_2O_5$ can yield high discharge capacities [44, 45]. In Li-ion batteries, $Li_xV_2O_5$ is explored extensively [46–49], but it suffers significant capacity loss dur-

ing repetitive charging/discharging for an intercalated amount of $x > 1$ [45]. Analogous to $Li_xV_2O_5$ in Li-ion, $Na_xV_2O_5$ was also proposed for Na-ion battery, where the Na cations can be reversibly cycled along the $b$ axis in a similar manner as the Li counterpart. Within pentoxide based cathode, $Na_{0.33}V_2O_5$ has gained significant attention as a cathode material for Li-ion/Na-ion batteries because of high experimental discharge capacity [43, 50–52]. Bach, *et. al.* have first investigated the Na intercalation and de-intercalation behavior in $Na_{0.33}V_2O_5$ bronze [50] and then similarly, $NaV_6O_{15}$ [$(Na_{0.33}V_2O_5)_3$] have been studied as a cathode material in various morphologies such as nanoplates, nanoroads, nanoflowers, and nanoflakes for Li/Na-ion batteries by various research groups [39, 44, 53, 54]. Jiang et al. synthesized $NaV_6O_{15}$ nanoplates and reported a discharge capacity of 116 mAh $g^{-1}$ measured vs. $Na^+/Na$ with a cycle retention of 55% at a current density of 50 mA $g^{-1}$ [53]. However, note that the electrochemical performance of $Na_{0.33}V_2O_5$ critically depends on the synthesis conditions, size in nanometer and morphology [39, 44, 53, 54]. Therefore, nanostructures of $Na_{0.66}V_4O_{10}$ [$(Na_{0.33}V_2O_5)_2$] in different morphologies still need to be explored in detail as cathode material for Na-ion batteries. Also, it have been reported that partial substitution of Li with Na in $Li_{2-x}Na_xMnO_3$ improve the electrochemical performance and significantly enhance the cycling stability of Li-ion battery [55]. This further motivates to test our $Na_{0.66}V_4O_{10}$ cathode with Li as counter electrode.

In this paper, we synthesize nanorods of $Na_{0.66}V_4O_{10}$ and study the electrochemical performance for a promising cathode material in Na-ion batteries. The Rietveld refinement of x-ray diffraction pattern reveals the monoclinic phase with C2/m space group. The transmission electron microscopy (TEM) and scanning electron microscopy (SEM) studies reveal the rod shaped morphology with agglomeration. The size and the length of the rods were found to be 50-100 nm and few $\mu$m, respectively. The cyclic voltammetry (CV) curves of prepared half-cells exhibit redox peaks at $\approx$ 3.1 and 2.6 V, which indicate phase transition reaction in $V^{5+/4+}$ and can be assigned to the single step deintercalation/intercalation of Na-ion. Interestingly, we observed good cycling stability with specific discharge capacity of 80 ($\pm$2) and 30 ($\pm$2) mAh $g^{-1}$ at a current density of 3 and 50 mA $g^{-1}$, respectively, measured vs. $Na^+/Na$ anode. We have also tested the electrochemical performance with Li anode. The energetics of Na vacancy formation on discharging and Na vacancy filling on charging could form a potential descriptor for electrochemical performance. Since Na vacancy formation was relatively easier on the electrode surface, a surface diffusion pathway for Na transport could be hypothesized. For nanostructures, the available surface area for this pathway is likely to increase, which further alludes to the measured electrochemical performance.

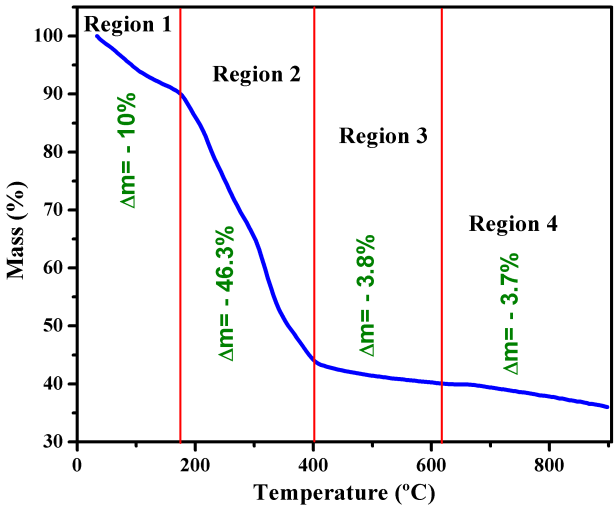



FIG. 1. Thermogravimetric analysis (TGA) of the precursors used to synthesize $Na_{0.66}V_4O_{10}$ samples.

## RESULTS AND DISCUSSION

### Structural Analysis

In order to decide the annealing temperature for stable phase formation, we performed the thermal analysis of the resulting gel precursor under air atmosphere using thermogravimetric analysis (TGA) profile. Figure 1 shows the variation in mass (%) of the gel precursor with increase in temperature. The mass loss of around 10% below 180°C (region 1) is due to the removal of physically adsorbed as well as some intercalated water molecules [56]. The steep mass loss ($\Delta$m = 46.3%) observed between 180-400°C in region 2 is attributed to the removal of acetate precursors and volatile impurities in the form of $NH_3$ and $CO_2$ gases. A vigorous oxidation and decomposition reactions take place in this region leading to the phase formation of NVO material. Further, negligible mass loss was observed for region 3 and 4, i.e. above 400°C in TGA profile. It can be observed from TGA profile that annealing at about 400°C would be sufficient for the phase formation of $Na_{0.66}V_4O_{10}$ (NVO) material. We have also tried annealing at higher temperatures and found good thermal stability of the prepared material.

The x-ray diffraction (XRD) pattern along with the corresponding Rietveld refinement of the prepared $Na_{0.66}V_4O_{10}$ (NVO) sample are shown in Figure 2(a). The XRD pattern was matched with the standard phase using X-pert high score plus software with PDF2 reference data file. The obtained XRD pattern of the synthesized NVO powder could be well-indexed with the pure phase of the pristine sodium vanadium bronze i.e. $\beta$-$Na_{0.33}V_2O_5$ having monoclinic structure with C2/m space group (JCPDS code: 48-0382), as also reported in

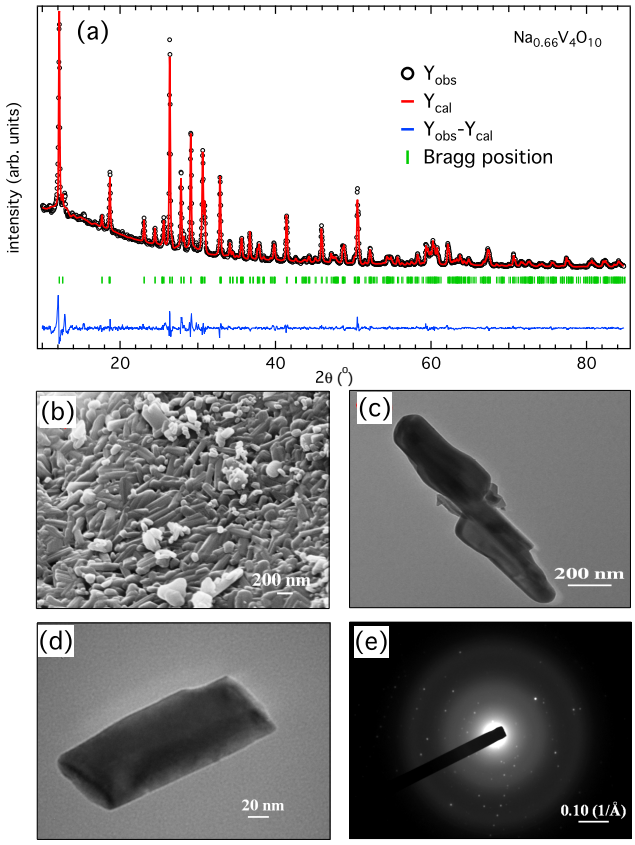



FIG. 2. Structural and morphological properties of prepared $Na_{0.66}V_4O_{10}$ electrode material: (a) XRD pattern with Rietveld refinement, (b) SEM micrograph, (c, d) TEM images at different scale, and (e) SAED pattern. The label 0.10 (1/Å) is the scale bar in reciprocal lattice is equivalent to 1 (1/nm).

ref. [57]. The sharp and intense peaks reveal the crystalline nature of the sample. The Rietveld refinement was performed to obtain the quantitative crystal structure information such as lattice parameters. The peaks were modelled using the Pseudo-Voigt function and the full width at half maximum (FWHM) was refined using Lorentzian broadening of the XRD peaks. The structural refinement results ($\chi^2$=2.47, $R_p$ = 4.83%, $R_{wp}$ = 6.6%) indicate that the fitting is reliable and good. The obtained lattice parameters from refinement are: $a$ = 15.426 Å, $b$ = 3.610 Å, $c$ = 10.073 Å and $\beta$= 109.55$^o$ are in good agreement with previous reports of $Na_{0.33}V_2O_5$ [58]. Further, we have provided the fractional coordinates and site occupancies in the following Table I.

We further characterize the prepared NVO cathode material to check the morphological properties. The SEM micrograph reveals the formation of uniformly distributed highly agglomerated rod-shaped morphology, as shown in Figure 2(b). However, the length of the rods varies dramatically, which give rise to some plate-like

TABLE I. Atomic coordinates, site occupancies and isotropic displacement parameters for as-prepared $Na_{0.66}V_4O_{10}$, obtained from the Rietveld refinement of room temperature x-ray diffraction spectrum.

| atom | site | x | y | z | $B_{iso}$(Å$^2$) | Occ. |
|---|---|---|---|---|---|---|
| Na | 4$i$ | 0.0045(12) | 0.00000 | 0.4100(16) | 0.5 | 0.775 |
| V1 | 4$i$ | 0.3382(3) | 0.00000 | 0.0987(5) | 0.5 | 1.337 |
| V2 | 4$i$ | 0.1196(3) | 0.00000 | 0.1198(5) | 0.5 | 1.334 |
| V3 | 4$i$ | 0.2885(3) | 0.00000 | 0.4059(6) | 0.5 | 1.334 |
| O1 | 2$a$ | 0.00000 | 0.00000 | 0.00000 | 0.5 | 0.977 |
| O2 | 4$i$ | 0.8156(10) | 0.00000 | 0.0509(17) | 0.5 | 1.202 |
| O3 | 4$i$ | 0.6345(9) | 0.00000 | 0.0796(16) | 0.5 | 1.295 |
| O4 | 4$i$ | 0.4334(9) | 0.00000 | 0.2192(15) | 0.5 | 1.394 |
| O5 | 4$i$ | 0.2666(9) | 0.00000 | 0.2175(17) | 0.5 | 1.299 |
| O6 | 4$i$ | 0.1078(9) | 0.00000 | 0.2599(15) | 0.5 | 1.357 |
| O7 | 4$i$ | 0.2391(11) | 0.00000 | 0.5709(16) | 0.5 | 1.229 |
| O8 | 4$i$ | 0.3981(9) | 0.00000 | 0.4597(16) | 0.5 | 1.254 |

particles over rod-shaped morphology. The TEM images in Figures 2(c, d) clearly confirm the formation of nano-sized rod-shaped morphology of the NVO sample, in agreement with the SEM analysis. The dark and slightly greyish regions correspond to the NVO nanorods, which display uniform dimensions with 50–100 nm wide and several micrometers long. The cathode materials in the nanosized particles/rods shape are expected to enhance the electrochemical performance of Na-ion batteries. Note that the characteristic diffusion length (L) of nanostructures and the diffusion coefficient (D) of Na-ion are related as: $\tau = L^2/D$, where $\tau$ is the time for Na-ion to diffuse through the electrode material. In this context, one can achieve better electrochemical performance by decreasing L, which effectively reduce $\tau$. The selected area electron diffraction pattern (SAED) for the synthesized NVO sample is shown in Figure 2(e), which reveals the existence of sharp diffraction spots along with diffused circles which represents polycrystllaine nature of the synthesized sample. The interplanar spacing/distance ($d$ = 0.47 nm) calculated using SAED pattern matches to the (200) peak of the XRD pattern.

### Electrochemical Performance

Now we present, in Figures 3(a–d), the electrochemical performance of the prepared NVO cathode in terms of cyclic voltammetry, galvanostatic charge/discharge curves, rate performance and Coulombic efficiency. Figure 3(a) shows the cyclic voltammograms (CV) of prepared NVO electrode in the potential range of 1.5–4.0 V at a scan rate of 0.05 mV $s^{-1}$, measured v/s $Na^+$/Na at room temperature. Interestingly, it exhibits redox peaks at around 3.1 and 2.6 V, which are due to two-phase transition reactions between $V^{5+/4+}$ and can be assigned

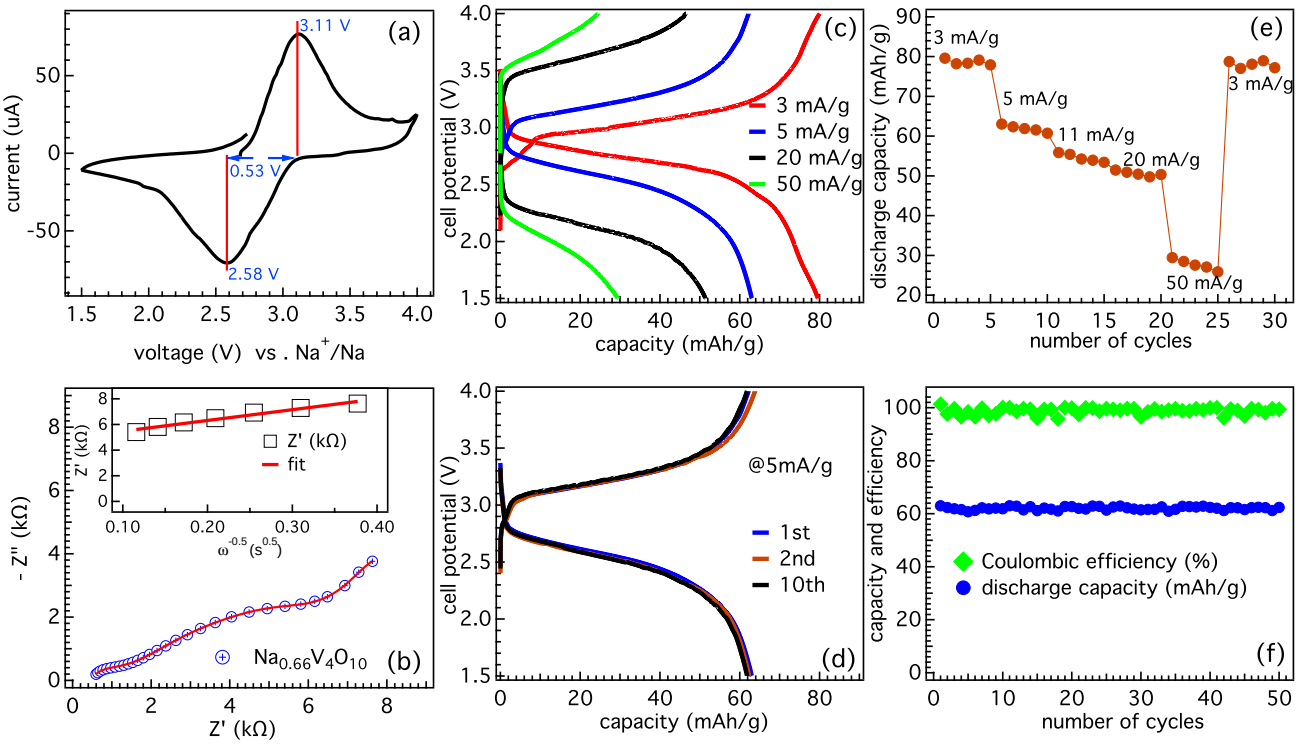



FIG. 3. Electrochemical properties of prepared NVO cathode with Na anode: (a) Cyclic voltammetry (CV) at a scan rate of 0.05 mV $s^{-1}$ vs. $Na^+$/Na at room temperature, (b) Electrochemical impedance spectra of as prepared NVO sample at AC amplitude of 5 mV, inset shows a plot between Z' and $\omega^{-0.5}$. (c) Initial charge/discharge curves at different current densities, (d) 1st, 2nd and 10th charge/discharge curve at 5 mA $g^{-1}$ current density, (e) The rate performance at different current densities, (f) The cycling performance and Coulombic efficiency for 50 cycles measured at 5 mA $g^{-1}$.

to the single step de-intercalation/intercalation of Na-ion in the $Na_{0.66}V_4O_{10}$ electrode. It is interesting to note that no such characteristics have been observed in vanadium based oxide cathodes. Liu *et al.* reported the charge/discharge profile of $NaV_6O_{15}$ $[(Na_{0.33}V_2O_5)_3]$, however, that have multiple steps/plateaus [39]. The obtained CV results are consistent with charge-discharge profiles, as shown in Figure 3(c). The insertion and extraction behaviour of Na-ion can be tentatively given as:

$$Na_{0.66}V_4O_{10} \longleftrightarrow Na_{0.66-x}V_4O_{10} + xNa^+ + xe^- \quad (1)$$

Figure 3(c) shows the galvanostastic first charge-discharge curves of prepared CR2016 half-cells using NVO as cathode and recorded at various current densities from 3 mA $g^{-1}$ to 50 mA $g^{-1}$ in the potential range of 1.5–4.0 V vs. $Na^+$/Na. The specific discharge capacity of the NVO sample is observed to be 80 (±2), 64 (±2), 56 (±2), 52 (±2) and 30 (±2) mAh $g^{-1}$ at a current density of 3, 5, 11, 20 and 50 mA $g^{-1}$, respectively. Assuming no electrolyte decomposition during the first charge, $Na_{0.66}V_4O_{10}$ charged to $Na_{0.09}V_4O_{10}$ (corresponding to the charge capacity of 80 mAh $g^{-1}$ at 3 mA $g^{-1}$) and then, back to $Na_{0.66}V_4O_{10}$. Note that the discharge capacity of about 80 mAh $g^{-1}$ (1C equated to 141 mAh $g^{-1}$) corresponds to +4.86 oxidation state of vanadium, consistent with the CV results. It can also be seen from the Figure 3(c) that as current density increases from 3 mA $g^{-1}$ to 50 mA $g^{-1}$, the specific discharge capacity value decreases from 80 (±2) to 30 (±2) mAh $g^{-1}$ with an apparent drop in potential plateau, which is due to the increase in polarization and resistance or iR drop of the electrode at high current densities. Note that the OCV for NVO cathode was observed to be around 2.8 V. The GCD results in Fig. 3(c) show the charge/discharge voltage plateau around 2.9/2.8 V, respectively at 3 mA $g^{-1}$ vs. Na/$Na^+$. The relatively high over potential indicates that the activation energy of the redox event is very high because of which the potential necessary to transfer an electron from NVO cathode is relatively high even at low current density. The comparison of 1st, 2nd and 10th cycle discharge capacity curves are shown in Figure 3(d). The discharge capacity remains fairly constant for the ten cycles, which elaborates the structural stability of the electrode material. The cycling stability and rate performance for an electrode are very important parameters to determine the stability in long-term applications. Figsure 3(e, f) show the stepwise electrochemical rate and cycling performance, respectively. As observed

from Figure 3(e), the capacity decreases with increase in current density and goes to minimum for 50 mA $g^{-1}$. Figure 3(f) shows the cycling performance of the synthesized NVO sample at 5 mA $g^{-1}$ current density at room temperature. It shows an initial discharge capacity 64 ($\pm$2) mAh $g^{-1}$ and retains nearly 100% capacity even after 50 charge/discharge cycles. The Coulombic efficiency of the NVO electrode for 50 charge/discharge cycles at 5 mA $g^{-1}$ current density is also shown in Figure 3(f). The Coulombic efficiency for all the charge/discharge cycles is observed to be more than 95% [Figure 3(f)].

Furthermore, the electrochemical impedance spectra (EIS) of fresh cells were performed to measure the electrode/electrolyte resistance. Figure 3(b) shows the impedance spectra of NVO sample measured with an AC voltage pulse having amplitude of 5 mV. The EIS curve, measured in the frequency range of 10 mHz to 100 kHz, shows a depressed semi-circle and a straight line in the high and low-frequency regions, respectively. In general, the ohmic resistance ($R_s$) of the cell, i.e. the resistance due to electrolyte and electrode material, can be estimated by an intercept at the $Z^{'}$-axis in the high-frequency region. The information about the electrochemical reactions taking place at the electrode/electrolyte interface, which indicate the charge transfer resistance ($R_{ct}$), can be obtained from the semi-circle in the middle-frequency range. Also, the Warburg impedance ($Z_w$), which is associated with Na-ion diffusion in the electrode active material, can be represented by the inclined line in the low-frequency region. The value of cell parameters were obtained to be $R_s$= 594 $\Omega$ and $R_{ct}$= 778 $\Omega$.

The chemical diffusion coefficient of the Na ions inside an electrode material is calculated using the following equation:

$$D = \frac{T^2R^2}{2n^4A^2F^4C^2\sigma_w^2} \quad (2)$$

where, D, T, R, $n$, A, C, and F are diffusion coefficient ($cm^2$ $s^{-1}$), absolute temperature (K), gas constant (8.314 J $mol^{-1}$ $K^{-1}$), number of electrons involved in the redox process, electrode area (2.2 $cm^2$), the Na-ion concentration ($3.139\times10^{-3}$ mol $cm^{-3}$), and the Faraday constant (96486 C $mol^{-1}$), respectively. The $\sigma_w$ is the Warburg impedance coefficient, which is related to $Z^{'}$ by the following equation:

$$Z' = R_s + R_{ct} + \sigma_w\omega^{-0.5} \quad (3)$$

In the inset of Figure 3(b), we show the obtained value of $\sigma_w$ from the slope of $Z^{'}$ and $\omega^{-0.5}$, which found to be 7739 $\Omega$ $s^{-0.5}$. The value of $n$ (=59 mV/$\Delta V$) can be calculated by taking the difference between oxidation and reduction peaks ($\Delta V$ = 0.53 V) in CV [see Figure 3(a)], which found to be 0.1 (number of electrons participating in charging and discharging process). Finally by using these values in equation 2, the obtained value of the diffusion coefficient is $1.239\times10^{-13}$ $cm^2$ $s^{-1}$, which is in good agreement with the one reported in ref. [59].

It is important to compare the observed capacity in this manuscript with the other related cathode materials reported in the literature [60–62]. For example, Liu *et al.*, observed the discharge capacity of about 100 $mAhg^{-1}$ at 0.1 C current rate in $Na_3V_2(PO_4)_3$/C nanofibers cathode materials [60]. Zhang *et al.*, found the reversible capacity of about 120 $mAhg^{-1}$ for $NaVO_3$ cathode after the activation at high-voltage [61]. More recently, Shinde *et al.*, used the ultrasonic sonochemical synthesis method to prepare $Na_{0.44}MnO_2$ cathode, which shows the reversible capacity of about 110 $mAhg^{-1}$ at a current rate of C/10 with good cycling stability [62]. By changing the preparation method to self-combustion, the $NaV_6O_{15}$ nanoplates show the capacity of about 150 $mAhg^{-1}$ at a current rate of 20 $mAg^{-1}$, which remains about 82 $mAhg^{-1}$ after 30 cycles [53]. On the other hand, $NaV_6O_{15}$ nanorods prepared by PVP-modulated synthesis route shows high initial capacity of about 157 $mAhg^{-1}$ at 20 $mAg^{-1}$ current density [41]. It should be noted that the discharge capacity of NVO cathode crucially depend on the preparation method, amount of Na content in the material and morphology of the nanoparticles as well as type of intercalation $Li^+/Na^+$.

In order to test the feasibility of prepared $Na_{0.66}V_4O_{10}$ (NVO) as a cathode material for Li-ion batteries, we assembled the half-cells vs. $Li/Li^+$ as an anode and performed the electrochemical measurements with two electrolytes i.e. 1 M $NaClO_4$ and $LiPF_6$ (EC:DMC 1:1 V/V) to find which one is better suited for this purpose. These cells are abbreviated as NVON (for $NaClO_4$) and NVOL (for $LiPF_6$), respectively. The Cyclic voltammograms of the prepared NVO vs. $Li/Li^+$ cells are shown in Figure 4(a) for both the electrolytes. The data was recorded at a scan rate of 0.1 $mVs^{-1}$. Both the half-cells show well defined oxidation/reduction peaks [Figure 4(a)]. However, the intensity of redox peaks in the NVOL half-cell is higher than those of NVON redox peaks, which could be attributed to the better Na-ion diffusion. The redox pairs observed at 3.5/2.7 V, 3.8/3.1 V (for NVOL) and 3.5/2.7 V, 3.9/2.8 V (for NVON) can be due to intercalation of Li-ions in the host NVO structure [Figure 4(a)]. The EIS results of both the half-cells are presented in Figure 4(b), where it can be observed that NVOL half-cell possess lower charge-transfer resistance as compared to the NVON half-cell, which is well consistent with the CV results. The charge-resistance was found to be around 800 and 1300 $\Omega$ for the NVOL and NVON half-cells, respectively [Figure 4(b)]. The CV and EIS results indicate that NVOL half-cell would display the better electrochemical properties as compared to the NVON half-cell.

Figures 4(c, d) show the initial charge/discharge curves for both the half-cells (NVOL and NVON) and the results are in well agreement with CV and EIS results. One slopping voltage plateau around 3.5 V and other flat plateau

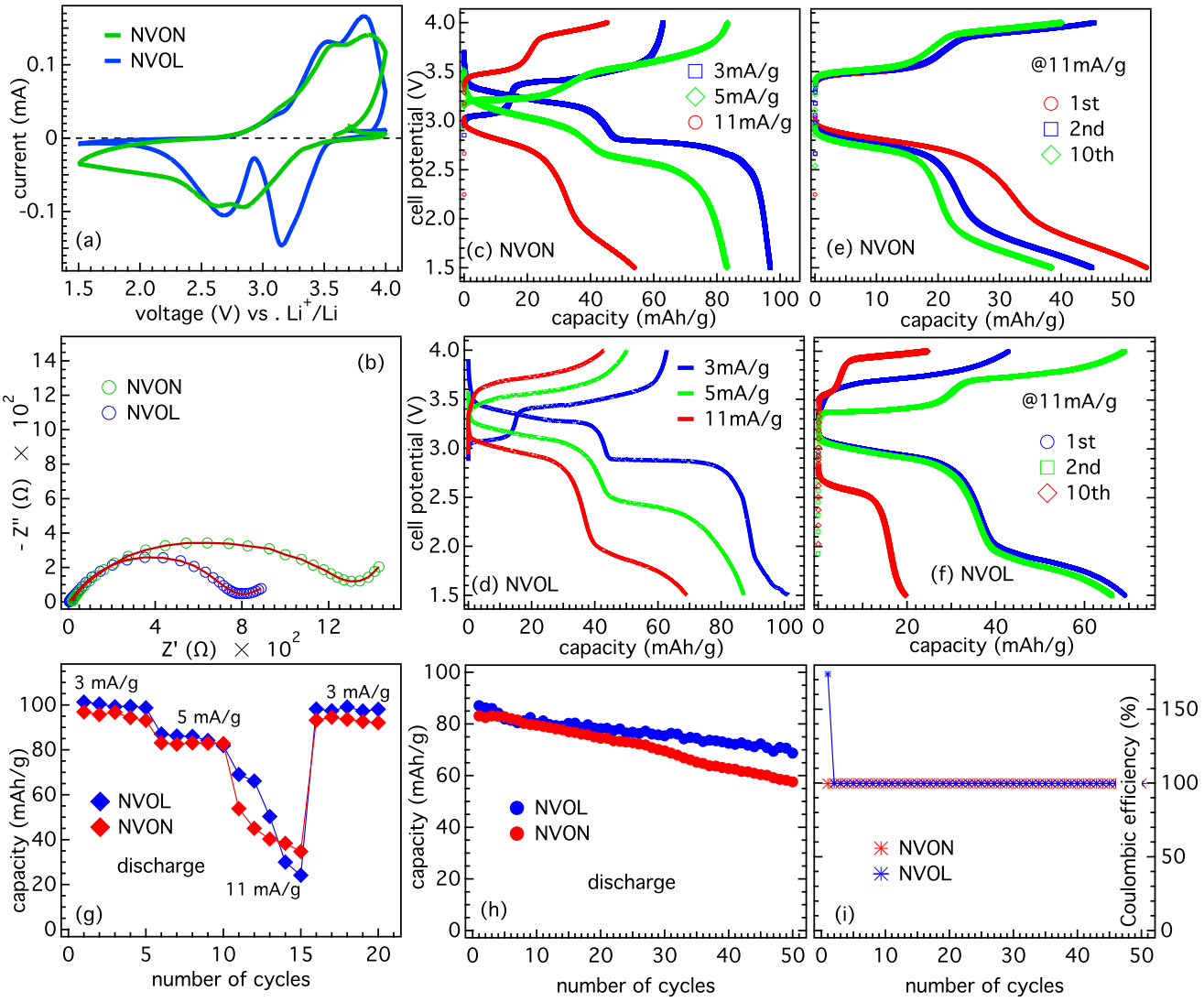



FIG. 4. (a) Cyclic voltammetry, and (b) EIS results of the NVO half-cell assembled with $NaClO_4$ (NVON) and $LiPF_6$ (NVOL) electrolyte vs. $Li/Li^+$. Initial charge/discharge comparison at different current densities for (c) NVON (d) NVOL. The charge/discharge comparison of 1st, 2nd and 10th cycles measured at 11 mA $g^{-1}$ for (e) NVON (f) NVOL. (g) The rate performance of the NVOL and NVON half-cells at different current densities. (h) The cycling performance of the half-cells at current density of 5 mA $g^{-1}$ for 50 charge/discharge cycles, and (i) the Columbic efficiency of the half-cells.

around 2.7 V can be clearly seen. However, as the current density increases from 3 mA $g^{-1}$ to 11 mAh $g^{-1}$, the voltage plateaus shift to the lower side due to an increase in the polarization and the cell resistance. The specific discharge capacity was found to be around 101, 87 and 69 mAh $g^{-1}$ at the current density 3, 5 and 11 mA $g^{-1}$, respectively for NVOL half-cell. On the other hand, for NVON half-cell the obtained values are around 97, 83 and 55 mAh $g^{-1}$ at the current density 3, 5 and 11 mA $g^{-1}$, respectively. The intercalation and deintercalation of the Na-ion in the host structure can be understood in the following reaction: During first discharge process, Li-ion moves in to the host NVO electrode:

$$Na_{0.66}V_4O_{10} + xLi^+ + xe^- \rightarrow Li_xNa_{0.66}V_4O_{10} \quad (4)$$

During the charging cycle, $Na^+$ ions together with $Li^+$ ions will be removed from the host structure:

$$Li_xNa_{0.66}V_4O_{10} - xLi^+ - yNa^+ - (x+y)e^- \rightarrow Na_{(0.66-y)}V_4O_{10} \quad (5)$$

Thus obtained $Na_{(0.66-y)}V_4O_{10}$ will again combined with Li-ion during discharge:

$$Na_{(0.66-y)}V_4O_{10} + zLi^+ + ze^- \rightarrow Li_zNa_{(0.66-y)}V_4O_{10} \quad (6)$$

Therefore, during repeated charging/discharging, both the Na and Li-ion will contribute in the electrochemical performance of the battery.

In Figures 4(e, f), we show the charging/discharging comparison of the $1^{st}$, $2^{nd}$ and $10^{th}$ cycles for both the half-cells recorded at the current density of 11 mA $g^{-1}$. The NVOL half-cell displays a large capacity fading over consecutive cycles as compared to the NVON cell. The specific discharge capacity of NVOL half-cell decrease from 69 mAh $g^{-1}$ to 20 mAh $g^{-1}$ ( 72% decrease), while in NVON half-cell the decrease is 31% (from 53 mAh $g^{-1}$ to 38 mAh $g^{-1}$) in $10^{th}$ cycle. Figures 4(g, h) show the rate and cycling performance of the two half-cells vs. $Li/Li^+$, which show comparable capacities up to 5 mA $g^{-1}$ and for 11 mA $g^{-1}$, the capacity in NVOL cells decays faster as compared to NVON cells, see Figure 4(g). The specific discharge capacity decays 76% for NVOL, while for NVON half-cell the capacity decays 64%, when the current density changes from 3 mA $g^{-1}$ to 11 mA $g^{-1}$. Now we show in Figure 4(h) the cycling performance measured at the current density of 5 mA $g^{-1}$ up to 50 charge/discharge cycles. Both the cells show a gradual decrease in the discharge capacity as capacity decreases from 106 to 69 mAh $g^{-1}$ for the NVOL and 83 to 57 mAh $g^{-1}$ for the NVON half-cell. The Columbic efficiency for both the half-cells is measured at 5 mA $g^{-1}$ current density [Figure 4(i)]. For both the half-cells, we observed 100% Coulombic efficiency for all the cycles except the first one for NVOL, which is unexpectedly higher 176%. The possible explanation for such high value in first cycle is that during the discharge there is an extra charge (or electrons) because of the presence of Li in the host structure in the alloy form ($Li_xNa_{0.66}V_4O_{10}$). This extra charge contributes in the capacity during the discharge, while during charging only $Na_{0.66}V_4O_{10}$ remains (see equations 4 and 5) and therefore, the charge capacity is lower, which reflect in the calculation of the Coulombic efficiency.

### Density Functional Theory

The density functional theory (DFT) is used to calculate charging and discharging energy of NVO structures with different Na content. The optimized NVO crystal structures, with different Na content are shown in the upper panel of Figure 5. Here, the structures of $Na_{1.0}V_4O_{10}$, $Na_{0.66}V_4O_{10}$, $Na_{0.33}V_4O_{10}$, and $V_4O_{10}$, were obtained by sequential removal of Na atoms, whereas structure of $Na_{1.66}V_4O_{10}$ was obtained by adding one Na atom to $Na_{1.33}V_4O_{10}$. The sequential removal of Na atom from $Na_{1.0}V_4O_{10}$ to $Na_{0.66}V_4O_{10}$, was done keeping in mind the highest separation of Na vacancy. The other structure obtained by sequential removal of Na atoms are unique as the Na atoms are equivalent. The charging, represents the energy gain by the

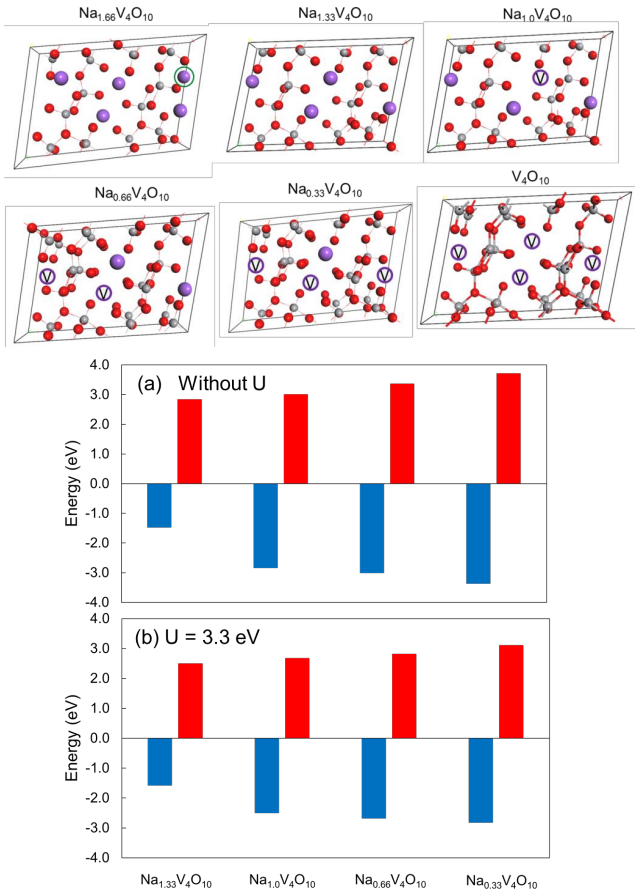



FIG. 5. DFT optimized geometry of NVO structures with different Na content, showing the position of Na vacancy (violet circle with V) and extra Na (green circle). DFT calculated Na vacancy formation energies (red) on discharging and Na filling energies (blue) on charging in the bulk of the NVO structures with different Na content, (a) without U, and (b) with $U_{eff}$ = 3.3 eV for Vanadium.

addition of Na; whereas discharging represents the energy required to remove a Na ion. Thus, charging/discharging process can be defined as:

$$Na_xV_4O_{10} + Na \rightarrow Na_{x+1}V_4O_{10} \quad \text{(charging)} \qquad (7)$$

$$Na_xV_4O_{10} \rightarrow Na_{x-1}V_4O_{10} + Na \quad \text{(discharging)} \qquad (8)$$

As the number of Na atoms in $Na_xV_4O_{10}$ decreases from 1.33 to 0.33, a gradual increase in both the energy gain during charging and also the energy required for discharging, is observed. The $Na_{1.33}V_4O_{10}$ has the lowest Na vacancy formation energy ($E_v^{Na}$ = 2.84 eV), but the energy gain during the charging is low ($E_f^{Na}$ = - 1.48 eV). Hence, $Na_{1.33}V_4O_{10}$ cannot be a good candidate for the battery operation. Whereas, $Na_{0.33}V_4O_{10}$ shows a very high energy storage capability due to its high energy gain during the Na addition ($E_f^{Na}$ = - 3.37 eV ), but it also

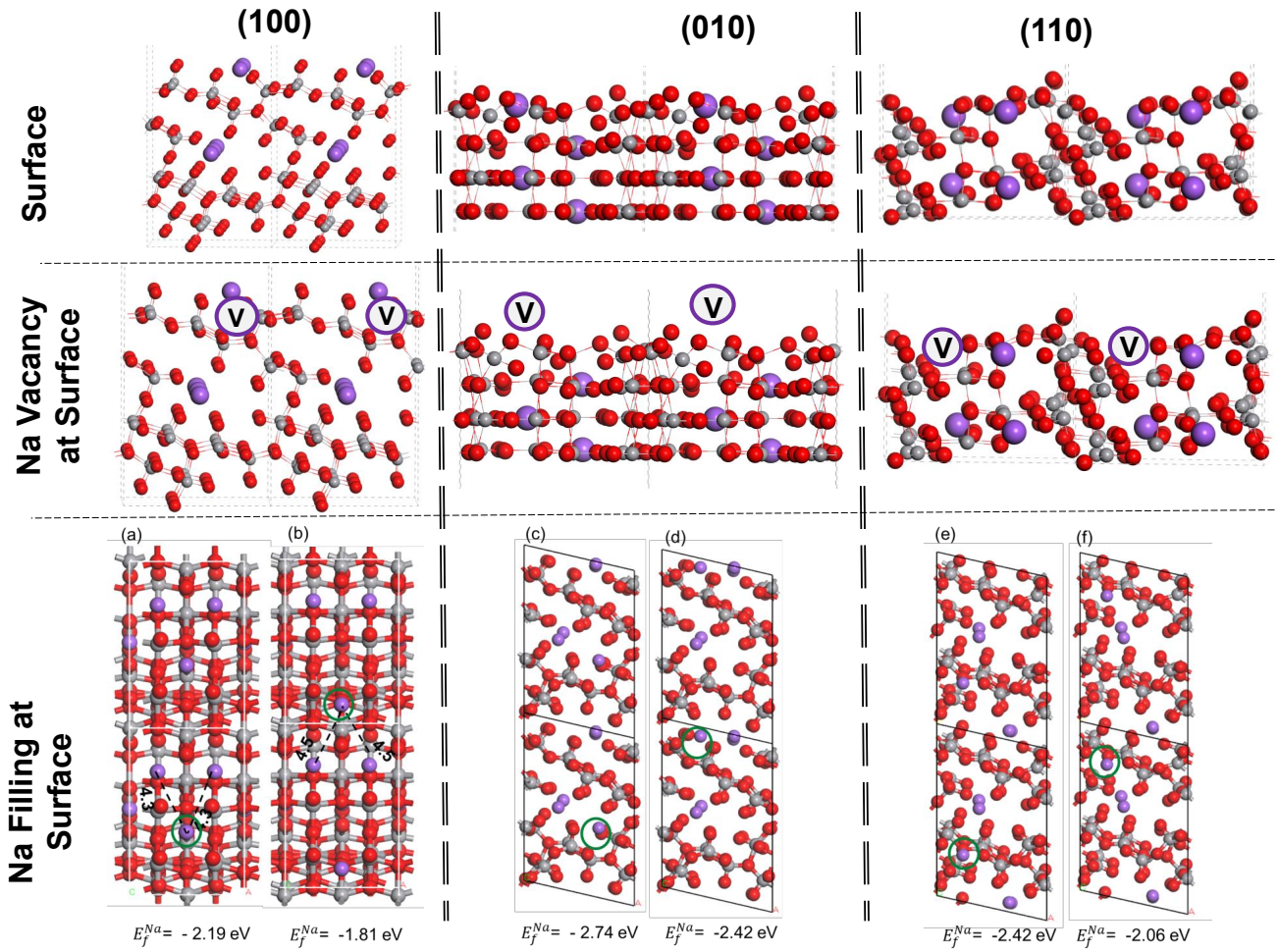



FIG. 6. DFT optimized geometry of (100), (010) and (110) surfaces of $Na_{0.66}V_4O_{10}$, showing the position of Na vacancy (violet circle with V) and extra Na (green circle). Lower panels (a–f) show the position extra Na atom during the Na filling (green circle) for different surfaces, energies are indicated at the bottom. The distances between the Na atoms in (a) and (b) are given in Angstrom unit.

have a very high Na vacancy formation energy ($E_v^{Na}$ = 3.71 eV), hence also not a good candidate for battery application. The $Na_{0.66}V_4O_{10}$ samples have the moderate values for both the energy gains in Na addition ($E_f^{Na}$ = - 3.01 eV) and the energy required for creating Na vacancy ($E_v^{Na}$ = 3.37 eV), see Figure 5(a) for calculations performed without considering U for Vanadium. These two characteristics directly point towards the potential of $Na_{0.66}V_4O_{10}$ as a good battery candidate. However, the energy required for creating Na vacancy at the bulk of $Na_{0.66}V_4O_{10}$ is still very high ($E_v^{Na}$= 3.37 eV, Table II) and will be difficult at the normal battery operating conditions.

In order to check the effect of U on the Na vacancy formation energy and Na filling energies were recalculated with $U_{eff}$ = 3.3 eV applied for Vanadium [63], see Figure 5(b). We observed the similar trend on the Na vacancy formation energy and filling energy of the $Na_xV_4O_{10}$ sample with $U_{eff}$ = 3.3 eV for vanadium.

TABLE II. The calculated energy gain due to the addition of Na at the different surfaces.

| NVO ($Na_{0.66}V_4O_{10}$) | Surface energy (eV/Å$^2$) | Na vacancy formation energy (eV) | Na filling energy (eV) |
|---|---|---|---|
| Bulk | — | 3.37 | -3.01 |
| (100) | 0.031 | 2.52 | -2.19 |
| (010) | 0.039 | 3.36 | -2.74 |
| (110) | 0.046 | 3.01 | -2.47 |

The nanostructured morphologies of the cathode materials are in general known to show better electrochemical activities [64–66], as also shown in our experimental results. It is understood that the nano-structuring opens up the surface diffusion channels, where the formation

of the Na vacancy and the diffusion of Na are known to be easy as compared to the bulk. In order to understand the nano-structuring phenomena for our target NVO material $Na_{0.66}V_4O_{10}$, we studied the surface Na vacancy formation energy and Na filling energies at three different surfaces of NVO, (100), (010) and (110) as shown in Figure 6. The DFT calculated energies for Na vacancy formation and Na filling are listed in Table II, for all the three surfaces. The surface energy calculation shows that the (100) surface is more stable than the (010) and (110) surface, and hence will cover most of the exposed surface of the nano scale material. The Na vacancy formation energy is calculated to be lowest for the (100) surface ($E_v^{Na}$ = 2.52 eV); which is lower by 0.85 eV, compared to the bulk ($E_v^{Na}$ = 3.37 eV). The (010) surface has the similar Na vacancy formation energy as the bulk; whereas on (110) surface it is a little lower ($E_v^{Na}$= 3.01 eV). The lower values of Na vacancy formation energy at the (100) surface, will enhance the concentration of Na vacancy at the surface; which will result in a higher diffusion and hence better electrochemical performance.

The Na filling energy also follows the same trends as the Na vacancy formation energy. The energy gain due to the addition of Na at the (100) surface is calculated to be the lowest ($E_f^{Na}$ = - 2.19 eV); whereas the energies are also lower for (010) and (110) surfaces ($E_f^{Na}$ = -2.74 eV and - 2.47 eV, respectively) as compared to the bulk ($E_f^{Na}$ = - 3.01 eV), as given in Table II. Furthermore, for the calculations of Na vacancy formation energy and Na filling energy, we have tested different possible configurations and the most stable configuration was used to calculate the Na vacancy formation energy and Na filling energy. The two Na atoms present at the (100) surface are equivalent, hence unique geometry was obtained for Na vacancy formation energy at the (100) surface. For the Na atom filling energy at the (100) surface, two configurations were studied as shown in the lower panel of Figure 6 (a, b) along with the filling energy. For the (010) surface there is only one Na atom present at the surface, hence the geometry obtained for Na vacancy formation energy at the (010) surface is unique. Similar to the (100) surface, two different configurations were obtained for the Na atom filling energy. On the other hand, the two Na atom present at the (110) surface are equivalent, hence only one unique geometry was obtained for Na vacancy formation energy. Also, two different configurations were obtained for the Na atom filling energy at the (110) surface, see Figure 6 (e, f).

## CONCLUSIONS

In summary, the rod-shaped $Na_{0.66}V_4O_{10}$ (NVO) nanostructured cathode material was prepared using sol-gel method. The XRD patterns confirm the formation of pure monoclinic phase with the C2/m space group and no impurity or secondary phases were observed. The SEM and TEM studies reveal the rod-shaped agglomerated morphology of the prepared cathode material with size in 50–100 nm range and the length a few $\mu$m. The observed selected area electron diffraction pattern is in good agreement with the XRD pattern. The cyclic voltammetry (CV) results showed the two-phase transition reaction between $V^{5+/4+}$, which is in good agreement with the galvanostatic charge/discharge curves. The electrochemical performance indicates that the NVO cathode exhibits specific discharge capacity of 80 ($\pm$2), 64 ($\pm$2), 56 ($\pm$2), 52 ($\pm$2) and 30 ($\pm$2) mAh $g^{-1}$, measured vs. $Na^+$/Na at the current densities of 3, 5, 11, 20 and 50 mA $g^{-1}$, respectively. The electrochemical performance of $Na_{0.66}V_4O_{10}$ electrode with Li anode is found promising, but decays faster as compared to the Na-anode. The energetics of Na vacancy formation and filling on discharging and charging, respectively, which could form a potential descriptor of the electrochemical performance. The available surface area for Na transport is likely to increase for nanostructures, which further supports the experimental results.

## MATERIALS AND METHODS

**Experimental details**. The nanostructured $Na_{0.66}V_4O_{10}$ (NVO) samples were synthesized by sol-gel route using sodium acetate ($C_2H_3NaO_2$) and ammonium metavanadate ($H_4NO_3V$) in 1 to 6.06 molar ratio, from Sigma Aldrich, $\geq$99.0%. The starting precursors were homogeneously mixed in deionized water with continuous stirring. Citric acid (Anhydrous, $C_6H_8O_7$, from Fisher Scientific, 99.5%) was used as a chelating agent in a molar ratio of 1:1 with metal ions. The resulting mixture of citric acid and precursors was heated at 80°C with continuous stirring at 400 rpm until a homogenous gel was obtained. Then, the gel was dried overnight at 120°C to obtain the powder followed by heat treating in muffle furnace in air at 400°C (with 5°C/minute rate) for 5 hrs. The final powder was stored in a vacuum desiccator for physical and electrochemical characterizations. The thermal decomposition of the resulting gel precursors was examined by thermogravimetric analysis (NETZSCH, TG 209 F3 Tarsus) at a ramp rate of 10°C/min in air from RT to 900°C.

The crystallographic structure and phase purity of the prepared samples were determined by room temperature powder x-ray diffraction (from Pananlytical) with CuK$\alpha$ radiation ($\lambda$=1.540 Å) operating at 40 kV voltage and 40 mA current. The XRD pattern was recorded in $2\theta$ angular range between 10 and 70° with a step size of 0.02°. The XRD data have been analyzed by Rietveld refinement using FullProf package and the background was fitted using linear interpolation between the data points. We study the morphology of the prepared cathode us-

ing Zeiss EVO 50 scanning electron microscope (SEM) working at 20 kV in scattered electron mode. A JEOL JEM-1400 Plus microscope coupled with energy dispersive x-ray spectroscopy (EDX) facility has been used for the transmission electron microscopy (TEM) measurements at 120 keV.

For the electrochemical measurements, cathodes were prepared by mixing prepared active material, PVDF (Polyvinylidene difluoride) as a binder and super P (conductive carbon) in a weight ratio of 80:15:5 in NM2P (N-methyl-2-pyrrolidinone) as a solvent. Then we stirred the resulting mixture for 5 hrs to have homogenous mixing and then coated over Al foil (as a current collector) using doctor blade technique. The coated Al sheet was then dried overnight in vacuum oven at 80°C to evaporate the solvent. The sheet was then rolling pressurized and punched into circular electrode with 12 mm diameter. The active mass loading on the electrodes was between 2-3 mg. Before bringing inside the glove box workstation, we dried the electrodes in vacuum at 60°C to remove any moisture. Na chips (16 mm diameter) were extracted from Na cubes (Sigma Aldrich, 99.9%) by cutting, rolling and pressing sodium cubes into thin sheets inside the glove box and used as both counter and reference electrode. For electrochemical measurements, the CR2016 coin half-cells were assembled in an argon-filled glove box (Jacomex, $\leq$0.5 ppm of $O_2$ and $H_2O$ level) with the cathode electrode, Na metal, and a glass fiber (Advantec, GB-100R) as the separator. The electrolyte used was 1 M $NaClO_4$ dissolved in ethylene carbonate (EC)/dimethyl carbonate (DMC) in a volume ratio of 1:1. Electro-chemical Impedance Spectroscopy (EIS) using an AC voltage pulse of 5 mV in the frequency range of 100 kHz – 10 mHz was also performed to measure the electrode resistance. We have also assembled and tested the half-cells vs. $Li^+$/Li as an anode to study the feasibility of the prepared NVO cathode as a material for Li-ion batteries. The assembled cells were tested with both electrolytes like 1 M $NaClO_4$ and $LiPF_6$ (EC:DMC 1:1 V/V) and abbreviated as NVON (for $NaClO_4$) and NVOL (for $LiPF_6$). All the electrochemical measurements were carried out at room temperature using VMP3 (Biologic) instrument. The cyclic voltmmetry (CV) have been performed in the potential window of 1.5 – 4.0 V vs. Na/$Na^+$ at the scan rate of 0.05 $mVs^{-1}$. The charging/discharging characteristics were studied in galvanostatic mode at different current densities.

**Computational Method**. Periodic plane-wave based density functional module implemented in Viena ab-initio Simulation Package (VASP) [67] is applied to calculate the energy of the cathode material with six different $Na_xV_4O_{10}$ compositions, where $x$ = 1.66, 1.33, 1, 0.66, 0.33 and 0. The model structure of $Na_{1.33}V_4O_{10}$ [$Na_2V_6O_{15}$] composition was obtained from The Materials Project database https://materialsproject.org/ (materials ID. mp-778594) [63], having monoclinic cyrstal structure with C2/m space group. The lattice parameter of $a$ = 15.7845 Å, $b$ = 3.6662 Å, $c$ = 10.6182 Å, and $\alpha$ = 90° and $\beta$ = 103.093°, matches well with the experimentally obtained cell parameters. The sodium bulk metal geometry is modelled as cubic *bcc* bulk with lattice parameter, $a$ = 4.2906 Å and $\alpha$ = 90°. Geometry optimization is performed on the structures, with energy and force convergence criteria set to $1\times10^{-4}$ eV and 0.05 eV/Å, respectively. RPBE [68] GGA exchange correlation functional and Ultra-soft pseudopotentials (USSP) [69] are used. Plane wave basis sets are expanded to an energy cut-off value of 396 eV. Monkhorst pack [70] k-points grid of 1×3×1 is utilized for all the calculations on different NVO structures, whereas a 5×5×5 grid is used for Na bulk lattice geometry optimization. The Na vacancy formation energy ($E_v^{Na}$) of NVO structure $Na_xV_4O_{10}$ is calculated using the formula below;

$$E_v^{Na} = E[Na_{(x-1)}V_4O_{10}] + E(Na) - E(Na_xV_4O_{10}), \quad (9)$$

where E($Na_xV_4O_{10}$), E($Na_{(x-1)}V_4O_{10}$) and E(Na) are the energies of NVO structure $Na_xV_4O_{10}$, $Na_{(x-1)}V_4O_{10}$ formed after Na removal and energy of bulk Na, respectively. Similarly the energy gain from Na addition ($E_f^{Na}$) to NVO structure $Na_xV_4O_{10}$ can be calculated as;

$$E_f^{Na} = E(Na_{(x+1)}V_4O_{10}) - E(Na_xV_4O_{10}) - E(Na) \quad (10)$$

In order to study the nano-structuring effect on the Na vacancy formation and Na filling energies; three different low-index surfaces, (100), (010) and (110) are created for the NVO structure studied experimentally, $Na_{0.66}V_4O_{10}$. A 2×1 super cell was used for (100) surface, whereas for both the (010) and (110) surfaces, a 1×1 unit cell is used for the surface calculation. For all the three surfaces, a four layer surface slab is used; where two bottom layers are kept fixed to their bulk positions and the top two layers are allowed to relax. A vacuum of 20 Å is used to model the surface slabs. Calculations of Na vacancy formation energy and Na filling energy are performed by removing and adding Na at the corresponding surface, respectively. For the addition of Na to the surface two distinct adsorption sites were studied and the geometry and energy of the most stable adsorption site was reported. Equations 9 and 10 are used to calculate the surface Na vacancy formation energy and Na filling energy, respectively.

## ACKNOWLEDGMENTS

We acknowledge the financial support from IIT Delhi through the FIRP project (IRD no. MI01418). RS thanks IRD, IIT Delhi for postdoctoral fellowship through FIRP project. MC, and RS thank SERB-DST (NPDF, no PDF/2016/003565), and MHRD, respectively, for the fellowship. The authors thank the

IIT Delhi for providing central research facilities: XRD, EDX, SEM, and TEM. We also thank the physics department, IIT Delhi for support. RSD acknowledges the financial support from SERB-DST through Early Career Research (ECR) Award (project reference no. ECR/2015/000159) and BRNS through DAE Young Scientist Research Award (project sanction no. 34/20/12/2015/BRNS). The authors thank the HPC facility of IIT Delhi for computational resources.

## REFERENCES

* rsdhaka@physics.iitd.ac.in

[1] J. B. Goodenough, and K. S. Park, The Li-ion rechargeable battery: a perspective, J. Am. Chem. Soc. **2013**, 135, 1167 – 1176.

[2] M. V. Reddy, G. V. Subba Rao, and B. V. R. Chowdari, Metal oxides and oxysalts as anode materials for Li ion batteries, Chem. Rev. **2013**, 113, 5364 – 5457.

[3] D. W. Han, S. J. Lim, Y. I. Kim, S. H. Kang, Y. C. Lee, and Y. M. Kang, Facile lithium ion transport through superionic pathways formed on the surface of $Li_3V_2(PO_4)_3$/C for high power Li ion battery, Chem. Mater. **2014**, 26, 3644 – 3650.

[4] R. Marom, S. F. Amalraj, N. Leifer, D. Jacob, and D. Aurbach, A review of advanced and practical lithium battery materials, J. Mater. Chem. **2011**, 21, 9938 – 9954.

[5] J. B. Goodenough, and Y. Kim, Challenges for rechargeable Li batteries, Chem. Mater. **2010**, 22, 587 – 603.

[6] S. W. Kim, D. H. Seo, X. H. Ma, G. Ceder, and K. Kang, Electrode materials for rechargeable sodium-ion batteries, Adv. Energy Mater. **2012**, 2, 710 – 721.

[7] H. Pan, Y.-S. Hu, and L. Chen, Room-temperature stationary sodium-ion batteries for large-scale electric energy storage, Energy Environ. Sci. **2013**, 6, 2338 – 2360.

[8] M. D. Slater, D. Kim, E. Lee, and C. S. Johnson, Sodium-ion batteries, Adv. Funct. Mater. **2013**, 23, 947 – 958.

[9] K. Kim, H. Kim, Z. Ding, M. H. Lee, K. Lim, G. Yoon, and K. Kang, Recent progress in electrode materials for sodium-ion batteries, Adv. Energy Mater. **2016**, 6, 1600943.

[10] D. Larcher, and J. M. Tarascon, Towards greener and more sustainable batteries for electrical energy storage, Nat. Chem. **2015**, 7, 19 – 29.

[11] C. Vaalma, D. Buchholz, M. Weil, and S. Passerini, A cost and resource analysis of sodium-ion batteries, Nature Reviews Materials **2018**, 3, 18013.

[12] J. -Y. Hwang, S. -T. Myung, and Y. -K. Sun, Sodium-ion batteries: present and future, Chem. Soc. Rev. **2017**, 46, 3529 – 3614.

[13] M. Sawicki, and L. L. Shaw, Advances and challenges of sodium ion batteries as post lithium ion batteries, RSC Adv. **2015**, 5, 53129 – 53154.

[14] P. K. Nayak, L. Yang, W. Brehm, and P. Adelhelm, From lithium-ion to sodium-ion batteries: advantages, challenges, and surprises, Angew. Chem. Int. Ed. **2018**, 57, 102 – 120.

[15] D. Kundu, E. Talaie, V. Duffort, and L. F. Nazar, The emerging chemistry of sodium ion batteries for electrochemical energy storage, Angew. Chem. Int. Ed. **2015**, 54, 3431 – 3448.

[16] L. Chen, M. Fiore, J. E. Wang, R. Ruffo, D. -K. Kim, and G. Longoni, Readiness level of sodium-ion battery technology: a materials review, Adv. Sustainable Syst. **2018**, 2, 1700153 (1 – 37).

[17] A. J. Toumar, S. P. Ong, W. D. Richards, S. Dacek, and G. Ceder, Vacancy ordering in O3-type layered metal oxide sodium-ion battery cathodes, Phys. Rev. Applied **2015**, 4, 064002 (1 – 9).

[18] C.-J. Yu, S.-H. Choe, G.-C. Ri, S.-C. Kim, H.-S. Ryo, and Y.-J. Kim, Ionic diffusion and electronic transport in eldfellite $Na_xFe(SO_4)_2$, Phys. Rev. Applied **2017**, 8, 024029 (1 – 9).

[19] C. Zheng, B. Radhakrishnan, Iek-H. Chu, Z. Wang, and S. P. Ong, Effects of transition-metal mixing on Na ordering and kinetics in layered P2 oxides, Phys. Rev. Applied **2017**, 7, 064003 (1 – 12).

[20] M. Debbichi, and S. Lebégue, Crystal and electronic structures of nitridophosphate compounds as cathode materials for Na-ion batteries, Phys. Rev. B **2015**, 92, 085127 (1 – 6).

[21] L. E. Rush, Jr. Zachary D. Hood, and N. A. W. Holzwarth, Unraveling the electrolyte properties of $Na_3SbS_4$ through computation and experiment, Phys. Rev. Materials **2017**, 1, 075405.

[22] N. Yabuuchi, M. Kajiyama, J. Iwatate, H. Nishikawa, S. Hitomi, R. Okuyama, R. Usui, Y. Yamada, and S. Komaba, P2-type $Na_x[Fe_{1/2}Mn_{1/2}]O_2$ made from earth-abundant elements for rechargeable Na natteries, Nat. Mater. **2012**, 11, 512 – 517.

[23] X. Xiang, K. Zhang, and J. Chen, Recent advances and prospects of cathode materials for sodium-ion batteries, Adv. Mater. **2015**, 27, 5343 – 5364.

[24] Q. Ni, Y. Bai, F. Wu, and C. Wu, Polyanion-type electrode materials for sodium-ion batteries, Adv. Sci. **2017**, 4, 1600275 (1 – 24).

[25] P. Gao, C. Wall, L. Zhang, M. A. Reddy, and M. Fichtner, Vanadium Oxychloride as electrode material for sodium ion batteries, Electrochem. Commun. **2015**, 60, 180 – 184.

[26] Y. Cao, L. Xiao, W. Wang, D. Choi, Z. Nie, J. Yu, L. V. Saraf, Z. Yang, and J. Liu, Reversible sodium ion insertion in single crystalline manganese oxide nanowires with long cycle life, Adv. Mater. **2011**, 23, 3155 – 3160.

[27] M. Chandra, R. Shukla, R. Saroha, A. K. Panwar, A. Gupta, S. Basu, and R. S. Dhaka, Physical properties and electrochemical performance of Zn-substituted $Na_{0.44}Mn_{1-x}Zn_xO_2$ nanostructures as cathode in Na-ion batteries, Ceramics International **2018**, 44, 21127 – 21131.

[28] M. Sathiya, K. Hemalatha, K. Ramesha, J. M. Tarascon, and A. S. Prakash, Synthesis, structure, and electrochemical properties of the layered sodium insertion cathode material: $NaNi_{1/3}Mn_{1/3}Co_{1/3}O_2$, Chem. Mater. **2012**, 24, 1846 – 1853.

[29] S. Lulu, Y. Zhengqiu, H. Lei, Z. Yongchun, and Q. Yitai, Uniform and continuous carbon coated sodium vanadium phosphate cathode materials for sodium-ion battery, J. of Power Sources **2014**, 272, 880 – 885.

[30] S. M. Oh, S. T. Myung, J. Hassoun, B. Scrosat, and Y. K. Sun, Reversible $NaFePO_4$ electrode for sodium secondary

batteries, Electrochem. Commun. **2012**, 22, 149 – 152.
[31] P. Barpanda, J. N. Chotard, N. Recham, C. Delacourt, M. Ati, L. Dupont, M. Armand, and J. M. Tarascon, Structural, transport, and electrochemical investigation of novel $AMSO_4F$ (A = Na, Li; M = Fe, Co, Ni, Mn) metal fluorosulphates prepared using low temperature synthesis routes, Inorg. Chem. **2010**, 49, 7401 – 7413.
[32] V. Palomares, M. C. Cabanas, E. C. Martnez, M. C. Han, and T. Rojo, Update on Na-based battery materials. a growing research path, Energy Environ. Sci. **2013**, 6, 2312 – 2337.
[33] Y. Zhu, Y. Xu, Y. Liu, C. Luo, and C. Wang, Comparison of electrochemical performances of olivine $NaFePO_4$ in sodium-ion batteries and olivine $LiFePO_4$ in lithium-ion batteries, Nanoscale **2013**, 5, 780 – 787.
[34] M. Chandra, R. Shukla, M. Rashid, A. Gupta, S. Basu, and R. S. Dhaka, Synthesis and physical properties of $Na_xTO_2$ (T=Mn, Co) nanostructures for cathode materials in Na-ion batteries, Mater. Res. Bull. **2018**, 105, 178 – 183.
[35] S. Ni, X. Lv, J. Ma, X. Yang, and L. Zhang, Electrochemical characteristics of lithium vanadate, $Li_3VO_4$ as a new sort of anode material for Li-ion batteries, J. of Power Sources **2014**, 248, 122 – 129.
[36] S. Ni, J. Zhang, J. Ma, X. Yang, and L. Zhang, Superior electrochemical performance of $Li_3VO_4$/ N-doped C as an anode for Li-ion batteries, J. of Materials Chemistry A **2015**, 3, 17951 – 17955.
[37] S. Ni, J. Zhang, J. Ma, X. Yang, L. Zhang, X. Li, and H. Zeng, Approaching the theoretical capacity of $Li_3VO_4$ via electrochemical reconstruction, Adv. Mater. Interfaces **2016**, 3, 1500340 (1 – 10).
[38] S. Yuan, Y. Zhao, and Q. Wang, Layered $Na_2V_6O_{16}$ nanobelts as promising cathode and symmetric electrode for Na-ion batteries with high capacity, J. Alloys Compd. **2016**, 688, 55 – 60.
[39] H. Liu, H. Zhou, L. Chen, Z. Tang, and W. Yang, Electrochemical insertion/deinsertion of sodium on $NaV_6O_{15}$ nanorods as cathode material of rechargeable sodium-based batteries, J. Power Sources **2011**, 196, 814 – 819.
[40] Y. Lee, S. M. Oh, B. Park, B. U. Ye, N.-S. Lee, J. M. Baik, S.-J. Hwang, and M. H. Kim, Unidirectional growth of single crystalline $\beta$-$Na_{0.33}V_2O_5$ and $\alpha$-$V_2O_5$ nanowires driven by controlling the pH of aqueous solution and their electrochemical performances for Na-ion batteries, CrystEngComm **2017**, 19, 5028 – 5037.
[41] X. Wang, Q. Liu, H. Wang, D. Jiang, Y. Chang, T. Zhang, B. Zhang, H. Mou, and Y. Jiang, PVP-modulated synthesis of $NaV_6O_{15}$ nanorods as cathode materials for high-capacity sodium-ion batteries, J. Mater. Sci. **2016**, 51, 8986 – 8994.
[42] V. Palomares, P. Serras, H. E. A. Brand, T. Rojo, and N. Sharma, Structural evolution of mixed valent ($V^{3+}/V^{4+}$) and $V^{4+}$ sodium vanadium fluorophosphates as cathodes in sodium-ion batteries: comparisons, overcharging and mid-term cycling, J. Mater. Chem. A **2015**, 3, 23017 – 23027.
[43] Y. Lu, J. Wu, J. Liu, M. Lei, S. Tang, P. Lu, L. Yang, H. Yang, and Q. Yang, Facile synthesis of $Na_{0.33}V_2O_5$ nanosheet-graphene hybrids as ultrahigh performance cathode materials for lithium ion batteries, ACS Appl. Mater. Interfaces **2015**, 7, 17433 – 17440.
[44] F. Hu, W. Jiang, Y. Dong, X. Lai, L. Xiao, and X. Wu, Synthesis and electrochemical performance of $NaV_6O_{15}$ microflowers for lithium and sodium ion batteries, RSC Adv. **2017**, 7, 29481 – 29488.
[45] A. Q. Pan, J. G. Zhang, Z. M. Nie, G. Z. Cao, B. W. Arey, G. S. Li, S. Q. Liang, and J. Liu, Facile synthesized nanorod structured vanadium pentoxide for high-rate lithium batteries, J. Mater. Chem. **2010**, 20, 9193 – 9199.
[46] J. Jiang, Z. Wang, and L. Chen, Structural and electrochemical studies on $\beta$-$Li_xV_2O_5$ as cathode material for rechargeable lithium batteries, J. Phys. Chem. C **2007**, 111, 10707 – 10711.
[47] P.-P. Wang, C.-Y. Xu, W.-D. Li, L. Wang, and L. Zhen, Low temperature electrochemical performance of $\beta$-$Li_xV_2O_5$ cathode for lithium-ion batteries, Electrochim. Acta **2015**, 169, 440 – 446.
[48] N. M. Asl, J.-H. Kim, W. C. Lee, Z. Liu, P. Lu, and Y. Kim, A new chemical route for the synthesis of $\beta$-$Li_xV_2O_5$ for use as a high performance cathode, Electrochim. Acta **2013**, 105, 403 – 411.
[49] D. A. Semenenko, D. M. Itkis, E. A. Pomerantseva, E. A. Goodilin, T. L. Kulova, A. M. Skundin, and Y. D. Tretyakov, $Li_xV_2O_5$ nanobelts for high capacity lithium-ion battery cathodes, Electrochemistry Commun. **2010**, 12, 1154 – 1157.
[50] S. Bach, N. Baffier, J. P. Pereira-Ramos, and R. Messina, Electrochemical sodium intercalation in $Na_{0.33}V_2O_5$ bronze synthesized by a sol-gel process, Solid State Ionics **1989**, 37, 41–49.
[51] J. K. Kim, B. Senthilkumar, S. H. Sahgong, J.-H. Kim, M. Chi, and Y. Kim, New chemical route for the synthesis of $\beta$-$Na_{0.33}V_2O_5$ and its fully reversible Li intercalation, ACS Appl. Mater. Interfaces **2015**, 7, 7025 – 7032.
[52] S. Liang, J. Zhou, G. Fang, C. Zhang, J. Wu, Y. Tang, and A. Pan, Synthesis of mesoporous $\beta$-$Na_{0.33}V_2O_5$ with enhanced electrochemical performance for lithium ion batteries, Electrochim. Acta **2014**, 130, 119 – 126.
[53] D. L. Jiang, H. Wang, G. P. Li, G. Q. Li, X. Z. Lan, M. H. Abib, Z. P. Zhang, and J. G. Yang, Self-combustion synthesis and ion diffusion performance of $NaV_6O_{15}$ nanoplates as cathode materials for sodium-ion batteries, J. Electrochem. Soc. **2015**, 162, A697 – A703.
[54] P. P. Wang, C. Y. Xu, F. X. Ma, L. Yang, and L. Zhen, In situ soft-chemistry synthesis of $\beta$-$Na_{0.33}V_2O_5$ nanorods as high-performance cathode for lithium-ion batteries, RSC Adv. **2016**, 6, 105833 – 105839.
[55] X. Dong, Y. Xu, L. Xiong, X. Sun, and Z. Zhang, Sodium substitution for partial lithium to significantly enhance the cycling stability of $Li_2MnO_3$ cathode material. J. of Power Sources, **2013**, 243, 78 – 87.
[56] P. Manikandan, D. Ramasubramonian, and M. Shaijumon, Layered P2-type $Na_{0.5}Ni_{0.25}Mn_{0.75}O_2$ as a high performance cathode material for sodium-ion batteries, Electrochim. Acta **2016**, 206, 199 – 206.
[57] A. Grzechnik, Y. Ueda, T. Yamauchi, M. Hanfland, P. Hering, V. Potapkin, and K Friese, Pressure-induced non-superconducting phase of $\beta-Na_{0.33}V_2O_5$ and the mechanism of high-pressure phase transitions in $\beta-Na_{0.33}V_2O_5$ and $\beta-Li_{0.33}V_2O_5$ at room temperature, J. Phys.: Condens. Matter **2016**, 28, 035401.
[58] R. Baddour-Hadjean, S. Bach, N. Emery, and J. Pereira-Ramos, The peculiar structural behaviour of $\beta$-$Na_{0.33}V_2O_5$ upon electrochemical lithium insertion, J. Mater. Chem. **2011**, 21, 11296 – 11305.

[59] Y. Hu, L. Wu, G. Liao, Y. Yang, F. Ye, J. Chen, X. Zhu, and S. Zhong, Electrospinning synthesis of $Na_2MnPO_4F/C$ nanofibers as a high voltage cathode materials for Na-ion batteries, Ceramic Internationals, **2018**, 44, 17577 – 17584.

[60] J. Liu, K. Tang, K. Song, P. A. van Aken, Y. Yu and J. Maier, Electrospun $Na_3V_2(PO_4)_3/C$ nanofibers as stable cathode materials for sodium-ion batteries, Nanoscale, **2014**, 6, 5081– 5086.

[61] J. Zhang, B. Su, A. Kitajou, M. Fujita, Y. Cui, M. Oda, W. Zhou, P. H. -L. Sit, and D. Y. W. Yu, Activating abnormal capacity in stoichiometric NaVO3 as cathode material for sodium-ion battery, J. of Power Sources, **2018**, 400, 377 – 382.

[62] G. S. Shinde, P. D. Nayak, S. P. Vanam, S. K. Jain, A. D. Pathak, S. Sanyal, J. Balachandran, and P. Barpanda, Ultrasonic sonochemical synthesis of $Na_{0.44}MnO_2$ insertion material for sodium ion battery, J. of Power Sources, **2019**, 416, 50 – 55.

[63] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, The materials project: a materials genome approach to accelerating materials innovation, APL Materials **2013**, 1, 011002 (1 – 11).

[64] J. Kang, S. Baek, V. Mathew, J. Gim, J. Song, H. Park, E. Chae, A. K. Rai, and J. Kim, High rate performance of a $Na_3V_2(PO_4)_3/C$ cathode prepared by pyro-synthesis for sodium-ion batteries, J. Mater. Chem. **2012**, 22, 20857 – 20860.

[65] Y. Qi, L. Mu, J. Zhao, Y.-S. Hu, H. Liu, and S. Dai, Superior Na-storage performance of low-temperature-synthesized $Na_3(VO_{1-x}PO_4)_2F_{1+2x}$ $(0 \leq x \leq 1)$ nanoparticles for Na-ion batteries, Angew. Chem. Int. Ed. **2015**, 54, 9911 – 9916.

[66] M. Hayashi, H. Uemura, K. Shimanoe, N. Miura, and N. Yamazoe, Enhanced electrocatalytic activity for oxygen reduction over carbon-supported $LaMnO_3$ prepared by reverse micelle method, Electrochem. Solid-State Lett. **1998**, 1, 268 – 270.

[67] G. Kresse, and J. Furthmüller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, Computational Materials Science **1996**, 6, 15 – 50.

[68] B. Hammer, L. B. Hansen, and J. K. Norskov, Improved adsorption energetics within density-functional theory using revised perdew-burke-ernzerhof functionals, Phys. Rev. B **1999**, 59, 7413 – 7421.

[69] D. Vanderbilt, Soft self-consistent pseudopotentials in a generalized eigenvalue formalism, Phys. Rev. B **1990**, 41, 7892 – 7895.

[70] H. J. Monkhorst, and J. D. Pack, Special points for brillouin-zone integrations, Phys. Rev. B **1976**, 13, 5188 – 5192.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/333208713

# Experimental Investigation on the Thermal Behavior of Untreated and Alkali-Treated Pineapple Leaf and Coconut Husk Fibers

**Article** · June 2019

CITATIONS
0

READS
18

**2 authors:**



Mohit Mittal
Delhi Technological University
**14** PUBLICATIONS **8** CITATIONS
SEE PROFILE



Rajiv Chaudhary
Delhi Technological University
**64** PUBLICATIONS **134** CITATIONS
SEE PROFILE

**Some of the authors of this publication are also working on these related projects:**



Thermogravimetric analysis of natural fiber View project



Working in the area of Tribology (Engine tribology, bearing design, tribology by nano particles and tubes. View project

All content following this page was uploaded by Mohit Mittal on 20 May 2019.

The user has requested enhancement of the downloaded file.

*International Journal of Applied Science and Engineering,* 76(1): 01-16, June 2019

DOI:

**©2019 New Delhi Publishers. All rights reserved**

# Experimental Investigation on the Thermal Behavior of Untreated and Alkali-Treated Pineapple Leaf and Coconut Husk Fibers

**Mohit Mittal* and Rajiv Chaudhary[2]**

*Department of Mechanical, Production and Industrial Engineering, Delhi Technological University, Delhi-110042, India*

*Corresponding author: mohit.30mittal@gmail.com

**ABSTRACT**

The primary objective of this research work is to study the thermal behavior of untreated and alkali treated pineapple leaf (PALF) and coconut husk fibers (COIR). In this context, firstly the cellulosic fibers were treated with an alkali solution of various concentration (2%, 4%, 6%, 8%, and 10 wt%) then secondly the change in their surface chemistry and thermal stability were investigated by means of a Fourier transform infrared spectroscopy (FTIR) and Thermogravimetric analysis technique (TGA) respectively. A significant variation in the chemical composition of cellulosic fibers was observed after an alkaline treatment. Changes in the peak at 1745, 1525, and 1250 $cm^{-1}$ in FTIR spectra corresponds to the partial removal of hemicellulose and lignin components. The thermal degradation of lignocellulosic fibers consists of two major steps i.e. moisture loss (below 150˚C) and the decomposition of hemicellulose, cellulose, and lignin components (between 150˚C to 400˚C). An alkali treatment increases the thermal stability of pineapple leaf and coir fibers through physical and chemical changes. It was observed that the 4 wt% NaOH treated PALF and COIR fibers possess higher thermal stability as compared to other treated and untreated fibers. PALF exhibits a higher rate of decomposition than the COIR fiber but its main decomposition peak lies at a higher temperature of about $40^0$C.

**Keywords:** Pineapple leaf fiber, Coir fiber, Alkali treatment (NaOH), Thermal stability, Thermogravimetric analysis (TGA), Natural fiber, Infrared spectroscopy

Recently, most of the industries are looking for those materials which have environment-friendly characteristics, contribute to the lower emission of pollutants, good physical properties, and economical in nature[1-3]. To fulfill the market demands, researchers and technologists are putting their attention towards biocomposite materials. These materials generally made up of lignocellulosic biomass and polymeric resin.

Amongst all the different types of lignocellulosic materials, the coconut husk (COIR) and pineapple leaf fibers (PALF) have high potential to develop high strength biocomposite materials since they possess a large number of advantageous properties such as low cost, lightweight, easy processing, easy availability, high specific strength and stiffness, and non-abrasive nature[4-5]. The COIR and PALF can obtain from the husk of coconut fruit (*Cocos nucifera*) and leaves of pineapple fruit (*Ananas-comosus*) respectively. The inherent properties of COIR and PALF are mentioned in Table 1[6]. COIR fiber contains more amount of lignin as compared to other typically used natural fibers which result it has good resistance to salt water, microbial, and fungus attack. In comparison to COIR, the PALF is a smooth, flexible, thin, and long fiber. PALF can be employed as an electrical, thermal, and acoustic insulator; for feedstock, and biofuel production[7-10]. According to the statistical database (2016) of "Food and Agriculture Organization," the total production of coconut and pineapple fruit in the whole world is 59 and 25.8 millions of tonnes respectively (Table 2)[11]. Inspite of the above mentioned profitable properties, they exhibit some limitations like hydrophilicity, poor wettability, low thermal stability, and poor interfacial adhesion with synthetic polymers which results that all the capabilities of developed biomaterials cannot be exploited to full extent. In this regard, researchers have done a lot of work to impart hydrophobicity in natural fibers by chemical treatment process such as maleic anhydride, peroxide, permanganate, sodium hydroxide, and organosilanes etc. But very little work has been done to improve the thermal stability of lignocellulosic fibers[12-15]. Previous work[16-18] has reported that the natural fibers can be subjected to thermal degradation during the biocomposite processing. So, it is practically significant to understand and overcome the thermal decomposition of natural polymeric materials. Their thermal behavior mainly depends on the chemical composition and internal structure. The constituents of natural fibers such as hemicellulose, cellulose, lignin, pectin, and waxes are not thermally stable even at a relatively low temperature (below 400˚C). The initial decomposition of natural fibers starts at around 200˚C and it was due to the cellular breakdown of hemicellulose. The previous research work[19-21] concluded that the major decomposition of rice husk, natural hemp, and bagasse fibers occur at 351.5˚C, 312˚C, and 327˚C respectively due to the decomposition of α-cellulose.

Thermogravimetric analysis (TGA) is one of the most popular and widely used techniques to analyze the decomposition process of solid material, kinetic analysis of de-volatilization, and to study the effects of heating rate, temperature, pressure, atmosphere gas, gas flow rate, biomass composition, and particles size on mass loss of a sample[22]. Yao Fei *et al.*[23] have studied the decomposition kinetics of 10 types of fibers and found that the activation energy of Bagasse, Bamboo, Cotton Silk, Hemp, Jute, Kenaf, Rice husk, Rice straw, Wood maple, and Wood pine were 161.1, 161,6, 146, 171.1, 165.6, 157.7, 167.4, 176.2, 153.7, 159.3 respectively. The thermal decomposition of natural fibers consists of two to three steps based on their type and chemical composition. Sergio *et al.*[24] stated that the thermal decomposition of NFRC consists of moisture evaporation, hemicellulose degradation, and decomposition of α-cellulose, and lignin. A. Rachini *et al.*[25] reported the order of thermal stability of hemp fiber: NaOH treated fibers> silane-treated fibers> solvent extracted fibers (water/ethanol mixture, 20/80 v/v)> untreated hemp fibers. M. Tajvidi *et al.*[26] studied the thermal degradation of wood flour, kenaf fiber, newsprint, and rice hulls - polypropylene composites. Na Lu *et al.*[27] reported that the hemp fiber was decomposed in two steps, starting with the loss of moisture around 100˚C and followed by successive decomposition of hemicellulose, cellulose, and lignin in the range of 150-400˚C. Jayamani Elammaran *et al.*[28] reported that the thermal decomposition of Betel nut-Polyester composite was shifted to a higher temperature after an alkaline treatment of betel fibers. Alkaline treatment is one of the commonly used cost-effective

techniques to develop high-performance biocomposite material. The previous research work reported in Table 3 also assured the researchers about the positive attributes of an alkaline treatment. However, very little work has been done towards the effect of alkaline treatment on thermal stability of natural fibers. Therefore, in the present communication, an experimental investigation has been carried out to study the effect of alkaline treatment of various concentrations (2%, 4%, 6%, 8% and 10%) on the thermal behavior of pineapple leaf and coir fibers. FTIR spectroscopy was also used to characterize the surface chemistry of untreated and alkali treated (NaOH) fibers. The thermo-chemical decomposition reaction of untreated and NaOH treated cellulosic fibers were studied by employing a TGA instrument.

**Table 1:** Physical properties of pineapple leaf and coir fiber[6]

| Physical properties | Pineapple leaf fiber | Coir fiber |
|---|---|---|
| Density (g/cm$^3$) | 0.98 | 1.2 |
| Cellulose (%) | 70-82 | 32-43 |
| Hemicellulose (%) | 18.8 | 0.15-0.25 |
| Lignin (%) | 5-12.7 | 40-45 |
| Pectin (%) | 1.1 | 3-4 |
| Moisture content (%) | 11.8 | 8 |
| Microfibrillar angle (deg) | 14 | 30-49 |
| Diameter (μm) | 20-80 | 100-460 |
| Tensile strength (MPa) | 413-1627 | 131-220 |
| Young's modulus (GPa) | 34.5-82.5 | 4-6 |
| Elongation at break (%) | 1.6 | 15-40 |

**Table 2:** Total annual production of pineapple leaf and coir fiber

| Pineapple Production-2016 | | Coconut Production-2016 | |
|---|---|---|---|
| **Country** | **Millions of tonnes** | **Country** | **Millions of tonnes** |
| Costa Rica | 2.9 | Indonesia | 17.7 |
| Brazil | 2.7 | Philippines | 13.8 |
| Philippines | 2.6 | India | 11.1 |
| Thailand | 1.9 | Brazil | 2.6 |
| India | 1.9 | Sri Lanka | 2.5 |
| Indonesia | 1.3 | | |
| World | 25.8 | World | 59.0 |

**Table 3:** Previous research work reported on the alkaline treatment of cellulosic fibers

| Authors | Factors explored by researchers | Results | Reference |
|---|---|---|---|
| M. Mittal *et al.* (2018) | Effect of alkaline treatment on the mechanical properties of PALF/Glass hybrid composite. | The 15/15 (v/v) alkaline treated PALF/Glass composite exhibits 35% higher flexural strength than that of the pure Glass-Epoxy composite. | [29] |
| R. Ranjan *et al.* (2013) | Effect of alkali treatment on the mechanical properties of Banana/Sisal fiber reinforced PLA composites. | The 2 wt% NaOH treated composite has better mechanical properties (tensile, flexural, and impact strength) than the untreated composites. | [30] |

| | | | |
|---|---|---|---|
| L. Uma Devi *et al.* (2011) | Effect of chemical treatment on the water absorption behavior of PALF/GF fiber reinforced polyester composites. | The alkali treated composites possess higher resistance to water sorption than the untreated composites. | [31] |
| G. Goud *et al.* (2010) | Effect of alkali treatment on the mechanical properties of *Roystonea regia* fiber reinforced epoxy composites | Alkali-treated fiber reinforced composite showed better tensile and flexural properties than that of the untreated composite. | [32] |
| P. Anand *et al.* (2014) | Effect of alkali treatment on the mechanical properties of hemp fiber reinforced epoxy composite | Alkali-treated hemp fiber composite yields better mechanical properties than that of untreated composites. | [33] |
| S.C Venkateshappa *et al.* (2011) | Effect of alkali treatment on the mechanical properties of areca fiber reinforced epoxy composites | The tensile, flexural, and compressive strength of composites were increased after the alkali treatment of areca fiber. | [34] |

## EXPERIMENTAL DETAILS

### Materials

Two different types of lignocellulosic fiber were used in this study, *viz.,* coconut husk (*Coco Nucifera)* and pineapple leaf fiber (*Ananas comosus*). These cellulosic fibers were obtained from M/s Go Green Products, Chennai (India). Table 1 shows the chemical composition of PALF and COIR fibers[6]. Sodium hydroxide (NaOH) used for an alkaline treatment was of laboratory reagent (LR) grade and obtained from a local supplier.

### Fiber treatment

PALF and COIR fibers were soaked in various concentrations (2%, 4%, 6%, 8% and 10 wt%) of alkaline solution for 24 hr at room temperature, followed by washing with deionized water and drying in an oven at 60°C for 24 hr. The untreated cellulosic fibers were also washed and dried to remove the surface impurities and absorbed water molecules.

### Fourier-Transform Infrared Spectroscopy (FTIR)

Fourier-transform infrared spectroscopy (Perkin Elmer 2000) was used to analyze the surface chemistry of untreated and alkali treated PALF and COIR fibers. FTIR spectroscopy of the cellulosic fibers had done in Analytical Instrumentation Laboratory, CSIR-CSIO, Chandigarh (India).

### Thermogravimetric Analysis (TGA)

Thermal analysis of COIR and PALF (untreated and alkali treated) was carried out by using a thermogravimetric analyzer (Perkin Elmer, TGA 4000). A sample of mass 5 mg was evenly and uniformly

distributed in the alumina crucible which was supported by a precision balance. This small amount of fiber was taken for uniformity of temperature throughout the sample. The variation in mass of a sample with respect to temperature and time was monitored and recorded by Perkin Elmer thermal software (Pyris). The temperature was ranged from 25˚C to 700˚C with a heating rate of 10˚C/min under the high purity nitrogen (inert) atmosphere at a flow rate of 20 ml/min. The heating rate of 10˚C/min was selected for better resolution of transition.

## RESULTS AND DISCUSSION

### Fourier-Transform Infrared Spectroscopy (FTIR)

Fig. 1 and 2 illustrates the FTIR spectra of COIR and PALF respectively. It was observed that the surface chemistry of cellulosic fibers was changed after an alkaline treatment. An increase in intensity around 3300 $cm^{-1}$ and 1000 $cm^{-1}$ after an alkaline treatment corresponds to the higher accessibility of -OH functional group. This was due to the removal of waxy and gummy substances. The increase in peak intensity at around 1600-1650 $cm^{-1}$ and 1250 $cm^{-1}$ after an alkaline treatment also confirmed the removal of wax, adhesives, pectin, and gummy substances from the fiber surface. The absorption peak at 1745 $cm^{-1}$ was observed in raw fibers (PALF and COIR) but disappears after NaOH treatment. It is related with C=O stretch and confirmed the removal of hemicellulose.

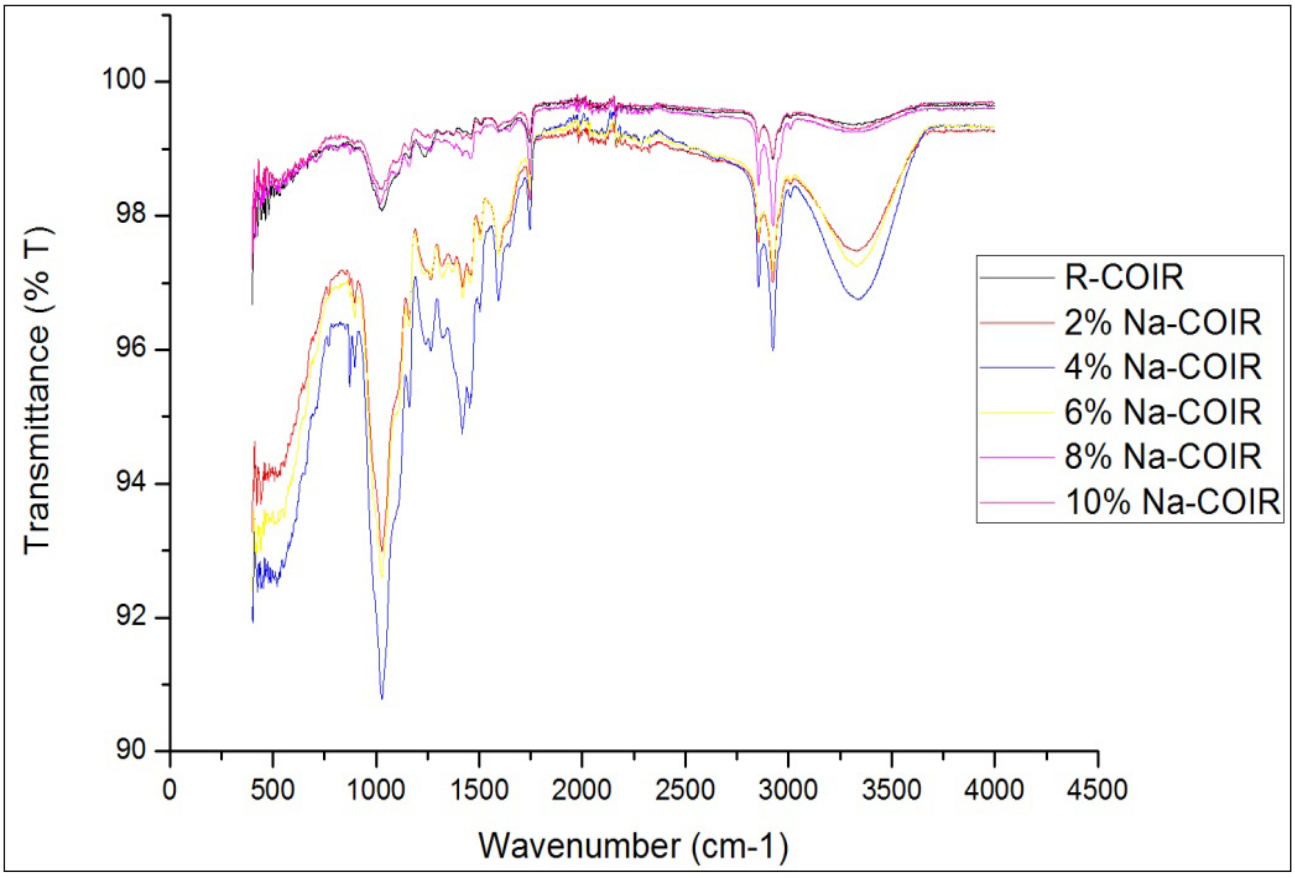



**Fig. 1:** IR spectra of untreated and alkali treated coir fiber

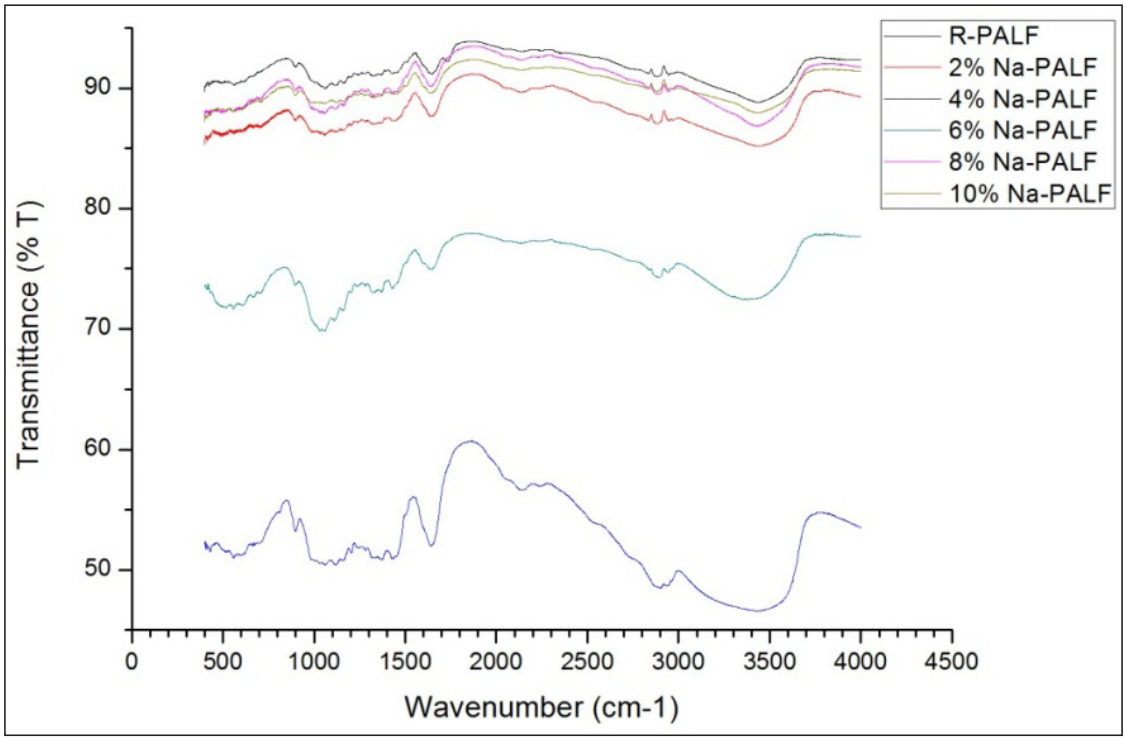



**Fig. 2:** IR spectra of untreated and alkali treated pineapple leaf fiber

The IR spectra of raw PALF and COIR fibers are shown in Fig. 3. It was observed that the PALF consists a large amount of hydroxyl (-OH) group than that of COIR fiber. The broad peak in the region 3200-3500 $cm^{-1}$ is attributed to the vibration of –OH group. The peak at 1244 $cm^{-1}$ is much smaller in PALF than COIR. This peak corresponds to the C=O stretch of an acetyl group of lignin.

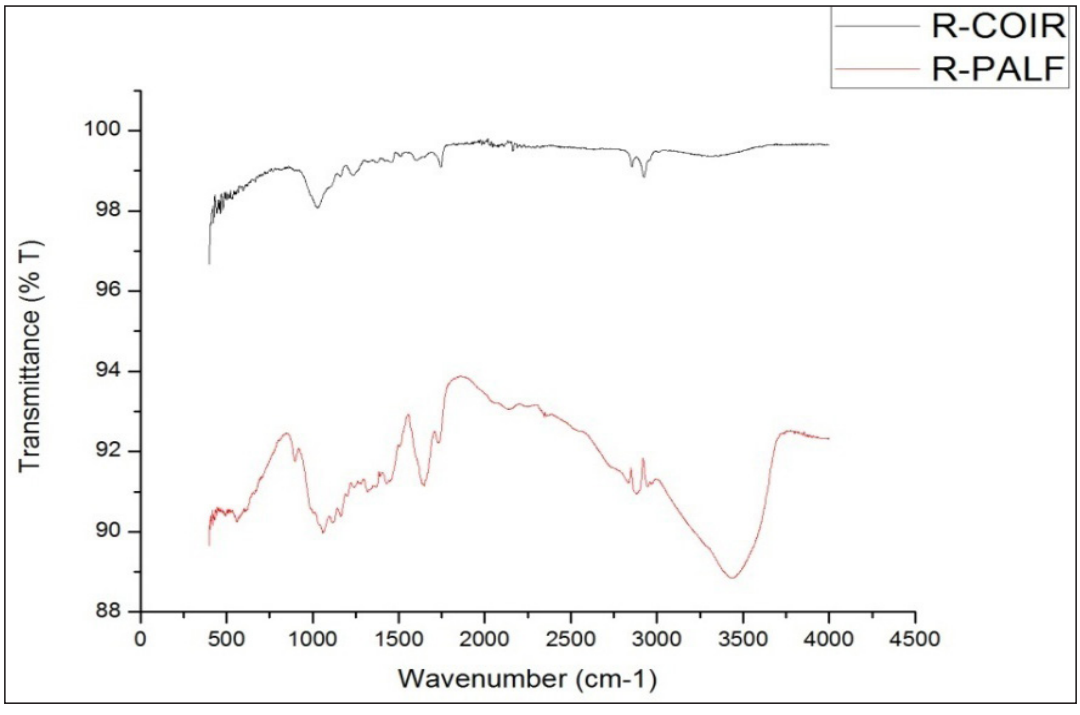



**Fig. 3:** IR spectra of coir and pineapple leaf fiber

## Thermal Decomposition of Untreated and Alkali-Treated Pineapple Leaf and Coir Fibers

TGA and DTG curves of untreated and alkali treated (2%, 4%, 6%, 8%, and 10%) COIR and PALF fibers are shown in Figs 4-9 and the corresponding weight loss as a function of temperature is mentioned in Table 4 and 5. In all cases, the initial stage of decomposition occurs between 30˚C to 185˚C which indicate the release of bound water and volatile extractives. The previous work (Slopiecka *et al.*[16], Ndazi Bwire *et al.*[35]) are also reported that all the types of natural fiber exhibit endothermic decomposition (within 5-10% range) due to an evaporation of moisture, but the rate of evaporation depends on the hydrophilicity of natural fibers. The peak temperature corresponds to the moisture loss appear at 48˚C and 31˚C for untreated COIR and PALF respectively. The PALF exhibits a higher rate of moisture loss than that of COIR fiber. This might be due to the high moisture content in PALF. Table 3 depicts that the 4 wt% NaOH treated PALF exhibits less weight loss (around 48%) due to moisture evaporation as compared to the untreated fiber. This was attributed to the removal of hemicellulose, pectin, and waxy substances after an alkaline treatment. Arifuzzaman Khan *et al.*[36] also reported that the percentage of initial weight loss due to moisture evaporation is higher in untreated COIR fiber than that of NaOH treated one. At 100˚C, 8 wt% alkaline treated COIR fibers possess less weight loss (approx. 40%) as compared to raw fibers. Ndazi *et al.*[35] concluded that the structurally bound water is resistant to complete water removal during drying.

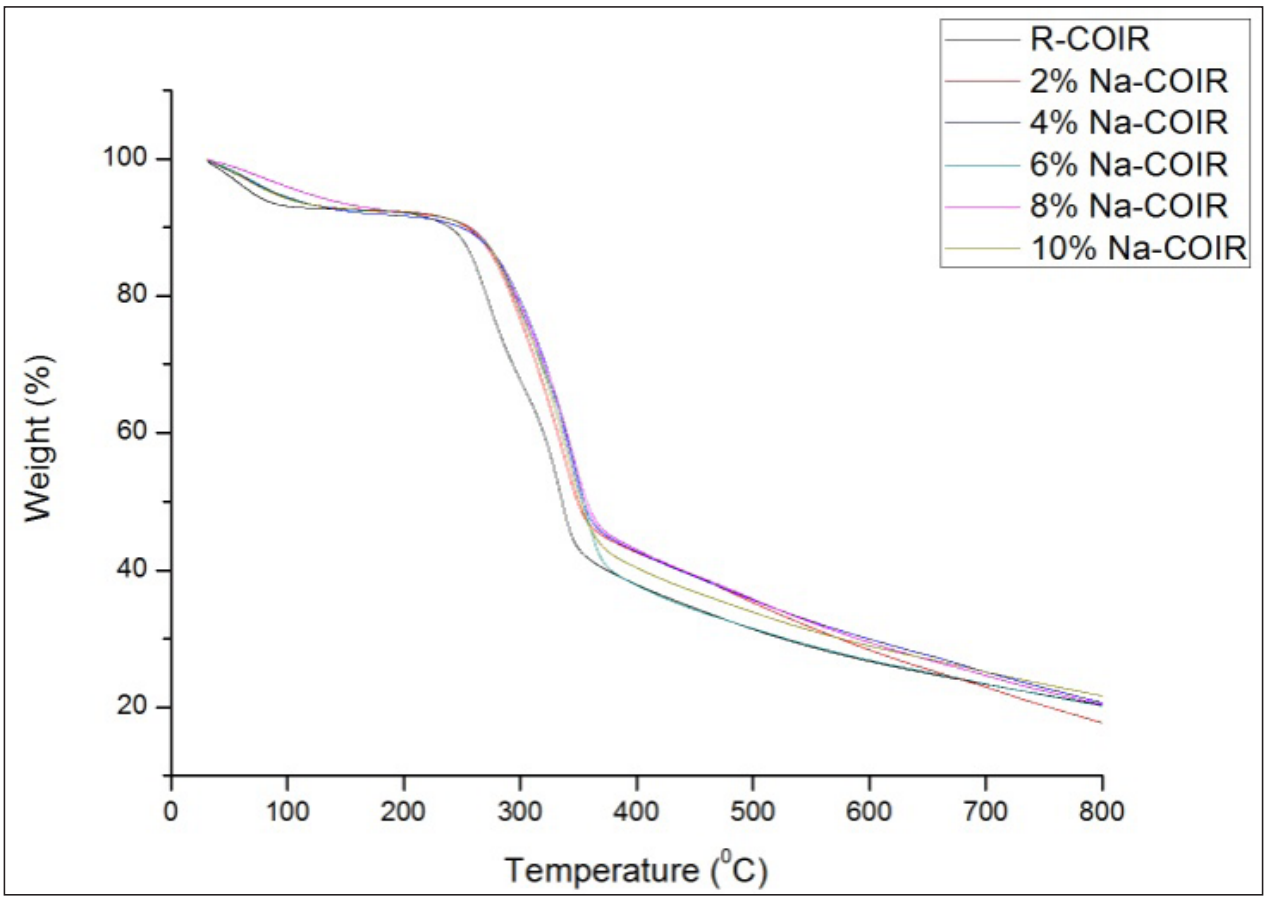



**Fig. 4:** TGA curves of untreated and alkali treated coir fibers

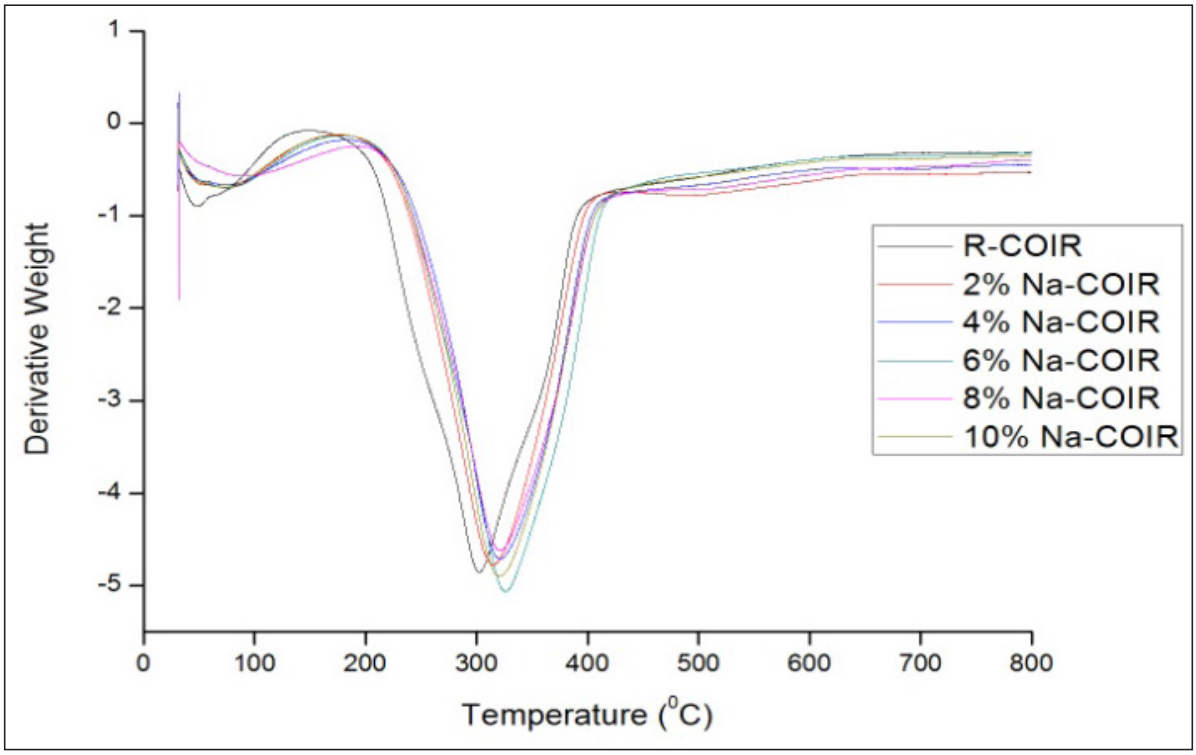



**Fig. 5:** DTG curves of untreated and alkali treated coir fibers

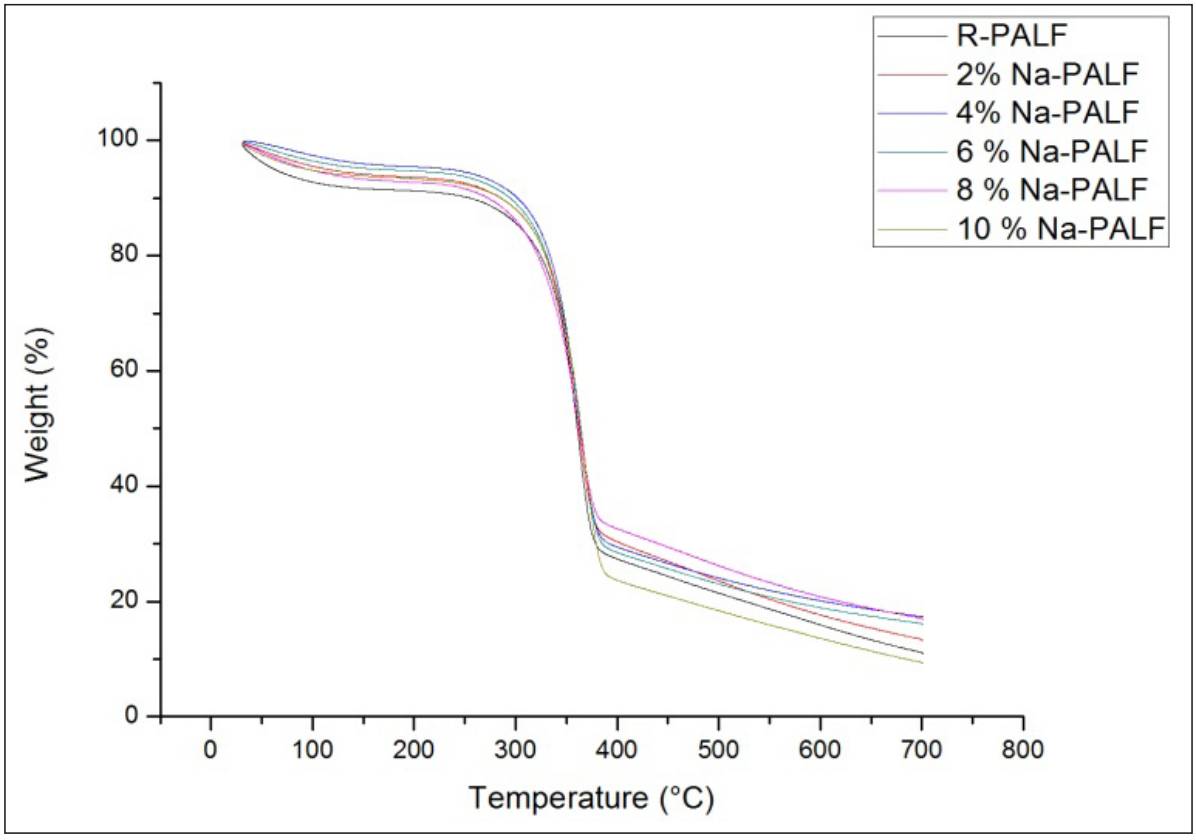



**Fig. 6:** TGA curves of untreated and alkali treated pineapple leaf fibers

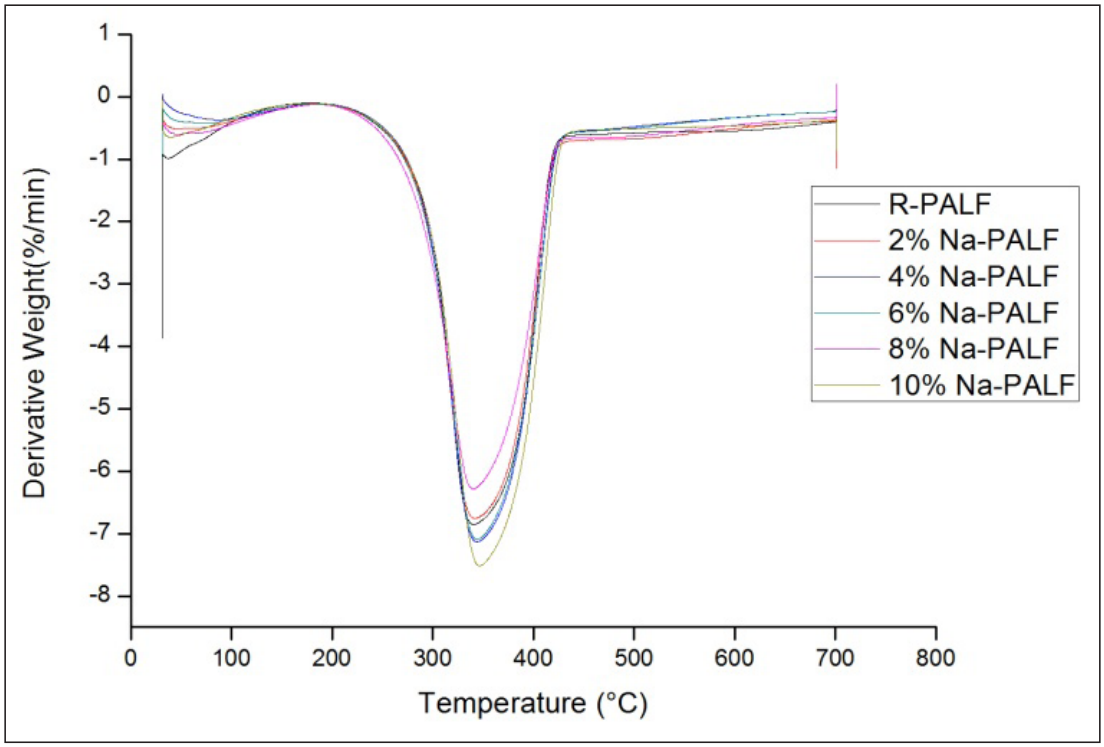



**Fig. 7:** DTG curves of untreated and alkali treated pineapple leaf fibers

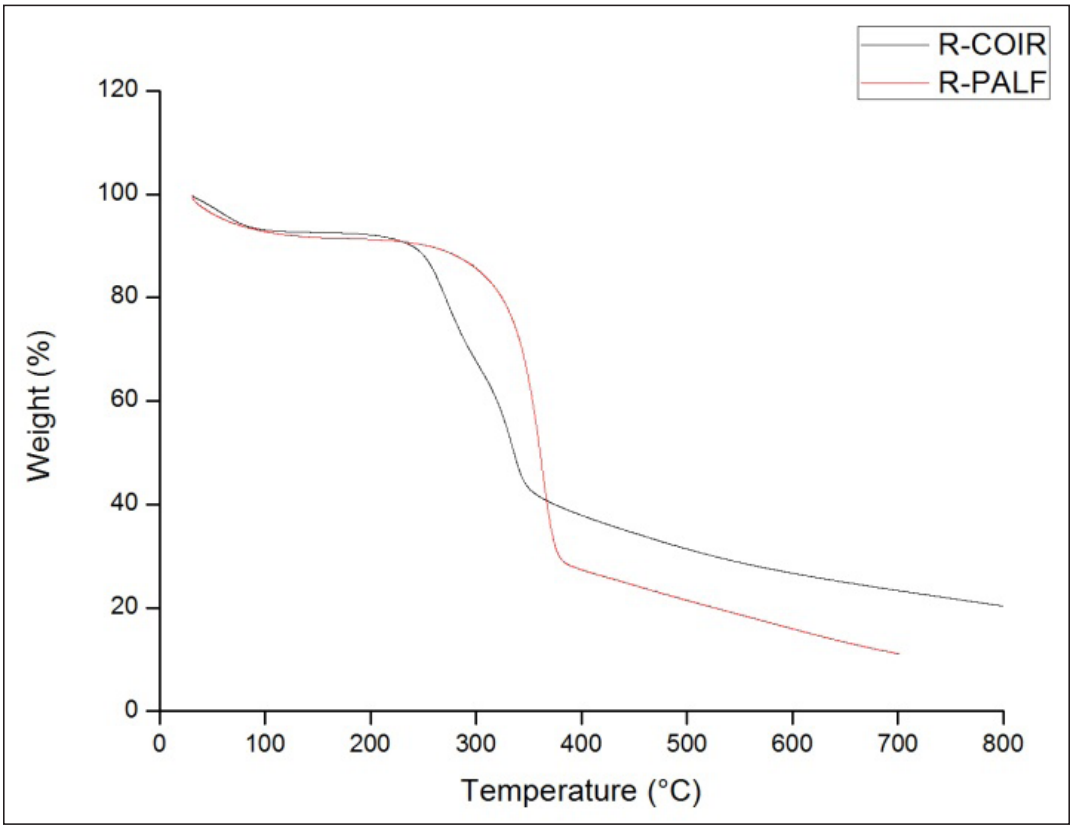



**Fig. 8:** TGA curve of untreated coir and pineapple leaf fiber

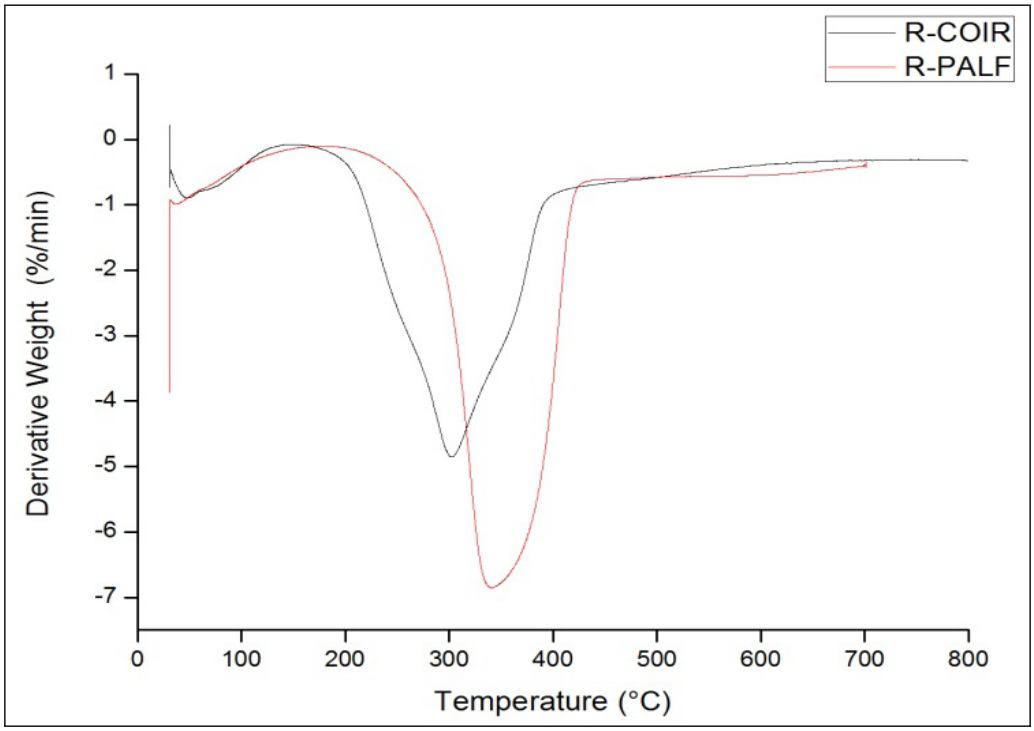



**Fig. 9:** DTG curve for untreated coir and pineapple leaf fiber

Table 4 and 5 depicts that the weight loss in all cases (Untreated and NaOH treated PALF and COIR fibers) is negligible below 290˚C; major mass loss (30-50 wt%) occur above this temperature and it was due to the degradation of crystalline cellulose. M. Siti. Alwani *et al.*[37] also reported that the major mass loss (30-50%) in SCB, PALF, COIR, and BPS fibers occurred at a temperature above than 300˚C. Hemicellulose decomposition is the first stage of fiber degradation. As the temperature increases, the cellular breakdown of hemicellulose takes place. The coconut husk fibers started to decompose at 150˚C; however, the pineapple leaf fibers decompose at 185˚C. In both cases, the peak temperature of maximum degradation was shifted to a higher temperature after an alkaline treatment. The 4 wt% NaOH treated COIR fiber exhibits higher peak temperature of thermal degradation (about 20˚C) than that of untreated one. This was due to the removal of hemicellulose after an alkaline treatment of cellulosic fibers. Yang *et al.*[38] reported that the presence of a carbonyl group (C=O) in hemicellulose be the cause of its lower thermal stability than that of α-cellulose, and lignin. Ndazi Bwire S *et al.*[19] reported the reduction in a temperature of maximum weight loss after an alkaline treatment of rice husk fiber and it was due to the excess removal of cementing materials like hemicellulose and lignin. The main degradation peak lies within 290˚C-380˚C temperature with 50% degradation. The PALF possess higher mass loss (around 10%) as compared to COIR fiber. This was because of the high cellulose content in PALF. However, the degradation of PALF occurs at a higher temperature than that of COIR fiber. M. Siti. Alwani *et al.*[37] reported that the sugarcane bagasse fiber exhibits higher mass loss than that of pineapple leaf, banana, and coconut fibers. This was attributed to the presence of high cellulose content in SCB fiber.

Lignin was the most thermally stable component of a cellulosic fiber due to its highly cross-linked complex aromatic structure. Mackay and Nassar[39] reported that the thermal stability of lignocellulosic fibers is in the order: lignin>α-cellulose>hemicellulose. Lignin starts to decompose at a lower temperature (typically 160-175˚C) as compared to cellulose but it decomposes slowly under the whole temperature range and its

temperature extends as high as 900˚C. According to Paiva *et al.*[40], the decomposition of lignin occurred in a wider temperature range as compared to the cellulose and hemicellulose components. Table 4 and 5 clearly revealed that the remaining mass of COIR was higher than PALF fiber after 700˚C. This was attributed to the presence of a large amount of lignin compound in COIR fiber. Williams *et al.*[41] reported that the high lignin content in lignocellulosic fibers leads to a high amount of residues during pyrolysis. It was observed that the 4 wt% NaOH treated COIR and PALF fibers possess higher residual mass than that of untreated fibers. This was due to the slow rate of chemical decomposition reaction and dominance of the formation of condensed cellulose as the solid residue.

Table 6 and 7 illustrate the DTG results of untreated and alkali treated fibers. In case of untreated COIR fiber, the two decomposition peaks at 228˚C and 302˚C at the rate of 1.475 and 4.852 %/min correspond to the decomposition of hemicellulose and cellulose components. It was observed that the rate of hemicellulose decomposition is more prominent in untreated fibers (PALF and COIR) as compared to alkali treated fibers. Pineapple leaf fiber exhibits a higher rate of decomposition than coconut husk fiber. Thus from DTG results, we can conclude that the alkaline treated fibers possess degradation peak at a higher temperature than that of untreated fibers.

**Table 4:** TGA of untreated and treated coir fibers

| Fiber | Temperature Range ˚C | Weight Loss (%) | Residual mass (%) at (700˚C) |
|---|---|---|---|
| | 0-152.4 | 7.396 | |
| R-COIR | 152.4-242.85 | 3.05 | 23.359 |
| | 242.85-352.94 | 46.49 | |
| | 0-161.9 | 7.4 | |
| 2% Na-COIR | 161.9-262.01 | 3.68 | 22.94 |
| | 262.01-361.86 | 42.73 | |
| | 0-176.06 | 8.07 | |
| 4% Na-COIR | 176.06-277.08 | 5.7 | 25.135 |
| | 277.08-363.15 | 39.39 | |
| | 0-176.19 | 7.66 | |
| 6% Na-COIR | 176.19-271.44 | 4.8 | 23.409 |
| | 271.44-376.18 | 47.29 | |
| | 0-152.38 | 6.65 | |
| 8% Na-COIR | 152.38-266.72 | 4.7 | 24.64 |
| | 266.47-376.25 | 43.43 | |
| | 0-171.52 | 7.55 | |
| 10% Na-COIR | 171.52-276.12 | 5.69 | 25.15 |
| | 276.12-361.95 | 40.92 | |

**Table 5:** TGA of untreated and treated pineapple leaf fibers

| Fiber | Temperature Range $^0$C | Weight Loss (%) | Residual Char (%) |
|---|---|---|---|
| R-PALF | 0-184.78 | 8.63 | 11.103 |
| | 184.78-330.4 | 13.63 | |
| | 330.4-382.66 | 48.75 | |
| 2% Na-PALF | 0-164.31 | 6.071 | 13.356 |
| | 164.31-261.85 | 1.97 | |
| | 261.85-361.97 | 41.98 | |
| 4% Na-PALF | 0-184.04 | 4.465 | 17.34 |
| | 184.04-324.15 | 11.18 | |
| | 324.15-376.7 | 49.59 | |
| 6% Na-PALF | 0-175.1 | 5.12 | 16.14 |
| | 175.1-321.7 | 11.16 | |
| | 321.7-378.32 | 50.57 | |
| 8% Na-PALF | 0-186.88 | 7.162 | 17.01 |
| | 186.88-315.15 | 10.54 | |
| | 315.15-378.23 | 46.56 | |
| 10% Na-PALF | 0-186.95 | 6.545 | 9.386 |
| | 186.95-326.09 | 12.03 | |
| | 326.09-382.54 | 54.058 | |

**Table 6:** DTG of untreated and treated coir fibers

| Fiber | Temperature $^0$C | Rate of Weight Loss (%/min) |
|---|---|---|
| R-COIR | 48.43 | 0.897 |
| | 228.43 | 1.475 |
| | 302.56 | 4.852 |
| 2% Na-COIR | 72.47 | 0.697 |
| | 243.35 | 1.162 |
| | 314.3 | 4.774 |
| 4% Na-COIR | 73.61 | 0.67 |
| | 250.13 | 1.156 |
| | 321.74 | 4.708 |
| 6% Na-COIR | 74.43 | 0.428 |
| | 252.74 | 0.565 |
| | 325.9 | 5.061 |
| 8% Na-COIR | 90.80 | 0.571 |
| | 252.96 | 1.413 |
| | 321.09 | 4.62 |
| 10% Na-COIR | 72.77 | 0.693 |
| | 258.48 | 1.719 |
| | 321.95 | 4.894 |

**Table 7:** DTG of untreated and treated pineapple leaf fibers

| Fiber | Temperature ˚C | Rate of Weight Loss (%/min) |
|---|---|---|
| R-PALF | 31.16 | 1.47 |
| | 310.82 | 3.54 |
| | 341.58 | 6.85 |
| 2% Na-PALF | 46.37 | 0.52 |
| | 247.81 | 0.49 |
| | 341.52 | 6.75 |
| 4% Na-PALF | 85.79 | 0.37 |
| | 308.69 | 3.39 |
| | 343.43 | 7.13 |
| 6% Na-PALF | 70.83 | 0.43 |
| | 302.33 | 2.73 |
| | 343.12 | 7.091 |
| 8% Na-PALF | 46.49 | 0.60 |
| | 300.08 | 2.71 |
| | 339.72 | 6.28 |
| 10% Na-PALF | 39.7 | 0.65 |
| | 313.05 | 3.66 |
| | 346.85 | 7.51 |

## CONCLUSION

In this work, an experimental investigation was carried out to study the thermal behavior of untreated and alkali treated PALF and COIR fibers. From this study, the following conclusions can be made:

- The thermal decomposition of PALF and COIR fibers proceeded in two steps. In 1st step, a minor weight loss occurs due to an evaporation of water and extraction of volatile compounds and in the 2nd step, major degradation occurs due to the decomposition of hemicellulose, cellulose, and lignin components.
- The alkaline treatment of PALF and COIR fibers leads to a lower weight loss and shifting of degradation peak to a higher temperature.
- Amongst all the samples, the 4 wt% NaOH treated PALF and COIR fibers exhibit maximum thermal stability. This improvement was due to the removal of less thermally stable hemicellulose component after an alkaline treatment. The temperature of maximum decomposition was shifted from 302˚C to 321˚C after the 4 wt% alkaline treatment of COIR fiber.
- Pineapple leaf fiber exhibits a higher rate of decomposition (6.85%/min) than coir fiber (4.85%/min). It was due to the presence of more cellulosic content in PALF.
- After 700˚C, the remaining mass of COIR (23.35%) was higher as compared to PALF (11.10%). This might be attributed to the presence of high lignin content in COIR fiber.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge all support received from Analytical Instrument Laboratory at CSIR-CSIO, Chandigarh.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Funding:** This research work received no specific grant from any funding agency in the public, commercial, and nonprofit sectors.

## REFERENCES

1. Maguno, A. 1999. *2nd International wood and natural fiber composites symposium*, Kassel, Germany.
2. Qureshi, H.A. 1999. *2nd International wood and natural fiber composites symposium*, Kassel, Germany.
3. Nickel, J. and Riedel, U. 2003. Activities in biocomposites. *Materials Today,* pp. 44-48.
4. Anuar, H., Ahmad, S., Rasid, R., Ahmad, A., and Busu, W. 2008. Mechanical properties and dynamic mechanical analysis of thermoplastic natural rubber reinforced short carbon fiber and kenaf fiber hybrid composites. *J. Appl. Polym. Sci.,* **107**(6): 4043-4052.
5. Barreto, A., Esmeraldo, M., Rosa, D., Fechine, P. and Mazzeto, S. 2010. Cardanol biocomposites reinforced with jute fiber: microstructure, biodegradability, and mechanical properties. *Polym. Compos.* **31**(11): 1928-1937.
6. Mohanty, A.K., Misra, M. and Drzal, L.T. 2005. Natural fibers, biopolymers and biocomposites. Boca Raton: CRC Press.
7. Asim, M., Abdan, K., Jawaid, M., Nasir, M., Dashtizadeh, Z., Ishak, M.R. and Hoque, M. 2015. A review on pineapple leaves fibre and its composites, *Int. J. Polym. Sci.* http://dx.doi.org/10.1155/2015/950567.
8. Mittal, M. and Chaudhary, R. 2018. Experimental study on the water absorption and surface characteristics of alkali treated pineapple leaf fiber and coconut husk fiber. *Int. J. Appl. Eng. Res.,* **13**(15): 12237-12243.
9. Mittal, M. and Chaudhary, R.2018. Experimental investigation on the mechanical properties and water absorption behavior of randomly oriented short pineapple/coir fiber-reinforced hybrid epoxy composites, *Mater. Res. Express.*
10. Mittal, M. and Chaudhary, R. 2018. Effect of fiber content on thermal behavior and viscoelastic properties of PALF/Epoxy and COIR/Epoxy composites, *Mater. Res. Express.*
11. Statistical databases of "*Food and Agriculture Organization of the United Nations*", www.fao.org/statistics/databases/en/.
12. Mwaikambo, L.Y. and Ansell, M.P. 2002. Chemical modification of hemp, sisal, jute, and kapok fibers by alkalization. *J. Appl. Polym. Sci.,* **84**: 2222-2234.
13. John, M.J., Francis, B., Varughese, K.T. and Thomas, S. 2008. Effect of chemical modification on properties of hybrid fiber biocomposites. *Comp. Part A.,* **39**: 352-363.

14. Li, X., Tabil, L.G. and Panigrahi, S. 2007. Chemical treatments of natural fiber for use in natural fiber-reinforced composites: A Review. *J. Polym. Environ.*, **15**: 25-33.

15. John, M.J. and Anandjiwala, R.D. 2008. Recent developments in chemical modification and characterization of natural fiber reinforced composites. *Polym. Comp.*, pp187-207.

16. Slopiecka, K., Bartocci, P. and Fantozzi, F. 2012. Thermogravimetric analysis and kinetic study of poplar wood pyrolysis. *Appl. Energy,* **97**: 491-497.

17. Oushabi, A., Sair, S., Oudrhiri Hassani, F., Abboud, F., Tanane, O. and Bouari, A. 2017. The effect of alkali treatment on mechanical, morphological, and thermal properties of date palm fibers: Study of the interface of DPF-Polyurethane composite. *Sou. Afr. J. Chem. Eng.,* **23**: 116-123.

18. Panyasart, K., Chaiyut, N., Taweechai, A. and Onuma, S. 2014. Effect of surface treatment on the properties of pineapple leaf fibers reinforced polyamide 6 composites. *11th Eco-Energy and Materials Science and Engineering,* **56**: 406-413.

19. Ndazi, B.S., Christian, N. and Joseph, T. 2007. Chemical and thermal stability of rice husks against alkali treatment. *Bioresources,* **3**(4): 1267-1277.

20. Rachini, A., Troedec, M.L., Peyratout, C. and Smith, A. 2009. Comparison of the thermal degradation of natural, alkali treated, and silane treated hemp fibers under air and an inert atmosphere. *J. Appl. Polym. Sci.,* **112**: 226-234.

21. Mortari, D.A., Britto, M.C. and Crnkovic, P.M. Correlation between activation energy and thermal decomposition yield of sugar cane bagasse under $CO_2/O_2$ and $N_2/O_2$. *Chem. Eng. Trans.,* **37**: 31-36.

22. White, J.E., Catallo, W.J. and Legendre, B.L. 2011. Biomass pyrolysis kinetics: A comparative critical review with relevant agricultural residue case studies. *J. Anal. Appl. Pyrol.,* **91**(1): 1-33.

23. Yao, F., Wu, Q., Lei, Y., Guo, W. and Xu, Y. 2008. Thermal decomposition kinetics of natural fibers: Activation energy with dynamic thermogravimetric analysis. *Polym. Degr. Stab.,* **93**: 90-98.

24. Sergio, N.M., Veronica, C. and Rodriguez, R.J.S. 2012. The thermogravimetric behavior of natural fibers reinforced polymer composites, *Mater. Sci. Eng.,* **557**: 17-28.

25. Rachini, A., Troedec, M.L., Peyratout, C. and Smith, A. 2009. Comparison of the thermal degradation of natural, alkali treated, and silane treated hemp fibers under air and an inert atmosphere. *J. Appl. Polym. Sci.*, **112**: 226-234.

26. Tajvidi, M. and Takemura, A. 2010. Thermal degradation of natural fiber reinforced polypropylene composites, *J. Thermoplast. Compos. Mater.,* **23**: 281-298.

27. Na Lu and Oza, S. 2013. Effect of surface treatment of hemp fibers on the thermal stability of Hemp-PLA composites. *Adv. Mat. Res*., **651**: 499-504.

28. Jayamani, E., Hamdan, S., Rahman, Md., Khusairy, M. and Bakri, B. 2014. Investigation of fiber surface treatment on mechanical, acoustical, and thermal properties of betel nut fiber polyester composites. *12th Global Congress on Manufacturing and Management,* **97**: 545-554.

29. Mittal, M. and Chaudhary, R. 2018. Development of PALF/Glass and COIR/Glass fiber reinforced hybrid epoxy composites, *J. Mat. Sci. Surf. Eng.,* **6**(5): 851-861.

30. Ranjan, R., Bajpai, P.K. and Tyagi, R.K. 2013. Mechanical characterization of Banana/Sisal fiber reinforced PLA hybrid composites for structural applications, *Eng. Int.*, **1**(1): 39-49.

31. Devi, L.U., Bhagawan, S.S., Nair, K.C.M. and Thomas, S. 2011. Water absorption behavior of PALF/GF hybrid polyester composites, *Polym. Compo.,* doi: 10.1002/pc.21034.

32. Goud, G. and Rao, R.N. 2010. Effect of fiber content and alkali treatment on mechanical properties of *Roystonea regia*-reinforced epoxy partially biodegradable composites, *B. Mater. Sci.,* **34**(7): 1575-1581.

33. Anand, P. and Anbumalar, V. Mechanical properties of cellulose-filled epoxy hybrid composites reinforced with alkali-treated hemp fiber, *Polym.,* **39**(1): 46-55.

34. Venkateshappa, S.C., Jayadevappa, S.Y. and Puttiah, P.K.W. 2011. Mechanical behavior of areca fiber reinforced epoxy composites, *Adv. Polym. Technol.,* **31**(4): 319-330.

35. Ndazi, B., Tesha, J. and Bisanda, E. 2006. Some opportunities and challenges of producing bio-composites from non-wood residues. *J. Mat. Sci.***, 41**(21): 6984-6990.

36. Arifuzzaman Khan, G.M. and Shamsul Alam Md. 2012. Thermal characterization of chemically treated coconut husk fibre. *Ind. J. Fib. Tex. Res.,* **37**: 20-26.

37. Alwani, M.S., Khalil, H.P.S., Sulaiman, O., Islam Md, and Dungani, R. 2014. An approach to using agricultural waste fibres in biocomposite application: Thermogravimetric analysis and activation energy study. *Bioresources.,* **9**: 218-230.

38. Yang, H., Yan, R., Chen, H., Lee, D.H. and Zheng, C. 2007. Characteristics of hemicellulose, cellulose, and lignin pyrolysis. *Fuel.,* **86**(12):1781-1788.

39. Nasar, M.N. and MacKay, G.D.M. 1984. Mechanism of thermal decomposition of lignin. *Wood. Fib. Sci.,* **16**(3): 441-453.

40. Paiva, J.M. and Frollini, E. 2006. Unmodified and modifies surface sisal fibers as reinforcement of phenolic and lignophenolic matrices composites: Thermal analysis of fibers and composites. *Macromol. Mater. Eng.,* **291**: 405-417.

41. Williams, P.T. and Reed, A.R. 2004. Thermal processing of biomass natural fibre wastes by pyrolysis. *Int. J. Energy. Res.,* **28**(2): 131-145.



View publication stats

Research on Chemical Intermediates
https://doi.org/10.1007/s11164-019-03847-8

# Facile one-pot synthesis of 5-amino-1*H*-pyrazole-4-carbonitriles using alumina–silica-supported $MnO_2$ as recyclable catalyst in water

**Poonam**[1] · **Ram Singh**[1]

Received: 21 February 2019 / Accepted: 28 April 2019
© Springer Nature B.V. 2019

## Abstract

A novel, facile, one-pot, multicomponent protocol for the synthesis of 5-amino-1*H*-pyrazole-4-carbonitrile derivatives has been developed using alumina–silica-supported $MnO_2$ as recyclable catalyst in water and sodium dodecyl benzene sulphonate at room temperature. The cyclo-condensation of substituted benzaldehydes, malononitrile and phenyl hydrazine gave the 5-amino-1,3-diphenyl-1*H*-pyrazole-4-carbonitriles in 86–96% yield.

**Keywords** Pyrazole · Benzaldehyde · Malononitrile · Phenyl hydrazine · Alumina–silica–$MnO_2$

## Introduction

One-pot multicomponent reactions (MCRs) are always recommended for the synthesis of complex molecules [1–5]. The combined utilization of green solvents and heterogeneous catalyst makes MCRs an environmentally benign method. Highly functionalized pyrazole derivatives are potential biologically active scaffolds due to their wide applications in pharmaceuticals [6–9]. These scaffolds have been used as analgesic [10], anti-bacterial [11–13], anti-convulsant [14], anti-pyretic [15], anti-depressant [16], anti-fungal [17], anti-inflammatory [18, 19], antimalarial [20], anti-microbial [21, 22], anti-parasitic [23, 24] and anti-tumour compounds [25, 26]. Further, these compounds are reported to have appreciable anti-hypertensive activity in vivo, and also exhibit properties such as human cannabinoid receptors (hCB1 and hCB2), inhibitors of p38 kinase and CB1 receptor antagonists [27].

✉ Ram Singh
ramsingh@dtu.ac.in

[1] Department of Applied Chemistry, Delhi Technological University, Delhi 110042, India

Published online: 07 May 2019

The most well-known method for the synthesis of the 5-aminopyrazole-4-carbonitriles scaffold is the three-component cyclocondensation (3-CC) of aldehydes, phenyl hydrazine derivatives and malononitrile [28–41]. A variety of catalysts and reagents like sodium ascorbate [6], molecular iodine [28], ionic liquids [29], nanoparticles [30, 31], rhodium catalyst with sodium acetate (NaOAc) [32], piperidine [33], piperidinium acetate [34], $Cu(OAc)_2$ [35], $CuO/ZrO_2$ [36], cerium (IV) ammonium nitrate [37], graphene oxide–$TiO_2$ [38], oxone [39], palladium and copper [40] and alum [41] have been studied for this reaction.

To our knowledge, metal oxide with inorganic oxide surfaces using MCRs has not been used as the catalyst in the synthesis of pyrazoles under aqueous conditions. Heterogenous catalysts using inorganic oxide surfaces like zeolites, clays, silica and alumina have received attention due to ease of the working process [42–45]. Considering the utility of alumina and silica, we report herein an efficient, eco-friendly, one-pot, three-component synthesis of 5-amino-1,3-diphenyl-1*H*-pyrazole-4-carbonitrile derivatives in water using alumina–silica-supported $MnO_2$ as the heterogeneous catalyst.

## Experimental section

All solvents and reagents were purchased from commercial suppliers and were used without further purification. Reactions were performed in oven-dried glassware. Melting points were determined on a laboratory open-glass melting apparatus and were uncorrected. Thin-layer chromatography (TLC) was performed on aluminium-coated silica plates purchased from Merck. Infrared (IR) spectra were recorded in KBr on a Perkin-Elmer Spectrum II spectrophotometer, and proton nuclear magnetic resonance ($^1$H NMR) spectra were recorded on a Bruker Avance-300 spectrophotometer using tetramethylsilane (TMS) as an internal reference with chemical shift values being reported in ppm.

### General procedure for the synthesis of catalytic $MnO_2$

In a solution of *n*-propanol and de-ionised water (70 mL, 1:1 v/v), manganese acetate tetrahydrate [$Mn(OAc)_2{\cdot}4H_2O$, 6 g] and silica and/or alumina (3 g) were added. The reaction mixture was stirred at 80 °C for 3 h. Further, an aqueous solution of $KMnO_4$ (1.5 g in 70 mL water) was added to the mixture and kept stirring for 2 h. A black precipitate was obtained. It was centrifuged and separated and then dried at 70 °C under vacuum for 5 h.

### General procedure for the synthesis of pyrazoles

Aldehyde (1 mmol), malononitrile (1 mmol) and catalyst (150 mg) were taken in distilled water (10 mL) into a round-bottomed flask. To the reaction mixture, sodium dodecyl benzene sulphonate (SDBS, 150 mg) was added and stirred at room temperature for 20 min till a white precipitate was obtained. At this point,

phenyl hydrazine (1 mmol) was added, and the reaction mixture was further stirred at room temperature (35 °C) for 3–12 min (Table 2). The progress of the reaction was monitored by TLC (hexane/ethyl acetate 7:3, v/v). After completion of the reaction, water and ethyl acetate (1:1, v/v, 30 mL) was added and filtered. The filtrate was put in a separating funnel to separate the ethyl acetate layer. The ethyl acetate was dried over anhydrous sodium sulphate ($Na_2SO_4$), filtered and evaporated to give the final product **4** that was further recrystallized with hot ethanol.

***5-Amino-1,3-diphenyl-1H-pyrazole-4-carbonitrile*** **(4a)** IR (KBr, $cm^{-1}$): 3470, 3312, 2925, 2193, 1634, 1591, 1493, 1257, 1133, 1085, 982.$^1$H NMR ($CDCl_3$): $\delta$ 7.85 (s, 2H), 7.64 (d, 2H, $J=7.2$ Hz), 7.41–7.36 (m, 2H), 7.31–7.24 (m, 3H), 7.06 (d, 2H, $J=7.5$ Hz), 6.74 (t, 1H, $J=7.2$ Hz).

***5-Amino-3-(4-nitrophenyl)-1-phenyl-1H-pyrazole-4-carbonitrile*** **(4b)** IR (KBr, $cm^{-1}$): 3454, 3300, 3105, 2926, 2196, 1594, 1555, 1536, 1493, 1324, 1241, 1164, 1100, 915.$^1$H NMR ($CDCl_3$): $\delta$ 8.24–8.21 (m, 2H), 7.79–7.70 (m, 3H), 7.36–7.29 (m, 3H), 7.15 (dd, 2H, $J=9.0$ and 1.2 Hz), 6.95 (t, 1H, $J=7.2$ Hz).

***5-Amino-3-(2-nitrophenyl)-1-phenyl-1H-pyrazole-4-carbonitrile*** **(4c)** IR (KBr, $cm^{-1}$): 3449, 3294, 3023, 2970, 2855, 2196, 1601, 1572, 1533, 1491, 1334, 1254, 1132, 903.$^1$H NMR ($CDCl_3$): $\delta$ 8.36 (s, 2H), 8.26–8.14 (m, 2H), 8.05–7.95 (m, 2H), 7.59–7.53 (m, 2H), 7.15–7.12 (m, 2H), 6.84 (t, 1H, $J=7.3$ Hz).

***5-Amino-3-(2-bromophenyl)-1-phenyl-1H-pyrazole-4-carbonitrile*** **(4d)** IR (KBr, $cm^{-1}$): 3455, 3331, 3298, 3053, 2967, 2226, 1602, 1579, 1512, 1438, 1255, 1142, 1021.$^1$H NMR ($CDCl_3$): $\delta$ 8.07–8.04 (m, 2H), 7.52 (dd, 2H, $J=8.1$ and 1.2 Hz), 7.30–7.24 (m, 4H), 7.25–7.15 (m, 2H), 6.86 (t, 1H, $J=7.2$ Hz).

***5-Amino-3-(2-fluorophenyl)-1-phenyl-1H-pyrazole-4-carbonitrile*** **(4e)** IR (KBr, $cm^{-1}$): 3458, 3344, 3026, 2926, 2855, 2209, 1594, 1493, 1354, 1257, 1138, 1091.$^1$H NMR ($CDCl_3$): $\delta$ 7.99–7.55 (m, 4H), 7.27–7.25 (m, 2H), 7.11–6.88 (m, 5H).

***5-Amino-3-(4-fluorophenyl)-1-phenyl-1H-pyrazole-4-carbonitrile*** **(4f)** IR (KBr, $cm^{-1}$): 3465, 3311, 3053, 2924, 2208, 1597, 1502, 1445, 1229, 1089.$^1$H NMR ($CDCl_3$): $\delta$ 7.87 (s, 2H), 7.61–7.56 (m, 2H), 7.28–7.23 (m, 2H), 7.09–7.02 (m, 4H), 6.86 (t, 1H, $J=7.2$ Hz).

***5-Amino-3-(4-bromophenyl)-1-phenyl-1H-pyrazole-4-carbonitrile*** **(4g)** IR (KBr, $cm^{-1}$): 3465, 3310, 3043, 2926, 2210, 1600, 1501, 1493, 1320, 1202, 1150, 1062.$^1$H NMR ($CDCl_3$): $\delta$ 7.83 (s, 2H), 7.71–7.69 (m, 2H), 7.65–7.59 (m, 3H), 7.32–7.21 (m, 3H), 6.89–6.83 (m, 1H).

***5-Amino-3-(4-chlorophenyl)-1-phenyl-1H-pyrazole-4-carbonitrile*** **(4h)** IR (KBr, $cm^{-1}$): 3445, 3342, 3033, 2986, 2203, 1612, 1595, 1544, 1493, 1090. $^1$H NMR

($CDCl_3$): $\delta$ 7.85 (s, 2H), 7.70–7.49 (m, 5H), 7.38–7.24 (m, 3H), 6.88 (t, 1H, $J = 7.3$ Hz).

***5-Amino-3-(2-chlorophenyl)-1-phenyl-1H-pyrazole-4-carbonitrile*** **(4i)** IR (KBr, $cm^{-1}$): 3455, 3341, 3022, 2921, 2853, 2204, 1591, 1497, 1355, 1257, 1130, 1091, 953.$^1$H NMR ($CDCl_3$): $\delta$ 7.98 (s, 2H), 7.62–7.55 (m, 2H), 7.40–7.34 (m, 3H), 7.25–7.15 (m, 3H), 6.76 (t, 1H, $J = 7.3$ Hz).

***5-Amino-3-(4-methylphenyl)-1-phenyl-1H-pyrazole-4-carbonitrile*** **(4j)** IR (KBr, $cm^{-1}$): 3455, 3310, 2924, 2205, 1595, 1505, 1256, 1128, 1069.$^1$H NMR ($CDCl_3$): $\delta$ 8.26 (s, 2H), 7.82–7.77 (m, 2H), 7.53–7.41 (m, 4H), 7.22–7.15 (m, 2H), 6.87 (t, 1H, $J = 7.2$ Hz), 2.33 (s, 3H).

***5-Amino-3-(4-methoxyphenyl)-1-phenyl-1H-pyrazole-4-carbonitrile*** **(4k)** IR (KBr, $cm^{-1}$): 3455, 3300, 2962, 2889, 2198, 1603, 1556, 1519, 1483, 1303, 1122, 1011, 993.$^1$H NMR ($CDCl_3$): $\delta$ 7.94 (s, 2H), 7.69–7.61 (m, 2H), 7.57–7.52 (m, 3H), 7.13–7.00 (m, 3H), 6.76 (t, 1H, $J = 7.3$ Hz), 3.74 (s, 3H).

***5-Amino-3-(2-methoxyphenyl)-1-phenyl-1H-pyrazole-4-carbonitrile*** **(4l)** IR (KBr, $cm^{-1}$): 3453, 3297, 2962, 2919, 2198, 1596, 1519, 1493, 1333, 1128, 1016, 997.$^1$H NMR ($CDCl_3$): $\delta$ 7.84 (s, 2H), 7.66–7.61 (m, 2H), 7.50–7.44 (m, 2H), 7.13–7.01 (m, 4H), 6.89–6.93 (m, 1H), 3.84 (s, 3H).

***5-Amino-3-(3,4-dimethoxyphenyl)-1-phenyl-1H-pyrazole-4-carbonitrile*** **(4m)** IR (KBr, $cm^{-1}$): 3463, 3294, 2972, 2939, 2201, 1596, 1511, 1494, 1330, 1128, 1016.$^1$H NMR ($CDCl_3$): $\delta$ 8.45 (s, 2H), 7.69–7.63 (m, 2H), 7.35–7.29 (m, 3H), 7.16–6.97 (m, 3H), 3.79–3.76 (m, 6H).

***5-Amino-3-(3-nitrophenyl)-1-phenyl-1H-pyrazole-4-carbonitrile*** **(4n)** IR (KBr, $cm^{-1}$): 3450, 3284, 3023, 2972, 2855, 2206, 1601, 1572, 1555, 1533, 1490, 1334, 1251, 1130, 943.$^1$H NMR ($CDCl_3$): $\delta$ 8.16 (s, 2H), 7.76–7.64 (m, 2H), 7.45–7.35 (m, 2H), 7.19–7.03 (m, 2H), 6.99–6.84 (m, 3H).

## Results and discussion

### Preparation of supported catalysts

The initial studies focussed on the development of heterogeneous catalysts based on inorganic oxides: silica gel (60–120 mesh) and acidic alumina [43–45]. The in-situ generation of manganese oxide ($MnO_2$) was done from manganese acetate tetrahydrate [$Mn(OAc)_2 \cdot 4H_2O$] by modification of the reported method [46]. A mixture of *n*-propanol, de-ionised water, manganese acetate tetrahydrate and silica and/or alumina was stirred at 80 °C for 3 h. To this, an aqueous solution of $KMnO_4$ was added, and the stirring was continued for another 2 h. The reaction mixture was filtered to

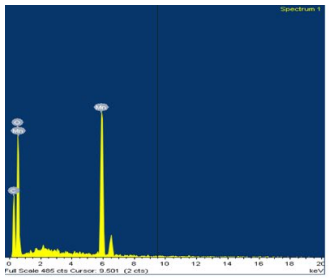
get a black solid which was washed with water and dried at 70 °C under vacuum for 5 h.

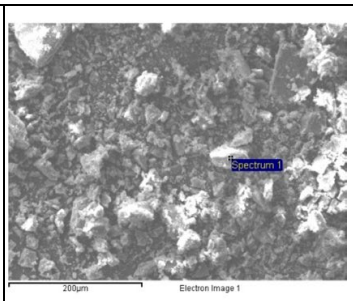
The elemental compositions of manganese acetate tetrahydrate, acidic alumina, silica gel, silica-supported $MnO_2$, alumina-supported $MnO_2$ and alumina–silica-supported $MnO_2$ were established by energy dispersive X-ray (EDX) spectroscopy (Fig. 1). The scanning electron microscopy (SEM) images showed the morphology of the prepared catalysts (Fig. 1). The EDX patterns clearly indicated the expected elemental components with no extra peaks observed for any other impurities. The corresponding analytical data derived from EDX analysis are summarized along with images in Fig. 1.

The images clearly confirm the presence of silica, alumina and $MnO_2$ in the catalyst. Moreover, the absence of C shows the conversion of manganese acetate tetrahydrate to manganese oxide.

**Manganese acetate tetrahydrate**

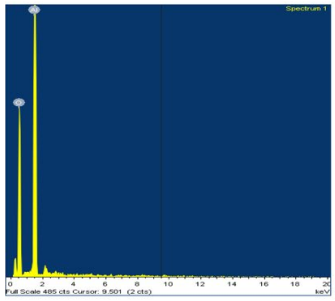
| Element | Weight% | Atomic% |
|---|---|---|
| C K | 2.23 | 22.21 |
| O K | 7.98 | 59.78 |
| Mn K | 8.26 | 18.01 |
| Totals | 18.47 | |

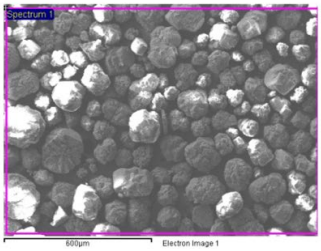



**Alumina**

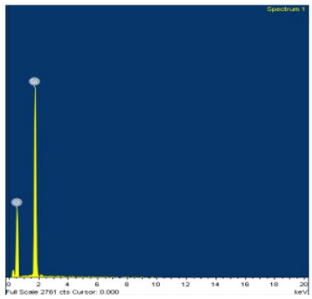
| Element | Weight % | Atomic% |
|---|---|---|
| O K | 11.25 | 72.08 |
| Al K | 7.35 | 27.92 |
| Totals | 18.59 | |




**Silica**

| Element | Weight% | Atomic% |
|---|---|---|
| O K | 22.91 | 65.83 |
| Si K | 20.88 | 34.17 |
| Totals | 43.79 | |



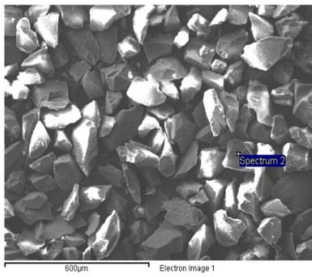
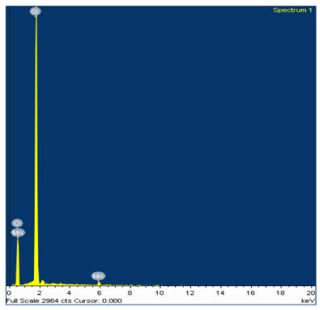
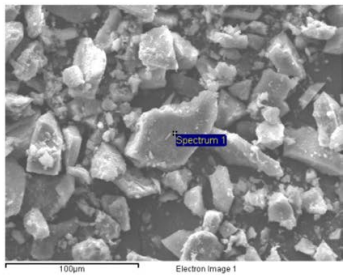




**Fig. 1** EDX images, quantitative analytical data and SEM images of catalysts

**Silica-supported $MnO_2$**

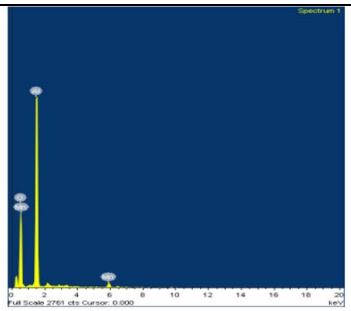
| Element | Weight% | Atomic% |
|---|---|---|
| O K | 25.81 | 66.18 |
| Si K | 22.59 | 33.00 |
| Mn K | 1.09 | 0.81 |
| Totals | 49.48 | |

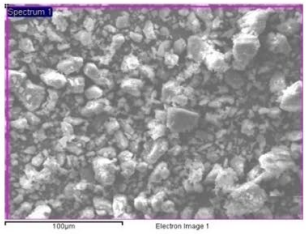


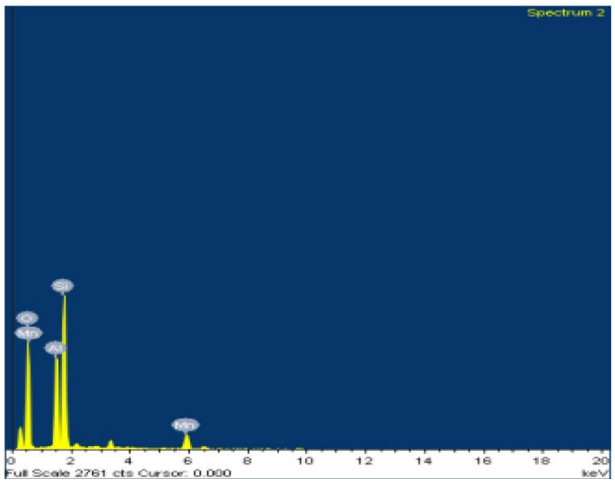

**Alumina-supported $MnO_2$**

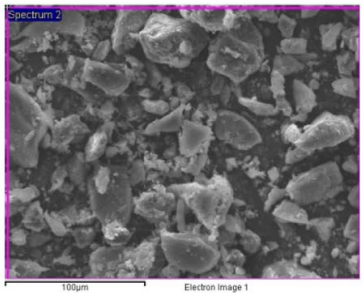
| Element | Weight% | Atomic% |
|---|---|---|
| O K | 26.73 | 73.42 |
| Al K | 15.65 | 25.48 |
| Mn K | 1.37 | 1.09 |
| Totals | 43.74 | |




**Alumina-silica-supported $MnO_2$**




| Element | Weight% | Atomic% |
|---|---|---|
| O K | 25.38 | 74.39 |
| Al K | 4.77 | 8.30 |
| Si K | 8.90 | 14.86 |
| Mn K | 2.87 | 2.45 |
| Totals | 41.92 | |

**Fig. 1** (continued)

The X-ray diffraction (XRD) results for the silica–$MnO_2$, alumina–$MnO_2$ and silica–alumina–$MnO_2$ catalyst are shown in Fig. 2. The features found in these diffractograms showed the presence of manganese oxide, silica and alumina in the catalyst samples and also compared well with the JCPDS data sheet.

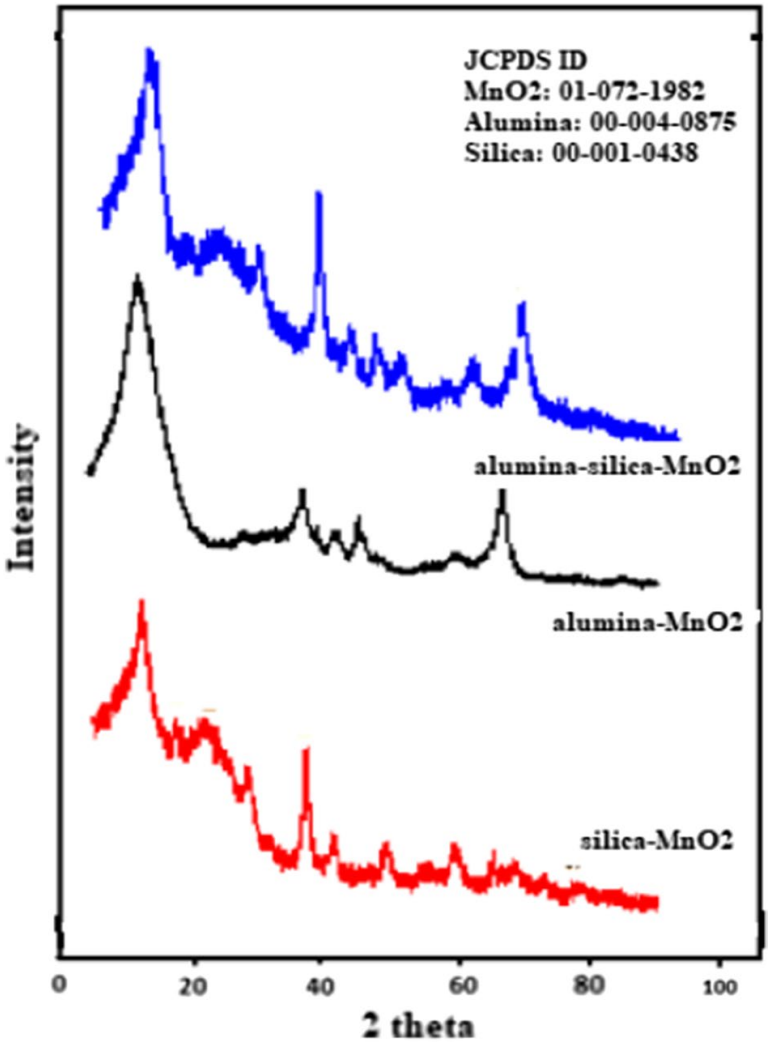
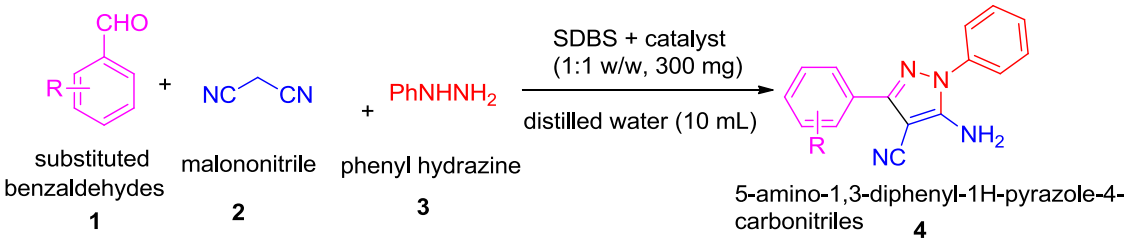



**Fig. 2** XRD results for silica–$MnO_2$, alumina–$MnO_2$ and silica–alumina–$MnO_2$ catalysts

## Synthesis of 5-amino-1,3-diphenyl-1*H*-pyrazole-4-carbonitriles

The use of water as solvents for organic synthesis in order to achieve environmentally friendly methods is in accordance with the principles of green chemistry [47–49]. Our objective was to develop a synthetic method for the synthesis of 5-amino-1,3-diphenyl-1*H*-pyrazole-4-carbonitriles (**4**) in water at room temperature. A reaction was attempted with reactants **1**–**3** (Scheme 1) in water using




**Scheme 1** Synthesis of 5-amino-1,3-diphenyl-1*H*-pyrazole-4-carbonitriles

**Table 1** Catalyst optimization for the synthesis of 5-amino-1,3-diphenyl-1*H*-pyrazole-4-carbonitrile (**4a**)

| S. no. | Catalyst[a] | Isolated yields (%) | S. no. | Catalyst[#] | Isolated yields (%) |
|---|---|---|---|---|---|
| 1 | $Mn(CH_3COO)_2{\cdot}4H_2O$ | 0 | 6 | Alumina–silica–$MnO_2$ | 30 |
| 2 | Alumina (neutral) | 0–5 | 7 | SDBS | 0 |
| 3 | Silica gel (100–200 mesh) | 0–5 | 8 | Alumina–$MnO_2$/SDBS | 50 |
| 4 | Alumina–$MnO_2$ | 25 | 9 | Silica–$MnO_2$/SDBS | 40 |
| 5 | Silica–$MnO_2$ | 20 | 10 | Alumina–silica–$MnO_2$/SDBS | 91 |

[a]Reactants 1 mmol; solvent (water) 10 mL; time 10 min; room temperature (35 °C); catalyst (1:1, w/w, 500 mg each)

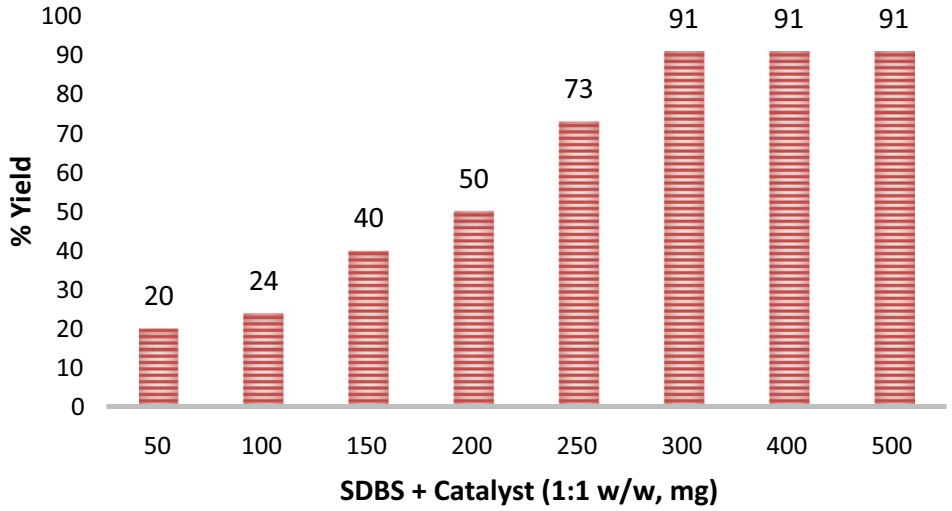



**Fig. 3** Yield optimization with product **4a**, 3 min after phenyl hydrazine (**3**) addition

manganese acetate tetrahydrate, a random selection of metal salt. Even after stirring for 24 h, we did not obtain any product (Table 1, entry 1). The reaction was repeated in silica gel and alumina with similar outcome. We changed the catalyst and prepared silica-supported $MnO_2$, alumina-supported $MnO_2$ and alumina–silica-supported $MnO_2$ [46]. In a typical case, the product **4a** formation showed the order silica-supported $MnO_2$ < alumina-supported $MnO_2$ < alumina–silica-supported $MnO_2$, though the yield was less than 30% (Table 1). We concluded solubility was the major issue. The literature shows the use of surfactant-combined catalysts can cater some of these problems [50, 51]. Hence, we tried the reaction in sodium dodecyl benzenesulfonate (SDBS) in combination with catalysts that gave better results (Table 1, entries 8–10). However, the reaction was also tried in SDBS only without the use of metal oxide to ascertain its role. The reaction did not produce any result in 10 min as shown by the TLC profile where reactant spots appeared unchanged. On prolonging this reaction for 40 min, formation of 10% **4a** was observed.

Then, we optimized the reaction conditions for the product **4a** with respect to yield by varying the amount of surfactant and catalyst (Fig. 3). Since the catalyst

alumina–silica-supported $MnO_2$ gave better results, it was chosen as a suitable catalyst for this reaction. The amount of SDBS and catalyst was optimized by changing their amount from 25 to 250 mg of each (Fig. 3). The reaction that used 150 mg of SDBS and catalyst each gave maximum isolated yield of 91% at room temperature in water.

A broad range of structurally diverse aromatic aldehydes, including electron-releasing and electron-withdrawing substituents, were used to check the prosperity in the reaction, and gave the corresponding products in high yields and short reaction time without the formation of any major by-products (Table 2). The formation of pyrazoles **4** was established on the basis of melting points, Fourier transform IR (FTIR), $^1$H NMR and their comparison with the literature.

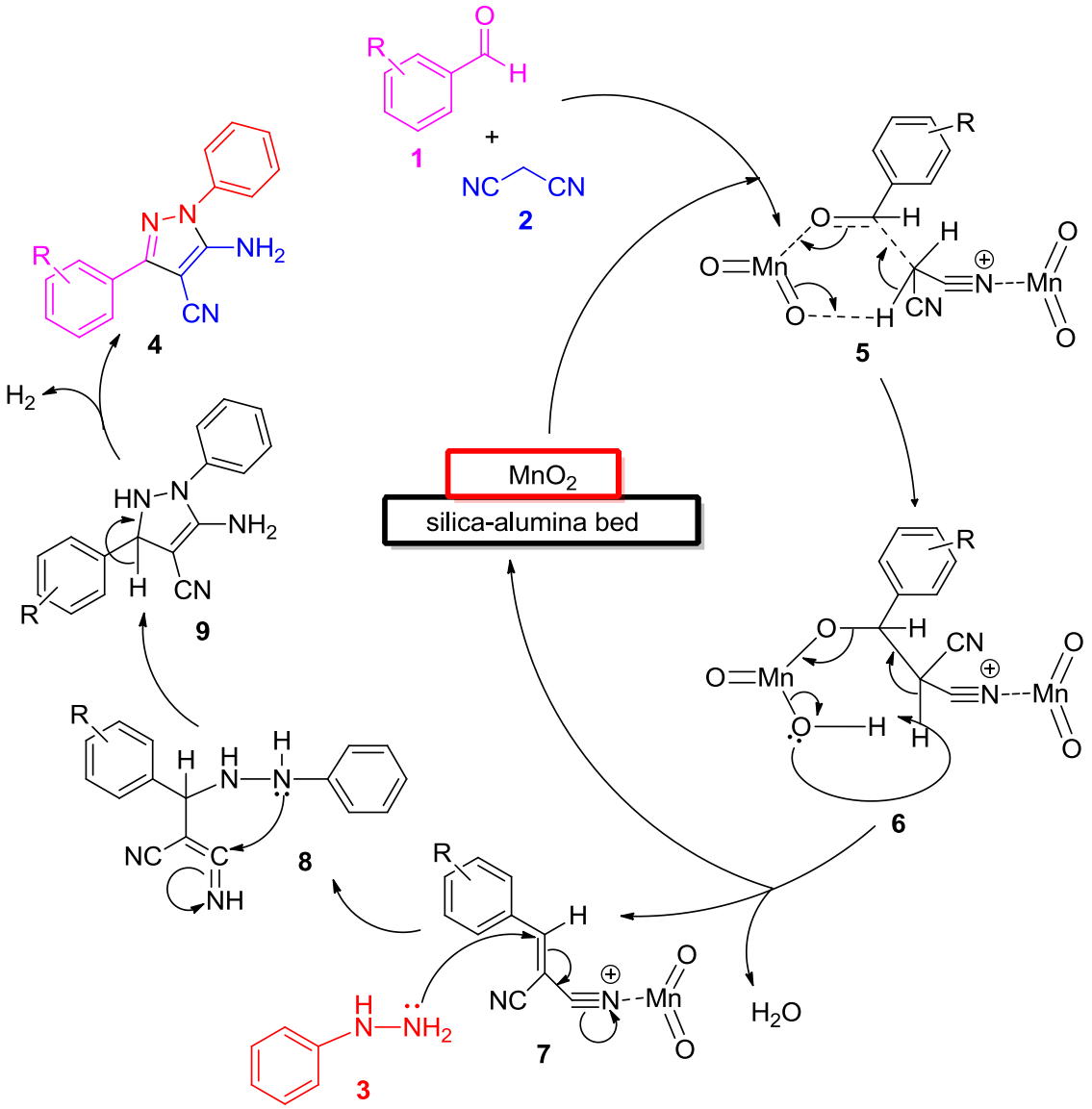
As per our results and reported literatures [6, 28, 54, 55], the product formation takes place through Knoevenagel condensation and Michael addition followed by uncommon hydride transfer and release of molecular hydrogen ($H_2$) [28]. A possible reaction mechanism is given in Scheme 2. The Knoevenagel condensation takes place between the substituted benzaldehyde (**1**) and malononitrile (**2**) which seems to be facilitated through a six-membered ring formation (**5**) using $MnO_2$. The Michael addition takes place between the Knoevenagel product (**7**) and phenyl hydrazine (**3**). The Michael addition product (**8**) undergoes intermolecular cyclization, and then air oxidation to give the final products (**4**).

For the prepared heterogeneous solid catalyst, its recycling status was also evaluated taking the reactants **1a**, **2** and **3** for the synthesis of **4a** under optimized conditions. The catalyst was recovered by simple filtration, and the obtained solid was washed with water (2 × 20 mL). The recovered catalyst was dried for 5 h

**Table 2** Synthesized 5-amino-1,3-diphenyl-1*H*-pyrazole-4-carbonitriles **(4)**

| Comp. no. | Structure of **4** R = | Time (min)[a] | Isolated yields (%) | Melting point (°C) | |
|---|---|---|---|---|---|
| | | | | Found m.p. | Literature. m.p. |
| 4a | H | 3 | 91 | 158–159 | 161 [52] |
| 4b | 4-$NO_2$ | 5 | 94 | 161–162 | 164–165 [53] |
| 4c | 2-$NO_2$ | 7 | 90 | 159 | 160–162 [6] |
| 4d | 2-Br | 7 | 86 | Semi-solid | – |
| 4e | 2-F | 7 | 88 | Oil | – |
| 4f | 4-F | 6 | 93 | Oil | Yellow oil [28] |
| 4g | 4-Br | 7 | 95 | 164–165 | 164–165 [6] |
| 4h | 4-Cl | 6 | 95 | 128–129 | 129 [52] |
| 4i | 2-Cl | 8 | 90 | Semi-solid | Semi-solid [28] |
| 4j | 4-$CH_3$ | 10 | 96 | 117–118 | 120 [52] |
| 4k | 4-$OCH_3$ | 6 | 96 | 107–108 | 110 [52] |
| 4l | 2-$OCH_3$ | 12 | 90 | 130–131 | 130–132 [28] |
| 4m | 3,4-$OCH_3$ | 10 | 94 | 120–122 | 120–123 [28] |
| 4n | 3-$NO_2$ | 11 | 93 | 127–129 | 128–130 [28] |

[a]Time after phenyl hydrazine addition




**Scheme 2** A proposed reaction mechanism for the synthesis of pyrazole derivatives

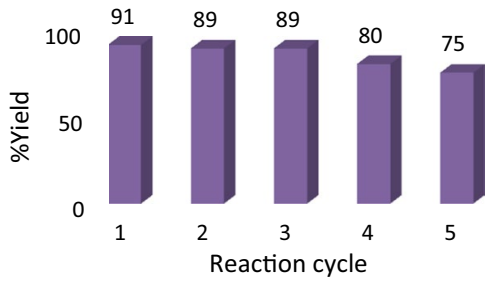



**Fig. 4** Recyclability of the alumina–silica-supported $MnO_2$ for the product **4a**

under vacuum at 70 °C. The performance of the recycled catalyst in reaction up to five successive runs is given in Fig. 4.

## Conclusion

We have successfully prepared heterogeneous solid catalyst alumina–silica-supported $MnO_2$, and its catalytic potential has been examined for the synthesis of highly functionalized pyrazoles. A novel, facile, one-pot, multicomponent protocol for the synthesis of 5-amino-1,3-diphenyl-1*H*-pyrazole-4-carbonitriles has been developed using this catalyst in water and SDBS at room temperature. The cyclocondensation of substituted benzaldehyde, malononitrile and phenyl hydrazine gave the pyrazoles in 86–96% yield. The catalytic activity of the catalyst remained similar till the third cycle with marginal decrease till the fifth cycle.

**Acknowledgements** The authors are grateful to Delhi Technological University for providing necessary facilities to carry out this work. The author Poonam is also thankful for financial support.

## References

1. M. Haji, Beilstein J. Org. Chem. **12**, 1269 (2016)
2. W. Zhao, F.-E. Chen, Curr. Org. Synth. **9**, 873 (2012)
3. A. Hasaninejad, S. Firoozi, Mol. Divers. **17**, 459 (2013)
4. Preeti, K.N. Singh, K.N. Preeti, Org. Biomol. Chem. **16**, 9084 (2018)
5. C.G. Neochoritis, T. Zhao, A. Dömling, Chem. Rev. **119**, 1970 (2019)
6. H. Kiyani, M. Bamdad, Res. Chem. Intermed. **44**, 2761 (2018)
7. M.F. Khan, M.M. Alam, G. Verma, W. Akhtar, M. Akhter, M. Shaquiquzzaman, Eur. J. Med. Chem. **120**, 170 (2016)
8. Z. Wu, S. Wu, Y. Ye, X. Zhou, P. Wang, W. Xue, D. Hu, J. Heterocycl. Chem. **54**, 325 (2017)
9. F.P.L. Lim, G. Luna, A.V. Dolzhenko, Synthesis **48**, 2423 (2016)
10. A. Gürsoya, S. Demirayakb, G. Çapana, K. Erolc, K. Vural, Eur. J. Med. Chem. **35**, 359 (2000)
11. E. Akbas, I. Berber, A. Sener, B. Hasanov, IL Farm. **60**, 23 (2005)
12. A. Tanitame, Y. Oyamada, K. Ofuji, Y. Kyoya, K. Suzuki, H.M. Kawasaki, K. Nagai, M. Wachi, J. Yamagishi, Bioorg. Med. Chem. Lett. **14**, 2857 (2004)
13. A.A. Bekhita, T. Abdel-Aziem, Bioorg. Med. Chem. **12**, 1935 (2004)
14. Z. Özdemir, H.B. Kandilci, B. Gümüsel, U. Calis, A.A. Bilgin, Eur. J. Med. Chem. **42**, 373 (2007)
15. G. Menozzi, L. Mosti, P. Schenone, M. D'Amico, A. Filippelli, F. Rossi, Farmaco **47**, 1495 (1992)
16. M.J. Naim, O. Alam, F. Nawaz, MdJ Alam, P. Alam, J. Pharm. Bioallied. Sci. **8**, 2 (2016)
17. A. Tanitame, Y. Oyamada, K. Ofugi, M. Fujimoto, N. Iwai, Y. Hiyama, K.M. Suzuki, J. Yamagishi, J. Med. Chem. **47**, 3693 (2004)
18. A.M. Badawey, I.M. El-Ashmawey, Eur. J. Med. Chem. **33**, 349 (1998)
19. A.K. Tewari, A. Mishra, Bioorg. Med. Chem. **9**, 715 (2001)
20. A.M. Gilbert, A. Failli, J. Shumsky, Y. Yang, J. Med. Chem. **49**, 1202 (2006)
21. H. Foks, D. Pancechowska-Ksepko, A. Kedzia, Z. Zwolska, M. Janowiec, E. Augustynowicz-Kopec, Il Farm. **60**, 513 (2005)
22. Z. Dardari, M. Lemrani, A. Sebban, A. Bahloul, M. Hassar, S. Kitane, M. Berrada, B. Boudouma, Arch. Pharm. (Weinheim, Ger.) **339**, 291 (2006)
23. P. Rathelot, N. Azas, H. El-Kashef, F. Delmas, Eur. J. Med. Chem. **37**, 671 (2002)
24. A.M.R. Bernardino, A.O. Gomes, K.S. Charret, A.C.C. Freitas, Eur. J. Med. Chem. **41**, 80 (2006)
25. G. Daidone, B. Maggio, S. Plescia, D. Raffa, C. Musiu, C. Milia, G. Perra, M.E. Marongiu, Eur. J. Med. Chem. **33**, 375 (1998)
26. E.C. Taylor, H. Patel, H. Kumar, Tetrahedron **48**, 8089 (1992)

27. S. Palle, J. Vantikommu, R. Redamala, M. Khagga, J. Appl. Chem. **4**, 1190 (2015)
28. M. Srivastava, P. Rai, J. Singh, J. Singh, New J. Chem. **38**, 302 (2014)
29. M. Srivastava, P. Rai, J. Singh, J. Singh, RSC Adv. **3**, 16994 (2013)
30. J. Rakhtashah, S. Salehzadeh, E. Gowdini, F. Maleki, S. Baghery, M.A. Zolfigol, RSC Adv. **6**, 104875 (2016)
31. A. Saha, S. Payra, S. Banerjee, Green Chem. **17**, 2859 (2015)
32. D.Y. Li, X.F. Mao, H.J. Chen, G.R. Chen, P.N. Liu, Org. Lett. **16**, 3476 (2014)
33. A. Saeed, P.A. Channar, J. Heterocycl. Chem. **54**, 780 (2017)
34. A. Kamal, K.N. Visweswara Sastry, D. Chandrasekhar, G.S. Mani, P.R. Adiyala, J.B. Nanubolu, K.K. Singarapu, R.A. Maurya, J. Org. Chem. **80**, 4325 (2015)
35. B. Chen, C. Zhu, Y. Tanga, S. Ma, Chem. Commun. **50**, 7677 (2014)
36. S. Maddila, S. Rana, R. Pagadala, S. Kankala, S. Maddila, S.B. Jonnalagadda, Catal. Commun. **61**, 26 (2015)
37. F.J. Meng, T. Sun, W.Z. Dong, M.H. Li, Z.Z. Tuo, Arch. Pharm. Chem. Life Sci. **349**, 168 (2016)
38. S. Kumari, A. Shekhar, D.D. Pathak, New J. Chem. **40**, 5053 (2016)
39. M. Kashiwa, Y. Kuwata, M. Sonoda, S. Tanimori, Tetrahedron **72**, 304 (2016)
40. C. Ma, P. Wen, J. Li, X. Han, Z. Wu, G. Huang, Adv. Synth. Catal. **358**, 1073 (2016)
41. M. Ubale, M. Shioorkar, Heterocycl. Lett. **6**, 365 (2016)
42. W. Zhang, J. Chen, S. Xu, Y. Chu, Y. Wei, Y. Zhi, J. Huang, A. Zheng, X. Wu, X. Meng, F. Xiao, F. Deng, Z. Liu, ACS Catal. **8**, 10950 (2018)
43. C.E. Pompe, M. Slagter, P.E. de Jongh, K.P. de Jong, J. Catal. **365**, 1 (2018)
44. S.M.S. Chauhan, R. Singh, Geetanjali. Synth. Commun. **33**, 1179 (2003)
45. R. Singh, G. Bhasin, S.R. Geetanjali, J. Chem. Pharm. Res. **6**, 776 (2014)
46. Y. Zhao, G. Chen, D. Wang, C. Song, L. Xie, L. Chang, R. Wang, N. Zhong, Int. J. Electrochem. Sci. **11**, 2525 (2016)
47. R. Mamgain, R. Singh, D.S. Rawat, J. Heterocycl. Chem. **46**, 69 (2009)
48. C.J. Li, Chem. Rev. **105**, 3095 (2005)
49. C.J. Li, T.H. Chan, *Organic reaction in aqueous media* (Wiley, New York, 1997)
50. K. Manabe, Y. Mori, T. Wakabayashi, S. Nagayama, S. Kobayashi, J. Am. Chem. Soc. **122**, 7202 (2000)
51. P.K. Sahu, P.K. Sahuab, D.D. Agarwal, RSC Adv. **4**, 40414 (2014)
52. D. Mishra, R. Singh, C. Rout, J. Chem. Pharm. Res. **9**, 16 (2017)
53. A. Mishra, C. Verma, S. Chauhan, M. Quraishi, E. Ebenso, V. Srivastava, J. Bio Tribo Corros. **4**, 53 (2018)
54. H. Kiyani, M. Bamdad, Heterocycl. **94**, 276 (2017)
55. K. Badhe, V. Dabholkar, S. Kurade, Curr. Organocat. **5**, 3 (2018)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

2019 6th International Conference on Signal Processing and Integrated Networks (SPIN)

# High Isolation EBG Based MIMO Antenna for X-Band Applications

Gaurav Saxena[1,2], Priyanka Jain[2] and Y K Awasthi[3]
[1]*Department of Electronics & Communication Engineering, Galgotia College of Engineering and Technology,*Greater Noida, India
gaurav.saxena@galgotiacollege.edu
[2]*Department of Electronics & Communication Engineering, Delhi Technological University*, Delhi, India
priyajain2000@rediffmail.com
[3]*Department of Electronics & Communication Engineering, Manav Rachna University*, Faridabad HR, India
yogendra@mru.edu.in

***Abstract*—In this article, a high isolation EBG based MIMO antenna is designed to achieve intended characteristics. Proposed antenna is tailored on FR-4 substrate with size of 55×49×1.6mm$^3$. This antenna has circular patches with extruded triangular and circular shapes which provides better radiation efficiency (>0.45) in the X-band (8.0 to 12GHz). High isolation (-22dB) is achieved in entire bandwidth by using mushroom shaped EBG structure near the microstrip feeding line. Diversity performance of antenna is also judged by the various parameters like ECC, Diversity gain, CCL and TARC, which has the practically accepted values <0.016, >9.96dB <0.3bits/sec/Hz and <-35dB correspondingly.**

**Keywords—** ***Electromagnetic band gap (EBG), Multiple input multiple output (MIMO), Envelope correlation coefficient (ECC), Total active reflection coefficient (TARC), Channel capacity loss (CCL).***

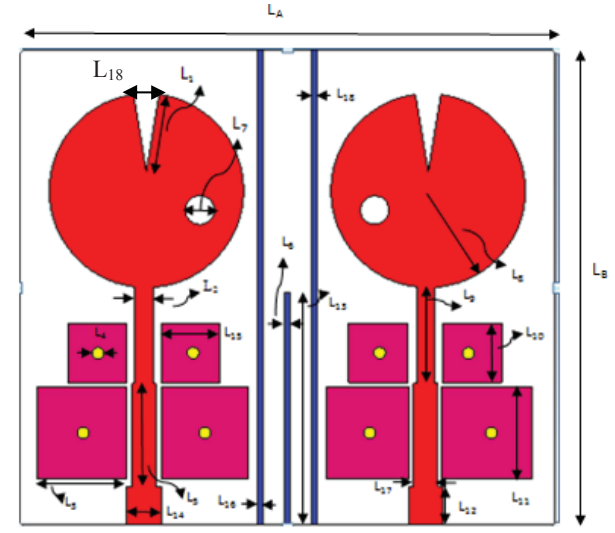
## I. INTRODUCTION

Wireless communication system is keep going different generations from single input single output (SISO), multiple input single output (MISO) system to multiple input multiple output (MIMO) system. In modern era, bandwidth is one of the dominant factors in wireless communication system for high data transmission in digital communication services like i.e. voice, data and video. At receiver end the capacity directly manipulate the quality of the communication systems. MIMO system provides these features in a challenging world to provide high data bit rate, low channel capacity loss (CCL) and high signal to noise ratio (SNR). MIMO antenna widely used in modern wireless RF communication system with improve diversity problem with SISO antenna [1]. MIMO antenna consist spatial diversity and spatial multiplexing capability for remarkable change in reliability and data rate. The antenna has four planar microstrip lines with open ended fed λ/4 slot for excite the fundamental radiating antenna mode over a prescribed frequency range for enhance the isolation between two radiating [2]. Decoupling method is introduced to improve the isolation between two radiating elements [3]. The use of SRR (Split Ring Resonator) to improve the isolation between two radiating patches [4]. Three open ended slots which are used as a decoupling structure to enhance the isolation between antenna elements [5]. U-shaped slot cut and embedded vertical slot to improve the impedance bandwidth [6]. To introduce H-shaped DGS should improve the impedance bandwidth with good isolation of referred antenna [7]. T–shaped decoupling structure in ground plane is extruded which is improved the isolation in the MIMO antenna [8]. Electronic band gap (EBG) and co-design approach structure is used to improve isolation of antenna which offers better diversity performance also [9-10].

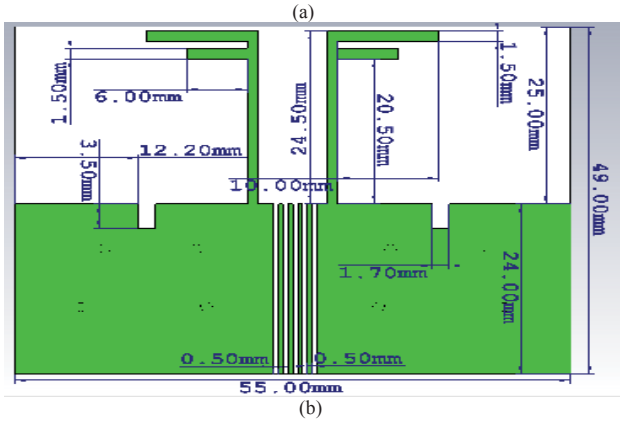
In this paper design EBG (Electronic Band Gap) antenna for isolation improvement in X-band range. Proposed antenna is design on FR-4 substrate with 1.6mm height (h). Section-II shows the antenna analysis and results of proposed MIMO antenna. Section-III shows the diversity performance of proposed MIMO antenna as ECC, DG, MEG, TARC and CCL. A conclusion of the proposed designing work is summarized in section-IV.

## II. ANTENNA ANALYSIS AND RESULTS

Front and bottom of patch antenna is presented in Fig.1 (a) and (b) including all dimensions of top of antenna as mentioned in table-I.




(a)




(b)

**Fig.1.** Proposed UWB-MIMO antenna structure (a) Front view (b) Bottom view.

**Table-I** Dimensions of designed antenna shown in Fig 1 (a).

| $L_1$ (mm) | $L_2$ (mm) | $L_3$ (mm) | $L_4$ (mm) | $L_5$ (mm) | $L_6$ (mm) | $L_7$ (mm) |
|---|---|---|---|---|---|---|
| 8.02 | 1.9 | 9.25 | 1.2 | 10.6 | 0.5 | 3 |

| $L_8$ (mm) | $L_9$ (mm) | $L_{10}$ (mm) | $L_{11}$ (mm) | $L_{12}$ (mm) | $L_{13}$ (mm) | $L_{14}$ (mm) |
|---|---|---|---|---|---|---|
| 10 | 9.93 | 6.1 | 9.5 | 4 | 24 | 3.5 |
| $L_{15}$ (mm) | $L_{16}$ (mm) | $L_{17}$ (mm) | $L_A$ (mm) | $L_B$ (mm) | $L_{18}$ (mm) | |
| 0.5 | 0.5 | 2.5 | 55 | 49 | 2.64 | |

The fundamental resonant frequency of MIMO antenna is given by Equation (1) [11]. Where, $l_1$ and $A_1$ are the length & area of ground plane, $d_2$ and $A_2$ are the diameter & area of radiating patch in that order.

$$f_r(GHz) = \frac{144}{l_1 + d_2 + h + \frac{A_1}{2\pi\sqrt{\epsilon_r + 1}} + \frac{A_2}{2\pi\sqrt{\epsilon_r + 1}}} \quad (1)$$

This MIMO antenna has circular patch with diameter of 19.84mm with four EBG slots near to the step feeding line for better isolation in the requisite bandwidth. F-shaped stub is introduced in ground plane to improve diversity performance as well as isolation. EBG structure is a periodic structure which can proscribe electromagnetic wave in the certain frequency band. Electromagnetic band gap structure is always used to suppress the surface wave to improve the isolation between the radiating patch and EBG is also influence the quality factor (Q) of the MIMO antenna. Therefore, the Q of this antenna is increase to introduce EBG structure; therefore it retains the more energy and produce narrow bandwidth not ultra wideband. Eventually, better isolation and return loss in particular band is achieved due to EBG structures but it is increases conductor loss and the size of antenna is also compact due to electromagnetic band gap.

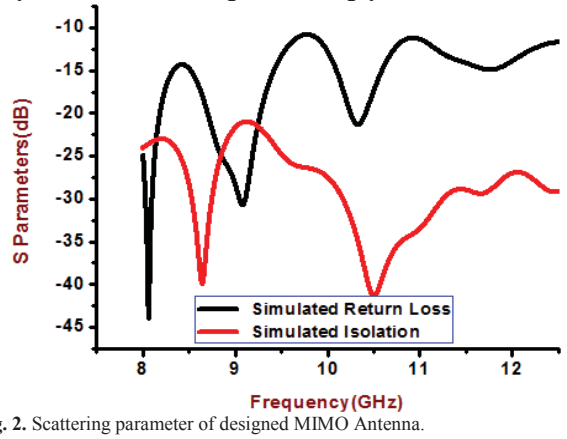
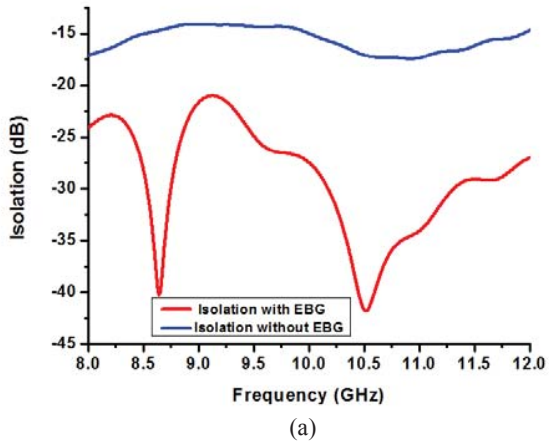
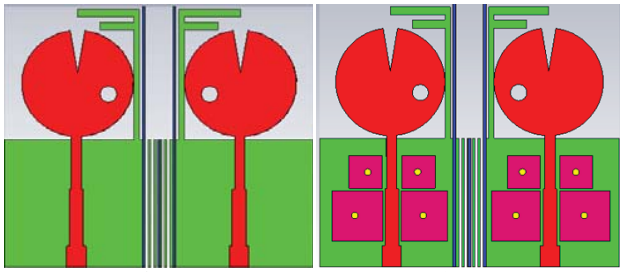



**Fig. 2.** Scattering parameter of designed MIMO Antenna.




(a)

Fig.2 explain about the S-parameter of proposed MIMO radiator, the return loss is less than -10dB in entire the X-band and isolation between two antennas is below than -22dB in entire band, which gives best fit in challenging environment. Fig. 3 (a) shows the comparison of proposed MIMO antenna with and without EBG structure. EBG structure in Fig. 3 (c) provides better isolation ($S_{21}$<-22dB) with comparison of without EBG structure ($S_{21}$<-15dB) as illustrated in Fig. 3 (b).



(b) (c)

**Fig. 3.** (a) Comparison of Isolation with and without EBG (b) MIMO structure without EBG (c) MIMO structure with EBG.

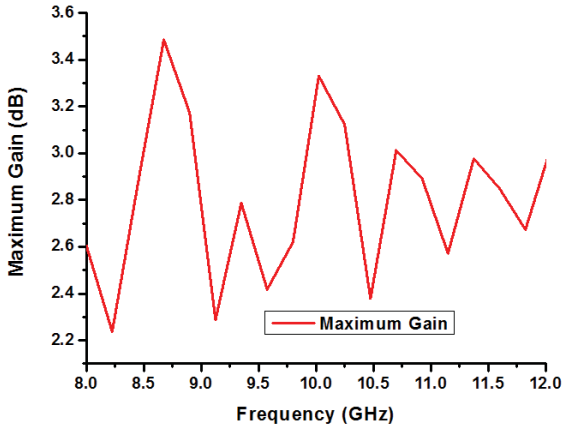
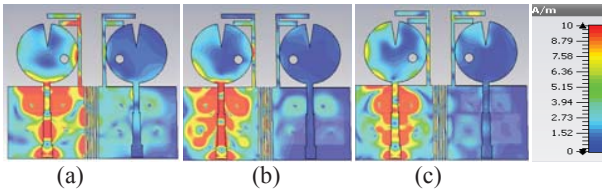



**Fig. 4.** Simulated maximum gain (MG) of MIMO antenna.



(a) (b) (c)

**Fig. 5.** Current density of proposed antenna at (a) 9.5 (b) 10.5 (c) 11.5GHz frequency when one port is energized and other port is terminated by 50Ω impedance.

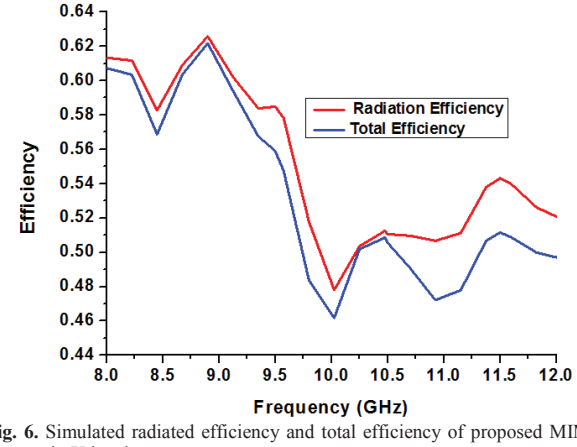



**Fig. 6.** Simulated radiated efficiency and total efficiency of proposed MIMO antenna in X-band.

The peak gain of proposed MIMO antenna is >2.2dB in the

range of 8-12GHz shown in Fig. 4. Gain bandwidth product of antenna is constant if bandwidth is high then gain should be low. So bandwidth is confined in X-band range. Percentage bandwidth of this proposed MIMO antenna is 43.9%. Current density of proposed antenna at 9.5, 10.5 and 11.5GHz frequency is shown in Fig. 5. At this frequency charge not accumulate because current density is low in the radiated patch so most of the power is radiated. Radiation and total efficiency of proposed MIMO antenna is depicted in Fig. 6. Maximum radiation efficiency meets at 8.75GHz near to 62% and minimum radiation efficiency meets on 10GHz frequency near to 47%.

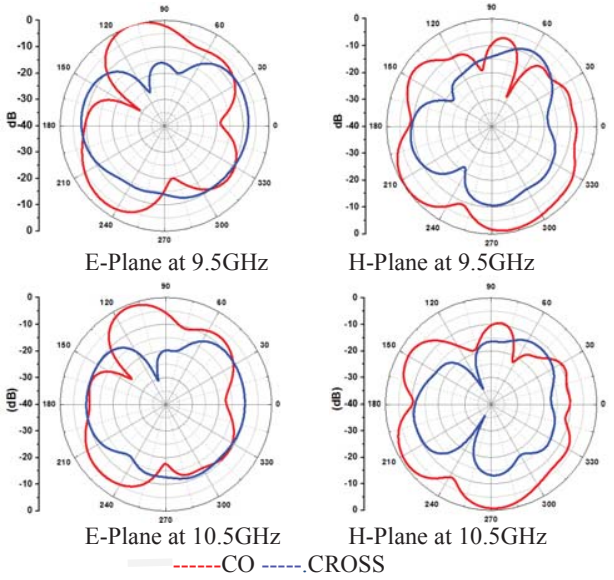



**Fig. 7.** Polar plot of E & H-Plane CO-CROSS at 9.5 and 10.5GHz when one port is excited by external source and another port terminated by 50Ω.

In Fig. 7 shows the 2D radiation pattern of E and H-planes at 9.5 and 10.5GHz frequency. In Co-polarization the radiation pattern is almost omni-direction which radiates the power in all the direction but in the cross-polarization which is orthogonal to co-polarization becomes least radiated power which is depicted in Fig. 7 and the entire graph is stable for best fit in MIMO antenna.

**Table-II.** Comparison table of proposed antenna with existing reference

| Size(mm²) | Isolation (dB) | ECC | Reference |
|---|---|---|---|
| 140×120 | 15 | <0.1 | [2] |
| 52.8×52.8 | 14 | <0.25 | [4] |
| 77×79 | 20 | - | [7] |
| 54.98×76 | 12 | <0.026 | [12] |
| **55×49** | **22** | **<0.016** | **Proposed** |

Above table-II shows that proposed MIMO antenna provides better isolation and ECC above referred articles.

## III. Diversity Performance of MIMO Antenna

Diversity performance of proposed MIMO antenna is judged by ECC "Envelope Correlation Coefficient" and mutual coupling between two antennas. It is determine the correlation factor of one antenna to another antenna and ECC is calculated by equation 2(a) under the assumption that the entire radiated field is uniformly distributed over the lossless antenna [12]. For the satisfactory performance of MIMO antenna, ECC should be <0.1 and the value of proposed MIMO antenna ECC and DG are less than 0.016 and >9.96dB respectively as shown in Fig. 8.

$$ECC = \frac{\left|s^*_{11}s_{12} + s^*_{21}s_{22}\right|^2}{\left(1-\left|s_{11}\right|^2-\left|s_{21}\right|^2\right)\left(1-\left|s_{22}\right|^2-\left|s_{12}\right|^2\right)} \quad 2(a)$$

Lossless antenna is a hypothetical term so we calculate the ECC in terms of radiated electric field by equation 2(b).

$$ECC = \frac{\left|\iint_{4\pi}\left[E_i(\theta,\phi)\times E_j(\theta,\phi)\right]d\Omega\right|^2}{\iint_{4\pi}\left|E_i(\theta,\phi)\right|^2 d\Omega \times \iint_{4\pi}\left|E_j(\theta,\phi)\right|^2 d\Omega} \quad 2(b)$$

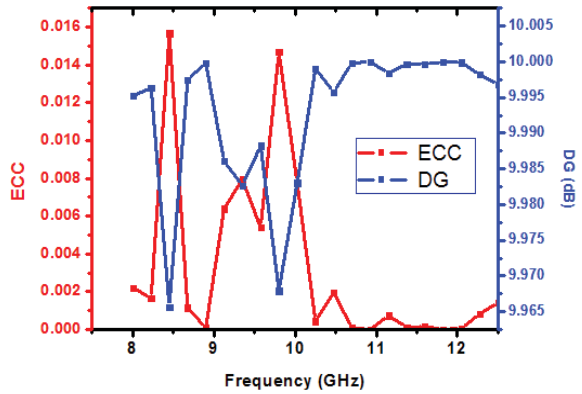



**Fig.8.** ECC and DG of proposed MIMO antenna.

DG (dB) = $10\sqrt{1-(ECC)^2}$, for ideally ECC = 0 and DG value should be 10dB. MEG ratio for proposed antenna lies between ±3dB. Mean effective gain is the "ratio of true mean effective received power ($P_{rec.}$) to the mean effective incoming power ($P_{inc.}$)" to the antenna along the same random route. The MEG can be calculated by $MEG_{i/j} = \varepsilon_{Total}/2$.

$\varepsilon_{Total}$ is the total efficiency associated with $i^{th}$ antenna, and $\varepsilon^i_{Total} = \varepsilon^i_{mis} \times \varepsilon^i_{rad}$, $\varepsilon^i_{mis} = 1 - \sum_{j=1}^{N}\left|S_{ij}\right|^2$, $\varepsilon^i_{rad} = \sum_{j=1}^{N}\left|S_{ij}\right|^2$ where, $\varepsilon^i_{mis}$ implies mismatch efficiency and $\varepsilon^i_{rad}$ radiation efficiency of $i^{th}$ antenna.

Mismatch efficiency can be calculated by following terms. Mean effective gain calculation for outdoor XPR = 0dB and indoor XPR = 6dB at isotropic and Gaussian medium by equation (3) [12] and is shown in Fig. 9. Where XPR = cross-polarization ratio and Ω = beam area of the antenna.

$$MEG_i = \frac{P_{rec.}}{P_{inc.}} = \oint\left[\frac{XPR\times G_{\theta i}(\Omega)+G_{\phi i}(\Omega)\times P_{\phi}(\Omega)}{1+XPR}\right]d\Omega \quad (3)$$

TARC is also an imperative parameter for calculating the MIMO antenna diversity performance. TARC is calculated by in terms of S parameters .TARC is the ratio of square root of total reflected power to the total incident power. It gives the information about apparent return loss of the antenna. The ideal value of TARC should be less than 0dB for MIMO antenna and it is calculated by equation (4) [8]. In this paper, the calculated value of designed antenna is less than -35dB, as depicted in Fig. 9.

$$TARC = \sqrt{\frac{\left(S_{11}+S_{12}\right)^2+\left(S_{22}+S_{21}\right)^2}{2}} \quad (4)$$

At XPR = 0dB (isotropic) and 6dB (indoor) in isotropic medium MEG should be -3.01dB throughout the X-band and -4.01 to -4.18dB for Gaussian medium -3.19 to -2.85dB and -4.8 to -5.18dB respectively, as revealed in Fig. 10.

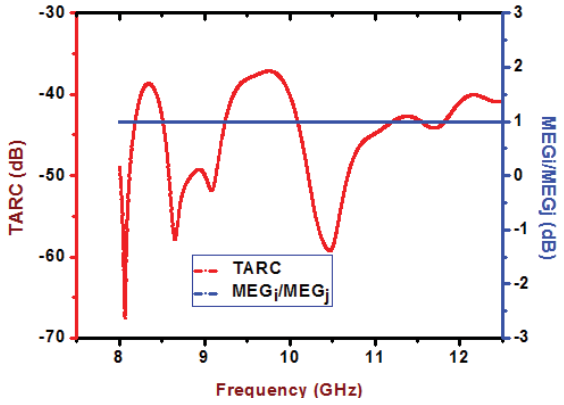



**Fig. 9.** TARC and MEG of proposed MIMO antenna.

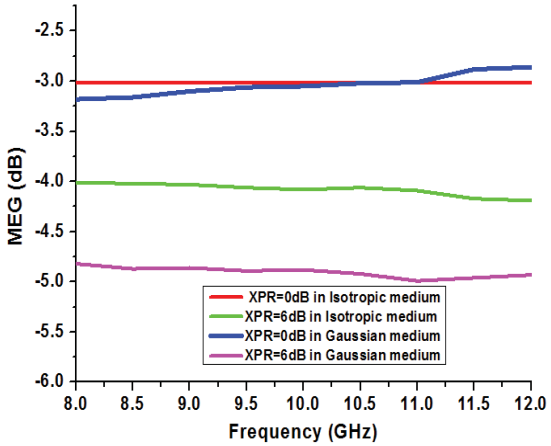



**Fig.10.** MEG calculation at XPR = 0 and 6dB in isotropic and Gaussian medium.

$$\mathrm{CCL(dB)} = -\log_2 \mathrm{Det}\left(\phi^{p}\right) \tag{5}$$

Where, $\phi^{p}$ is the 1×2 correlation matrix in terms of S-parameter $\phi_{ii} = 1 - |S_{ii}|^2 - |S_{ij}|^2$, $\phi_{ij} = -\left(S_{ii} \times S_{ij} + S_{ji} \times S_{jj}\right)$. Channel capacity loss should be determined by equation (5)[8].For MIMO antenna channel bandwidth should be improve but as well as channel capacity loss also increase for 1×2 proposed MIMO antenna, it should be less than 0.30bits/sec/Hz and CCL of proposed MIMO antenna is less than 0.30bits/sec/Hz which is made known in Fig.11.

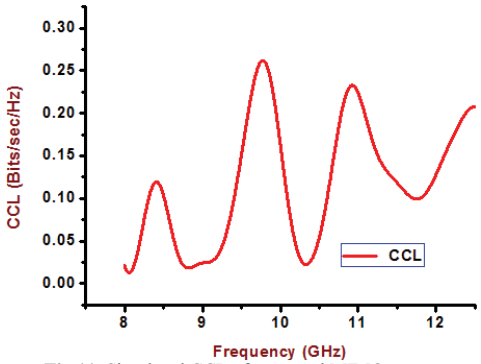



**Fig.11.** Simulated CCL of proposed MIMO antenna

## IV. Conclusion

MIMO antenna has been designed for X-band application with empirically accepted isolation i.e. <-22dB, by using F-shaped decoupling and mushroom shaped EBG structure. The diversity performance has been also verified in terms of ECC, DG, TARC, CCL and MEG, which are in excellent acceptable limit for wireless communication and also provide better channel capacity with low channel capacity loss including high data rate with respect to SISO antenna. Consequently, the designed antenna is high-quality for future communications as for sea based X-band RADAR sbx-1 application.

## References

[1] M. Sonkki, D. Pfeil, V. Hovinen, and K. R. Dandekar, "Wideband Planar Four-Element Linear Antenna Array," IEEE Antennas and Wireless Propagation Letters, vol. 13, pp. 1663-1666, 2014.

[2] A. MoradiKordalivand, T. A. Rahman, and M. Khalily, "Common Elements Wideband MIMO Antenna System for WiFi/LTE Access-Point Applications," IEEE Antennas and Wireless Propagation Letters, vol. 13, pp. 1601-1604, 2014.

[3] S. Soltani and R. D. Murch, "A Compact Planar Printed MIMO Antenna Design," IEEE Transactions on Antennas and Propagation, vol. 63, no. 3, pp. 1140-1149, 2015.

[4] R. Anitha, P. V. Vinesh, K. C. Prakash, P. Mohanan, and K. Vasudevan, "A Compact Quad Element Slotted Ground Wideband Antenna for MIMO Applications," IEEE Transactions on Antennas and Propagation, vol. 64, no. 10, pp. 4550-4553, 2016.

[5] S. Nandi and A. Mohan, "A Compact Dual-Band MIMO Slot Antenna for WLAN Applications," IEEE Antennas and Wireless Propagation Letters, vol. 16, pp. 2457-2460, 2017.

[6] H.-T. Hu, F.-C. Chen, and Q.-X. Chu, "A Wideband U-Shaped Slot Antenna and Its Application in MIMO Terminals," IEEE Antennas and Wireless Propagation Letters, vol. 15, pp. 508-511, 2016.

[7] L. Kong and X. Xu, "A Compact Dual-Band Dual-Polarized Microstrip Antenna Array for MIMO-SAR Applications," IEEE Transactions on Antennas and Propagation, vol. 66, no. 5, pp. 2374-2381, 2018.

[8] R. Chandel, A. K. Gautam, and K. Rambabu, "Design and Packaging of an Eye-Shaped Multiple-Input–Multiple-Output Antenna With High Isolation for Wireless UWB Applications," IEEE Transactions on Components, Packaging and Manufacturing Technology, vol. 8, no. 4, pp. 635-642, 2018.

[9] Q. Li, A. P. Feresidis, M. Mavridou, and P. S. Hall, "Miniaturized Double-Layer EBG Structures for Broadband Mutual Coupling Reduction Between UWB Monopoles," IEEE Transactions on Antennas and Propagation, vol. 63, no. 3, pp. 1168-1171, 2015.

[10] S. K. Dhar, M. S. Sharawi, O. Hammi, and F. M. Ghannouchi, "An Active Integrated Ultra-Wideband MIMO Antenna," IEEE Transactions on Antennas and Propagation, vol. 64, no. 4, pp. 1573-1578, 2016.

[11] N. Malekpour, M. Amin Honarvar, A. Dadgarpur, B. S. Virdee, and T. A. Denidni, "Compact UWB mimo antenna with band-notched characteristic," Microwave and Optical Technology Letters, vol. 59, no. 5, pp. 1037-1041, 2017.

[12] S. Chouhan, D. K. Panda, V. S. Kushwah, and P. K. Mishra, "Octagonal-shaped wideband MIMO antenna for human interface device and S-band application," International Journal of Microwave and Wireless Technologies, pp. 1-10, 2018.

2019 6th International Conference on Signal Processing and Integrated Networks (SPIN)

# *Inverse Document Frequency-Weighted Word2Vec model to recommend apparels*

Priyanka Meel
Information Technology
Delhi Technological University
New Delhi, India
priyankameel@dtu.ac.in

Agniva Goswami
Information Technology
Delhi Technological University
New Delhi, India
agniva03@gmail.com

***Abstract*— with the rapid growth of e-commerce markets the need for recommendation engine and efficient algorithms are becoming the need of the hour for business models of the companies to generate a huge amount of profit. This paper proposes a hybrid algorithm to benefit apparel retailing market which gives the benefits of both, semantics based search and frequency based search. Later this paper compares results of the proposed hybrid algorithm with the other known algorithms used to recommend products.**

***Keywords- e-commerce; recommendation engine; hybrid algorithm; semantics; frequency;***

## I. INTRODUCTION

Nowadays, with the popularity of e-commerce websites, online shopping has become a trend among people. With billions of users accessing the shopping websites every day, there arises a necessity to develop state of the art machine learning algorithms which can make the user buying experience hassle free. Companies such as Amazon invest a lot on research and development to improve the current existing algorithms. One such field of research is product recommendation, either content based or collaborative based.

According to statistics, Product recommendation generates about 35% of revenue of Amazon.com, which ranges to approximately more than 40billion USD in 2016. Internally Amazon uses both content based and collaborative filtering based recommendation; we have used content based (text and image) recommendation here because the collaborative data is very closely guarded by Amazon.

This paper suggests recommendations methods to women apparels. Basically apparel is an item that is worn on the body. Product recommendation makes the buying easier for the customer and also benefits the company as it lures the customer in buying more products. Researchers have developed many content based recommendation algorithms such as Term frequency, Bag of Words, Inverse Document Frequency, Word2Vec which was originally developed by Google in the year 2013, it is on their shoulder that we are standing and applying these algorithms and sometimes modifying to improve our query results accordingly. Some examples where product recommendation is used are, Amazon.com in "people who bought this also bought...", youtube.com suggestions, etc.

RELATED WORK:

Basically in BagofWords ,let's say we have N data and thus N titles, if for each title we can represent it as a D-dimensional vector where D is the total number of words in all the titles, therefore we can apply Euclidean distance method to it to predict similar products. Where each value of the vector denotes the frequency of the word in the particular title.

In tf-idf(term frequency-inverse document frequency), we replace the values with the tf-idf value, tf (Wj,Ti) is Number of times word Wj occurred in Ti divided by Number of words in T, and idf(Wj,D) is equal to log(number of titles in Corpus D divided by number of titles in D that contains word Wj),and finally we multiply the tf value with the idf value.

A recommendation system or engine is a kind of information system that suggests only important results to the user by eliminating redundant information by using different algorithms which are automated or computerized. Recommender systems are used in search engines, social tags, research article recommendation, movies etc.

A recommender system suggests the recommendations by one the two ways:

1. Collaborative filtering
2. Content based filtering.

Content-based Filtering: This approach is common while a recommender system is being designed. Content-based filtering methods are based on a description of the item and a profile of the user's preferences [10]. In content-based recommender system, items and user's preferences are used to recommend items. The algorithm recommends items based on the items the user has seen in the past. The user provides the data to the system either implicitly or explicitly, implicitly for example by clicking a link and explicitly by rating an item, the content based recommender system uses these data to further recommend products. In particular, various possible items are compared with the query item or item that was rated by the user. The recommendation suggestion is updated with every activity of the user, for example when the user gives good rating to a particular product, preference of that product on future recommendations increases.

Collaborative Filtering: in this type of filtering the algorithm that is used, takes into consideration the behaviour of other user and items, in terms of what rating the other user has given to the same item, what closest items other user has preferred to the current item etc. For example if item A is liked by three users and item B is liked by two users, then the recommender system will predict that item A will also be liked by the current user and will be shown in the result

## II. METHODOLOGY

### *A. DATA ACQUISITION, PRE-PROCESSING & CLEANING*

1. Data Acquisition: The data has been acquired from Amazon product Advertisement API available at:

https://docs.aws.amazon.com/AWSECommerceService/latest/DG/Welcome.html [14]. I have acquired over 183000 data about women's tops in apparel section.

2. Data Cleaning: The initial number of data points after acquisitions are 183138, and the total number of features are 19. The 19 features contains features such as 'asin', 'author', 'titile', etc. But among these only a few are useful in uniquely identification of items, so we excluded all others and kept features such as 'asin', 'brand', 'color', 'product_type_name', 'medium_image_url', 'title', 'formatted_price'.

- **Handling missing data of various features:** The data was then checked to see if any null values are present or not. There were null values in features 'formatted_price' and 'color'.

After removing the data rows which has null 'formatted_price', the number of data points were reduced from 183138 to 28395.
Then the null values of 'color' features were handled, after the removal of rows which had null 'color', the data points reduced from 28395 to 28385.

- **Removing Near Duplicate Items:** Some titles had all duplicate words except the last few words.
  The algorithm:
  1. Split every title into words and for each title, check the number of similar words with the previous title and count it.
  2. Subtract the counter from the total number of words.
  3. If it is greater than 2, or simply the number of dissimilar words in the two title is great than 2, then consider it as a new title, otherwise ignore the title.

After de-duping, the number of data points reduced to around 16 thousand.

3. Data Pre-Processing:

Initially stop words have been downloaded from NLTK, such as an', 'down', 'during', 'about', 'was', 'couldn', 'of', 'few', 'my', 'weren', 'off', 'all', 'because', 'had', 'both', 'after', 'having', 'again', 'y' etc have been removed by simply iterating of the words present in the whole document corpus because words like these are not at all useful in recommendation

### *B. THE PROPOSED ALGORITHM:*

The Flowchart:

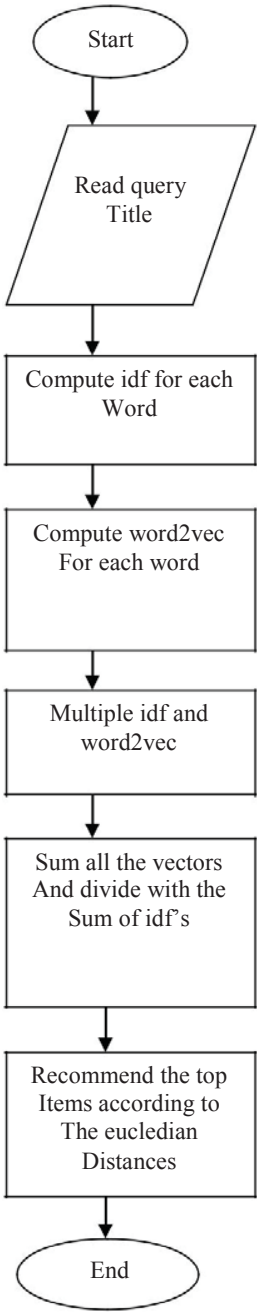



**Fig. 1 flow chart of idf-weighted word2vec algorithm**

IDF-WEIGHTED WORD2VEC MODEL

The similarity of average Word2Vec with IDF weighted word2vec is that, in average word2vec we take IDF supposedly to be 1, but here for each word W(j) in a title T(i), we run Word2Vec and get a 300 dimensional vector and then multiply that vector with that word W(j)'s IDF(inverse document frequency) and then sum up all the corresponding cell of the each vector to get a new vector just like average Word2Vec, and then her we divide each cell of the resultant vector by the total sum of IDF 's of each word W(j) of the title, to get our final vector of each title T(i). The Algorithm also shown in flow chart (Fig 1)**:**

1. **Computing IDF:**
    - IDF is always measured for a given word Wj and a given document corpus D, IDF(Wj,D) is equal to log(number of titles in Corpus D divided by number of titles in D that contains word Wj), thus IDF measure how less frequent a word is.
2. **Computing Word2Vec:**
    - Let T(i) be a title ,which consists of 'k' words, w1,w2,…..,wk, for each word W(j), we get a 300 dimensional vector by simply running Word2Vec model on word W(j).
    - For every title T(i), take up every word W(j)'s 300 dimensional vector which we got from previous step, and multiply it with the IDF of that particular word W(j) which we got from step 1.
    - Now create vector of size 300, where the value of each cell is equal to the sum of the values of all the corresponding cell of all the vectors of each word W(j). Finally we divide each cell of the resultant vector by the total sum of IDF 's of each word W(j) of the title(calculated from step 1), to get our final Vector of each title T(i).

We can apply similar approach to make TF-IDF weighted Word2Vec, where we just multiply each word W(j)'s vector with its TF-IDF and the rest of the procedure is same.

III.ANALYSIS AND RESULTS

The results of Bag of Words and TF-IDF show that these algorithms are purely based on the frequency of words in the title and in the whole data corpus, whereas algorithms like average word2vec lays stress on semantic based recommendations. Thus when these two types of algorithms are merged, we get results which show products based on semantics as well as based on the texts present in the title.

In the results of IDF-Weighted Word2Vec, the two unique results we got are :

1. Anna kaci woman's asymmetrical sheer brown leopard cheetah print long sleeve top multicoloured medium.
2. Merona woman's printed blouse brown leopard print xxl

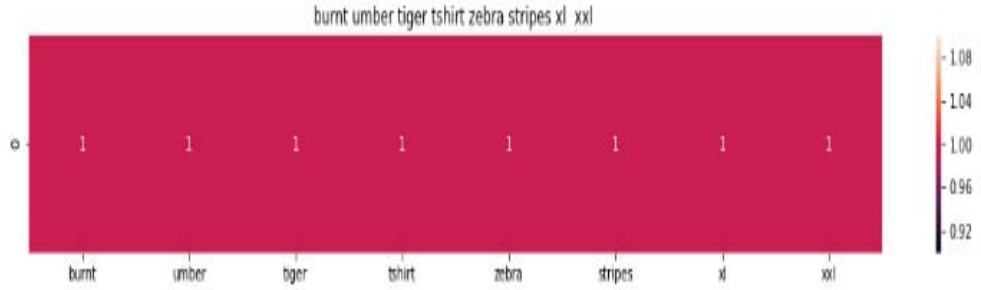
In both these results we see the word 'leopard', we get such type of recommendation because of Word2Vec being semantic based; leopard is semantically similar to tiger, which is contained in our query image. Also here, we get results based on the similarity of texts of our query results like 'women's' etc.

Thus this algorithm has computational complexity same as Average Word2Vec but has better results than Average Word2Vec because of the inclusion of both semantics and frequency of words.

**Bag of Words results:** As this is a frequency based technique so the results shown below (Fig. 2) shows that images which has more common words to that of the query title has lesser eucledian distance, hence given more preference in search results, for example in Fig.2.1 "pink tiger shirt zebra stripes xl xxl" has most number of common words hence lesser eucledian distance.

**Tf-Idf results:** It is also a frequency based algorithm, so the results are similar to BagofWords algorithm, but gives better results than BagofWords, because it also takes into account the presence of the word in the whole data corpus, thus it gives results shown in (Fig. 3)

**Idf Weighted Word2Vec Results:** In addition to the results of BagofWords and tf-idf, in idf weighted Word2Vec we also got the results shown below in (Fig. 4),which were not given by the above two algorithms at all because it combines both semantic based results as well as frequency based results. It gives results based on the fact that a word occurs many times in a title and very fewer titles in the data corpus should contain that particular word.





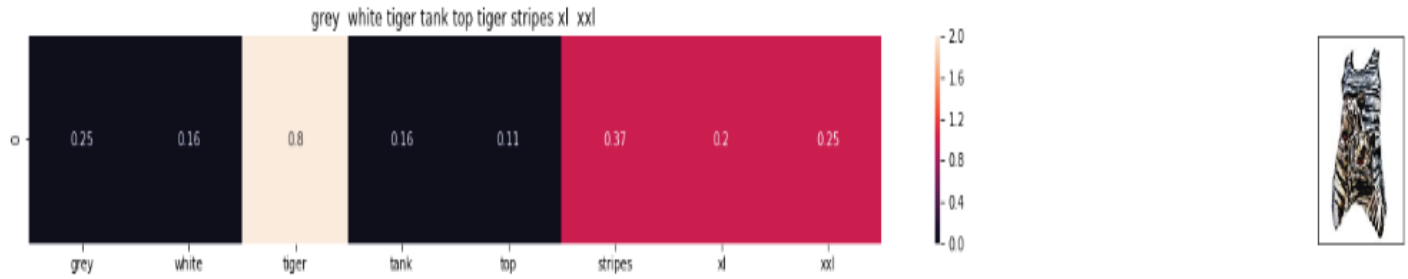

ASIN : B00JXQB5FQ
Brand: Si Row
Title: burnt umber tiger tshirt zebra stripes xl xxl
Euclidean similarity with the query image : 0.0
==========





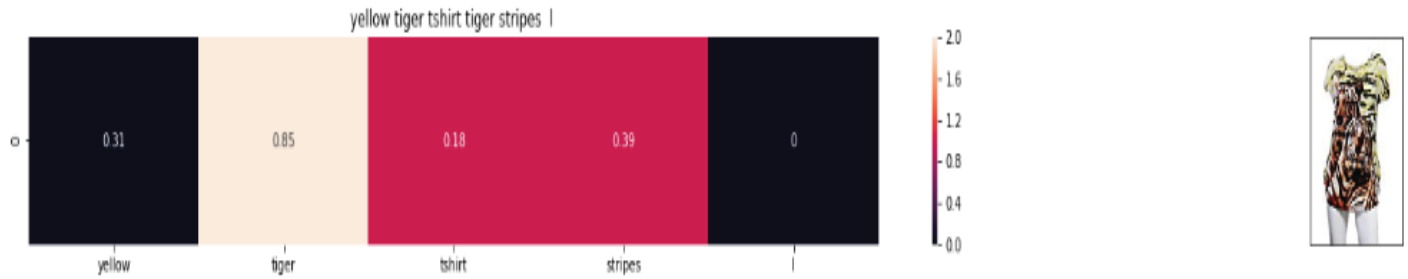

ASIN : B00JXQASS6
Brand: Si Row
Title: pink tiger tshirt zebra stripes xl xxl
Euclidean similarity with the query image : 1.73205080757

**Fig.2 BagofWords results**




ASIN : B00JXQAFZ2
BRAND : Si Row
Eucliden distance from the given image : 0.95861535242

==========




ASIN : B00JXQCUIC
BRAND : Si Row
Eucliden distance from the given image : 1.00007496145

**Fig.3 tf-idf results**

## *IDF-Weighted Word2Vec Results:*

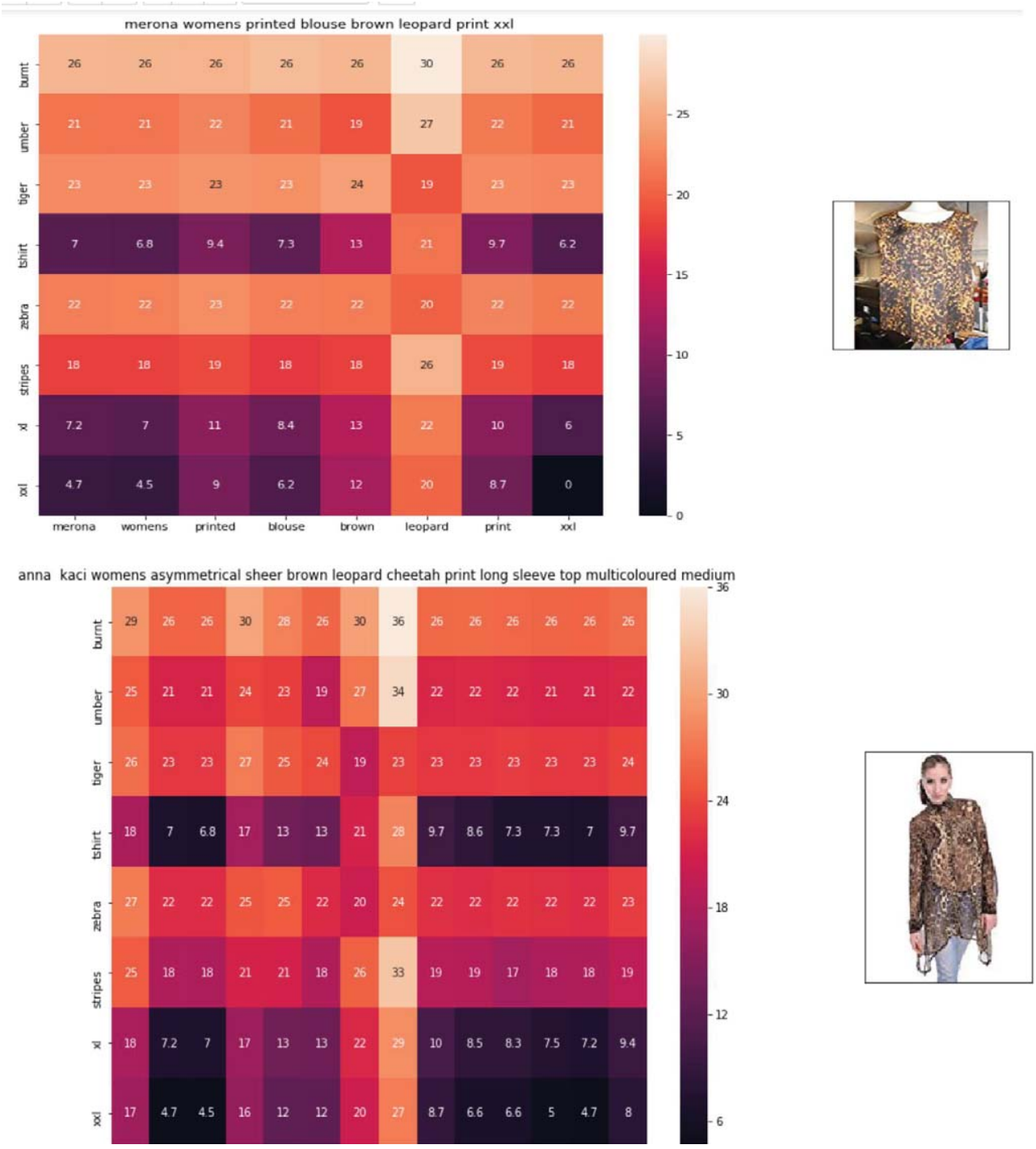



**Fig.4 idf-weighted word2vec results**

## IV. CONCLUSION AND FUTURE WORK

The results above shows the benefits of using idf-weighted word2vec algorithm over BagofWords, tf-idf. Because of the semantic based results it gives better predictions of the user's future needs and frequency based results gives predictions based on the users query product. More work can be done in this area, on addition to idf weighted word2vec we can increase the efficiency by reading the image and classifying the image property and then give results which are similar to the query image.

## REFERENCES


[1] R. C. Bagher, H. Hassanpour, and H. Mashayekhi, "User trends modeling for a content-based recommender system," *Expert Syst. Appl.*, vol. 87, pp. 209–219, 2017.

[2] M. S. Tajbakhsh and J. Bagherzadeh, "Microblogging hash tag recommendation system based on semantic TF-IDF: Twitter use case," *Proc. - 2016 4th Int. Conf. Futur. Internet Things Cloud Work. W-FiCloud 2016*, pp. 252–257, 2016.

[3] G. Carullo, A. Castiglione, and A. De Santis, "Friendship recommendations in online social networks," *Proc. - 2014 Int. Conf. Intell. Netw. Collab. Syst. IEEE INCoS 2014*, pp. 42–48, 2014.

[4] J. Hannon, M. Bennett, and B. Smyth, "Recommending twitter users to follow using content and collaborative filtering approaches," *Proc. fourth ACM Conf. Recomm. Syst. - RecSys '10*, p. 199, 2010.

[5] M. Razghandi and S. A. H. Golpaygani, "A Context-Aware and User Behavior-Based Recommender System with Regarding Social Network Analysis," *2017 IEEE 14th Int. Conf. E-bus. Eng.*, pp. 208–213, 2017

[6] S. Aral and D. Walker, "Identifying Influential and Susceptible Members of Social Networks," *Science,* vol. 337, no. 6092, pp. 337-341, June 2012.

[7] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 6, pp. 734–749, 2005.

[8] P. Alencar and D. Cowan, "The use of machine learning algorithms in recommender systems : A systematic review," *Expert Syst. Appl.*, vol. 97, pp. 205–227, 2018.

[9] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a Social Network or a News Media?", *Int. World Wide Web Conf. Comm.*, pp. 1–10, 2010.

[10] Aggarwal, Charu C. (2016). *Recommender Systems: The Textbook*. Springer-ISBN *9783319296579.*

[11] Yoav Goldberg and Omer Levy, "word2vec Explained: Deriving Mikolov et al.'s Negative-Sampling Word-Embedding Method", arXiv:1402.3722v1, Feb 2014.

[12] Appliedaicourse.com

[13] https://deeplearning4j.org

[14] https://docs.aws.amazon.com/AWSECommerceService/latest/DG/Welcome.html

Analog Integrated Circuits and Signal Processing
https://doi.org/10.1007/s10470-019-01466-2

MIXED SIGNAL LETTER

# LECTOR incorporated differential cascode voltage swing logic (L-DCVSL)

Pratibha Bajpai[1] · Neeta Pandey[1] · Kirti Gupta[2] · Jeebananda Panda[1]

Received: 30 November 2016 / Revised: 19 March 2019 / Accepted: 2 May 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

**Abstract**
This contribution aims at incorporating leakage control transistors (LCTs) in differential cascode voltage swing logic (DCVSL) to reduce leakage and is named as L-DCVSL. The concept multi threshold is also introduced in LCTs and resulting static, dynamic and enhanced DCVSL variants are referred as MTL-DCVSL, MTL-dyDCVSL and MTL-EDCVSL. The performance at 90 nm, 65 nm, and 45 nm technology nodes is investigated using SymicaDE tool. Two input exclusive OR/NOR (XOR2/XNOR2) gate is used to illustrate the proposed technique due to its extensive use in arithmetic cores. The leakage current is examined in all the examples and it is found that leakage current decreases with scaling down of geometry. Maximum saving in leakage current is observed to be 81.27% while minimum is 44.18% for MTL-DCVSL. Similar observations for MTL-dyDCVSL and MTL-EDCVSL are (12.24%, 9.52%)/(33%, 66.04%) and (16.06%, 61.67%)/(37.11%, 78.45%) respectively for precharge/evaluate state. Effect of temperature is also investigated and it is found that leakage current in both static and dynamic configurations of DCVSL follows a directly proportional behaviour with respect to temperature.

**Keywords** Differential cascode voltage switch logic · Dynamic differential cascode voltage switch logic · Enhanced dynamic differential cascode voltage switch logic · LECTOR

## 1 Introduction

Scaling down of device geometries has profound impact on the performance of systems constructed with these devices. This, however, leads to substantial increase in leakage current and affects the total power consumption. Leakage power is classified as active leakage and stand by leakage power. The former type exists when the processor is doing computation while the later one is concerned with the leakage when processor clock is turned off. It is important to mention here that even in nominal operating conditions, leakage contributes to 30–50% of the overall power consumption [1, 2]. Besides increased power consumption, leakage also reduced the available design margins as process variation influences leakage power. It is, therefore, highly desirable to to devise effective leakage control methodologies to make devices power efficient in deep submicron and nanometer technologies. This has motivated the researchers to explore leakage control techniques and various such techniques are available for conventional CMOS based circuits [3–8]. The DCVSL is yet another style that combines the advantages of both conventional CMOS logic and pseudo NMOS logic and thereby provides a high speed, area efficient and rail to rail swing logic design alternative [9]. The logic is implemented using only NMOS logic tree which results in reduction of parasitic capacitance and hence improved speed. The DCVSL allows sharing of the common transistors in the logic network for both the differential outputs resulting in compact structure which provides significant saving of area as compared to conventional CMOS logic. This version is referred as static DCVSL and its clocked versions are termed as dynamic DCVSL (dyDCVSL) [10] and enhanced differential cascode voltage switch logic (EDCVSL) [10].

✉ Neeta Pandey
n66pandey@rediffmail.com

[1] Department of Electronics and Communication Engineering, Delhi Technological University, Rohini, Delhi, India

[2] Department of Electronics and Communication Engineering, Bharati Vidyapeeth's College of Engineering, New Delhi, Delhi, India

Published online: 24 May 2019

The NMOS logic tree is greatly simplified in EDCVSL [10] leading to significant reduction in the interconnect count. Study on leakage power aspect of DCVSL family is somewhat limited and is open for exploration.

Three techniques are available for handling leakage current in DCVSL. The first technique [11] uses multi threshold transistors in pull down network (PDN) while the second is based on employing sleep transistors. The name MT-DCVSL is given to first technique and it requires criticality of threshold voltage selection for the transistors in PDN. Second technique [12] is known as sleep embedded technique and requires external control circuitry to generate sleep signal. The third technique uses self-controlled leakage power reduction technique (LECTOR) [13] as well as reduces stacked levels in PDN by implementing the logic using transmission gates of DCVSL gates [14]. Alternatively, it is well known that stacking in PDN results in leakage reduction, therefore, the authors applied LECTOR in DCVSL wherein stacked levels are used in PDN. Two input XOR gates which are an essential fundamental building unit in a lot of VLSI applications like adders and microprocessors are simulated and verified for illustrating effectiveness of the proposed configurations.

The organisation of rest of the paper is given as follows: Sect. 2 describes briefly the architecture and operation of DCVSL. Available leakage power reduction techniques in DCVSL and their limitations are also included in this section. LECTOR incorporated DCVSL is proposed in Sect. 3 and is termed as L-DCVSL. The concept of multi threshold is introduced in leakage control transistors (LCTs) employed in LECTOR technique and resulting static, dyDCVSL and EDCVSL circuits are referred as MTL-DCVSL, MTL-dyDCVSL and MTL-EDCVSL circuits respectively. Section 4 presents simulation results for all proposed configurations by taking two input XOR/XNOR (XOR2/XNOR2) gates as examples. The discussion on results is placed to substantiate efficacy. Paper is finally concluded in Sect. 5.

## 2 DCVSL and leakage control techniques

This section briefly describes the architecture of DCVSL and available leakage control techniques used therein.

### 2.1 DCVSL architectures

#### 2.1.1 Static DCVSL

Generic architecture of static DCVSL is shown in Fig. 1(a) [9] which consists of a pull up circuitry in the form back to back connected PMOS transistors which form a positive feedback latch and a pull down circuitry in the form of NMOS logic tree. Logic is implemented using pull down transistors (PDN1 and PDN2) and pull up transistors (PUTs) help in attaining correct logic state by the differential branches. A DCVSL based XOR2/XNOR2 gate is depicted in Fig. 1(b).

#### 2.1.2 Dynamic DCVSL

Generic structure of dynamic DCVSL (dyDCVSL) is shown in Fig. 2. The structure consists of pull down networks (PDN1 and PDN2) to implement the logic, pull up transistors (M3, M4) which assist in attaining the correct logic state and pre-charge transistors (M1, M2) to pre-charge both the output nodes to logic high before evaluation phase. Transistors M5 and M6 between PUTs and PDNs here work as accelerators. Evaluation transistor M8 driven by clock turns OFF in pre-charge phase and turns ON in evaluation phase. Operation of dyDCVSL is detailed in [10] and omitted from here for the sake of brevity.

### 2.2 Leakage control techniques

To control leakage currents in static DCVSL circuits two techniques have been presented in literature [11, 12]. First technique uses both high and low threshold ($V_{TH}$) devices in DCVSL circuit and is called multi-threshold DCVSL (MT-DCVSL) [11]; low $V_{TH}$ devices improve performance whereas high $V_{TH}$ devices work on leakage reduction. Generic static MT-DCVSL circuit is shown in Fig. 3. As proper selection of threshold of device is critical for performance, MT-DCVSL uses algorithm for this purpose. It may be emphasized here that this would be circuit dependent.

Another technique proposed in the concerned domain is Sleep-Embedded DCVSL [12] circuit and its generic architecture is shown in Fig. 4. This circuit operates in two modes; (i) Active mode where sleep signal is logic low which turns off the complete sleep circuitry and the logic circuit undergoes its normal operation (ii) Sleep mode where sleep signal is logic high which turns on the transistors (P1, P2, N1, N2) and turns off the high VTH transistors (PHV1, PHV2). Transistors P1 and N1 in ON state drive nodes X1 and X2 to voltage VDD and ground respectively. Since (VDD and X1) and (X2 and ground) are equipotential now, voltage balancing is achieved across PUN and PDN so that no leakage current flows through these paths. The only possible path for leakage left is through sleep transistor PHV1. Since PHV1 is high threshold (high threshold transistors are denoted by broad channel line) and turned off during sleep mode, the direct path between power supply and ground is cut off which minimizes leakage significantly.

Comparison of existing leakage control techniques in static DCVSL reveals that circuit specifies optimization is

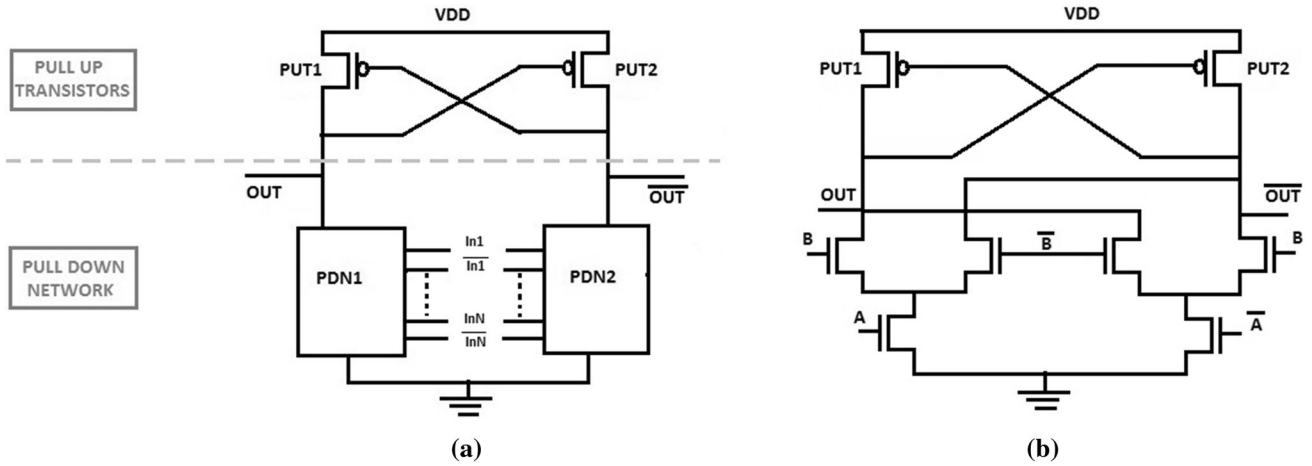



**Fig. 1** **a** Generic architecture of static DCVSL [9], **b** DCVSL XOR2/XNOR2 gate

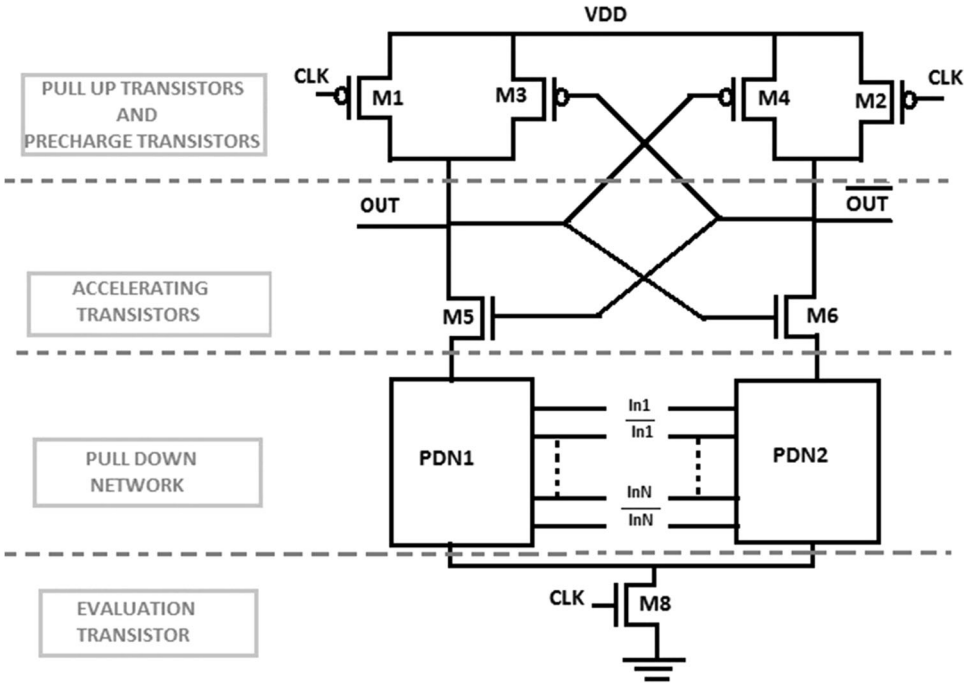



**Fig. 2** Generic architecture of dyDCVSL [10]

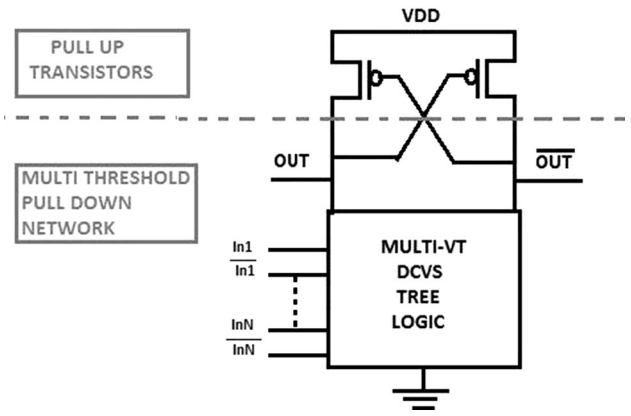



**Fig. 3** Generic architecture static MT-DCVS circuit [11]

needed in MT-DCVSL while sleep based technique does not require any such algorithm. The later technique, however, suffers serious limitations like additional signal requirement to drive sleep transistors, loss of actual state in sleep mode, high wake up time, energy requirement to recharge lost state and area overhead [12]. Further, it may be noted that both MT-DCVSL and sleep based techniques exist in context of static CMOS. LECTOR [13] is yet another technique which is introduced for leakage control in CMOS circuits. It is a self-controlled technique and works well in both active and standby mode. It retains the actual output state in standby mode. Considering the advantages of LECTOR technique, this work aims at adapting this scheme in DCVSL circuits.

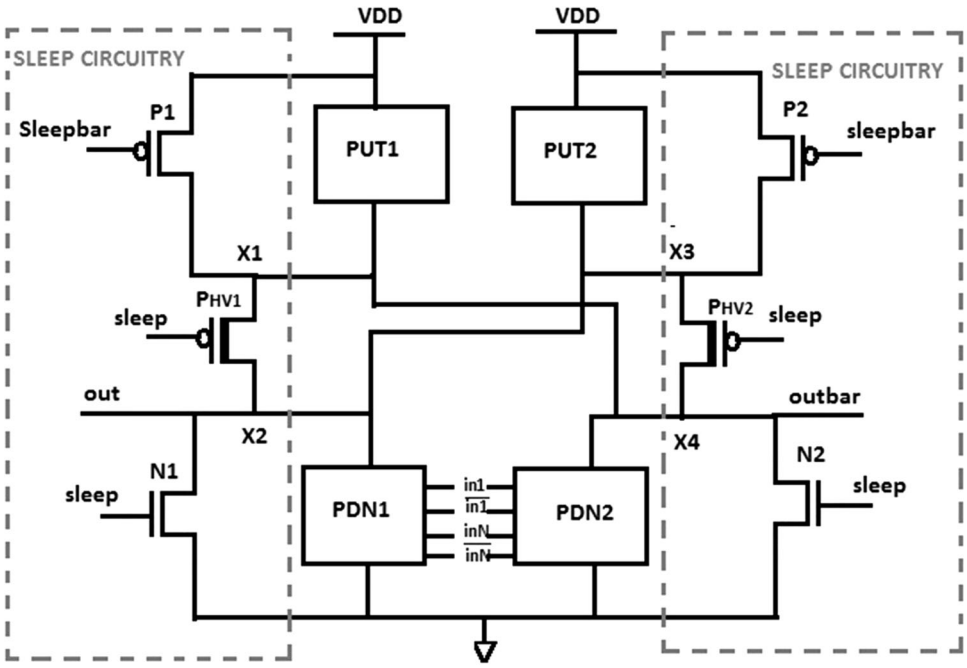



**Fig. 4** Generic architecture of sleep embedded DCVSL circuit [12]

## 3 Proposed configurations for leakage power reduction in DCVSL

In this section, LECTOR technique is incorporated static DCVSL is proposed and is termed as L-DCVSL. The concept of multithreshold is introduced in leakage control transistors (LCTs) employed in L-DCVSL and resulting static, dyDCVSL and EDCVSL circuits are referred as MTL-DCVSL, MTL-dyDCVSL and MTL-EDCVSL circuits respectively.

### 3.1 Proposed LECTOR incorporated DCVSL circuit

#### 3.1.1 Architecture

Generic architecture of proposed L-DCVSL circuit is shown in Fig. 5. The proposed leakage controlled network is akin to DCVSL of Fig. 1 except the insertion of two PMOS and NMOS leakage control transistors (LCTs) in both the true (PUT1 and PDN1) and complementary (PUT2 and PDN2) logic branches as leakage control circuitry.

As high $V_{TH}$ for LCTs usage culminates into achieving further leakage saving therefore configurations proposed in remaining part of this paper employ these transistors.

#### 3.1.2 Leakage control mechanism in proposed L-DCVSL circuit

Concept behind the proposed L-DCVSL architecture of Fig. 5 is as follows: One of the two LCTs in the path from VDD to ground in either of the differential branches is in near cut-off state for any input combination applied. This increases the overall resistance offered to the current flow along the path and thereby reducing the leakage current through the path. One of the most significant advantage of this technique is that it works well in both active and standby mode. Thus in active mode the technique very efficiently cuts down the static current through the circuit too. Static current flows through the branch when the input combination applied turns on the PDN causing output to discharge and transition period is over after the stabilisation of inputs and output.

Leakage loss occurs in the time duration when the output is stable i.e. output low or output high. Undesired current flowing when output is at logic high is called leakage current and the unwanted current flowing when the output is stable at logic low is called static current. As PDN1 and PDN2 are mutually exclusive so either PDN1 or PDN2 remains OFF and other is ON. Here the OFF branch would experience leakage loss due to subthreshold leakage current offered by OFF transistors. Incorporation of leakage control transistors (LCTs) diminishes leakage current as well as static current. To understand leakage control mechanism, operating regions of LCTs need be examined. Let's consider the case where $OUT$ and $\overline{OUT}$ are at '1' and '0' respectively and the input combination is such that PDN 1 is OFF and provides no discharge from Out. By analysing the voltages at intermediate nodes (N1, N2 and N3), it can be inferred that LCT2 is ON and LCT1 is in near cut off state. Subsequently, it reduces the leakage current by offering more impedance along the path. Similarly, the case of static current reduction can be analysed

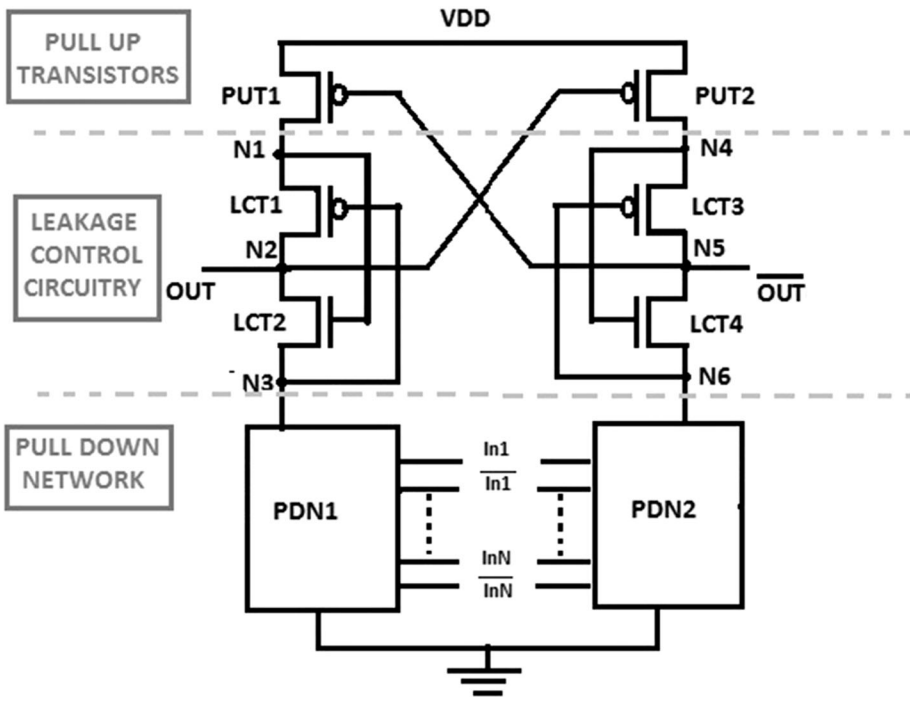



**Fig. 5** Generic architecture of proposed L-DCVSL

when the given input combination is such that it turns on the PDN1. In this case LCT1 comes out to be in ON state and LCT2 in near cut off state contributing in reduction of unnecessary power loss because of static currents. Hereby it can be comprehended that proposed L-DCVSL circuitry proves to be an effective approach to regulate unnecessary power loss in both active and standby modes. For the comparison purpose, inverters are implemented in DCVSL and L-DCVSL styles. The circuit diagrams are drawn in Fig. 6(a, b). The circuits are simulated with $V_{DD}$ = 1.2 V at 90 nm technology node. Their voltage transfer characteristics (VTC) as well as the switching current has been plotted in Fig. 6(c), (d). The values of five critical points namely maximum output voltage $V_{OH}$, minimum output voltage $V_{OL}$, maximum input voltage which can be interpreted as logic '0' $V_{IL}$, minimum input voltage which can be interpreted as logic '1' $V_{IH}$, threshold voltage $V_{TH}$ are noted for both the inverters. For DCVSL inverter, the observed values of $V_{OH}$, $V_{OL}$, $V_{IH}$, $V_{IL}$ and $V_{TH}$ are 1.2 V, 0.325 mV, 306 mV, 351 mV and 336 mV respectively. Analogously, the respective values of 1.2 V, 0.32 mV, 232 mV, 349 mV and 336 mV are noted for $V_{OH}$, $V_{OL}$, $V_{IH}$, $V_{IL}$ and $V_{TH}$ parameters of a L-DCVSL inverter. The two noise margins defined for low logic levels $NM_L$ (= $V_{IL} - V_{OL}$) and high logic levels $NM_H$ (= $V_{OH} - V_{IH}$) are evaluated. The $NM_H$ values for DCVSL and L-DCVSL inverters are 849 mV and 851 mV respectively. The obtained vales are relatively same for both the inverters. Similarly, the respective $NM_L$ values for DCVSL and L-DCVSL inverter are 305.7 mV and 231 mV. The values signify a reduction in $NM_L$ of L-DCVSL inverter in comparison to DCVSL inverter. It can be observed that both the inverters maintain the same switching threshold voltage. Also, while switching between the logic levels, the L-DCVSL based inverters results in low switching currents. Their state matrix has also been evaluated and listed in Table 1. Since the circuit symmetric the transistors for the half circuit are considered for evaluating the matrix.

### 3.2 Proposed LECTOR incorporated multi threshold DCVSL configuration

Here, a variant of Fig. 5 is proposed wherein concept of Multi Threshold transistor is introduced in LCTs. The resultant structure is shown in Fig. 7 and is named as LECTOR incorporated Multi Threshold DCVSL (MTL-DCVSL). It is well known that high $V_{TH}$ transistors are slow but less leaky [15] and is also motivation behind this proposal. The transistors used in DCVS logic tree are standard $V_{TH}$ for not to hamper speed of logic evaluation. Only LCTs are chosen to be high $V_{TH}$ so as to achieve improved leakage saving without much hampering the logic evaluation speed.

As high $V_{TH}$ for LCTs usage culminates into achieving further leakage saving therefore configurations proposed in remaining part of this paper employ these transistors.

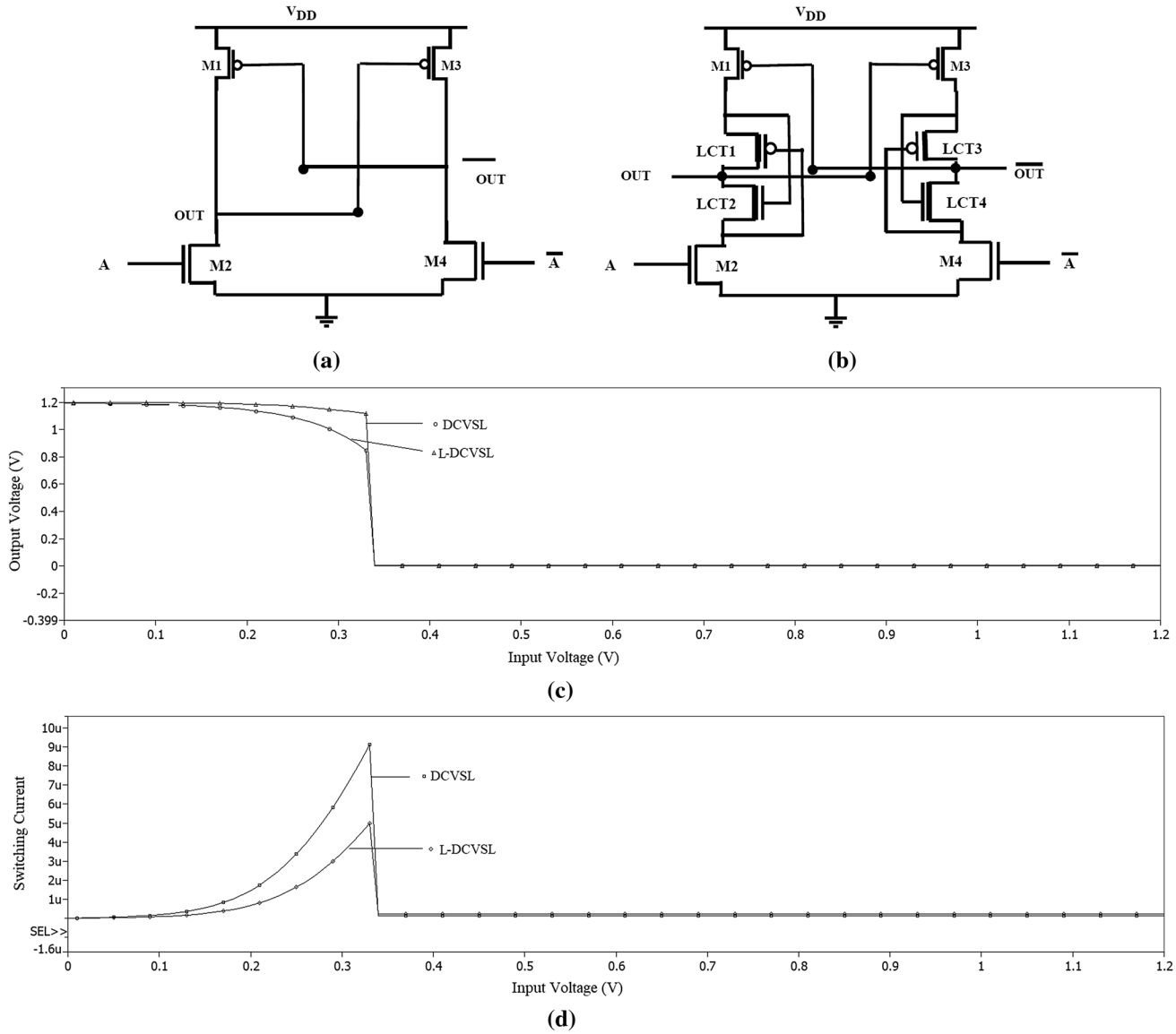



**Fig. 6** **a** DCVSL inverter, **b** L-DCVSL inverter, **c** VTC comparison, **d** switching current comparison

**Table 1** State matrix of DCVSL and L-DCVSL inverters

| Transistor | DCVSL | | L-DCVSL | |
|---|---|---|---|---|
| | A = 0 | A = 1 | A = 0 | A = 1 |
| M1 | On | Off | On | Off |
| M2 | Off | On | Off | On |
| LCT1 | – | – | Near cut off | On |
| LCT2 | – | – | On | Cut off |

## 3.3 Proposed LECTOR incorporated multi-threshold dynamic DCVSL configurations

It is well known that DyDCVS logic designs show significant dynamic power saving. However, it suffers from leakage power loss, though less due to dynamic configuration. Since feedback transistors in dy-DCVSL maintain the connection between VDD and output node, there are chances of leakage currents along the path from VDD to ground. Therefore, a LECTOR incorporated dynamic DCVSL and EDCVSL structures have been proposed to combat the leakage currents flowing from VDD to ground.

### 3.3.1 Architecture

General architecture of proposed LECTOR incorporated multi-threshold dynamic DCVSL (MTL-dyDCVSL) is depicted in Fig. 8. It is exactly same as that of DyDCVSL except the inclusion of four leakage control transistors (LCT1, LCT2, LCT3, and LCT4), two on each branch (true logic and complementary).

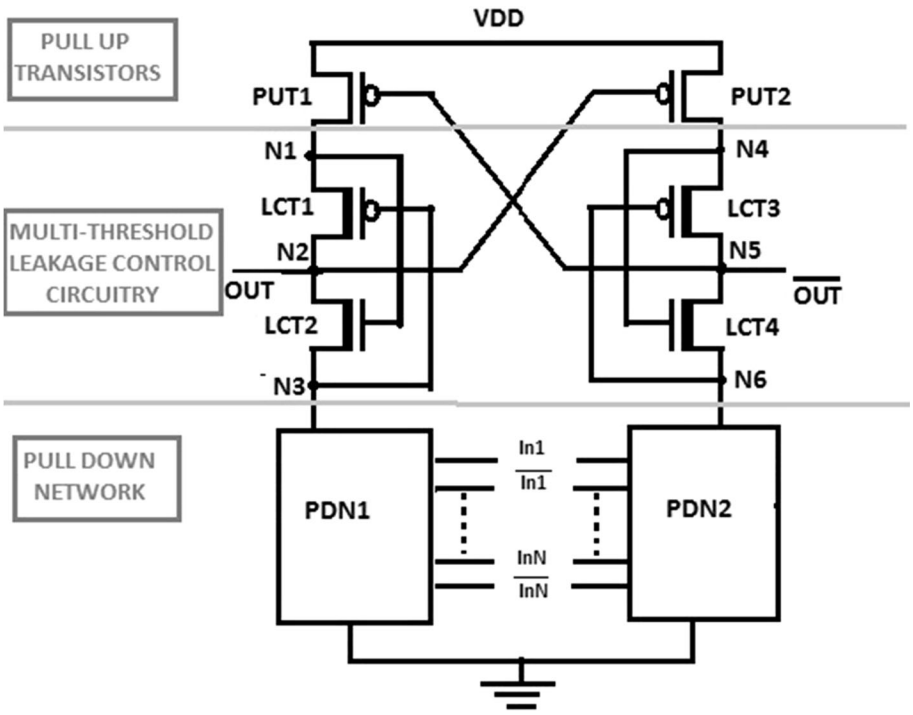



**Fig. 7** Generic architecture of proposed MTL-DCVSL

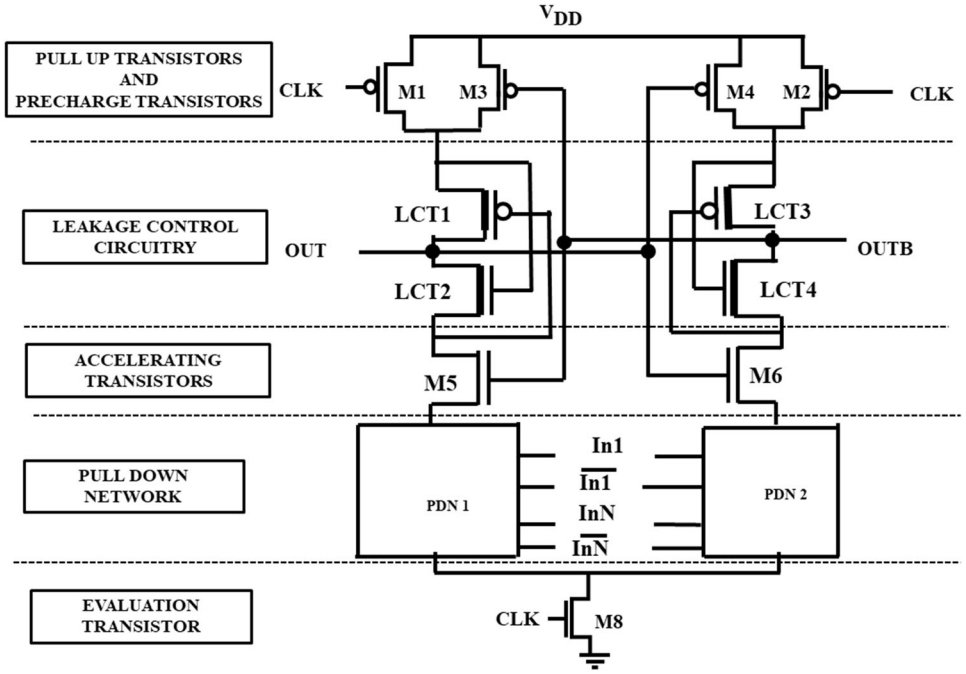
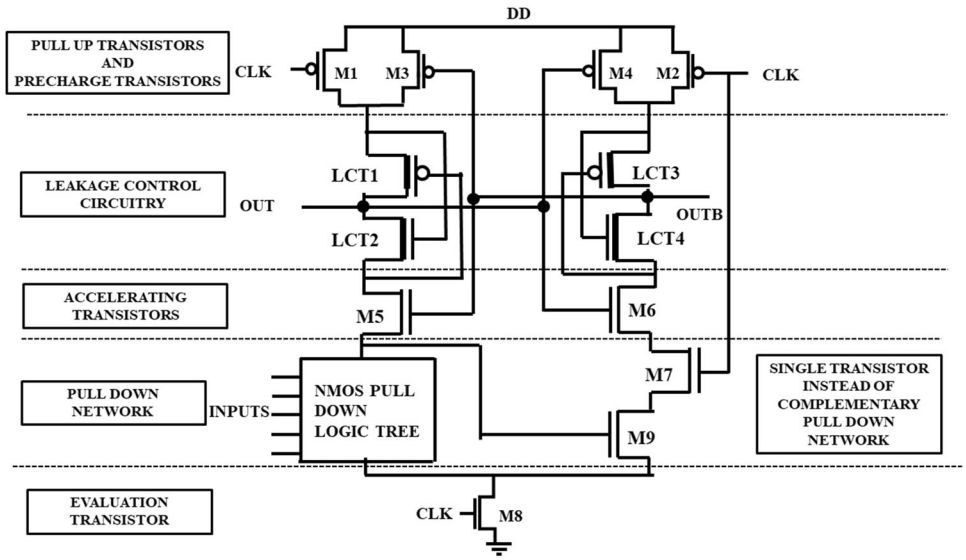



**Fig. 8** Generic architecture of proposed MTL-dyDCVSL

### 3.3.2 Leakage conditions and control mechanism

Leakage currents are those currents which flow through the branch when PDN doesn't conduct and output node is supposed to be at logic high. These situations arise in both pre-charge and evaluation phases.

Consider Fig. 8, the output remains at logic high during pre-charge phase (CLK = 0) irrespective of the input conditions since evaluate transistor M7 is OFF and blocks the path for current to ground. Ideally the current through power supply should be zero in such condition, but the subthreshold leakage currents through OFF transistors along the path manage to flow and reach the ground terminal. During evaluation phase, for input conditions which don't turn ON the PDN, the output node is preserves its logic high state as there is no current path provided by the



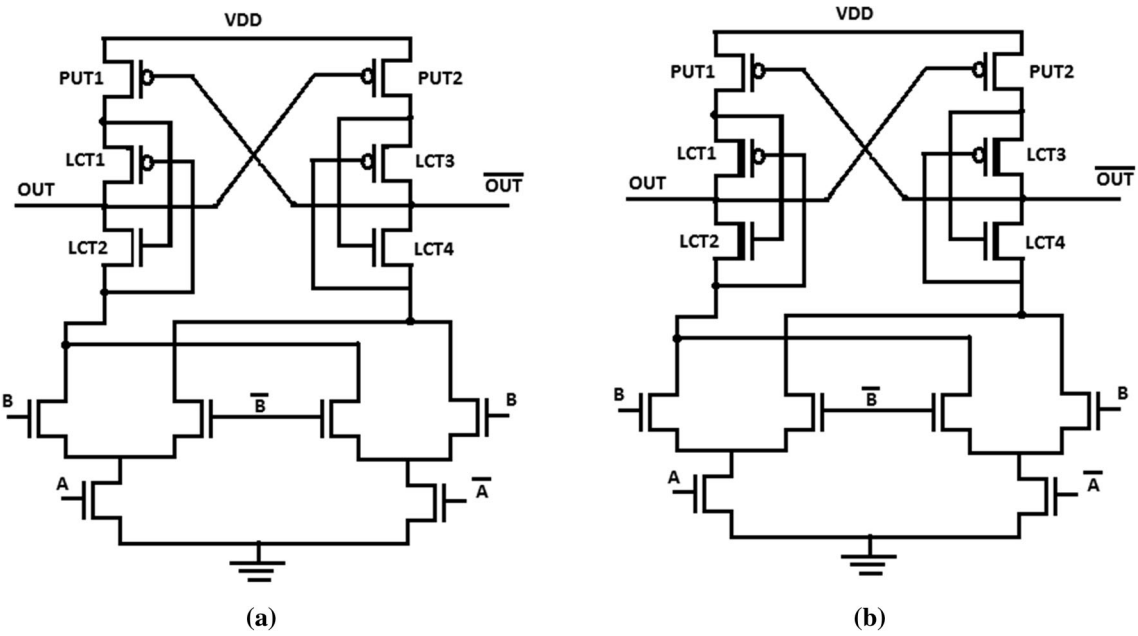




**Fig. 9** Generic architecture of proposed MTL-EDCVSL

**Fig. 10** Lector incorporated proposed XOR2/XNOR2 gates **a** L-DCVSL, **b** MTL-DCVSL

logic tree. Nevertheless, there is a leakage current flow through OFF transistors in the evaluation logic tree. Henceforth, LECTOR configuration has been adapted for dyDCVSL as well which efficiently controls the leakage current flow.

### 3.4 Proposed LECTOR incorporated multi-threshold EDCVSL configuration

The EDCVSL [10] is yet another variant of dyDCVSL where complementary logic tree is replaced with just one simple NMOS transistor driven by voltage at true output node to obtain inversion of true logic to implement complementary logic [10]. It eliminates inputs needed in complementary branch and reduces the number of interconnects and transistor count. Leakage current reduction in EDCVSL is achieved by proposed LECTOR incorporated multi-threshold EDCVSL (MTL-EDVSL). Generic architecture of proposed MTL-EDVSL is depicted in Fig. 9.

Leakage control mechanism is same as discussed for MTL-DyDCVSL.

**Table 2** Leakage current measurement for basic and LECTOR incorporated DCVSL XOR2/XNOR2 gate ($V_{DD}$ = 1.2 V)

| Technology/input states (A, B) | Static DCVSL | L-DCVSL | MTL-DCVSL | Percentage saving with MTL-DCVSL |
|---|---|---|---|---|
| *90 nm* | | | | |
| OFF states | Leakage current (nA) | | | |
| (1, 0) | 15.55 | 11.37 | 8.68 | 44.18 |
| (0, 1) | 15.55 | 11.37 | 8.68 | 44.18 |
| ON states | Static current (nA) | | | |
| (1, 1) | 6.05 | 3.42 | 2.29 | 62.14 |
| (0, 0) | 6.05 | 3.42 | 2.29 | 62.14 |
| *65 nm* | | | | |
| OFF states | Leakage current (nA) | | | |
| (1, 0) | 19.94 | 13.62 | 9.42 | 52.75 |
| (0, 1) | 19.94 | 13.62 | 9.42 | 52.75 |
| ON states | Static current (nA) | | | |
| (1, 1) | 8.64 | 5.30 | 2.51 | 70.94 |
| (0, 0) | 8.64 | 5.30 | 2.51 | 70.94 |
| *45 nm* | | | | |
| OFF states | Leakage current (nA) | | | |
| (1, 0) | 31.32 | 15.56 | 11.18 | 64.30 |
| (0,1) | 31.32 | 15.56 | 11.18 | 64.30 |
| ON states | Static current (nA) | | | |
| (1, 1) | 16.61 | 10.21 | 3.11 | 81.27 |
| (0, 0) | 16.61 | 10.21 | 3.11 | 81.27 |

## 4 Simulation results

The performance of proposed LECTOR incorporated circuits is verified by observing leakage current through SymicaDE simulations using PTM's 90 nm, 65 nm and 45 nm technology models. The choice of lower technology nodes is based on the fact that leakage effect has predominance there. Supply voltage of 1.2 V and 1.2 V are considered throughout the simulations. Moreover, the effect of temperature variation on leakage current is also investigated. The simulation results are arranged in three sections one each for proposed static DCVSL architecture, dyDCVSL and EDCVSL architectures; and their temperature dependence.

### 4.1 Performance of static DCVSL configurations

Under this category, performance in terms of leakage current reduction through incorporation of LECTOR technique in static DCVSL circuits is examined. A XOR2/XNOR2 gate is being used as the test bench due to its wide range of application. Simulations are performed for Static DCVSL, L-DCVSL and MTL-DCVSL based two-input XOR-XNOR gate (XOR2/XNOR2) of circuits shown in Figs. 1(b) and 10 respectively. Leakage and static current measurement were performed for true logic branch and every possible input combination for technologies 90 nm, 65 nm and 45 nm at 27 °C is considered. The observed results are comprehended in Table 2. The OFF states and ON states in Table 2 represent input states corresponding to which output is logic high and logic low respectively. For example, there should be no current from output node to ground for input combinations (1, 0) and (0, 1) (known as OFF states). Nonetheless, leakage current makes its path to ground through off transistors for these inputs. Similarly, for input combinations (1, 1) and (0, 0) (known as ON states) PDN1 turns on and discharges "out" to logic "0". Current flow through PDN after output stabilisation is static current which is undesirable.

An obvious saving in leakage and static currents due to incorporation of leakage control transistors can be observed from Table 2. Further, percentage saving obtained in leakage and static current is 44–62% for 90 nm, 52–71% for 65 nm and 64–81% respectively.

### 4.2 Performance of dynamic DCVSL configurations

This category delineates about the leakage current reduction through incorporation of LECTOR technique in dynamic DCVSL circuits (Figs. 2, 8, 9). The XOR2/XNOR2 gate based on dyDCVSL, proposed MTL-

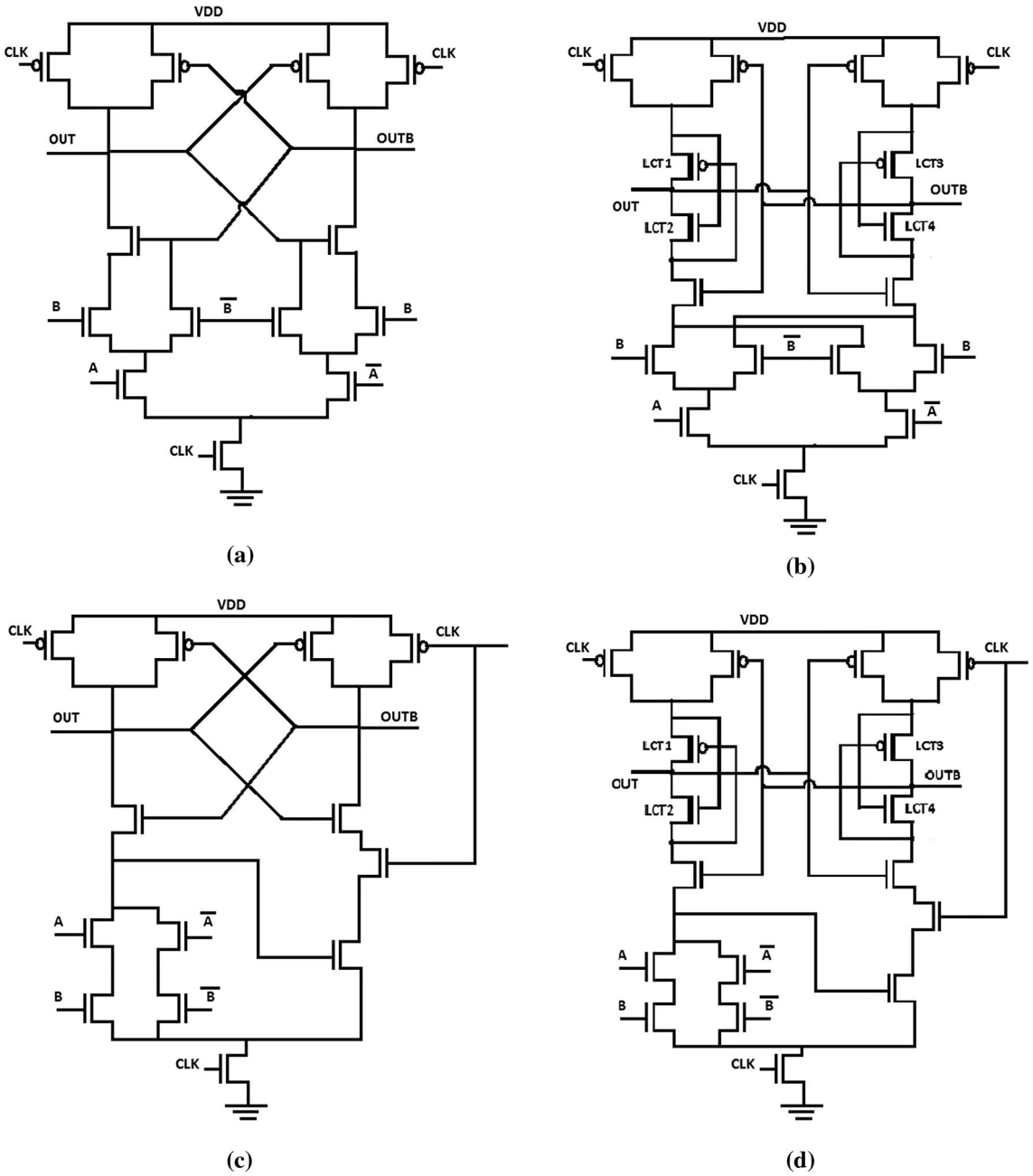



**Fig. 11** Dynamic XOR2/XNOR2 gates **a** dyDCVSL, **b** proposed MTL-dyDCVSL, **c** EDCVSL, **d** proposed MTL-EDCVSL

dyDCVSL, EDCVSL and proposed MTL-EDCVSL are given in Fig. 11(a)–(d) respectively. As the operation of dynamic circuits is based on precharge—evaluate basis, leakage and static current are observed in both pre-charge and evaluate phase and Table 3 lists the findings for power supply of 1.2 V. It can be seen that for power supply of 1.2 V, the percentage saving obtained in leakage and static current is 22–65% for 90 nm, 9–66% for 65 nm and 21–76% for 45 nm for proposed MTL-DyDCVSL XOR2/XNOR2 gate in comparison to its dyDCVSL counterpart. Similarly, percentage saving obtained in leakage and static current for proposed MTL-EDCVSL is 16–64% for 90 nm, 17–67% for 65 nm and 19–78% for 45 nm in comparison to its EDCVSL counterpart. Performing a comparison of leakage reduction efficiency of MTL-EDCVSL with respect to dyDCVSL reveals that former is 42–90%, 4–67% and 34–79% more efficient for 90 nm, 65 nm and 45 nm technologies respectively.

Following are the important findings from Tables 2 and 3:

**Table 3** Leakage current measurement for basic and LECTOR incorporated dynamic DCVSL and EDCVSL XOR2/XNOR2 GATE ($V_{DD}$ = 1.2 V)

| Technology/input states (A, B) | Dynamic DCVSL | | | EDCVSL | | | Percentage saving of EDCVSL w.r.t. dyDCVSL |
|---|---|---|---|---|---|---|---|
| | dyDCVSL | MTL-dyDCVSL | Percentage saving | EDCVSL | MTL-EDCVSL | Percentage saving | |
| *90 nm* | | | | | | | |
| OFF states | Leakage current | | | | | | |
| Precharge | | | | | | | |
| (0, 0) | 2.83 | 2.2 | 22.26 | 1.93 | 1.62 | 16.06 | 42.75 |
| (0, 1) | 0.93 | 0.62 | 33 | 0.229 | 0.144 | 37.11 | 84 |
| (1, 1) | 2.83 | 2.2 | 22.26 | 1.93 | 1.62 | 16.06 | 42.75 |
| (1, 0) | 0.93 | 0.62 | 33 | 0.229 | 0.144 | 37.11 | 84 |
| Evaluation | | | | | | | |
| (1, 0) | 0.48 | 0.17 | 64.58 | 0.128 | 0.045 | 64.84 | 90 |
| (0, 1) | 0.48 | 0.17 | 64.58 | 0.120 | 0.046 | 61.67 | 90 |
| ON states | Static current | | | | | | |
| Evaluation | | | | | | | |
| (1, 1) | 13.04 | 5.86 | 55 | 13.04 | 5.16 | 60.42 | 60.40 |
| (0, 0) | 13.04 | 5.86 | 55 | 13.04 | 5.16 | 60.42 | 60.40 |
| *65 nm* | | | | | | | |
| OFF states | Leakage current | | | | | | |
| Precharge | | | | | | | |
| (0, 0) | 1.96 | 1.72 | 12.24 | 2.28 | 1.87 | 17.98 | 4.5 |
| (0, 1) | 0.53 | 0.41 | 22.6 | 0.445 | 0.297 | 33.25 | 43.9 |
| (1, 1) | 1.96 | 1.72 | 12.24 | 2.28 | 1.87 | 17.98 | 4.5 |
| (1, 0) | 0.53 | 0.41 | 22.6 | 0.445 | 0.297 | 33.25 | 43.9 |
| Evaluation | | | | | | | |
| (1, 0) | 0.021 | 0.019 | 9.52- | 0.015 | 0.005 | 66.67 | 23 |
| (0, 1) | 0.021 | 0.019 | 9.52- | 0.016 | 0.005 | 66.68 | 23 |
| ON states | Static current | | | | | | |
| Evaluation | | | | | | | |
| (1, 1) | 20.21 | 6.58 | 66.04 | 19.20 | 6.52 | 66.04 | 67 |
| (0, 0) | 20.21 | 6.58 | 66.04 | 19.20 | 6.52 | 66.04 | 67 |
| *45 nm* | | | | | | | |
| OFF states | Leakage current | | | | | | |
| Precharge | | | | | | | |
| (0, 0) | 3.24 | 2.54 | 21.6 | 2.66 | 2.14 | 19.54 | 34 |
| (0, 1) | 0.52 | 0.41 | 21.6 | 0.452 | 0.306 | 32.30 | 41 |
| (1, 1) | 3.24 | 2.54 | 21.6 | 2.66 | 2.14 | 19.54 | 34 |
| (1, 0) | 0.52 | 0.41 | 21.6 | 0.452 | 0.306 | 32.30 | 41 |
| Evaluation | | | | | | | |
| (1, 0) | 0.032 | 0.015 | 66 | 0.020 | 0.007 | 66 | 78 |
| (0, 1) | 0.032 | 0.015 | 66 | 0.020 | 0.007 | 66 | 78 |
| ON states | Static current | | | | | | |
| Evaluation | | | | | | | |
| (1, 1) | 35.2 | 8.26 | 76 | 34.72 | 7.48 | 78.45 | 78.75 |
| (0, 0) | 35.2 | 8.26 | 76 | 34.72 | 7.48 | 78.45 | 78.75 |

**Table 4** Leakage current measurement for basic and LECTOR incorporated DCVSL XOR2/XNOR2 gate with temperature variation ($V_{DD}$ = 1.2 V)

| Technology/input states (A, B) | Static DCVSL | | | MTL-DCVSL | | |
|---|---|---|---|---|---|---|
| | − 25 °C | 27 °C | 100 °C | − 25 °C | 27 °C | 100 °C |
| OFF states | Leakage current (nA) | | | | | |
| (1, 0) | 5.68 | 15.5 | 48.20 | 3.27 | 8.68 | 37.59 |
| (0, 1) | 5.68 | 15.5 | 48.20 | 3.27 | 8.68 | 37.59 |
| ON states | Static current (nA) | | | | | |
| (1, 1) | 6.2 | 16.61 | 52.30 | 1.35 | 2.29 | 9.6 |
| (0, 0) | 6.2 | 16.61 | 52.30 | 1.35 | 2.29 | 9.6 |

**Table 5** Leakage current measurement for basic and LECTOR incorporated dynamic DCVSL and EDCVSL XOR2/XNOR2 gate with temperature variation ($V_{DD}$ = 1.2 V)

| | EDCVSL | | | MTL-EDCVSL | | |
|---|---|---|---|---|---|---|
| | − 25 °C | 27 °C | 100 °C | − 25 °C | 27 °C | 100 °C |
| Precharge | | | | | | |
| (0, 0) | 0.87 | 1.93 | 4.8 | 0.53 | 1.62 | 4.1 |
| (0, 1) | 0.27 | 0.229 | 0.35 | 0.11 | 0.144 | 0.19 |
| (1, 0) | 0.87 | 1.93 | 4.8 | 0.53 | 1.62 | 4.1 |
| (1, 1) | 0.27 | 0.229 | 0.35 | 0.11 | 0.144 | 0.19 |
| Evaluation | | | | | | |
| (1, 0) | 0.042 | 0.128 | 4.2 | 0.033 | 0.045 | 1.8 |
| (0, 1) | 0.042 | 0.128 | 4.2 | 0.033 | 0.045 | 1.8 |

1. The leakage current follows an upward trend with lowering down of the technology for all the experimented circuits.
2. Percentage saving also increases towards lower technology nodes in case of static DCVSL and for dyDCVSL and EDCVSL remains more or less same.
3. Input states (1, 1) and (0, 0) achieve maximum percentage saving whereas input states (0, 1) and (1, 0) achieve minimum percentage saving for static DCVSL.
4. The leakage effect is inherently less in clocked configurations (dynamic DCVSL and EDCVSL).
5. For dynamic DCVSL and EDCVSL in pre-charge phase, maximum saving is achieved for input states (0, 1) and (1, 0), whereas minimum saving is achieved for (0, 0) and (1, 1).
6. Similar analysis for evaluation phase shows that input states (1, 1) and (0, 0) achieves maximum saving.

### 4.3 Effect of temperature and process variation on leakage current

Temperature variation of average leakage current for XOR/XNOR2 DCVSL gate and XOR/XNOR2 EDCVSL gate for 90 nm technology and $V_{DD}$ = 1.2 V is listed in Tables 4 and 5 respectively. The pattern followed verifies that leakage current is directly proportional to temperature. Leakage current increases with temperature. A 10,000-point Monte Carlo simulation under 3σ global and 1σ independent local process variations is evaluated on the 90 nm technology node. The VTC variations in the LECTOR incorporated multi-threshold DCVSL inverter is shown in Fig. 12. A 3σ variation of 30 mV is obtained in the threshold voltage of the inverter. The results for leakage current measurement for all the proposed XOR2/XNOR2 are summarized in Table 6. It may be observed that the mean value of the leakage current is always larger than 3σ variations.

**Table 6** Summary of Monte Carlo simulations

| Gate | Mean (nA) | 3σ (nA) |
|---|---|---|
| Static DCVSL | 15.55 | 11.5 |
| L-DCVSL | 11.37 | 8 |
| MTL-DCVSL | 8.68 | 7.2 |
| EDCVSL | 1.93 | 1.4 |
| MTL-EDCVSL | 1.62 | 1.2 |

## 5 Conclusion

A novel leakage power reduction technique for DCVSL circuits has been presented. The technique employs incorporation of leakage control circuitry between PUT and PDN which significantly reduced the leakage power in DCVSL circuits. Simulations have been performed for 90 nm, 65 nm and 45 nm technologies that show an increase in leakage current with lowering down of the technology. LECTOR technique has been adapted for DCVS logic structure for leakage power reduction and thereby proposing LECTOR incorporated circuits for static DCVSL, EDCVSL and dynamic DCVSL 2 input XOR/XNOR gates. These newly introduced LECTOR incorporated DCVSL 2-input XOR/XNOR gate achieves average 71.33% saving in leakage current for power supply of 1.2 V where percentage saving increases with lowering

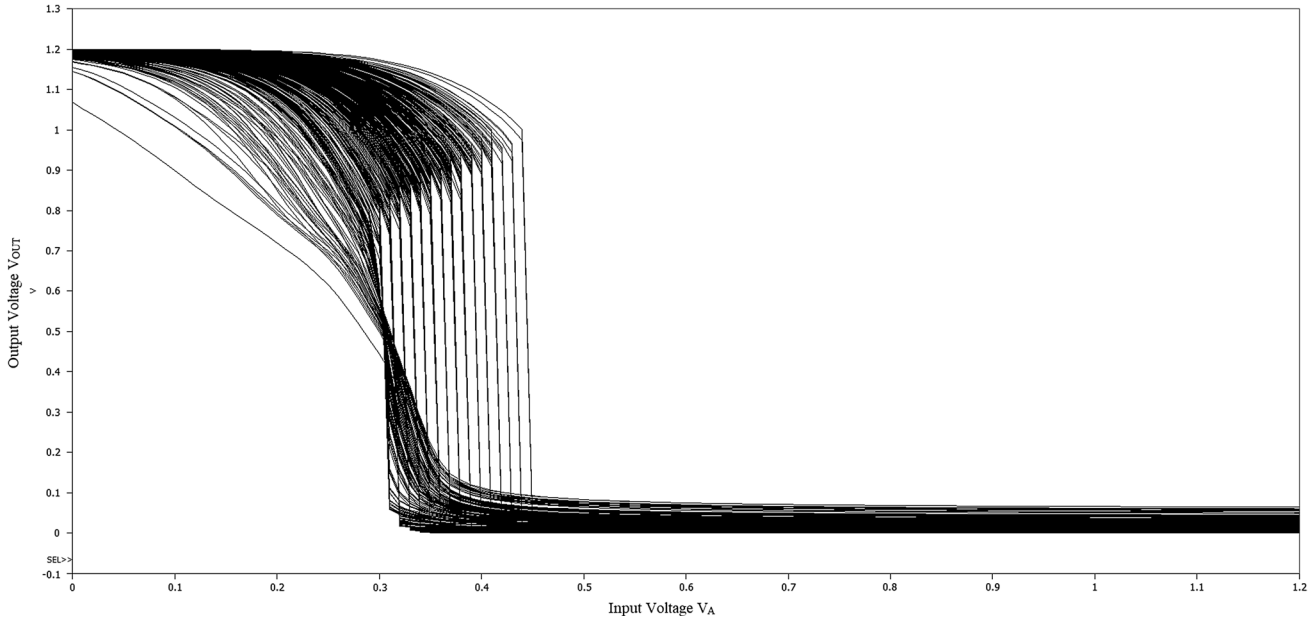



**Fig. 12** LECTOR incorporated multi-threshold DCVSL inverter VTC variations simulation under 3σ global and 1σ independent local process variations

down of technology from 90 to 45 nm. LECTOR incorporated dynamic DCVSL 2-input XOR/XNOR gate achieves average saving of 69% in leakage current. LECTOR incorporated EDCVSL 2-input XOR/XNOR gate achieves average 70% saving in leakage current. Temperature variation performed for temperature values 25 °C, 27 °C and 100 °C shows that leakage current is directly proportional to temperature.

## References


1. Narendra, S. G., & Chandrakasan, A. (2006). *Leakage in nanometer CMOS technologies*. Berlin: Springer.
2. Fallah, F., & Pedram, M. (2005). Standby and active leakage current control and minimization in CMOS VLSI circuits. *IEICE Transactions on Electronics, 88-C*(4), 509–519.
3. Dadoria, A. K., Khare, K., & Singh, R. P. (2016). Sleepy lector: A novel approach for leakage reduction in DSM technology. In *International conference on cloud system and big data engineering*.
4. Johnson, M. C., Somasekhar, D. L., Chiou, Y., & Roy, K. (2002). Leakage control with efficient use of transistor stacks in single threshold CMOS. *IEEE Transactions on VLSI Systems, 10*(1), 1–5.
5. Sirichotiyakul, S., Edwards, T., Oh, C., Panda, R., & Blaauw, D. (2002). Duet: An accurate leakage estimation and optimization tool for dual-V circuits. *IEEE Transactions on VLSI System, 10*(2), 79–90.
6. Sarkar, H., & Kundu, S. (2015). Standby leakage current estimation model for multi threshold CMOS inverter circuit in deep submicron technology. In *International symposium on VLSI design and test*.
7. Kumar, M., Hussain, M. A., & Paul, S. K. (2013). New hybrid digital circuit design techniques for reducing subthreshold leakage power in standby mode. *Circuits and Systems, 4*(1), 75–82.
8. Narayan, S. L., Korah, R., & Kumar, N. K. (2013). A novel sleepy stack 6-T SRAM cell design for reduced leakage power in submicron technologies. In *International conference on communication and signal processing* (pp. 1–5).
9. Rabaey, J. M. (1996). *Digital integrated circuits: A design perspective*. New York: Prentice Hall.
10. Kang, D. W., & Kim, Y. B. (2002). Design of enhanced differential cascode voltage switch logic (EDCVSL) circuits for high fan-in gate. In *IEEE international conference on ASIC/SOC* (pp. 309–313).
11. Chen, W., Hwang, W., Kudva, P., Gristede, G. D., Kosonocky, S. & Joshi, R. V. (2001). Mixed multi threshold differential cascode voltage switch (NIT-DCVS) circuit styles and strategies for low power VLSI design. In *International symposium on low power electronics and design* (pp. 1–5).
12. Lakshmikanthan, P., & Nũnez, A. (2006). A novel methodology to reduce leakage power in differential cascode voltage switch logic circuits. In *3rd International conference on electrical and electronics engineering*.
13. Hanchate, N., & Ranganathan, N. (2004). LECTOR: A technique for leakage reduction in CMOS circuits. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 12*(2), 196–205.
14. Bajpai, P., Pandey, N., Gupta, K., Bagga, S., & Panda, J. (2016). On improving the performance of dynamic DCVSL circuits. *Journal of Electrical and Computer Engineering, 2017,* 8207104.
15. Park, J. C., & Mooney, V. J. (2006). Sleepy stack leakage reduction. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 14*(11), 1250–1263.



**Pratibha Bajpai** received her B.Tech. in Electronics and Communication from JSS Academy of Technical Education, Noida and M.Tech. in VLSI and Embedded Systems from Delhi Technological University, Delhi. Her area of research is low power VLSI. Specifically, leakage power reduction in VLSI circuits. At present, she is working as Junior Executive (Electronics) at Airports Authority of India.



**Neeta Pandey** received her M.E. in Microelectronics from Birla Institute of Technology and Sciences, Pilani and Ph.D. from Guru Gobind Singh Indraprastha University Delhi. She has served in Central Electronics Engineering Research Institute, Pilani, Indian Institute of Technology, Delhi, Priyadarshini College of Computer Science, Noida and Bharati Vidyapeeth's College of Engineering, Delhi in various capacities. At present, she is Professor in ECE Department, Delhi Technological University. A life member of ISTE, and senior member of IEEE, USA. She has published papers in international, national journals of repute and conferences. Her research interests are in analog and digital VLSI Design.



**Kirti Gupta** received B.Tech. in Electronics and Communication Engineering from Indira Gandhi Institute of Technology, Delhi in 2002, M.Tech. in Information Technology from School of Information Technology in 2006. She received her Ph.D. in Electronics and Communication Engineering from Delhi Technological University, in 2016. Since 2002, she is with Bharati Vidyapeeth's College of Engineering, New Delhi and is presently serving as Professor in the same institute. A life member of ISTE, and senior member of IEEE, she has published papers in international, national journals and conferences. Her teaching and research interest is in digital VLSI design.



**Jeebananda Panda** received his M.E. in Applied Electronics from Bharathiyar University, Coimbatore and completed his Ph.D. from Delhi University Delhi. At present, he is working as Professor in ECE Department, Delhi Technological University. He is a life member of ISTE, IETE (India), and also a member of IEEE, USA. He has published papers in international, national journals of repute and conferences. His research interests are in Image Processing, Digital Watermarking and digital VLSI Design.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

2019 6th International Conference on Signal Processing and Integrated Networks (SPIN)

# Low Noise Amplifier in Smart Healthcare Applications

Malti Bansal
Department of Electronics and Communication Engineering
Delhi Technological University
Delhi -110042,India
maltibansal@gmail.com

Jyoti
Department of Electronics and Communication Engineering
Delhi Technological University
Delhi -110042,India
electronics jyoti91@gmail.com

***Abstract*— The smart health care monitoring systems are playing a pivotal role in the medical health industry. The wireless healthcare monitoring system is one of the leading innovative developments in the medical system. A smart wireless health monitoring system helps to measure the user's neural signals, heart rate, body temperature and any unusual activity in the patient's body. This paper reviews and surveys the various applications of a Low Noise Amplifier (LNA) in the healthcare domain, such as ECG monitoring systems, EEG monitoring Systems, neural recording system, biosensor applications, Medical implantable systems, etc. The design specifications of LNA for different applications have also been analyzed. LNA is a key element of a front-end receiver, which combines low noise figure, nominal supply voltage, high gain bandwidth, good input, and output impedance matching, etc. Finally, a comparison of the different parameters for various applications has been done and reviewed in this paper.**

***Keywords—Low Noise Amplifier (LNA), CMOS, EEG, ECG, healthcare systems.***

## I. INTRODUCTION

The use of wireless healthcare monitoring systems has gained popularity and has become a necessary requirement for user information to be retrieved anywhere in the healthcare system. A healthcare system provides service planning such as registration, scheduling, and collection and management of clinical data of patients [1]. The general architecture of a wireless healthcare monitoring system is as shown in Figure 1 [2]. Basically, the overall system consists of three sections; the first section is composed of intelligent physiological sensors. The first level of the healthcare system depends upon user applications; like the motion sensor, blood pressure sensor, ECG sensor, etc. The function of physiological sensors is to capture biological signals and convert them into a digital form for easier processing in the subsequent stages. In a Wireless Body Area Network (WBAN), physiological sensors capture these signals and interface with wireless platforms [3]. Furthermore, this wireless system provides storage, computational and communication abilities.

The second part of the wireless health monitoring system consists of network nodes which collect and process information from physiological sensors, store the data and send the information continuously to the distinctive server. The third level provides communication between the personal server and healthcare provider. The main purpose of the personal server is to process the raw data from the physiological sensors and also interface with healthcare services using internet resources. The third section of the system is a healthcare provider who receives the data continuously from the users or patients. The data of a patient is then analyzed and collaborated into the user's medical record. The benefit of these wireless systems is to provide emergency services and for research purposes without taking any time to accumulate samples [3].

The Low Noise Amplifier (LNA) is integral to a Radio frequency communication system design, as in both cases of homodyne as well as heterodyne concepts, it is the foremost building section. It also sturdily influences the linearity with the noise of the transceiver [4]. The common topology of the Low Noise Amplifier includes three stages. The first stage is input matching circuit or network (IMN), the second stage consists of the amplifier section itself, and the last stage includes the output matching circuit or network (OMN) [5].

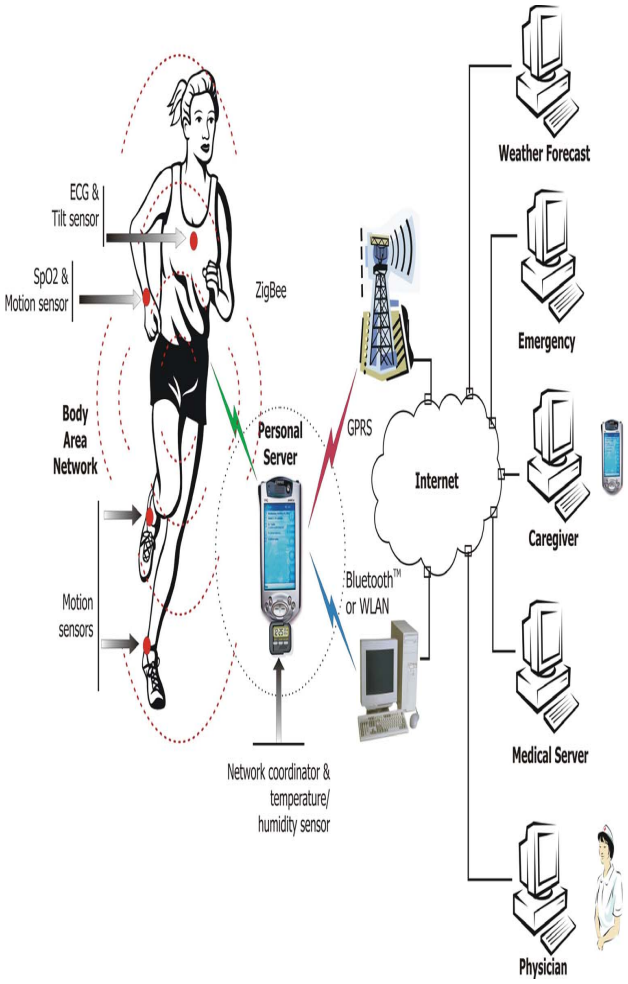



Fig. 1: The architecture of Wireless Healthcare Monitoring Systems [2]

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE
978-1-7281-1380-7/19/$31.00 ©2019 IEEE

## II. APPLICATIONS OF LOW NOISE AMPLIFIER

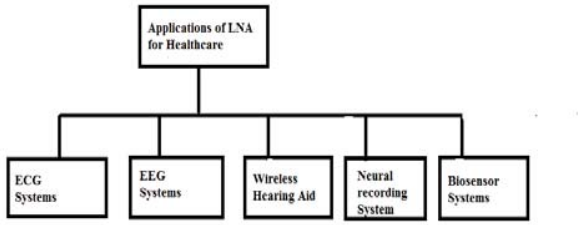



Fig. 2: Applications of LNA

### A. Electrocardiography Applications

In general, electrocardiography is the process of measuring the electrical signals of the heart. An ECG is a biopotential signal that is generated by the heart in the form of an electrical current, that can be recognized on the body surface. The monitoring and analysis of ECG signals become necessary for the diagnosis of cardiovascular diseases. However, standard devices are used to calculate a heartbeat which is basically to determine the end user situation. Any kind of unusual activity in ECG signal in a patient indicates that they suffer from cardiovascular diseases [6][7].

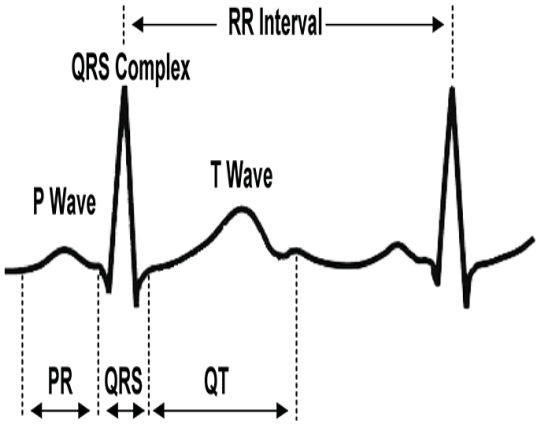



Fig. 3: ECG Signal [8]

TABLE I. CONSIDERATIONS FOR ECG SIGNAL LOW NOISE AMPLIFIER CIRCUITS [9]

| Specification | Value | Remarks |
|---|---|---|
| Supply voltage | 1V | 350nm technology |
| Supply current | < 1µA | System total current |
| Amplifier current | < 500nA | Low noise amplifier only |
| Nominal Gain | 200~1000 | Programmable Gain |
| Referred noise | < 3µVrms | Integrated in the signal band |
| corner frequency | < 0.5Hz | No external components |

*Low Noise Amplifier Topologies Employed*

The low noise amplifier is one of the major elements in a biopotential signals recording system. The bandpass filter is also integrated into an LNA because it improves the power efficiency of signals. One of the most widely used topologies adopted in biomedical signals is the capacitive feedback technique. The critical characteristics of biomedical sensor interface are low power consumption with low noise figure [9].

*Low Noise Operational Transconductance Amplifier (LNOTA)*

An LNOTA is utilized in a low noise amplifier. It comprises of a metal oxide semiconductor depletion pseudo-resistor, the feedback capacitor, load capacitor, and a coupling capacitor. Basically, the main function of a pseudo-resistor is to reject DC offset voltage at electrode-tissue interface. It provides higher value of resistance, i.e., tera ohm among smaller swings. A 2-stage miller compensation LNOTA is implemented because it provides high open loop gain as well as rail to rail output voltage swing [10].

Basically, LNA is a fully differential open trans-conductance amplifier composed of a common-mode feedback (CMFB) circuit. The initial module consists of a large positive MOS (PMOS) pair, M1 and M2. The PMOS circuitry is biased in threshold region to decrease the flicker noise. Hence, the output stage consists of cascode topology that is used to increase the open loop gain [11].

TABLE II. PERFORMANCE OF LNA FOR ECG APPLICATIONS

| References | [9] | [12] | [13] | [11] | [14] |
|---|---|---|---|---|---|
| CMOS process | 350nm | 350nm | 50nm | 150nm | 80nm |
| Power supply (V) | 1 | 1 | 2.8 | ±1.7 | 1.8 ~ 3.3 |
| Current (TB-FEA) | 8µA | 743nA | 330nA | 1µA | 33~337nA |
| High-pass corner frequency | 0.015Hz ~700Hz | 0.392Hz (Tunable) | 0.003Hz (Fixed) | 0.05/0.4/ 2.5Hz | 4.5mHz~ 3.6Hz |
| Low-pass corner frequency | 40/400/3k/ 4k Hz | 295Hz (Tunable) | 245Hz (Fixed) | 120Hz (Fixed) | 31~292Hz |
| Input Referred Noise | 3.6µVrms (20~10kHz) | 1.66µVrms (0.2~1kHz) | 2.7µVrms (0.05~250Hz) | 0.95µVrms (0.05~100Hz) | 2.5µVrms (0.05~460Hz) |

### B. Neural Applications

Nowadays, there has been an increase in the designing of neural recording systems, which can be consolidated with the smart wireless data transmission abilities for a restrained estimation of the brain activity [15] [16][17]. During the last few years, there has been an ever-increasing demand for an

optimal design of recording systems. Basically, neural recording systems use various interfaces which are required to play the main role simultaneously in clinical applications, i.e., as a component of therapeutic mechanism in sufferer with neurological diseases as well as neuroscience applications. The neural recording systems necessarily consist of a set of electrodes which are used in electrophysiology to capture or record neural signals as well as measure the neural activity. These electrodes are used for electrical simulation of nervous tissue, represented as a microelectrode, and is followed by a low-noise amplifier (LNAs). The main purposes of a low noise amplifier in neural recording system is signal conditioning and neural signal transfer of the mixed signal circuitry to enable the digitization and processing of the collected data before its wireless data transmission. A major element in this system is the LNA, which is enabled to enhance the weak signals acknowledged by the microelectrodes [18].

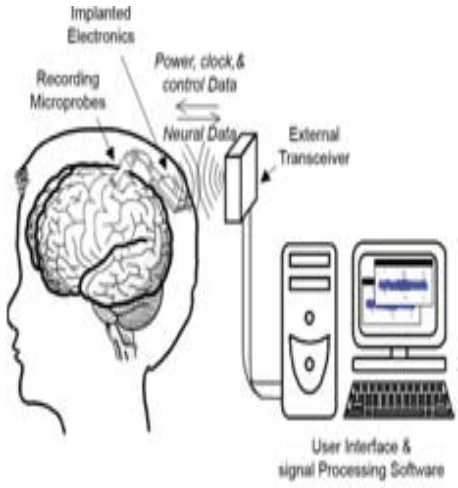



Fig. 4: Neural recording microsystem [15]

A building diagram of a wireless neural recording system is shown in Figure 5. The amplifiers used in the wireless recording device are for boosting the weak signals which are measured by every electrode. On the other hand, the function of differential amplifiers is to measurement the potential of an individual signal electrode with a lower value of impedance reference electrode. Typically, in wireless neural recording applications, every signal electrode must have its own enthusiastic LNA. In a multichannel recording system, an assemblage of various amplifiers can take up a larger chip area and more power consumption [19].

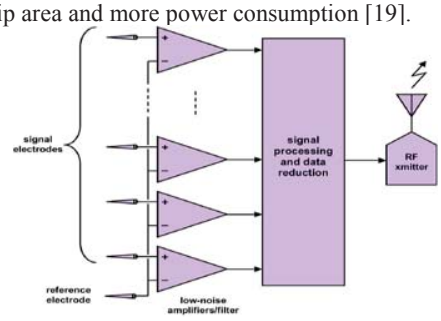



Fig. 5: Diagram of wireless neural recording system [19]

*Low Noise Amplifier Topologies used*

In an LNA, the trade-off between design parameters such as noise, power, and severe area is efficiently solved by distinct proposals which can be found in the literature. The methodology used in neural LNAs are referred to as Capacitive Feedback Network Amplifier (CFN), Capacitive Amplifier Feedback Network Amplifier (CAFN), Miller Integrator Feedback Network amplifiers [18], Logrithmic amplifier, etc. The important criteria for implementation of neural recording microsystems are smaller size, low power battery operation, acquire lesser invasion, avoid enormous heat consumption, etc. Sundarasaradula et al. [20] presented the implementation and analysis of low voltage, low noise, with a lower consumption of power in logarithmic amplifier, for a neural recording system application. The design schematic of a logarithmic amplifier is generally dependent upon the progressive-compression parallel-summation using direct current (DC) offset cancellation feedback loop. The architecture of the logarithmic amplifier utilized the progressive-compression parallel-summation technique as given in Figure 4. The proposed architecture of amplifier consisted of five modules whcih is fully differential amplifier using direct current compensation feedback loop [20] [21].

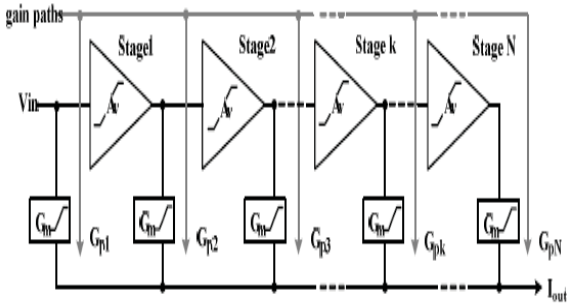



Fig. 6: Stages of logarithmic amplifier [20]

*Neural amplifier*

Figure 6 shows the architecture of a neural bioamplifier. The architecture generally consists of transistors $M_a$-$M_d$ which are the metal oxide semiconductor devices which act as pseudoa resistors where the midband gain is set to be C1/C2, and the bandwidth is $g_m$ (transconductance of OTA), where OTA is the operational transconductance amplifier (OTA). Negative $V_{GS}$ functions as diode-connected PMOS devices. The gate to source voltage having positive value($V_{GS}$), the parasitic source-well-drain pnp bipolar transistor is triggered. Despite the larger time- constant value, huge revision in the input side was observed [22].

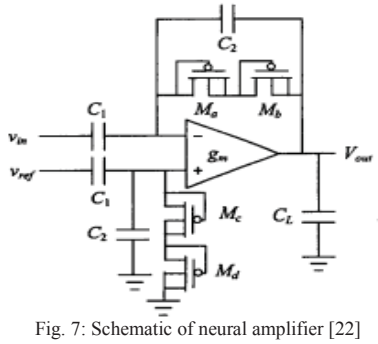



Fig. 7: Schematic of neural amplifier [22]

TABLE III. PERFORMANCE OF LNA FOR NEURAL APPLICATIONS

| References | [18] | [22] | [20] | [23] | [24] | [25] | [26] | [27] | [28] | [29] | [30] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Midband gain | 47dB | 39.8dB | 64dB | 39.8dB | 40dB | 50dB | 46dB | 40.85dB | 39.3dB | - | 39.4dB |
| Bandwidth | 250Hz to 7kHz | 0.1mHz to7.2kHz | 0.32 to 22kHz | 0.025Hz to7.2kHz | 2kHz-200kHz | 9.2kHz | 13-8.9kHz | 45Hz - 40.5kHz | 20Hz - 10kHz | - | 10Hz - 7.2kHz |
| CMOS process | 130nm | 1.5µm | .12µm | 1.5µm | 350nm | 180nm | 350nm | 500nm | 1.5µm | 500nm | 180nm |
| Power dissipation | - | 80µW | 2.86µW | 80µW | 6mW | 8.6µW | - | - | 27.2µW | 13.5mW | 7.92µW |

### C. Wireless Hearing Aids

The main purposes of wireless hearing aid devices are to amplify and improve a hearing problem and increasing audibility for proper communication. These devices receive sound signals from a microphone and then convert these signals into electrical signals. The amplifier used in a hearing device enhances the efficiency of signals. Thus, it sends the sound to the ear by a speaker. In wireless hearing aid systems, an amplifier consists of various transistors or an amalgamation of integrated circuits. The function of an amplifier is that it modulates the electrical signal and increases the amplitude of the sound waves [31].

The manufacturing of a hearing aid device is technical because it is a battery operated, electronic device. The main problem of hearing-impaired persons are different types of hearing losses. Therefore, each hearing aid device has to be customized to match the user's needs. Hence, every user's hearing loss is different, so every hearing aid device is unique [32].

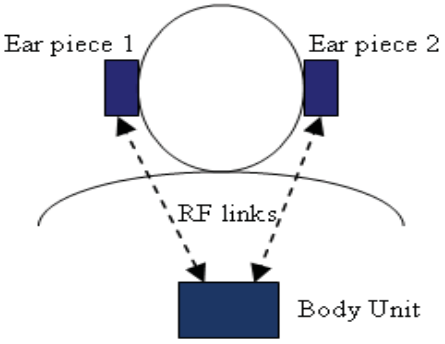



Fig. 8: Wireless Hearing Aid Principle [33]

The LNA is the first stage of receiver in the earpiece, which is powered by a battery. Reducing LNA power consumption parameter will improve the battery life [34]. LPLV LNA for wireless hearing aids was designed and described in this chapter according to the design target. In order to select the most suitable topology, five different LNA topologies are simulated and compared. The cascode inductor source degeneration topology was selected for the required design. The circuit used in wireless hearing aid implemented by using radio frequency mosfet, transistor, capacitor as well as inductor models. The simulation results demonstrated satisfaction with the design specification. A simulation results predicts a forward gain (S21) of 13.08 dB with a noise factor of 2.19 dB while drawing 0.95 mW from a 1.0 V power supply. The simulated IIP3 and P1dB are 7.5 dBm and -14.0 dBm [33].

### D. EEG Applications

An Electroencephalogram (EEG) is a device which is used to measure and record the electrical activities produced by discharging of neurons in the brain along the scalp [35]. Electroencephalography (EEG) is the most functional and significant test used to confirm a diagnosis of epilepsy because the epileptic activity of brain can produce clear unusual activity on a standard EEG study [36]. The low noise front-end amplifier employs capacitance feedback circuitry with a fully differential operational transconductance amplifier to obtain low signal distortion and high dynamic range using balanced tunable pseudo-resistors which is implement in feedback loop [37]. The low noise amplifier has high pass corner which are able to continuously tune in to the frequency of the front-end amplifier and is continuously tunable to meet up with the necessities of various applications. The main advantage of a capacitor circuit is that it consumes less area as compared to inductors. Hence, in the designing of multi-channel integrated circuits, the area of the distinct circuit module becomes more critical due to its low cost of production as well as using minimal size of chip. Miniaturization of integrated circuits has been growing tremendously. Therefore, to reduce the total coverage area of the system, the value of input capacitance of the low noise front-end amplifier should be as small as possible. Every aspect has its pros and cons; hence, smaller value of capacitance has lower common mode rejection ratio. Additionally, it results in poor impedance matching of positive as well as negative feedback paths throughout fabrication [9].

### E. Biosensor Applications

A biosensor is an analytical device that is used to detect the analyte, i.e., a chemical species with a combination of a biological component and physicochemical detector. The biological system comprises of a microorganism, receptor cell, and tissue, organelles. Healthcare systems such as communication technologies, medical sensors and mobile computing use a number of biosensors which are either implanted and worn by a patient to enable the diagnoses of diseases. The body area sensor network (BASN) performs the passive operation of the transmit the data for further processing . Secondly, a biosensor placed on or inn the

human body provides the available information while interacting on a specific and secure channel. The performance of BASN recognizes both, the psychological as well as behavioral characteristics of a patient [39].

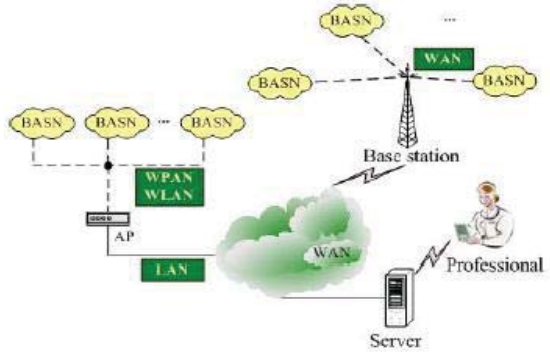



Fig. 9: A BASN combined health model [40]

The biosensor applications operate in 2.45GHz ISM radio bands [41][42]. The authors present a resistive termination low noise amplifier which is able to connect with any medical facility with an ISM receiver. The standard IEEE 802.15.6 indicates the wireless devices to be used for within, outside and around the body communication as well as their specific band for every country [43][44]. The architecture used in the biomedical field is composed of two main sections. The first section consists of a coherent receiver which is mainly based on the channel to create coincidental symbols.The outcome of this section is digital, having high precision, and has various consumable sections. Secondly, the non-coherent receiver is composed of an LNA, a mixer, an integrator and a comparator. The other applications of LNA are Bluetooth, Bluetooth low energy, GSM, etc. [45][46]. Advantages of this module are power consumption and small area, smaller than a coher coherent receiver which is mainly based on the channel to create coincidental symbols with the signal reception system. The outcome of this section is digital, having high precision, and has various consumable sections. Secondly, the non-coherent receiver is composed of an LNA, a mixer, an integrator and a comparator. The other applications of LNA are Bluetooth, Bluetooth low energy, GSM, etc. [45][46]. Advantages of this module are power consumption and small area, smaller than a coherent generation topology, resistive feedback topology, distributed topology and resistive termination topology. The main requirements of biosensor application are higher gain and small area. A comparison of the performance of these four basic LNA topologies indicates that resistive termination topology has high gain and small surface [47]

## III. Conclusion

Review of various smart healthcare applications of a Low Noise Amplifier has been done. The design considerations and specifications for various applications of LNA have been reviewed in this paper. An analysis of the performance of LNA in various applications has been done. The applications of Low Noise Amplifier such as wireless hearing aid system, EEG, ECG, etc. have been reviewed and surveyed in this paper. Finally, a comparison of the different parameters of LNA for the various applications has been done and reviewed in this paper.

## IV. Acknowldegement

One of the authors (Jyoti) acknowledges the guidance support from her thesis supervisor, Dr. Malti Bansal, Assistant Professor, Department of Electronics and Communication Engineering, Delhi Technological University, for carrying out this research work, as a part of her Ph.D. thesis work in the domain of RF Microelectronics.

## References

[1] E. Suter, N. Oelke, C. Adair, and G. Armitage, "Ten key principles for successful health systems integration", Healthcare Quarterly, vol. 13, no. sp, pp. 16–23, Oct. 2009.

[2] E. Jovanov, A. Milenkovic, C. Otto, and P. C. d. Groen, "A Wireless Body Area Network of Intelligent Motion Sensors for Computer Assisted Physical Rehabilitation", *Journal of NeuroEngineering and rehabilitation*, vol. 2, pp. 1-10 ,Mar. 2005

[3] A. C.-W. Wong, D. McDonagh, G. Kathiresan, O. C. Omeni, O. Ei-Jamaly, T. C.-K. Chan, P. Paddan, and A. J. Burdett, "A 1V, Micropower System-on-Chip for Vital-Sign Monitoring in Wireless Body Sensor Networks", in *ISSCC Dig. Tech. Papers*, pp. 138-139, Feb. 2008.

[4] J. Kebaisy, S. Domann, and B. Meinerzhagen, "A l0mW Low-Noise Amplifier Design for 5.5GHz Wireless Communication Systems", 2006 Information and Communication technologies pp. 2–7, 2006.

[5] C. K. Jha and N. Gupta, "Design of a front end Low Noise Amplifier for wireless devices," *2012 Students Conf. Eng. Syst. SCES 2012*, pp. 6–9, 2012.

[6] Z. Fei and L. Yong, "QRS Detection Based on Multiscale Mathematical Morphology for Wearable ECG Devices in Body Area Networks", Biomedical Circuits and Systems, IEEE Transactions on,vol. 3, pp. 220-228, 2009.

[7] Ooi chip pin, Asral bahari jambek, Sazali yacoob "circuit architecturers reviews for portable ECG Signal Analyzer", 2014 2nd internation conference on electronic design (ICED), pp. 19-21, Aug 2014.

[8] M. Young-Jae, K. Hoon-Ki, K. Yu-Ri, K. Gil-Su, P. Jongsun, andK.Soo-Won, "Design of Wavelet-Based ECG Detector for implantable actions on , vo cardiac pacmakers", Biomedical circuits and systems, IEEE transactions 7, pp. 426-436, 2013.

[9] Zou Xiaodan , "low power low noise analog front-end ic Design For Biomedical Sensor Interface", National University of Singapore, Singapore, pp. 1-148, 2010.

[10] E. Lopez-Morillo, R. G. Carvajal, J. Galan, J. Ramirez-Angulo, and E. Rodriguez-Villegas, "A Low-Voltage Low-Power QFG-Based Sigma-Delta Modulator for Electroencephalogram Applications", in Proc. of IEEE Biomedical Circuits and Systems Conference, pp. 571-574 , Nov. 2006.

[11] M. Yin and M. Ghovanloo, "A Low-Noise Preamplifier with Adjustable Gain and Bandwidth for Biopotential Recording Applications", in Proc. IEEE International Symposium on Circuits and Systems, pp. 321-324, May 2007.

[12] H. Wu and Y. P. Xu, "A 1V 2.3µw Biomedical Signal Acquisition IC", in *IEEE ISSCC Dig. Tech. Papers*, pp.58-59, Feb. 2006.

[13] T. Denison, K. Consoer, W. Santa, A. T. Avestruz, J. Cooley, and A. Kelly, "A 2 W 100 nV/rtHz Chopper-Stabilized Instrumentation Amplifier for Chronic Measurement of Neural Field Potentials", *IEEE Journal of Solid-State Circuits*, vol. 42, pp. 2934-2945, Dec. 2007

[14] W. Wattanapanitch, M. Fee, and R. Sarpeshkar, "An Energy-Efficient Micropower Neural Recording Amplifier", *IEEE Transactions on Biomedical Circuits and Systems*, vol. 1, pp. 136-147, Jun. 2007.

[15] A. M. Sodagar, G. E. Perlin, Y. Yao, K. Najafi, and K. D. Wise, "An implantable 64-channel wireless microsystem for single-unit neural recording," *IEEE J. Solid-State Circuits*, vol. 44, no. 9, pp. 2591–2604, 2009.

[16] R. R. Harrison *et al.*, "Low-Power Integrated Circuit," vol. 17, no. 4, pp. 322–329, 2009.

[17] M. S. Chae, Z. Yang, M. R. Yuce, L. Hoang, and W. Liu, "A 128-channel 6 mW wireless neural recording IC with spike feature

extraction and UWB transmitter," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 17, no. 4, pp. 312–321, 2009.

[18] J. Ruiz-Amaya, A. Rodríguez-Pérez, and M. Delgado-Restituto, "A comparative study of low-noise amplifiers for neural applications,"*Proc. Int. Conf. Microelectron. ICM*, pp. 327–330, 2010.

[19] Reid R. Harrison, "The Design of Integrated Circuits to Observe Brain Activity "' proceedings of ieee,vol-96, pp. 1203-1206, 2008.

[20] Y. Sundarasaradula and A. Thanachayanont, "A 0.7-V, 2.86-µW low-noise logarithmic amplifier for neural recording system," *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON*, pp. 0–3, 2013.

[21] B. Gilbert, "Logarithmic amplifier," U.S. Patent 4 990 803, Feb. 5,1991.

[22] R. R. Harrison, "A low-power, low-noise CMOS amplifier for neural recording applications," in *Proc. IEEE Int. Symp. Circuits and Systems*, vol. 5, pp. 197-200, 2002.

[23] R. R. Harrison and C. Charles, "A low-power low-noise CMOS amplifier for neural recording applications," IEEE J. Solid-State Circuits, vol. 38, no. 6, pp. 958–965, 2003.

[24] Moo Sung Chae, Zhi Yang, Mehmet R. Yuce, Linh Hoang, and Wentai Liu, " A 128-Channel 6 mW Wireless Neural Recording IC With Spike Feature Extraction and UWB Transmitter"' Ieee Transactions On Neural Systems And Rehabilitation Engineering, VOL. 17, pp. 312-321, 2009.

[25] Benoit Gosselin, Mohamad Sawan, C. Andrew Chapman," A Low-Power Integrated Bioamplifier With Active ,Low-Frequency Suppression," IEEE transactions on biomedical circuits and systems, vol. 1, pp. 184-192, 2007.

[26] Wei Zhao, Hongge Li*, Youguang Zhang, " A Low-Noise Integrated Bioamplifier with Active DC Offset Suppression", IEEE biomedical circuits and systems",pp 5-8,2009.

[27] Woradorn Wattanapanitch, Michale Fee, Rahul Sarpeshkar, "An Energy-Efficient Micropower Neural Recording Amplifier",'IEEE transactions on biomedical circuits and systems, vol. 1, no. 2, june 2007

[28] Ming Yin and Maysam Ghovanloo, "A Low-Noise Preamplifier with Adjustable Gain and Bandwidth for Biopotential Recording Applications". ieee symposiu on circuits and systems, pp.321-324, 2007.

[29] Reid R. Harrison, Paul T. Watkins, Ryan J. Kier, "A Low-Power Integrated Circuit for a Wireless 100-Electrode Neural recording system", ieee journal of solid-state circuits, vol. 42, pp. 2258-2267, 2007.

[30] Vahid Majidzadeh, Alexandre Schmid, Yusuf Leblebici, "Energy Efficient Low-Noise Neural Recording Amplifier With Enhanced Noise Efficiency Factor", ieee transactions on biomedical circuits and systems, vol. 5, no. 3, pp-136-147, 2011.

[31] Marion A. Hersh and Michael A. Johnson, "Hearing aid principles and technology"In:*Assistive Technology for the hearing impaired, Deaf and Deaf blind"*, London, Springer-Verlag, pp. 71- 115, 2003.

[32] Michael Valente, "*Hearing Aids: Standards, Options, and Limitation"s*, New York, Thieme Medical Publisher, 1996

[33] Zhang Liang, "Low Noise Amplifier Design And Noise Cancellation For Wireless Hearing Aids", National University Of Singapore, Singapore, pp. 1-103, 2005.

[34] Ram Singh Rana, Garg Hari Krishna, ZhangLiang and Tang Bin "Hearing Aid Devices A Few Selected Research Issues", *the 8th World Multi-Conference on Systemic, Cybernetics and Informatics*, pp.80-84, July 18-21, 2004, Orlando, Florida, USA.

[35] E. Niedermeyer and F. L. d. Silva, *"Electroencephalography: Basic Principles, Clinical Applications, and Related Fields"*, 4th Ed., Baltimore, Md: Williams & Wikins, 1999.

[36] B. Abou-Khalil and K. E. Misulis, *"Atlas of EEG & Seizure Semiology"*, Amsterdam, Elsevier Science, 2005.

[37] X. D. Zou, X. Y. Xu, L. B. Yao, and Y. Lian, "A 1-V 450-nW Fully Integrated Programmable Biomedical Sensor Interface Chip"*, IEEE Journal of Solid-State Circuits*, vol. 44, pp. 1067-1077, Apr. 2009.

[38] Shu-Di-Bao,Yuan-Ting Zhang and Lian-Feng sheh , "Physiological signal based entity auntentication for body area sensor networks and mobile health care systems",proc.of 2005 IEEE engineering in medicines and biology,September, pp. 1-4, 2005.

[39] Shih-Lun Chen,Ho-Yin Lee,Chiung-An Chen,Hong-Yi-Huang and Ching-Hsing Luo, "Wireless Body sensor network with adaptive low-power design for Biometric and healthcare applications", IEEE system journal,Vol.3,No.4,pp 398-409, december 2009.

[40] C.J.Deepu,X.Y.Xu,X.D.Zou and Y.Lian, "An ECGon- chip for wearable cardiac monitoring Devices", IEEE international symposium on electronic design,test and applications, pp 225-228,2010.

[41] Malti Bansal, Jyoti "A Review of Low Noise Amplifier for 2.4GHz Frequency Band", IEEE International conference on innovations in control, communication and information system (ICICCI),Greater Noida, pp.63-69, 12-13 august, 2017

[42] Malti Bansal, Jyoti "A Review of Various Applications of Low Noise Amplifiers", IEEE International conference on innovations in control, communication and information system (ICICCI),Greater Noida, 2017, pp.142-147, 12-13 august, 2017.

[43] Y.-H. Liu, X. Huang, M. Vidojkovic, G. Dolmans, and H. de Groot, "An Energy-Efficient Polar Transmitter for IEEE 802.15.6 Body Area Networks: System Requirements and Circuit Designs," IEEE Communications Magazine, vol. 49, pp. 426 - 438, October 2012.

[44] Robert F. Heile et.al, "IEEE Standard for local and metropolitan area networks - Part 15.6: Wireless Body Area Network", IEEE Computer Society, IEEE Std. 802.15.6, pp. 1-271, 2012, NY, USA.

[45] Malti Bansal, Jyoti "CMOS LNA for BLE applications"IETE International Conference on Science, Technology & Management (ICSTM-2017), Pune, pp. 654-661, 12 November 2017.

[46] Malti Bansal, Jyoti, "Utilizing CMOS Low-Noise Amplifier for Bluetooth Low Energy Applications" In: Malik H., Srivastava S., Sood Y., Ahmad A. (eds) Applications of Artificial Intelligence Techniques in Engineering. Advances in Intelligent Systems and Computing, vol 697. Springer, Singapore, pp. 239-251, 2019.

[47] M. Daoud, I. Student, H. Mnif, I. S. Member, and I. Member, "Resistive Termination Low Noise Amplifier for Bio-sensor Applications",2016 International design and test symposium, pp. 245–249, 2016.

[Downloaded free from http://www.jpbsonline.org on Sunday, June 2, 2019, IP: 14.139.251.100]

Original Article

# Management and Disposal of Mercury and Amalgam in the Dental Clinics of South India: A Cross-Sectional Study

*Karthik Krishna Ramesh, Maya Ramesh[1], Ramesh Krishnan[2]*

Department of Environmental Engineering, Delhi Technological University, [1]Department of Oral Pathology, Vinayaka Missions Sankarachariyar Dental College, Vinayaka Missions Research Foundation Deemed to be University, Salem, Tamil Nadu, [2]Registrar, Pedodontics, King Fahad Dental Centre, Armed Forces Hospitals Southern Region (AFHSR), Khamis Mushait, Kingdom of Saudi Arabia

**Abstract**

**Context:** Dental offices are known to be one of the largest users of inorganic mercury in the preparation of amalgam—a restorative material which, if not handled and disposed through scientific methods, can pose grave threats to the biosphere. **Aims:** The objective of this study was to assess and record the mercury management and disposal strategies of dental practitioners belonging to the two South Indian states, Kerala and Tamil Nadu. **Subjects and Methods:** A questionnaire regarding the usage and disposal of a filling material containing mercury (amalgam) was designed and distributed online. The 150 dental practitioners partaking in this study responded anonymously. **Results:** The results were statistically analyzed using chi-square test and the *P* value was evaluated. The usage of amalgam was correlated with the age of the practitioner and the nature of practice and it was found to be statistically significant. The number of fillings was correlated with the years of practice and locality of the clinic, which were found to be statistically significant. **Conclusions:** This study showed that mercury was widely preferred and utilized even today as a restorative material by dentists and dental specialists, despite the availability of other alternative strategies. Only a minor section of practitioners were found to be aware of the global changes in the guidelines pertaining to the handling and disposal of amalgam. We feel that Safe Mercury Amalgam Removal Technique, amalgam safety rules, and amalgam-free practice should be a part of the academic curriculum and continuing dental education.

**Key words:** *Biosphere hazards, dental amalgam, environmental strategies, occupational hazards, solid waste disposal*

## Introduction

Dental amalgam, a combination of silver alloy powder mixed with mercury, is one of the most preferred restorative materials among dentists because of its various advantages over other restorative materials. Amalgam has been used in dentistry for several decades owing to its low cost, durability, bacteriostatic effect, and strength. As a result, dental clinics are a major channel that serves as an outlet for mercury into the environment. The aim of this study was to assess and record the mercury management and disposal strategies of dental practitioners.

## Subjects and Methods

Institutional ethical committee clearance was obtained prior to the start of the study. This study was conducted among the general dentists and dental specialists practicing in South Indian states. An anonymous questionnaire regarding the usage and disposal of a filling material containing mercury (amalgam) was designed using Google Forms and circulated online by random sampling method in WhatsApp groups consisting of around 750 practitioners based out of Tamil Nadu and Kerala, India, out of which 150 responded to the survey within the stipulated deadline date. In the circulated questionnaire, participating practitioners who did not use silver amalgam consequently did not record their

Access this article online

Quick Response Code:

Website: www.jpbsonline.org

DOI: 10.4103/jpbs.JPBS_280_18

*Address for correspondence:* Mr. Karthik Krishna Ramesh, Delhi Technological University, Shahbad Daulatpur, Main Bawana Road, Delhi 110042, India. E-mail: karthik0197@gmail.com

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

**For reprints contact:** reprints@medknow.com

**How to cite this article:** Ramesh KK, Ramesh M, Krishnan R. Management and disposal of mercury and amalgam in the dental clinics of South India: A cross-sectional study. J Pharm Bioall Sci 2019;11:S151-5.

[Downloaded free from http://www.jpbsonline.org on Sunday, June 2, 2019, IP: 14.139.251.100]

responses to the questions pertaining to amalgam usage whereas general questions pertaining to mercury safety were answered by all the participants.

## Results

The responses thus received were compiled and tabulated in Microsoft Excel followed by statistical analysis using the SPSS software (version 11.5, IBM Corporation, Armonk, New York, USA). The data were subjected to chi-square test and the *P* value was evaluated to study the significance.

Of all the respondents, there were 13.33% practitioners belonging to the age group of younger than 25 years, 40% of practitioners belonging to the age group 26–35 years, 29.33% in the age group 36–45 years, and 17.33% older than of 45 years. The male to female ratio of the respondents was found to be 1.27 with 56% male and 44% females.

Of the participants of the survey, 50% were general practitioners whereas 42% were specialized in various branches of dentistry and 8% were endodontists—specialized in filling and conserving the decayed teeth.

When the total number of years in active dental practice was evaluated, 42.67% had less than 5 years of practice, 14.00% had 5–10 years of practice, and 22.67% had 11–20 years of practice whereas 20.67% had more than 20 years of clinical experience.

As far as the locality of the clinic of the respondents is concerned, rural practice accounted only for 9.33% of the participants, semi-urban practitioners were 46.67%, and urban practitioners making up 44% of the total practitioners.

Dental amalgam was used in their clinic by 57.33% of the partaking practitioners, and 42.67% of dentists refrained from dental amalgam in their practice.

Further, the usage of amalgam was correlated with age, sex, nature of practice, and experience of the practitioner and also with the locality of clinic. When amalgam usage was compared with age, the *P* value was 0.008 and hence was deemed significant. When amalgam usage was compared with the nature of the practice, the *P* value was 0.003, which is highly significant. With the other parameters tested, it was found to be not significant as in Table 1.

When the number of amalgam fillings performed per day by the dental practitioners was correlated with age, sex, nature of practice, and experience of the practitioner and also with the locality of clinic, it was observed that the experience of the practitioner on correlation generated a *P* value of 0.013, hence holding significance. With the other parameters mentioned earlier, it was found to be not significant as in Table 2

## Discussion

Despite the unparalleled strength and durability offered by dental amalgam over any other restorative material, health hazards posed by mercury present in amalgam to both the dental practitioner and mainly the patient, as well as the environmental threat arising from the

**Table 1: Comparison of the age, sex, nature of practice, years of practice, and locality of clinic with use of dental amalgam in clinic**

| | | Use dental amalgam in clinic | | | | Total | Chi-square test | *P* |
|---|---|---|---|---|---|---|---|---|
| | | No | | Yes | | | | |
| | | *N* | % | *N* | % | | | |
| Age in years | Up to 25 | 6 | 30 | 14 | 70 | 20 | 11.70 | 0.008* |
| | 26–35 | 25 | 42 | 35 | 58 | 60 | | |
| | 36–45 | 27 | 61 | 17 | 39 | 44 | | |
| | Above 45 | 6 | 23 | 20 | 77 | 26 | | |
| Sex | Male | 34 | 40 | 50 | 60 | 84 | 0.37 | 0.541 |
| | Female | 30 | 45 | 36 | 55 | 66 | | |
| Nature of practitioner | General practitioner | 24 | 32 | 51 | 68 | 75 | 11.66 | 0.003* |
| | Specialist | 37 | 59 | 26 | 41 | 63 | | |
| | Endodontist | 3 | 25 | 9 | 75 | 12 | | |
| Years of practice | Less than 5 | 24 | 38 | 40 | 63 | 64 | 5.34 | 0.149 |
| | 5–10 | 12 | 57 | 9 | 43 | 21 | | |
| | 11–20 | 18 | 53 | 16 | 47 | 34 | | |
| | More than 20 | 10 | 32 | 21 | 68 | 31 | | |
| Locality of clinic | Rural | 4 | 29 | 10 | 71 | 14 | 1.68 | 0.432 |
| | Semi-urban | 29 | 41 | 41 | 59 | 70 | | |
| | Urban | 31 | 47 | 35 | 53 | 66 | | |
| Total | | 64 | 43 | 86 | 57 | 150 | | |

*Significant at 5 %

[Downloaded free from http://www.jpbsonline.org on Sunday, June 2, 2019, IP: 14.139.251.100]

**Table 2: Comparison of the age, sex, nature of practice, years of practice, and locality of clinic with number of dental amalgam fillings performed per day in clinic**

| | | Number of dental amalgam fillings approximately performed in a day | | | | | | Total | Chi-square test | *P* |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Less than 5 | | 5–10 | | 11–20 | | | | |
| | | *N* | % | *N* | % | *N* | % | | | |
| Age in years | Up to 25 | 11 | 79 | 2 | 14 | 1 | 7 | 14 | 1.93 | 0.926 |
| | 26–35 | 25 | 71 | 5 | 14 | 5 | 14 | 35 | | |
| | 36–45 | 14 | 82 | 2 | 12 | 1 | 6 | 17 | | |
| | Above 45 | 16 | 80 | 3 | 15 | 1 | 5 | 20 | | |
| Sex | Male | 40 | 80 | 5 | 10 | 5 | 10 | 50 | 1.57 | 0.457 |
| | Female | 26 | 72 | 7 | 19 | 3 | 8 | 36 | | |
| Nature of practitioner | General practitioner | 35 | 69 | 11 | 22 | 5 | 10 | 51 | 7.55 | 0.109 |
| | Specialist | 22 | 85 | 1 | 4 | 3 | 12 | 26 | | |
| | Endodontist | 9 | 100 | | | | | 9 | | |
| Years of practice | Less than 5 | 31 | 78 | 6 | 15 | 3 | 8 | 40 | 16.21 | 0.013* |
| | 5–10 | 5 | 56 | | | 4 | 44 | 9 | | |
| | 11–20 | 13 | 81 | 3 | 19 | | | 16 | | |
| | More than 20 | 17 | 81 | 3 | 14 | 1 | 5 | 21 | | |
| Locality of clinic | Rural | 9 | 90 | 1 | 10 | | | 10 | 12.79 | 0.012* |
| | Semi-urban | 26 | 63 | 11 | 27 | 4 | 10 | 41 | | |
| | Urban | 31 | 89 | | | 4 | 11 | 35 | | |
| Total | | 66 | 77 | 12 | 14 | 8 | 9 | 86 | | |

*Significant at 5%

unscientific methods of disposal of extracted amalgam-containing tooth and other residue from the packaging of mercury, heavily outweigh the advantages it offers.

From the data compiled in our study, it can be surmised that 40% of the dental practitioners who participated in the survey were between the ages 26 and 35 years, which may be attributed to the fact that newer generation is technologically inclined. A lion's share of the responses was received from dentists in urban and semi-urban regions (90.67%). In stark contrast, only 9.33% were from the rural practitioners, which may be attributed to the fact that rural population lack technological awareness or have difficulty in accessing the Internet.

Of the practitioners giving the survey, 32% were of the opinion that amalgam poses no significant short-term or long-term health hazards to both the practitioner and the patient whereas another 27.33% were unsure about its safety or the lack thereof. This explains why dental amalgam is strongly preferred as a restorative material in South India by the dental practitioners. On the other side of the spectrum, 40.67% of the responses were that dental amalgam is unsafe for both patient and dentists, indicating that this can trigger a considerable number of mercury-free clinics in the region.

In a study conducted[1] in Saudi Arabia in 1996, it was found that though 75% of the dentists were aware of the existing controversy surrounding amalgam safety, 85% believed in its safety. Another study in Nigeria (Udoye and Aguwa[2]) reported that a greater number of specialists than general dentists (74.5%) affirmed that dental amalgam is safe, whereas more general dentists (14.9%) than specialists (4.3%) reported dental amalgam to be hazardous. Dental practitioners of our area had better awareness of amalgam compared to these studies.

Some of the possible reasons for 57.33% of the dental practitioners utilizing dental amalgam in their practice over other restorative materials might be the exceptional strength and durability it offers and also the fact that from an economic point of view, it is cheaper and affordable to a larger section of the society.

Our study revealed that of those who utilized amalgam, 27.33% used it in its powder and liquid form which involves the mixing of silver alloy powder and liquid mercury using mortar and pestle and squeezing them using thick cloth, which arguably is the most unsafe method of handling mercury because it can result in the expulsion of mercury droplets in the indoor environment—in turn leading to inhalation by the practitioner and patients. Of those who utilized amalgam, 14% resorted to the capsule method, which is the safest method wherein silver powder and mercury are forced together in a capsule by provision of a vibration. The remnants from this method can be

[Downloaded free from http://www.jpbsonline.org on Sunday, June 2, 2019, IP: 14.139.251.100]

recycled. Of the dental practitioners, 16% used both these forms of dental amalgam in their practice.

It can be surmised from the responses that 44% of the dental practitioners performed less than 5 dental amalgam fillings per day, 8% reported performing 5–10 fillings, and 5.33% performed 11–20 fillings per day. Findings from a similar study performed in Nigeria[3] reported that the majority of the practitioners (77.8%) reported to be performing less than 5 fillings per day whereas a study[4] at Pune, India, reported that 43% of dentists performed less than 5 dental amalgam fillings per day.

It is observed from this study that 19.33% of the total practitioners disposed amalgam remnants along with other routine wastes, 22% stored it in tight containers whereas only 16% disposed it separately. Studies conducted in Bangkok[5] and in Palestine[6] align with our findings, i.e., a major chunk of practitioners do not separate the wastes, resulting in it ending up as a part of regular refuse. It was reported in a study in the Indian capital city of New Delhi that dentists at a teaching hospital were not aware of biomedical waste management and needed training.[7] It was also reported that New Delhi releases approximately 51 kg of mercury as a result of improper disposal of amalgam annually, which is unregulated and reckless. Contrastingly, a particular study on similar lines in a dental school in Turkey[8] reported that in majority of the practices, waste collection protocols were being adhered to. The stark contrast here could be due to the fact that the study in Turkey was conducted in a dental school where exposure and awareness regarding current protocols are greater when compared to many of the other studies that were conducted among a population of practicing dentists.

In the case of extracted teeth containing dental amalgam fillings, 30.67% of dental practitioners disposed them along with routine wastes without any segregation, 12.67% kept them in closed containers while only 14% disposed of it separately. The fact that a large portion of the practitioners do not opt for waste segregation and scientific methods of mercury disposal highlights that they do not view it as a major hazard to health and environment. Only 16% of the practitioners reported the separate disposal of dental amalgam despite the presence of third party agencies in these states who segregate and dispose of these clinical wastes in a scientific manner.[9]

In our study, among the participating dental practitioners, 47% had heard of dental amalgam recycling whereas only 28% reported to being aware of the Safe Mercury Amalgam Removal Technique (SMART).

The SMART[10] as recommended by International Academy of Oral Medicine and Toxicology broadly involves reducing the risk of mercury exposure faced by the practitioner and patient by means of installing an amalgam separator to collect mercury amalgam waste; installation of an adequate filtration system such as a high-volume air filtration system in the clinic premises in order to remove the mercury vapor generated in the restorative procedure; provision of a suitable adsorbent such as charcoal slurry or chlorella to the patient to rinse and swallow; and making use of proper protective equipment for both the practitioner and patient such as face shields and hair/head coverings, non-latex nitrile gloves, a full-body, impermeable barrier and dental dams and saliva ejector to shield the patient from exposure and ingestion. In addition, it is also recommended that in case of removal of fillings, the dental amalgam should be sectioned and removed as chunks of largest possible size using a small-diameter carbide drill. The need for adhering to these standards is further emphasized in the dosimetry study,[11] which reported that dental practitioners were found to have on an average, urinary mercury levels that were four times that of normal. Consequently, dental practitioners were at an increased risk of suffering from kidney disorders.

The dental practitioners as well as their support staff should be made aware of the scientific guidelines of storing mercury and the mopping of mercury spills. Usage of fine mesh to trap amalgam waste as endorsed by ISO 14011 Environmental Audit Procedures is also effective in reducing the exposure. If the waste water emanating from the clinics contain a mercury concentration higher than 2000 ppm, treatment to reduce the mercury concentration is required before it is released into the general sewers. In addition, as recommended in the Food and Drug Administration guidelines, dental amalgam fillings should be used only in patients older than 6 years.[12] Our study indicates that 45% of the practitioners who gave the survey reported that they faced patient enquiry on the safety of dental amalgam fillings. The fact that a majority of the practitioners are either unsure about dental amalgam safety or feel that it poses no hazards increases the need for an increased awareness to be created among practitioners and in turn in patients.

## Conclusion

This study proves that silver alloy (mercury) dental amalgam is widely preferred as a restorative material by general practitioners as well as specialists despite the advent of hazard-free alternatives whereas a considerable chunk of practitioners who participated

[Downloaded free from http://www.jpbsonline.org on Sunday, June 2, 2019, IP: 14.139.251.100]

in the survey were of the opinion that dental amalgam was safe or were inconclusive of its safety. The matter at hand is further exacerbated by the unscientific disposal strategies followed by a considerable number of the practitioners, which can result in a wide spectrum of issues ranging from vapor ingestion to water table infiltration. Thus, it is imperative that the existing practitioners need to be made aware of the global changes in the guidelines pertaining to the handling and disposal of amalgam while the inclusion of SMART techniques, dental amalgam safety protocols, and the benefits of mercury-free practices should be made a part of the existing academic curriculum. Every individual has equal responsibility in saving and protecting the environment for the future generations.

### Acknowledgement

We would like to acknowledge the statistician Dr. Nantha Kumar who did the statistical analysis and dental practitioners who took part in this study.

### Financial support and sponsorship

Nil.

### Conflicts of interest

There are no conflicts of interest.

## References

1. Khairuldean N, Sadig WM. Amalgam safety and alternative restorative materials: A cross-sectional survey among dentists Saudi Dent J 1996;8:27-33.
2. Udoye C, Aguwa E. Amalgam safety and dentists' attitude: A survey among a subpopulation of Nigerian dentists. Oper Dent 2008;33:467-71.
3. Ajayi DM, Iwaola MO, Abiodun-Solanke IF. Perceived health effects of dental amalgam among dentists in selected South Western State of Nigeria. Oral Hyg Health 2015;3:175.
4. Srinidhi SR, Ranadive N. Mercury hygiene practices followed in dental clinics in Pune: A survey. World J Dent 2013;4:92-5.
5. Punchanuwat K, Drummond BK, Treasure ET. An investigation of the disposal of dental clinical waste in Bangkok. Int Dent J 1998;48:369-73.
6. Darwish RO, Al-Khatib IA. Evaluation of dental waste management in two cities in Palestine. East Mediterr Health J 2006;12(Suppl 2):S217-22.
7. Kishore J, Goel P, Sagar B, Joshi TK. Awareness about biomedical waste management and infection control among dentists of a teaching hospital in New Delhi, India. Indian J Dent Res 2000;11:157-61.
8. Ozbek M, Sanin FD. A study of the dental solid waste produced in a school of dentistry in Turkey. Waste Manag 2004;24:339-45.
9. Mumtaz R, Ali Khan A, Noor N, Humayun S. Amalgam use and waste management by Pakistani dentists: An environmental perspective. East Mediterr Health J 2010;16:334-9.
10. The Safe Mercury Amalgam Removal Technique (SMART). https://iaomt.org/resources/safe-removal-amalgam-fillings/ [Last accessed on 2019 February 28]
11. Ritchie KA, Gilmour WH, Macdonald EB, Burke FJ, McGowan DA, Dale IM, *et al*. Health and neuropsychological functioning of dentists exposed to mercury. Occup Environ Med 2002;59:287-93.
12. Ngim C, Ngim AD. Health and safety in the dental clinic—Hygiene regulations for use of elemental mercury in the protection of rights, safety and well-being of the patients, workers and the environment. Singapore Dent J 2013;34:19-24.

Wireless Personal Communications
https://doi.org/10.1007/s11277-019-06551-3

# Performance Evaluation of QAM for Improvement of BER

**Balwant Singh**[1] · **Vinod Shokeen**[2] · **Rishi Asthana**[3]

© Springer Science+Business Media, LLC, part of Springer Nature 2019

**Abstract**
This paper presents performance evaluation of the mary quadrature amplitude modulation (M-QAM). BER performance is evaluated on operation of The Binary coding, The Gray coding, Pulse shaping in raised cosine filter and Error correction using convolution codes. Simulation results shows the BER and SNR of M-QAM (M = 16, 64, 256 and 1024) for oversampling factor 1, 2, 3, 4 and 5. BER and SNR is related to spectral efficiency (SE) i.e. SE is the optimize use of spectrum for highest quantity of information can be transmitted with the least transmission errors. SNR and BER will increase as order of QAM increases. Improvement in BER shows that the increment in the number of errors.

**Keywords** M-ary quadrature amplitude modulation · Bit error rate · Spectral efficiency · Signal to noise ratio

## 1 Introduction

Modulation is basic of all fundamentals in wireless and mobile communication. Modulation is the process of impressing data to be transmitted on the radio carrier. Now days, aim of modulation is to squeeze much data into the least amount of spectrum. How hurriedly data can be transmitted in an assigned bandwidth, can be measured by spectral efficiency (b/s/Hz). The symbols can transmit more bits per symbol if they have some combination of amplitude and phase. Quadrature amplitude modulations (QAM) have this property. QAM is efficient of spectrum but difficult to demodulate in the presence of noise. Basically QAM have two orthogonal pulse amplitude modulation (PAM) so that spectral efficiency going to be almost double [1].

✉ Balwant Singh
balwant.23apr@gmail.com

Vinod Shokeen
vshokeen@amity.edu

Rishi Asthana
asthanarishi1973@gmail.com

[1] Department of ECE, ASET (AMITY University), Noida, India

[2] Amity School of Engineering, Noida, India

[3] IMS Engineering College, Ghaziabad, India

Published online: 15 May 2019

Establishment of a uniform metric for spectrum efficiency usage is not possible because of wide range of services and users for which spectrum is needed [2]. Basically there are six categories for application of different efficiency metrics and these categories are: Satellite broadcast systems, point to point satellite systems, terrestrial broadcast systems, terrestrial personal communication systems, terrestrial point to point systems and terrestrial hybrid system [2].

Now day's traffic of communication system increases. So that higher order modulation technique is required for the same purpose. But the problem is higher order modulation technique not provide efficient transmission for long distance due to optical signal to noise ratio (OSNR) [3]. Due to efforts of researchers various modulation techniques can provide high data rate, improved system capacity, more exposure, bandwidth flexible operations, improved spectral efficiency, short round trip time, small dormancy, cheap operation cost, multi antenna provision, smooth incorporation with internet and existing generation of communication systems [4]. Orthogonal property in communication system can provide higher data rate, improved spectral efficiency and power reduction [1].

This paper provides performance of 16QAM, 64QAM, 256QAM and 1024QAM in terms of BER (bit error rate). The binary coding scheme, the gray coding scheme, pulse shaping using a raised cosine filter technique and error correction using convolution coding, are operations where M-QAM will be implemented to show the BER on the oversampling factor 1, 2, 3, 4 and 5. Here SNR is also evaluated for 16QAM, 64QAM, 256QAM and 1024QAM for oversampling factor 1, 2, 3, 4 and 5. BER for wireless communication and wire line communication is:

$$\text{BER for wireless communication: } \frac{1}{2SNR} \tag{1}$$

$$\text{BER for wire line communication: } Q\sqrt{SNR} \tag{2}$$

$$\text{where } Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{-\frac{x^2}{2}} dx \tag{3}$$

## 2 Bit Error Rate

The ratio of 'Number of errors and total number of bits sent' is known as BER.

$$BER = \frac{Number\ of\ Errors}{Total\ number\ of\ bits\ transmitted} \tag{4}$$

BER of M-QAM (Rectangular QAM) can be calculated by:

$$BER = 1 - \left(1 - P_{bc}\right)^2 \tag{5}$$

where

$$P_{bc} = \frac{4}{k}\left(1 - \frac{1}{\sqrt{M}}\right) Q\left(\sqrt{\frac{3k}{M-1} \cdot \frac{E_b}{N_0}}\right) \tag{6}$$

and

$$k = log_2 M \tag{7}$$

As order of QAM increases bandwidth and power efficiency also improves. 4QAM, 16QAM, 64QAM, 256QAM, 1024QAM and 4096-QAM have power efficiency 1, 2, 3, 4, 5 and 6 respectively which shows improvement as order of QAM increases. Again $E_b/N_0$, for BER = $10^{-6}$ will be 10.5, 15, 18.5, 24, 28 and 33.5 for M-QAM (M = 4, 16, 64, 256, 1024 and 4096) respectively. Certain BER will be produced relative to SNR on each modulation scheme.

$$SNR = 10log_{10}\frac{Signal\ Power}{Noise\ Power} \tag{8}$$

Mathematically SNR can be calculated by:

$$SNR = \frac{E_b}{N_0} + 10log_{10}\left(\frac{Number\ of\ Bits\ Per\ Symbol}{Over\ Sampling\ Factor}\right) \tag{9}$$

In oversampling, the signals are sampled higher than Nyquist rate. Oversampling can improve resolution, reduce noise, avoid aliasing and phase distortion.

Physical layer protocol can correct transmitted bits by using coding known as forward error correction (FEC). Higher bit rate required more coding so more overhead required. Efficient modulation scheme is more beneficial than coding scheme for improvement in BER [5]. The system cannot recover the whole codeword if number of error is more than the error correcting capability. For this situation BER will degrade but interleaving procedure can reduce this problem [3]. IFFT technique for OFDM [6], use of band edge subcarriers of LTE system [6], use of modified regularized least sequence channel estimation [6], bit level scrambling in MIMO-OFDM system [7], antenna selection in MIMO system [8–10], use of ordered SCM (sub carrier multiplexing) [11], cancellation of inter cell interference in MIMO [12], proper selection of IMF (intrinsic mode function) [13], use of RZ-DPSK receivers [14] and many more techniques are able to improve performance of various systems in terms of BER.

If W is bandwidth of assigned channel, S is area of zone cluster (or cell cluster) which determine frequency reuse unit, $n_a$ is number of channels within a zone cluster and $a_c$ is call traffic/channel. Overall spectrum efficiency is expressed using spatial efficiency ($\eta_s$), temporal efficiency ($\eta_t$) and band efficiency ($\eta_b$) [15]:

$$\eta_W = \eta_S.\eta_t.\eta_b \tag{10}$$

Spectral efficiency is indicated as spatial repeat pattern of the same channel frequency, temporal efficiency is indicated by active time/channel ratio and band efficiency defines the number of band/channel. Where

$$\eta_s = \frac{1}{S}\left(1/\mathrm{m}^2\right) \tag{11}$$

$$\eta_t = a_c\ (\mathrm{erl/channel}) \tag{12}$$

$$\eta_b = \frac{n_a}{W}\ (\mathrm{channel/Hz}) \tag{13}$$

and

$$\eta_W = \frac{n_a a_c}{SW} \left(\text{erl/Hz m}^2\right) \tag{14}$$

$$\begin{aligned} \textit{Capacity of cellular coverage area} = (\textit{number of sectors/cell site}) \cdot (\textit{number of cells in sites}) \\ \cdot (\textit{amount of spectrum}) \cdot (\textit{spectral efficiency}) \end{aligned} \tag{15}$$

Small cells can improve cellular spectrum but cost is problem. Each small cell needs physical location with mounting, power and backhaul to the core network [1].

Link adaptation method [16], Soft resource reservation mechanism [17], Fog radio access network (F-RAN) [5], Reduction of spectrum partition factor [18], RF energy harvesting [19], Received signal strength (RSS) threshold [14, 20], Direct D2D discovery [21], and many more, are the methods which is used to improve the SE. BER performance is directly proportional to SE improvement so that this paper is going to show that performance evaluation of BER in various conditions and parameters.

## 3 Quadrature Amplitude Modulation

QAM is adaptive modulation technique. SE can improve by applying higher order QAM on communication system. Better SNRs (signal-to-noise ratios) required to overcome interference. This interference will be generated because of less inter-symbol distance in constellation as order of QAM increases. M-ary QAM, the transmitted signal is represented by [22]:

$$s_i = a_i \sqrt{\frac{2}{T_s}} \cos \omega_c t + b_i \sqrt{\frac{2}{T_s}} \sin \omega_c t \tag{16}$$

where$a_i \sqrt{\frac{2}{T_s}} \cos \omega_c t$ and $b_i \sqrt{\frac{2}{T_s}} \sin \omega_c t$ is basis functions for two dimensional space $\varphi_1(t)$ and $\varphi_2(t)$.

$$a_i = \pm\frac{d}{2}, \pm\frac{3d}{2}, \ldots\ldots \pm \frac{\left(\sqrt{M}-1\right)d}{2} \tag{17}$$

$$b_i = \pm\frac{d}{2}, \pm\frac{3d}{2}, \ldots\ldots \pm \frac{\left(\sqrt{M}-1\right)d}{2} \tag{18}$$

Conditional bit error probability of M-QAM is [23]:

$$P_b(k) = \frac{1}{\sqrt{M}} \sum_{j=0}^{(1-2^{-k})\cdot\sqrt{M}-1} \left[ (-1)^{\left\lfloor \frac{j\cdot 2^{k-1}}{\sqrt{M}} \right\rfloor} \cdot \left( 2^{k-1} - \frac{j\cdot 2^{k-1}}{\sqrt{M}} - 1 \right) \cdot erfc\left( (2\cdot j+1)\sqrt{\frac{3\left(log_2 M\right)\cdot r}{2(M-1)}} \right) \right] \tag{19}$$

$P_b(k)$ can be extanded by $j = 0, 1, 2, \ldots\ldots..\left(1-2^{-k}\right)\cdot\sqrt{M}-1$. where $\left(1-2^{-k}\right)\cdot\sqrt{M}-1$ is the actual number of term per conditional probability. Floor function $x$ is largest integer not greater than x. Now BER for M-QAM, where M is the square number and $M = 2^N$ is:

$$P_b = \frac{1}{log_2\sqrt{M}} \sum_{k=1}^{log_2\sqrt{M}} P_b(k) \tag{20}$$

For long distances lower modulation required (example: BPSK) but for closer distance higher order modulation (example: M- QAM) could be utilized to achieve higher throughput. So D2D (Device to device) communication technology can effectively implement M-QAM [24]. For the video and data modulation in cable system, 256-QAM can be used. 256 QAM provide 33% more efficient bandwidth than 64-QAM but 256-QAM has higher SNR and increased sensitivity with phase noise and interference [24].

BER expressions for digital systems in terms of SNR can also be presented by:

$$P_e(64QAM) = \frac{7}{12}Q\left(\frac{SNR}{21}\right)^{1/2} \tag{21}$$

$$P_e(256QAM) = \frac{5}{32}Q\left(\frac{SNR}{85}\right)^{1/2} \tag{22}$$

$$P_e(1024QAM) = \frac{31}{160}Q\left(\frac{SNR}{34121}\right)^{1/2} \tag{23}$$

As signal power is related to energy per symbol as $E_s = P_s \cdot T_s$ and $E_b = \frac{E_s}{log_2 M}$.

So

$$P_e(64QAM) = \frac{7}{12}Q\left(\frac{2E_b}{7N_0}\right)^{1/2} \tag{24}$$

$$P_e(256QAM) = \frac{15}{32}Q\left(\frac{8E_b}{85N_0}\right)^{1/2} \tag{25}$$

$$P_e(1024QAM) = \frac{31}{160}Q\left(\frac{10E_b}{341N_0}\right)^{1/2} \tag{26}$$

where $N_0$ is Noise power density and Q() is oft-tabulated function associated with the solution to the integral, under the upper tail of Gaussian probability density function (PDF), which is accounting statistical nature of AWGN [25]. If symbol rate is $f_s$ for a QAM than approximated RF bandwidth will be $f_s(1+\alpha)$, where $\alpha$ is the filter roll-off factor [26].

Raised cosine filter is providing lower ISI and determine the excess bandwidth occupies beyond the nyquist bandwidth (roll-off factor) [26]. So raised cosine filtering operation used for analysis of BER with QAM. As order of QAM increases, number of bits per symbol and efficiency increases:

## 4 Simulation and Results

Parameters for simulation are described below:

**Modulation Type:** 16QAM, 64QAM, 256QAM and 1024 QAM

**Operations for Simulation:** The Binary Coding, The Gray Coding, Pulse shaping using Raised Cosine filter and Error correction using Convolution Code

**Number of bits to Process:** 900,000

**Table 1** BER of M-QAM on the binary coding for various over sampling factor

| Over-sampling factor | 16QAM | | 64QAM | | 256QAM | | 1024QAM | |
|---|---|---|---|---|---|---|---|---|
| | BER | No of errors | BER | No of errors | BER | No of errors | BER | No of errors |
| 1 | 1.79e−03 | 1613 | 2.66e−02 | 23,935 | 7.86e−02 | 70,700 | 1.34e−01 | 120,212 |
| 2 | 1.73e−02 | 15,529 | 6.79e−02 | 61,092 | 1.24e−01 | 111,701 | 1.83e−01 | 164,739 |
| 3 | 3.86e−02 | 34,750 | 9.70e−02 | 87,266 | 1.55e−01 | 139,775 | 2.11e−01 | 189,984 |
| 4 | 5.91e−02 | 53,179 | 1.19e−01 | 107,361 | 1.79e−01 | 160,761 | 2.31e−01 | 207,819 |
| 5 | 7.76e−02 | 69,821 | 1.37e−01 | 123,569 | 1.97e−01 | 177,720 | 2.46e−01 | 221,197 |

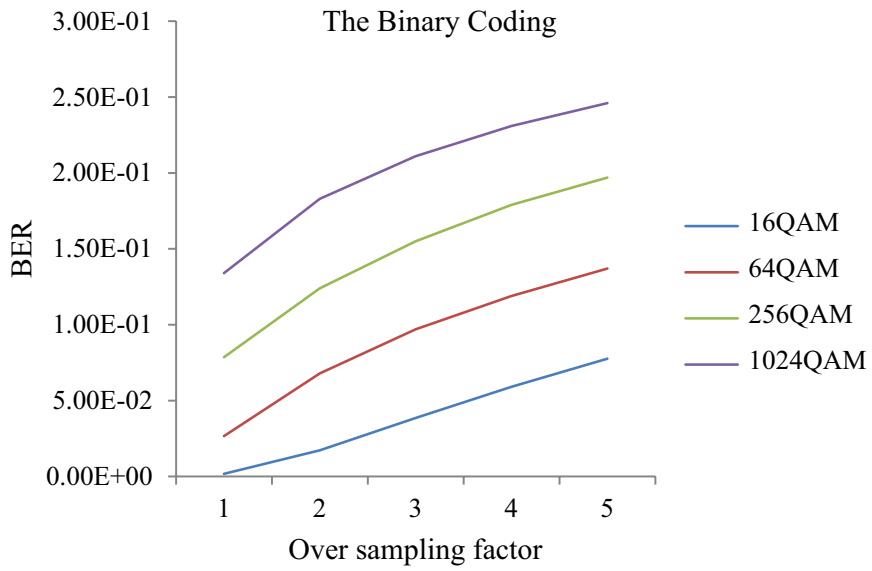



**Fig. 1** BER of M-QAM on the binary coding for various over sampling factor

**Oversampling Factor:** 1, 2, 3, 4 and 5
**Noise:** White Gaussian Noise
**The ratio of bit energy to noise power spectral density ($E_b/N_0$):** 10 dB
**Binary coding Phase offset**: 0
**Gray coding Phase offset:** 0
**Filter span in Filter:** 10
The simulation result shows on the various operations with presented parameters.

### 4.1 The Binary Coding

The digital data is represented, stored and transmitted as group of binary bits which is known as binary code. Simulation results for the binary coding are shown in Table 1 and Fig. 1. BER is increasing as values of over sampling factor increases for M-QAM. So performance of the system will decrease as oversampling factor and order of QAM increases.

### 4.2 The Gray Coding

The gray code is also known as a reflected code or the binary reflected gray code. BER performance is improving for gray coding in M-QAM rather than binary coding. It is an encoding of numbers where adjacent numbers have a single digit differing by 1. Simulation results shows BER is increasing as over-sampling factor increases for QAM. One symbol can cause one or more bit error. Symbol error is generated due to confusion of its closest neighbors. Coding Gray makes sure that those symbol error cause only one bit error. So, Gray coding minimizes the BER for given symbol error rate.

### 4.3 Pulse Shaping Using Raised Cosine Filter

The raised cosine filter uses pulse shaping in digital modulation for minimization of ISI. A pair of square root raised cosine (RRC) filters applied in the simulation to perform pulse shaping and matched filtering. Here BER is improving as over-sampling factor increases from 1 to 2. But over-sampling factor 3, 4 and 5 have constant BER for every M-QAM. Simulation result also shows that the BER is improving as order of QAM increase. Again performance of the system, which have pulse shaping using raised cosine filter, will degrade on the implementation of higher order QAM and higher value of over-sampling factor.

### 4.4 Error Correction Using Convolution Code

In convolution code, the block code of n digit code does not depend on only k messages digits within that time unit but also on the data digits within a previous span [26]. This code is used for correcting random errors, burst errors, or both. Here error correction is performed by convolution coding by introducing forward error correction (FEC) to improve BER performance. Simulation result (Fig. 4, Table 4) shows constant BER for value of over-sampling factor 2, 3, 4 and 5 for 16QAM, 64QAM, 256QAM and 1024QAM.

### 4.5 SNR

SNR depends on order of QAM and over sampling factor. SNR is increasing as order of QAM is increases. More number of bits can be transmit by applying higher order QAM in communication system. Increment in over-sampling factor will reduce the value of SNR for given $E_b/N_0$ (Eq. 9). Simulation result shows the best result for 1024QAM for oversampling factor 1 (Fig. 5 and Table 5).

## 5 Conclusions

As order of the QAM increases, the performance of BER evaluated for given operations and fading conditions i.e. The binary coding, The Gray coding, Pulse shaping using Raised cosine filter and Error correction using convolution code. BER is the ratio 'number of errors and total number of bits sent' (Eq. 4). Improvement in the BER means, number of errors increases which is shown in simulation result for M-QAM (M = 16,

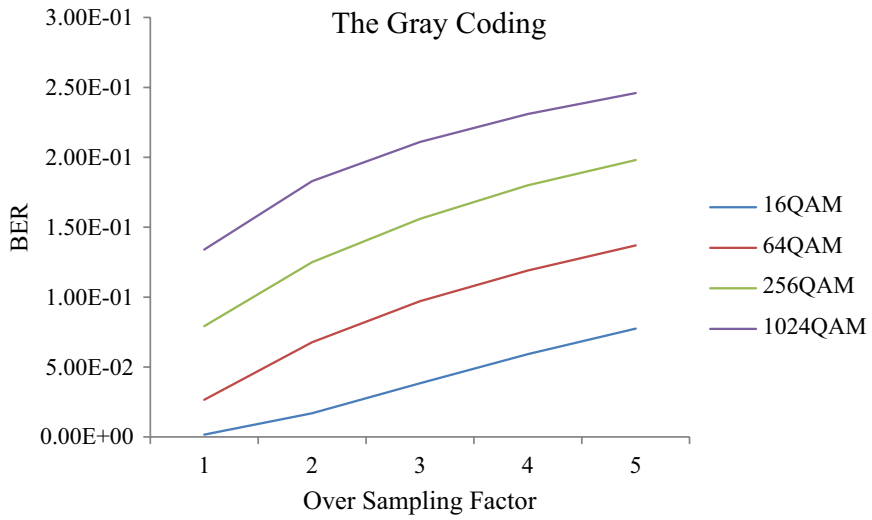



**Fig. 2** BER of M-QAM on the gray coding for various over sampling factor

**Table 2** BER of M-QAM on the gray coding for various over sampling factor

| Over-sampling factor | 16QAM | | 64QAM | | 256QAM | | 1024QAM | |
|---|---|---|---|---|---|---|---|---|
| | BER | No of errors | BER | No of errors | BER | No of errors | BER | No of errors |
| 1 | 1.70e−03 | 1534 | 2.66e−02 | 23,962 | 7.92e−02 | 71,273 | 1.34e−01 | 120,339 |
| 2 | 1.70e−02 | 15,339 | 6.78e−02 | 61,006 | 1.25e−01 | 112,565 | 1.83e−01 | 164,619 |
| 3 | 3.84e−02 | 34,552 | 9.71e−02 | 87,417 | 1.56e−01 | 140,382 | 2.11e−01 | 190,317 |
| 4 | 5.92e−02 | 53,240 | 1.19e−01 | 107,230 | 1.80e−01 | 161,630 | 2.31e−01 | 208,082 |
| 5 | 7.75e−02 | 69,736 | 1.37e−01 | 123,403 | 1.98e−01 | 178,579 | 2.46e−01 | 221,598 |

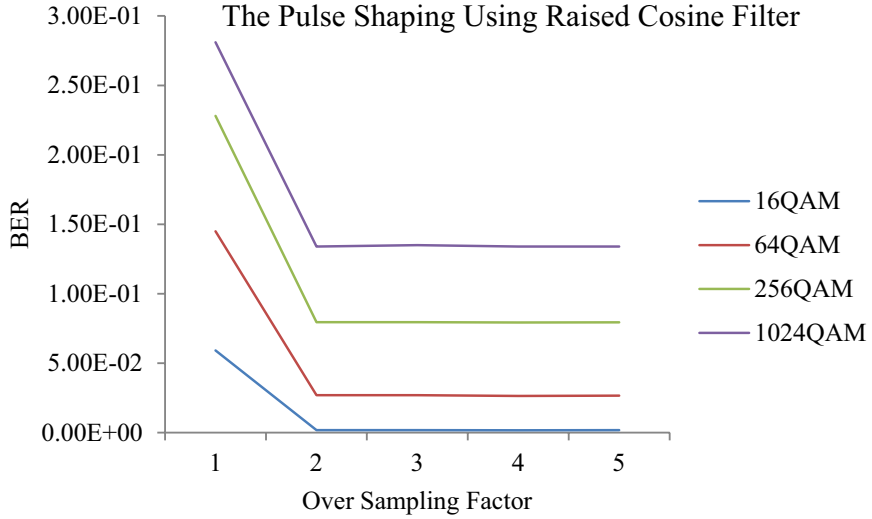



**Fig. 3** BER of M-QAM on the pulse shaping using raised cosine filter for various over sampling factor

64, 256 and 1024) and over sampling factor (1, 2, 3, 4, 5) in the Binary coding and Gray coding operation for given parameters (Fig. 1, Table 1 and Fig. 2, Table 2 respectively). Again Simulation results (Fig. 3, Table 3 and Fig. 4, Table 4) of Pulse shaping using

**Table 3** BER of M-QAM on the pulse shaping using raised cosine filter for various over sampling factor

| Over-sampling factor | 16QAM | | 64QAM | | 256QAM | | 1024QAM | |
|---|---|---|---|---|---|---|---|---|
| | BER | No of errors | BER | No of errors | BER | No of errors | BER | No of errors |
| 1 | 5.92e−02 | 53,298 | 1.45e−01 | 130,926 | 2.28e−01 | 205,346 | 2.81e−01 | 252,916 |
| 2 | 1.80e−03 | 1617 | 2.70e−02 | 24,303 | 7.96e−02 | 71,609 | 1.34e−01 | 120,987 |
| 3 | 1.79e−03 | 1608 | 2.70e−02 | 24,303 | 7.95e−02 | 71,558 | 1.35e−01 | 121,064 |
| 4 | 1.77e−03 | 1590 | 2.64e−02 | 23,778 | 7.93e−02 | 71,385 | 1.34e−01 | 120,820 |
| 5 | 1.81e−03 | 1632 | 2.67e−02 | 24,056 | 7.94e−02 | 71,446 | 1.34e−01 | 120,769 |

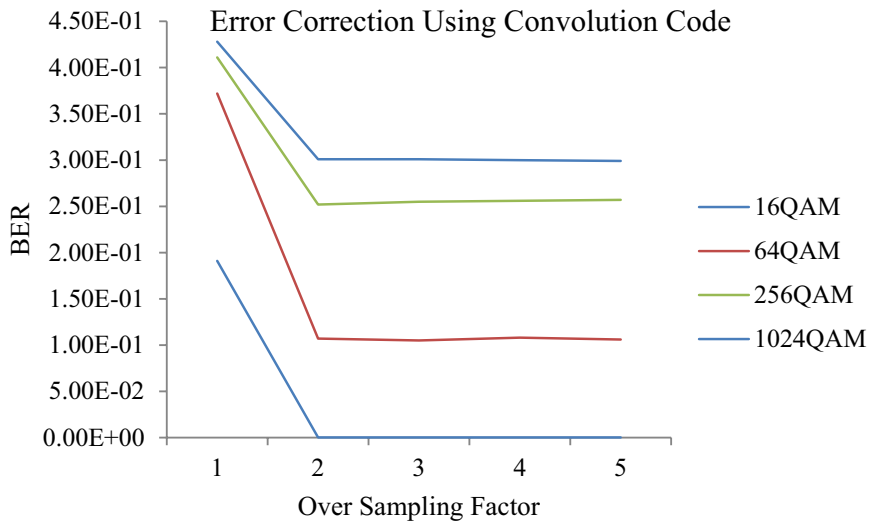



**Fig. 4** BER of M-QAM on the error correction using convolution codes for various over sampling factor

**Table 4** BER of M-QAM on the error correction using convolution code for various over sampling factor

| Over-sampling factor | 16QAM | | 64QAM | | 256QAM | | 1024QAM | |
|---|---|---|---|---|---|---|---|---|
| | BER | No of errors | BER | No of errors | BER | No of errors | BER | No of errors |
| 1 | 1.91e−01 | 172,323 | 3.72e−01 | 334,823 | 4.11e−01 | 370,061 | 4.28e−01 | 384,071 |
| 2 | 1.04e−04 | 94 | 1.07e−01 | 95,905 | 2.52e−01 | 226,641 | 3.01e−01 | 269,852 |
| 3 | 1.01e−04 | 91 | 1.05e−01 | 94,721 | 2.55e−01 | 229,142 | 3.01e−01 | 270,093 |
| 4 | 8.89e−05 | 80 | 1.08e−01 | 97,006 | 2.56e−01 | 230,590 | 3.00e−01 | 269,777 |
| 5 | 1.14e−04 | 103 | 1.06e−01 | 95,368 | 2.57e−01 | 230,849 | 2.99e−01 | 268,758 |

raised cosine filter and error correction using convolution code shown for the same parameters. Here simulation result shows, number of error is increases as order of QAM increases but number of error decreasing as oversampling factor increases. If fact, BER is constant for 16QAM, 64QAM, 256QAM and 1024QAM for the value of oversampling factor 2, 3, 4, and 5. Increment in the value of over-sampling factor will reduce SNR for given $E_b/N_0$ (Eq. 9).

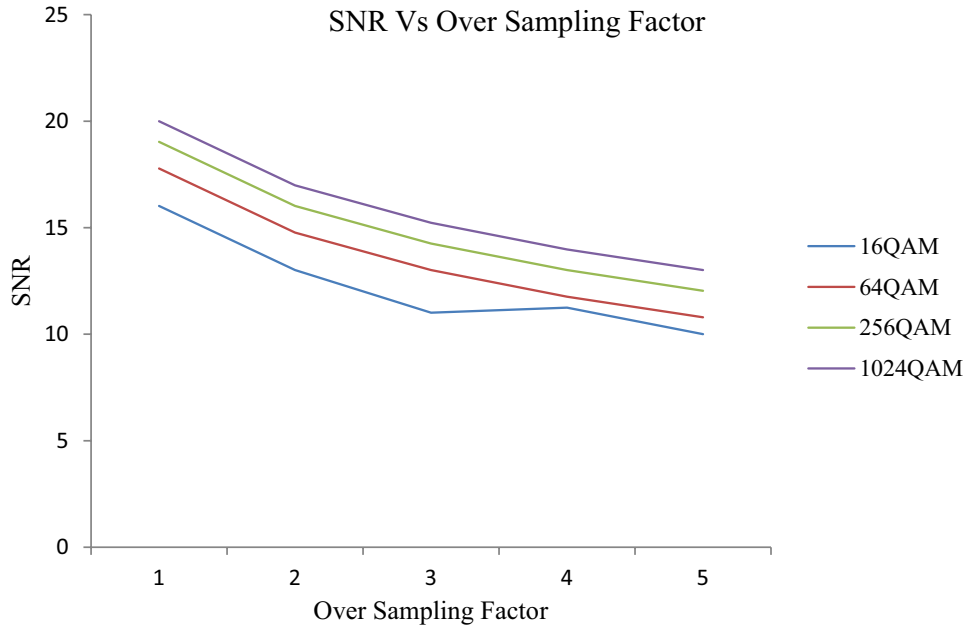



**Fig. 5** SNR of M-QAM for various over sampling factor

**Table 5** SNR of M-QAM for various over sampling factor

| Over-sampling factor | 16QAM | 64QAM | 256QAM | 1024QAM |
|---|---|---|---|---|
| 1 | 16.0206 | 17.7815 | 19.0309 | 20 |
| 2 | 13.0103 | 14.7712 | 16.0206 | 16.9897 |
| 3 | 11.0103 | 13.0103 | 14.2597 | 15.2288 |
| 4 | 11.2494 | 11.7609 | 13.0103 | 13.9794 |
| 5 | 10 | 10.7918 | 12.0412 | 13.0103 |

After the simulation, we get improvement in the SNR, as order of QAM increases for every value of over-sampling factor (Fig. 5, Table 5). So, in the communication, more number of bits can be transmitted on higher order QAM but BER will in increase due to more number of errors generated.

## References

1. Torres, J. J. G., Soto, A. M. C., & González, N. G. (2014). Evaluation and compensation of interchannel interference effects in a 16-QAM Nyquist-WDM system with LMS equalization. In *IEEE Latin America conference on communications (LATINCOM)* (pp. 1–6).
2. Sheikh, J. A., Uzma, Parah, S. A. G., & Bhat, M. (2015). Bit error rate (BER) improvement of multiple input multiple output orthogonal frequency division multiplexing (MIMO-OFDM) system using bit level scrambling. In *Annual IEEE India conference (INDICON)* (pp. 1–6).
3. Toor, G. S., Singh, H., & Bhandari, A. S. (2014). PAPR reduction and BER improvement by using logarithmic companding hybrid with SLM technique in bit interleved COFDM system. In *5th international conference: confluence the next generation information technology summit (confluence)* (pp. 604–608).
4. Araújo, D. C., Lucena, A. M. P., & Mota, J. C. M. (2014). Successive interference cancellation algorithm in M-QAM non orthogonal multicarrier systems. In *International telecommunications symposium (ITS)* (pp. 1–5).

5. Rysavy, P. (2014). Challenges and consideration in defining spectral efficiency. *Proceeding of IEEE, 102,* 386–392.
6. Li, C.-M., & Lai, J.-H. (2014). BER performance improvement for the band-edge subcarriers of a LTE system. In *International conference on information science, electronics and electrical engineering (ISEEE)* (Vol. 2, pp. 1262–1266).
7. Gangwar, V., & Chandra, D. (2013). Capacity and BER improvement through antenna selection in MIMO systems. In *International conference on optical imaging sensor and security (ICOSS)* (pp. 1–5).
8. Grover, A. (2013). BER improvements with multi-antennas in maximum likelihood decoder. In *Fifth international conference on computational intelligence, modelling and simulation (CIMSim)* (pp. 329–332).
9. Kocan, E., & Pejanovic-Djurisic, M. (2013). BER performance improvement for OFDM DF infrastructure relay links. In *16th international symposium on wireless personal multimedia communications (WPMC)* (pp. 1–5).
10. Patra, A., Sahu, P. K., Parija, S., & Mohapatra, B. (2012). Interference cancellation and BER improvement in a wireless network using network MIMO. In *Students conference on engineering and systems (SCES)* (pp. 1–6).
11. Han, S., Cheng, X., Wu, G., & Yue, G. (2011). A novel transmit beam forming scheme for BER improvement in EIRP constrained UWB systems. In *International conference on wireless communications and signal processing (WCSP)* (pp. 1–5).
12. Xu, J., Wang, Y., & Song, C. (2010). EMD-based BER improvement of TH-PPM UWB signal in AWGN channel. In *6th international conference on wireless communications networking and mobile computing (WiCOM)* (pp. 1–3).
13. Chaouch, H., Parsons, E., & Kueppers, F. (2010). BER improvement for RZ-DPSK receivers using saturated asymmetric filtering. In *Conference on optical fiber communication (OFC/NFOEC), collocated national fiber optic engineers conference* (pp. 1–3).
14. Khina, A., Pattersson, G. M., Kostina, V., & Hassibi, B. (2016). Multi-rate control over AWGN channels via analog joint source channel coding. In *IEEE 55th conference on decision and control (CDC)* (pp. 5968–5973).
15. Youssef, T., & Abdelfattah, E. (2013). Performance evaluation of different QAM techniques using Matlab/Simulink. In *IEEE long Island systems, applications and technology conference (LISAT)* (pp. 1–5).
16. Howald, R. L. (2002). QAM bulks up once again: Modulation to the power of ten. Proc. of SCTE Cable-tec EXPO. http://broadband.motorola.com/ips/pdf/QAM.pdf
17. Mohammed, U. R., & Habeeb (2014). QAM is rising: 1024 and beyond. https://www.mwf.com/systems/qam-rising-1024qam-and-beyond
18. Kwon, Y., Suh, H., Oh, J., & Hwang, T. (2017). Energy efficient communication for secure D2D underlaid cellular networks. *IEEE Transactions on Vehicular Technology, 66,* 9110–9123.
19. Condoluci, M., Mahmoodi, T., Steinbach, E., & Dohler, M. (2017). Soft resource reservation for low delayed teleoperations over mobile networks. *IEEE Access, 5,* 10445–10455.
20. Yan, S., Peng, M., Abana, M. A., & Wang, W. (2017). An evolutionary game for user access mode selection in fog radio access networks. *IEEE Access, 5,* 2200–2210.
21. Atat, R., Liu, L., Ashdown, J., Medley, M., & Matyjas, J. (2016). On the performace of relay assisted D2D networks under spatial correlated interference. In *IEEE global communication conference (GLOBECOM)* (pp. 1–6). https://doi.org/10.1109/GLOCOM.2016.7842060
22. Lathi, B. P., & Ding, Z. (2009). *Modern digital and analog communication systems* (4th ed.). New York: Oxford University Press.
23. Sicat, R. B., & Al-Naffouri, T. Y. (2009). Bit error probability computations for M-ar quadrature amplitude modulation. In *EE 242 Digital Communication and Coding*. King Abdullah University of Science and Technolgy. https://pdfs.semanticscholar.org/47e6/36eee36f76c2d1ee82b0d6fd5696dc32a040.pdf
24. Atat, R., Liu, L., Ashdown, J., Medley, M., Matyjas, J., & Yi, Y. (2016). Improving spectral efficiency of D2D cellular networks through RF energy harvesting. In *IEEE global communications conference (GLOBECOM)* (pp. 1–6). https://doi.org/10.1109/GLOCOM.2016.78411890
25. Yadv, A., Trivedi, A. (2016). Analysis of area spectral efficiency and energy efficiency based on RSS-threshold in device to device communication. In *International conference on signal processing and communication (ICSC)* (pp. 10–15).
26. Zhang, J., Deng, L., Li, X., Zhou, Y., Liang, Y., & Liu, Y. (2017). Novel device to device discovery scheme based on random backoff in LTE-Advance networks. *IEEE Transaction on Vehicular Technology, 66,* 11404–11408.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Balwant Singh** was born in Faizabad, Utter Pradesh, India, in 1982. He received B. Tech degree in Electronics and Communication Engineering from UPTU, Lucknow in 2005, and M.Tech in Electronics and Communication Engineering from AMITY University, Noida in 2013. Mr. Singh is currently Ph.D. Scholar in Department of Electronics and Communication Engineering from AMITY University since January 2015. His Research topic is "Improving spectral efficiency in device to Device Communication". He also did post graduate diploma in Embedded System and VLSI Design from CDAC Noida in 2006. In 2006, and 2007 he joined the Department of Electronics and Communication Engineering, SDEC Ghaziabad and Ideal Institute of Technology Ghaziabad Respectively as Lecturer. In 2013, He Joined as Assistant Professor in IMSEC Ghaziabad. His currently research interest include wireless communication, Mobile communication and Device to Device Communication.



**Dr. Vinod Shokeen** was born in Delhi, in 1975. He did BE in Electronics and Communication Engineering from SEC college of Engineering, North Maharashtra University in 1999. He completed his ME in Electronics and Communication Engineering from Delhi College of Engineering, Delhi University in 2006. He awarded Ph.D. degree in 2011 from Singhania University. He joined SPC Janakpuri as product Engineer. He was Lecturer and Sr, Lecturer science 2001–2008 in PDM Bahadurgarh and CRRIT Kanjhawala Delhi. Then he joined HIT bahadurgarh and BMCTM Farukhnagar as Assistant professor in 2008 and 2009 respectively. Presently, He is Assistant Professor since 2010 in AMITY University Noida. His Area of Interest in research is wireless and mobile communication, Antenna and wave propagation, Electronic Circuit.



**Dr. Rishi Ashathan** was born in Lucknow, India in 1973. He received the B.E., M.S., and Ph.D. degrees from Department of Electrical Engineering, MMM Engineering College Gorakhpur in 1996. He completed his M.S. from Department of Electrical Engineering, BITS Pilani in 2003. He was awarded Ph.D. Degree in 2012 from Department of Electrical Engineering, University Petroleum and Energy studies, Dehradun. He joined as a Guest Lecturer in department of Electrical Engineering, MMM Engineering College Gorakhpur in 1996. He had worked as Adoc Lecturer in the department of Electrical Engineering, G B Pant Engineering College Pauri Garwal India. He was Chief Coordinator of B.Tech in Babu Banarasi Das University, Lucknow since 1999–2014. He was HOD in Electrical and Electronics Engineering department, IMSEC Ghaziabad since 2014 to 2018. Now he is working as director in AIMT Grater Noida, India. His main areas of research interest are wireless and mobile communication, control system, soft computing and electromagnetic field. Dr Ashatha is Member of IEEE USA, ITE UK, Computer society of India and institutions of Engineers India.

Wireless Networks
https://doi.org/10.1007/s11276-019-02018-2

# Performance of energy-efficient cooperative cognitive radio system over erroneous Nakagami-*m* and Weibull fading channels

Kiran Kumar Godugu[1] · Ashok Kumar Nallagonda[2] · Srinivas Nallagonda[1]

© Springer Science+Business Media, LLC, part of Springer Nature 2019

**Abstract**
Cognitive radio (CR) is developed as one of the important techniques to improve the utilization of the radio spectrum. A CR node shares the radio spectrum with a licensed primary user opportunistically. In this paper, we study the performance of an energy-efficient cooperative cognitive radio system (CCRS) in the presence of noise plus different fading conditions under channel error constraints. Two fading environments namely, Nakagami-*m* and Weibull fading are considered and performance characteristics are evaluated. More precisely, Every CR node uses the similar energy detectors to sense a primary user, and forwards their knowledge to fusion center (FC) as one bit binary information. Different hard-decision fusion operations are carried out at FC to take the global decision about the status of the primary user. The analytical and simulation models related to noise and fading for calculating missdetection performance, total error performance, and sensing time and throughput/energy efficiency trade-off are developed. Analytical frameworks are validated via computer based simulations. For comparison purposes, we also investigate the performance of CCRS in Rayleigh fading with and without channel errors constraints. Further, the performance CCRS is investigated for various hard-decision fusion rules under several parameters of the system. Finally, an optimal values of the number of CR nodes and sensing threshold for different fusion rules, and sub-optimal hard-decision fusion rule are also investigated. The energy efficiency is enhanced by 19% when CCRS is operated over Nakagami-*m* fading environment as compared to Rayleigh fading environment for a fixed parameter values of the network.

**Keywords** Cognitive radio · Channel error · Fading models · Total error probability · Throughput · Energy efficiency

## 1 Introduction

Cognitive radio (CR) technology is developed to alleviate the conflicts between spectrum scarcity and poor utilization of the primary users (PUs). In CR networks, the CR users are unlicensed and can access the spectrum of PUs without creating any harmful interferences. Moreover, this technology supports many emerging wireless applications such as TV bands for smart grid, broadband cellular, military and public safety [1, 2]. Both the CR network and 5th generation (5G) cellular network are going to be the next generation technologies. The CR network incorporates both radio and networking technologies to present effective utilization of radio spectrum [3, 4]. On the other hand, the CR network detects the available channels in a wireless spectrum automatically and also improves the utilization efficiency of the wireless spectrum [5]. Specific and

✉ Srinivas Nallagonda
srinivas.nallagonda@gmail.com

Kiran Kumar Godugu
gkk145@gmail.com

Ashok Kumar Nallagonda
ashokmtech2k8@gmail.com

[1] Department of Electronics and Communication Engineering, Marri Laxman Reddy Institute of Technology and Management, Dundigal, Hyderabad, Telangana 500043, India

[2] Department of Electronics and Communication Engineering, Kasireddy Narayan Reddy College of Engineering and Research, Abdullapur, Hyderabad, Telangana 501505, India

Published online: 20 May 2019

consistent detection of available spectrum bands are necessary for the implementation of CR functions.

The energy detection method is an appropriate preference of taking a valid decision on the status of a licensed primary user [6]. In case of heavily faded environment, a network with a one CR node may not capture a valid decision. To capture correct detection in presence of severe fading, a network with several CR nodes and the mutual aid among them is required to make a proper decision about a primary user, which is known as cooperative spectrum sensing (CSS) [7]. A fusion center (FC) is considered in CR network which collects sensing information from all the CR nodes and makes a global decision by using hard or soft fusion operations. In hard fusion operations, sensing information is considered as one-bit binary [8] or in soft fusion operation, sensing information is considered as energies and also treating them as soft data [9].

### 1.1 Related work

In [10], the impact of threshold-based censoring using improved energy detectors on the overall detection performance of a cooperative spectrum sensing scheme is investigated and voting rule (i.e., $k$-out-of$N$ rule) is performed at the FC. In [11], author investigated the performance of unlicensed devices in a hidden node scenario between empty and occupied TV channels, where signal power measurements in the UHF frequency band have been carried out on rural, sub-urban, and urban environments practically. In [8], the detection performance of CSS is studied based on hard-decision fusion assuming noisy-faded sensing channels and noise-free reporting channels. In a practical scenario, errors may also present in both sensing and reporting channels along with noise and fading, however, in [8], studies on CSS performance through erroneous channels are not discussed. In [12], the throughput performance is investigated using erroneous channels in presence of Rayleigh fading only. In [13], authors studied the sensing time and throughput trade-off over faded sensing channel and ideal or noise less reporting channel. In [14, 15], both sensing and reporting channels are assumed as AWGN channels without considering any types of fading. However, studies [13–15] on an energy efficiency are not discussed. In [16], throughput is evaluated in $\kappa - \mu$ and $\eta - \mu$ fading channels. In [17], missed detection and total error performances, affecting log-normal shadowing in erroneous sensing channels and erroneous reporting are evaluated. The analysis presented in [17] is extended for different fading environments (like Nakagami-$m$ and Weibull) and also evalued throughput and energy efficiency of CSS. In [18], the performance of CSS with perfect and imperfect channel state information (CSI) is evaluated over Nakagami-$q$ (Hoyt) and Nakagami-$n$ (Rician) fading affecting both reporting and sensing channels. The Nakagami-$m$ and Weibull distributions have been proved to exhibit an excellent fitting for indoor and outdoor environments. These distributions also include one-sided Gaussian distribution ($m = 0.5$ and $v = 1$) and Rayleigh distribution ($m = 1$ and $v = 2$) as special cases [8]. In [19], the sensing and reporting channels affects on the trade-off performance of the CSS under Rayleigh faded channel are studied. The average channel throughput-sensing time trade-off is studied over Nakagami-$m$ fading under different outage constraints for a CSS network [20]. In [21, 22], authors studied outage performance of underlay and overlay CR network in Nakagami-$m$ fading channel with energy harvesting. An effective capacity of Nakagami-$m$ and Rayleigh fading channels under interference-power constraint is investigated in [23]. However, studies in [21–23] on throughput and energy-efficiency of cognitive radio network over Nakagami-$m$ fading channels are not studied. In [24], authors investigated an energy-efficiency of multi-channel and multi-hops decode and forward (DF) relay cognitive radio sensor networks (CRSNs) under outage probability and sensing performance constraints. In [25], the spectral-energy efficiency trade-off for CR networks is analyzed at both link and system levels in different fading environments. In [26] a comprehensive survey of wireless multimedia cognitive radio networks (WMCRNs) is presented and various multimedia applications supported by CRNs and various CR based wireless networks are surveyed. In [27, 28], authors investigated an effect of primary user traffic on the sensing-throughput tradeoff of secondary network in cognitive radio network.

In summary, the efforts are made to develop both numerical and simulation models for investigating miss-detection, total error, sensing threshold-throughput trade-off, and sensing time-energy efficiency trade-off performances in Nakagami-$m$ or Weibull fading environment under channel error constraints.

### 1.2 Summary and contributions

The present work inspects the studies on a considered cooperative cognitive radio system (CCRS) with hard-decision fusion in terms of missdetection, total error, sensing time-energy efficiency trade-off in Nakagami-$m$ fading or Weibull fading environment under channel error constraints. Based on the studies discussed in the literature, we think that the investigation on the energy efficiency of CCRS under channel error constraints over above said fading environments is the considerable area of research. The energy efficiency is enhanced by 19% when CCRS is operated over Nakagami-$m$ fading environment as compared to Rayleigh fading environment for a fixed parameter

values of the network. The important and novel contributions are given below.

(1) Design of analytical framework and validation through MATLAB based simulations for calculating probability of detection of a network with one CR node in non fading AWGN and Nakagami-*m* or Weibull fading environments under erroneous sensing channels.
(2) The framework for investigating the performance through overall missdetection and total error probability (TEP) for various fusion rules are presented under erroneous sensing and reporting channels.
(3) The investigation of missdetection and TEP performances for different fading parameters and network parameters.
(4) An optimal values of the number of CR nodes and sensing threshold where TEP is minimized for different fusion rules are calculated. Also, suboptimal hard-decision fusion rule is identified.
(5) The sensing time-energy effciency/throughput tradeoff of the proposed network in the presence of Nakagami-*m* or Weibull fading environments under erroneous sensing and reporting channels constraints is studied. In [12, 16], throughput analysis of CCRN is only presented. However, studies on energy efficiency are not discussed.

### 1.3 Splitting of the paper

The reminder of the paper is splitted in to four sections: Sect. 2 presents a brief description of a proposed system framework. Section 3 discusses a MATLAB based simulation model for validating analytical framework. The results are discussed in Sect. 4. Finally, the major observations are presented in conclusion Sect. 5.

## 2 Considered system model

The present section describes a CCRS (Fig. 1) considering a single PU and multiple CR nodes along with one FC. The channels between PU and CR nodes are referred as sensing channels while wireless channels between CR nodes and FC are referred as reporting channels. We consider that both the channels are erroneous channels with an error probability $q$. The proposed network is also cooperative and operated over additive white Gaussian noise (AWGN) plus Nakagami-*m* fading or Weibull fading environment. More precisely, sensing performance can be improved by allowing different CR nodes to cooperate by sharing their information. Every CR node uses an ED to sense a PU, and forwards their knowledge to FC as one bit-binary

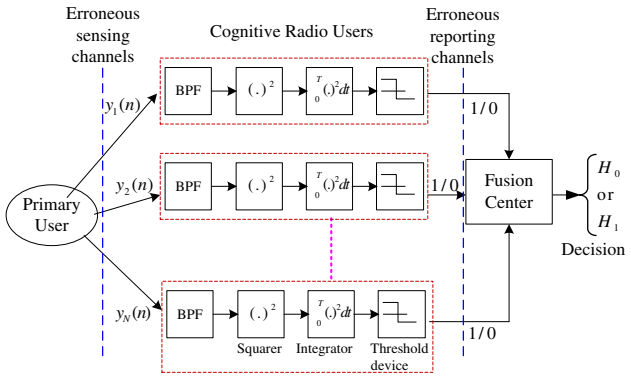



**Fig. 1** Proposed model: cooperative cognitive radio system with $N$ CR nodes, one of each PU and FC, and hard-decision fusion is considered at FC

information. The FC takes a global decision on the status (active or inactive) of a PU using hard-decision combining rules ( i.e., $k$-out-$N$ fusion rule) and sends final decision back to all the CR nodes (noiseless feedback channel is assumed). The OR rule, AND rule, and MAJORITY rules are special rules and can be determined for the specific $k$ values. We consider a fixed detection frame time of $T_f$ ms as shown in Fig. 2. Each CR node senses the PU's spectrum in $\tau_s$ ms of $T_f$ ms ($\tau_s < T_f$) and $T_f - \tau_s$ is the transmission time denoted as $T_r$. More precisely, as $\tau_s$ increases, $T_r$ decreases and $\tau_s + T_r = T_f$. The number ($N_s$) of samples received can be given as $N_s = \tau_s f_s$ where $f_s$: sampling rate and $\tau_s$: sensing time. The signal samples received at $i$th CR node, $y_i(n)$ where $(i \in \{1, 2, \ldots, N\})$, is expressed as

$$y_i(n) = \begin{cases} u_i(n) & : \mathcal{H}_0 \\ h_i s(n) + u_i(n) & : \mathcal{H}_1 \end{cases} \quad (1)$$

where $s(n)$: primary signal with mean zero and variance$\sigma_s^2 = \gamma \sigma_u^2$ where $n \in \{1, 2, \ldots, N_s\}$ are the samples, $u(n)$ indicates noise with zero mean and $\sigma_u^2 = 1$ variance, $h$ indicates the fading constant, hypothesis $\mathcal{H}_0$ represents PU inactive status, and hypothesis $\mathcal{H}_1$ represents PU active status. A two state Markov model [14] is used to model the PU activity (active state and inactive state). The duration of active state and inactive state of the PU are exponentially distributed with mean values $\alpha_1$ and $\alpha_0$, respectively. The probability density function (PDF) of the active state and inactive state can be written as [14],

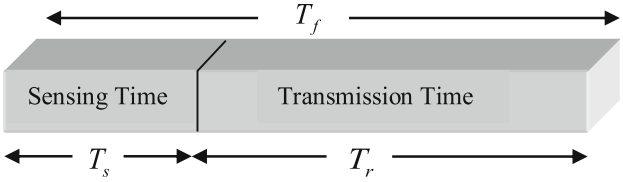



**Fig. 2** A fixed detection frame-time structure where $T_s = \tau_s$ and $T_r = T_f - T_s$

$$p_{act}(t) = \alpha_1^{-1} e^{-t/\alpha_1}; \quad p_{inact}(t) = \alpha_0^{-1} e^{-t/\alpha_0} \tag{2}$$

where $p(.)$ denotes the probability. The probabilities, $p(\mathcal{H}_0)$ and $p(\mathcal{H}_1)$ are modeled as $p(\mathcal{H}_0) = \alpha_0/(\alpha_0 + \alpha_1)$ and $p(\mathcal{H}_1) = \alpha_1/(\alpha_0 + \alpha_1)$, where $p(\mathcal{H}_0)$ and $p(\mathcal{H}_1)$ are the prior probabilities of the PU being inactive and PU being active, respectively. Then $p(\mathcal{H}_0) + p(\mathcal{H}_1) = 1$ or $\alpha_0 + \alpha_1 = 1$. It is assumed that $s(n)$, $u(n)$ and $h$ samples are independent with each other. The test statistic for ED, denoted as $\mathcal{T}$ can be written as [29]:

$$\mathcal{T}(y) = \frac{1}{N_s} \sum_{n=1}^{N_s} |y(n)|^2. \tag{3}$$

under $\mathcal{H}_0$ and $\mathcal{H}_1$, the probability density function (PDF) of $\mathcal{T}(y)$ can be approximated by using central limit theorem. The approximated PDF can be shown as Gaussian distributions as follows [29]:

$$\mathcal{T}(y)|\mathcal{H}_0 \sim \mathcal{N}\left(\sigma_u^2, \frac{2}{N_s}\sigma_u^4\right), \tag{4}$$

$$\mathcal{T}(y)|\mathcal{H}_1 \sim \mathcal{N}\left((1+\gamma)\sigma_u^2, \frac{2(1+\gamma)^2\sigma_u^4}{N_s}\right). \tag{5}$$

### 2.1 The AWGN (non-fading) environment

In a $i$th CR node, false alarm and detection probabilities in AWGN are given from [29]

$$P_f = P_r(\mathcal{T} > \lambda|\mathcal{H}_0) = \mathcal{Q}\left(\left(\frac{\lambda}{\sigma_u^2} - 1\right)\sqrt{\frac{\tau_s f_s}{2}}\right), \tag{6}$$

$$P_d = P_r(\mathcal{T} > \lambda|\mathcal{H}_1) = \mathcal{Q}\left(\left[\frac{(\lambda/\sigma_u^2)}{1+\gamma} - 1\right]\sqrt{\frac{\tau_s f_s}{2}}\right). \tag{7}$$

where $\lambda$ sensing threshold, $\gamma(= \sigma_s^2/\sigma_u^2)$ instantaneous signal to noise ration (SNR), and $\mathcal{Q}(.)$ the standard Gaussian complementary distribution function. At each CR node, the same average SNRs ($\bar{\gamma}_i = \bar{\gamma}; \forall i$) and probability of false alarm ($P_{f,i} = P_f; \forall i$) are assumed.

When sensing channels affected by fading $h_i$ varies and Eq. (7) contains the parameter $\gamma$, then at $i$th CR node average probability of detection ($\bar{P}_{d,i}$) is obtained as

$$\bar{P}_{d,i}^{x} = \int_0^\infty \mathcal{Q}\left(\left[\frac{(\lambda/\sigma_u^2)}{1+\gamma} - 1\right]\sqrt{\frac{\tau_s f_s}{2}}\right) f_\gamma(\gamma) d\gamma. \tag{8}$$

where $x$ denotes type of fading (Nakagami-$m$ or Weibull) and $f_\gamma(\gamma)$: $\gamma$ PDF under fading at CR node level. Under $\mathcal{H}_0$ case, from (6), for all CR nodes, it is observed that $P_f$ does not depend on $\gamma$. Hence, the expression of $\bar{P}_f$ is same as (6) under any fading environment.

### 2.2 Nakagami-*m* environment

When Nakagami-$m$ fading is considered, the PDF of $\gamma$ can be written as [8]:

$$f_\gamma(\gamma) = \left(\frac{m}{\bar{\gamma}}\right)^m \frac{\gamma^{m-1}}{\Gamma(m)} \exp\left(-\frac{m\gamma}{\bar{\gamma}}\right); \quad \gamma \geq 0 \tag{9}$$

where $m(0.5 \leq m \leq \infty)$ is the fading severity parameter. For the special case of $m = 1$, (9) reduces to the Rayleigh PDF. Following appendix A, the final expression for $\bar{P}_{d,i}^{Naka}$ can be obtained as

$$\begin{aligned}\bar{P}_{d,i}^{Naka} = \frac{1}{2} - \sum_{m=0}^{2z+1} (-1)^{2z+m+1} \binom{2z+1}{m} \\ \times \int_0^{\pi/2} \left(\frac{\lambda/\sigma_u^2}{1+\tan\theta}\right)^m \tan\theta^{m-1} \\ \exp\left(-\frac{m\tan\theta}{\bar{\gamma}}\right) \sec^2\theta d\theta\end{aligned} \tag{10}$$

### 2.3 Weibull fading environment

When Weibull fading is considered in the sensing channels then $\gamma$ PDF under Weibull fading is given as [8]:

$$f_\gamma(\gamma) = \frac{v}{2}\left[\frac{\Gamma(p)}{\bar{\gamma}}\right]^{\frac{v}{2}} \gamma^{\frac{v}{2}-1} \exp\left[-\left\{\frac{\gamma\Gamma(p)}{\bar{\gamma}}\right\}^{\frac{v}{2}}\right]; \quad \gamma \geq 0, \tag{11}$$

where $v$ is the fading severity parameter and $p = 1 + 2/v$. Equation (11) includes the Rayleigh PDF for $v = 2$ and negative exponential PDF for $v = 1$. Following appendix B, the final expression for $\bar{P}_{d,i}^{Wieb}$ can be obtained as

$$\begin{aligned}\bar{P}_{d,i}^{Wieb} = \frac{1}{2} - \sum_{m=0}^{2z+1} (-1)^{2z+m+1} \binom{2z+1}{m} \\ \times \int_0^{\pi/2} \left(\frac{\lambda/\sigma_u^2}{1+\tan\theta}\right)^m \tan\theta^{\frac{v}{2}-1} \\ \exp\left[-\left\{\frac{\tan\theta\Gamma(p)}{\bar{\gamma}}\right\}^{\frac{v}{2}}\right] \sec^2\theta d\theta\end{aligned} \tag{12}$$

Assumed that each CR node has similar performance:

$$\bar{P}_{f,i} = pr\{d_i = 1|\mathcal{H}_0\} = \bar{P}_f, \tag{13}$$

$$\bar{P}_{d,i} = pr\{d_i = 1|\mathcal{H}_1\} = \bar{P}_d, \tag{14}$$

$$\bar{P}_{m,i} = pr\{d_i = 0|\mathcal{H}_1\} = \bar{P}_m. \tag{15}$$

where $\bar{P}_m(= 1 - \bar{P}_d)$ is the missdetection probability and $d$ is the decision. The revised expressions of $P_f$ and $P_m$ with channel error ($q$) are written as [30]:

$$P_{fe} = (1 - \bar{P}_f)q + \bar{P}_f(1 - q), \tag{16}$$

$$P_{me} = (1 - \bar{P}_m)q + \bar{P}_m(1 - q). \tag{17}$$

where $\bar{P}_f$ and $\bar{P}_m$ are shown in (13) and (15), respectively. From (16) and (17), it is seen that erroneous sensing channel ((channel error $q$)) degrades the single CR node performance. The total error probability (TEP) in a CR node is

$$\bar{P}_e = p(\mathcal{H}_1)P_{me} + p(\mathcal{H}_0)P_{fe}. \tag{18}$$

In this paper, TEP is evaluated for various values of $p(\mathcal{H}_0)$ and $p(\mathcal{H}_1)$.

### 2.4 Hard-decision combining rules

The decision may not be guaranteed by a single CR node when the erroneous sensing channels affected by high amount of noise and fading. To overcome this, $N$ number of cooperative CR nodes are used to make the valid decision about a PU. At last, the performance of the system can be enhanced. The analytic expressions of several performance metrics under channel error constraints can be obtained from the generalized $k$-out-of-$N$ rule given in [30]:

$$Q_{fe} = \sum_{k=n}^{N} \binom{N}{k} (P_{fe})^k (1 - P_{fe})^{N-k}, \tag{19}$$

$$Q_{de} = \sum_{k=n}^{N} \binom{N}{k} (P_{de})^k (1 - P_{de})^{N-k}, \tag{20}$$

$$Q_{me} = 1 - Q_{de}. \tag{21}$$

Further, for a specific values of $k$ different sub fusion rules such as OR rule for $k = 1$, AND rule for $k = N$, and MAJORITY rule for $k = \frac{N}{2} + 1$. Finally, expression of cooperative TEP is written as

$$Q_e = p(\mathcal{H}_1)Q_{me} + p(\mathcal{H}_0)Q_{fe}. \tag{22}$$

On the other hand, for $q = 0$, an alternate expression of TER is obtained from (16) and (17), and (22).

### 2.5 Sensing threshold-throughput trade-off

In the considered CCR system, CR nodes access the PU's channel only when the channel is not in use. In this section, the throughput analysis, following [12] is presented. Let us define $\mathcal{B}_p$ and $\tilde{\mathcal{B}}_p$ are the PU network throughput when CCRS is inactive and active, respectively. Also, $\mathcal{B}_s$ and $\tilde{\mathcal{B}}_s$ are the CCR system throughput when PU is inactive and active, respectively. Let $\mathcal{B}_{avg}$ represents the average channel throughput for the generalized $k$-out-of-$N$ rule as

$$\mathcal{B}_{avg}(N) = r_0 + r_1(1 - Q_{me}(N)) - r_2 Q_{fe}(N). \tag{23}$$

where

$$r_0 = p(\mathcal{H}_1)\left(\tilde{\mathcal{B}}_p + \tilde{\mathcal{B}}_s\right) + p(\mathcal{H}_0)\mathcal{B}_s \tag{24}$$

$$r_1 = p(\mathcal{H}_1)\left(\mathcal{B}_p - \tilde{\mathcal{B}}_p - \tilde{\mathcal{B}}_s\right) \tag{25}$$

$$r_2 = p(\mathcal{H}_0)\mathcal{B}_s. \tag{26}$$

### 2.6 Sensing time and energy efficiency trade-off

In this section, analysis of energy efficiency, denoted as $\eta$ is presented. Let us assume that each CR node consumes energy $E_1$ to perform sensing operation and energy $E_2$ to send sensing data to FC. Similarly, let us assume that PU consumes energy $E_p$ to transmit its data and CR node consumes the energy $E_s$ to transmit its data. Then, the total energy consumption, denoted as $\mathcal{E}$, which is given as

$$\begin{aligned}\mathcal{E}(N) = {} & e_0 p(H_1)Q_d(N) + e_1 p(H_1)Q_m(N) \\ & + e_2 p(H_0)\left(1 - Q_f(N)\right) + e_3 p(H_0)Q_f(N).\end{aligned} \tag{27}$$

where

$$e_0 = N(E_1 + E_2) + E_p \tag{28}$$

$$e_1 = N(E_1 + E_2) + E_p + E_s \tag{29}$$

$$e_2 = N(E_1 + E_2) + E_s \tag{30}$$

$$e_3 = N(E_1 + E_2) \tag{31}$$

and $\eta$ can be estimated by taking the ratio between $\mathcal{B}_{avg}$ given from (23) and $\mathcal{E}$ given from (27) as [31]:

$$\eta = \frac{B_{avg(N)}}{\mathcal{E}(N)}. \tag{32}$$

## 3 Simulation model

The present sections describes the framework for the step by step simulation process developed using MATLAB based on the theoretical analysis discussed in the Sect. 2. In Algorithm 1, we have used square root of a gamma random variable (RV) $\mathcal{G}(a, b)$ with parameters $a = m$; $b = \Omega/m$ to generate Nakagami-$m$ RVs for arbitrary $m$ values. Gamma RVs can be generated with MATLAB using the *gamrnd*() function. Following Algorithm 2, Weibull fading envelop can be generated, the fading parameter $v$ varies from 1 to $\infty$. Next, standard deviation $\sigma$ (for in-phase as well as quadrature Gaussian process) is calculated using the step 2.

**Algorithm 1** Nakagami-$m$ fading coefficient generation

$\Omega \leftarrow 1$
$a \leftarrow m$
$b \leftarrow \frac{\Omega}{m}$
**generate** $\mathbf{h_{sq}} \sim \mathcal{G}(a, b)$
**return** $\mathbf{h} = \sqrt{\mathbf{h_{sq}}}$

**Algorithm 2** Weibull fading coefficient generation

$\Omega \leftarrow 1$
$\sigma \leftarrow \sqrt{\frac{\Omega}{2}}$
**generate** $\mathbf{h_I} \sim \mathcal{N}(0, \sigma^2)$
**generate** $\mathbf{h_Q} \sim \mathcal{N}(0, \sigma^2)$
$\mathbf{R} = \sqrt{\mathbf{h_I^2} + \mathbf{h_Q^2}}$
**return** $\mathbf{h} = \mathbf{R}^{2/v}$

The simulation process flow chart for the evaluation of detection probability, TEP, throughput, and energy efficiency is shown in Fig. 3. Also, the description of Fig. 3 is given as a step by step procedure and is implemented using MATLAB.

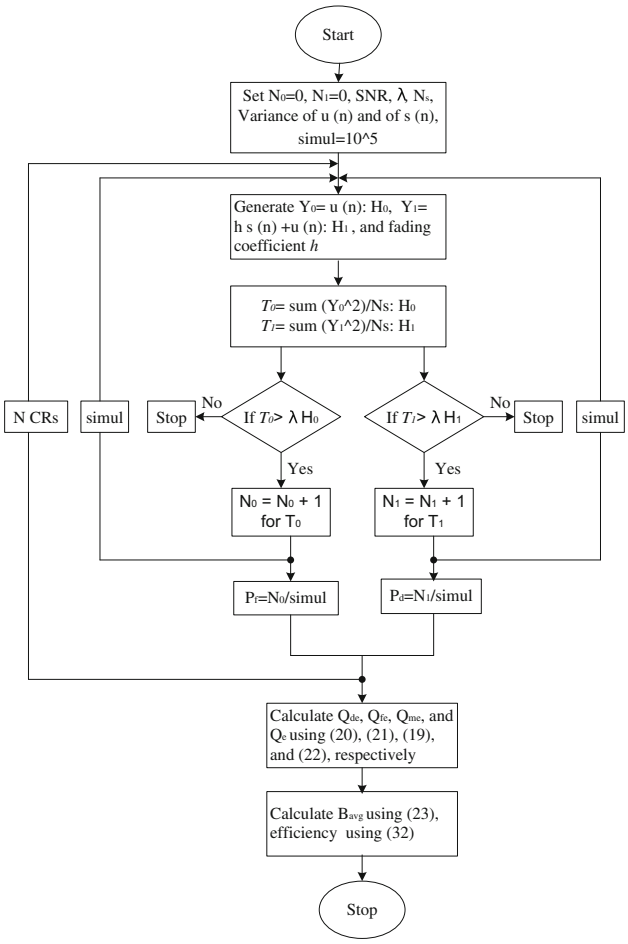



**Fig. 3** The simulation process flow chart for the evaluation of detection probability, TEP, throughput, and energy efficiency

(1) Obtain a binary phase shift keying (BPSK) signal ($s(n)$ of mean zero and variance $\sigma_s^2 = \gamma\sigma_u^2$ ) and noise signal (AWGN with mean zero and variance $\sigma_u^2 = 1$, $u(n)$) of $N_s$ samples.
(2) Generate fading coefficient for Nakagami-$m$ using Algorithm 1 or generate a Weibull fading coefficient using Algorithm 2.
(3) Calculate $y(n) = hs(n) + u(n)$ received at a CR node as input.
(4) Using (3) estimate an energy ($\mathcal{T}$).
(5) Estimate $\lambda$ for a specific $P_f$ value and $\lambda$ is set in a CR node. Identical $\lambda$ is assumed in each CR node.
(6) Compare $\mathcal{T}$ with $\lambda$. If $\mathcal{T} > \lambda$, the output of ED in a CR node is either '1' (active) or '0' (inactive).
(7) For $N$ CR nodes steps from 1 to 6 are repeated and $N$ decisions in terms of 1's and 0's can be obtained and are forwarded to FC for fusing operations.
(8) Steps from one to six should be in a loop so that a large number of iterations can be performed and $\bar{P}_d$ at a single CR node can be determined.
(9) Using (20), (21), (19) and (22) calculate cooperative detection, missdetection, false alarm, and TEP, respectively.
(10) Using (23) calculate throughput and using (32) calculate energy efficiency of CCRS.
(11) Generate various simulation and numerical plots under various channel as well as network conditions: $N$, $m$ $v$, $q$, $\bar{\gamma}$, $P_f$, $N_s$, and $T_s$.

## 4 Results and discussions

This section presents various results and their descriptions. Performance of cooperative detection, missdetection, false alarm, total error probability, sensing threshold-throughput

**Table 1** Parameters used in numerical evaluation

| Parameter | Value |
|---|---|
| $T_f$ | 100 ms |
| $f_s$ | 6 MHz |
| $\mathcal{B}_s$ | 10 |
| $\tilde{\mathcal{B}}_s$ | 5 |
| $\mathcal{B}_p$ | 20 |
| $\tilde{\mathcal{B}}_p$ | 10 |
| $E_p$ | 40 |
| $E_s$ | 10 |
| $E_1$ | 0.1 |
| $E_2$ | 0.05 |

and sensing time-energy efficiency tradeoffs are discussed subject to hard decision fusion rule for several parameters values such as $m$, $v$, $N$, $N_s$, $\bar{\gamma}$, $P_f$, and $q$. Also, various values of $p(\mathcal{H}_0)$ and $p(\mathcal{H}_1)$ are considered. Table 1 shows the parameters considered for throughput and energy efficiency analysis.

Figures 4, 5 and 6 show the $P_{de}$ variation as a function of $N_s$ for a non-cooperative CCRS for various values of $q$ in AWGN, Nakagami-$m$, and Weibull fading, respectively. It is seen from these figures that as $N_s$ goes up the value of $P_{de}$ raises. The behavior of $P_{de}$ for $q = 0$ is also depicted for comparison purpose. It is seen that $P_{de}$ reduces when $q$ raises. In Fig. 4, the numerical results are plotted using (7) and compared with the results plotted based on the simulation testbed for AWGN channel and different $q$ values. It can be clearly noticed that both simulation results and numerical results are matched perfectly for a given parameter values. This confirm our simulation model discussed in Sect. 3. From Fig. 5 the impact of $m$ on the performance of non-cooperative CCRS is shown. It can be noted that $P_{de}$ increases when $m$ increases. When $m = 1$, for any value of $q$, the Nakagami-$m$ fading reduces to Rayleigh fading where the value of $P_{de}$ increases as the value of $N_s$ increases. The performance in the presence of Nakagami-$m$ is lower than the AWGN case as shown in Fig. 5. Similarly, the impact of $v$ on the performance of non-cooperative CCRS is shown in Fig. 6. It can be noted that $P_{de}$ increases when $v$ increases. The Rayleigh performance characteristics can be obtained when $v = 2$, for any value of $q$. Also, the performance in Weibull fading is degraded when compare to performance in AWGN case. When $N_s = 2500$, $q = 1.5$ and $m$ increases from 0.5 to 1.5,

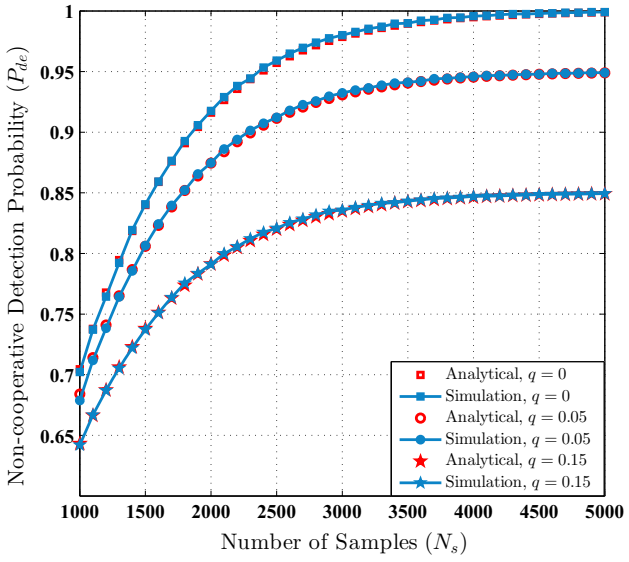



**Fig. 4** The $P_{de}$ variation as a function of $N_s$ is shown over AWGN channel for different $q$ values ($\bar{\gamma} = -10$dB and $Pf = 0.05$)

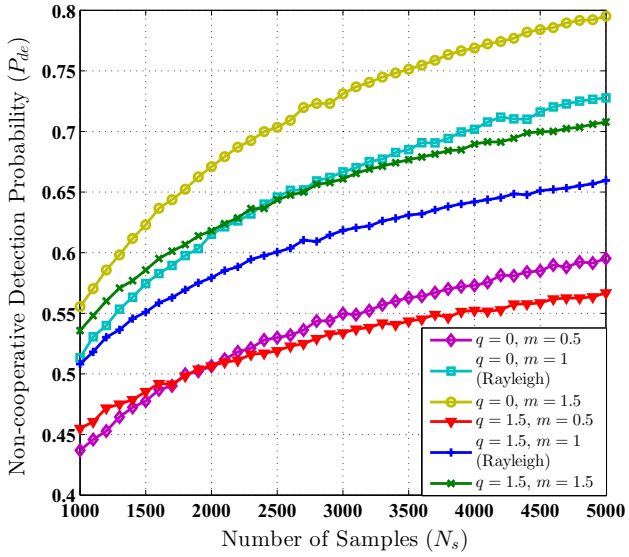



**Fig. 5** The $P_{de}$ is shown as a function of $N_s$ is shown over Nakagami-$m$ fading channel for different $q$ and $m$ values ($\bar{\gamma} = -10$dB and $Pf = 0.05$)

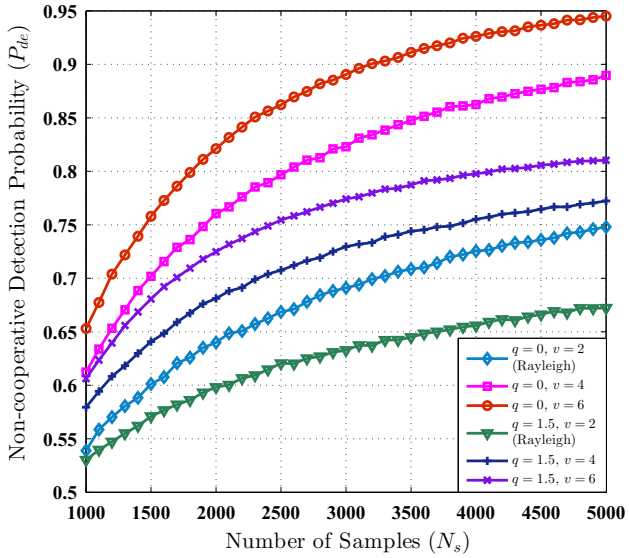



**Fig. 6** The $P_{de}$ is shown as a function of $N_s$ is shown over Weibull fading channel for different $q$ and $v$ values ($\bar{\gamma} = -10$dB and $Pf = 0.05$)

the $P_{de}$ increases by 23.98% in the presence of Nakagami-$m$ fading and increases by 21.59% in the presence of Weibull fading as $v$ increases from 2 to 6.

The $Q_{de}$ variation as a function of $\bar{\gamma}$ and $q$ under different fusion rule in Nakagami-$m$ fading is shown in Fig. 7 and in Weibull fading is shown in Fig. 8. From both the figures, for different SNR values, $Q_{de}$ values are determined. It is seen that as SNR raises $Q_{de}$ raises significantly even for bigger $q$ values present in sensing and reporting channels. It is an obvious that bigger values of SNR reduces the

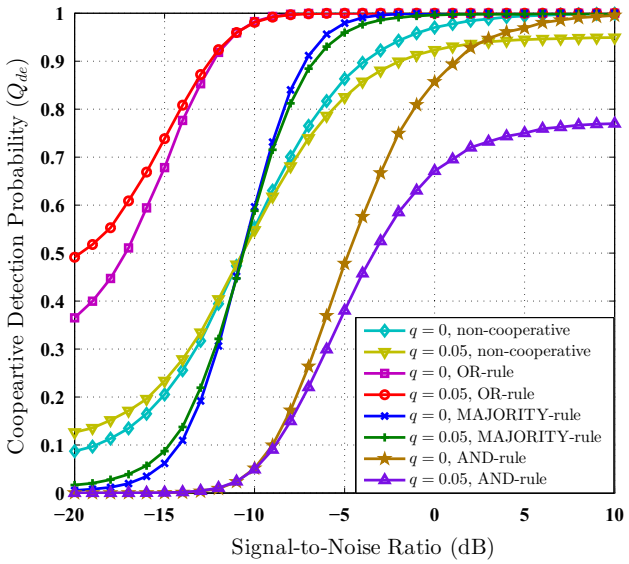



**Fig. 7** The $Q_{de}$ variation as a function of $\bar{\gamma}$ and $q$ is shown over Nakagami-$m$ fading channel ($m = 1.5$, $N = 5$, $N_s = 1000$, and $P_f = 0.05$)

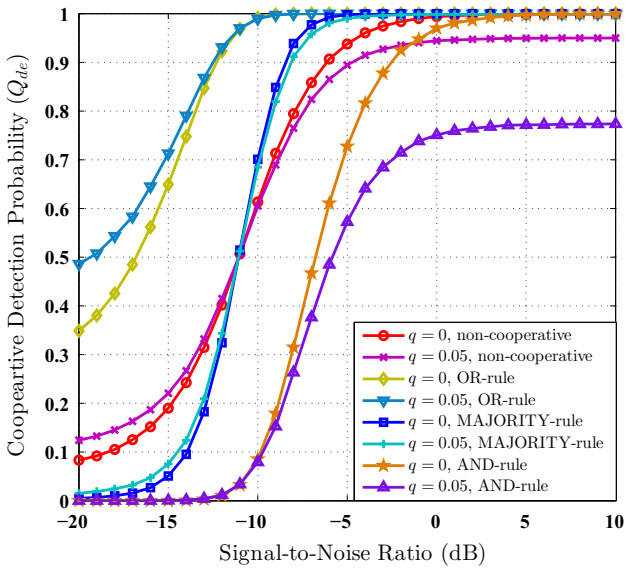



**Fig. 8** The $Q_{de}$ variation as a function of $\bar{\gamma}$ and $q$ is shown over Weibull fading channel ($v = 4$, $N = 5$, $N_s = 1000$, and $P_f = 0.05$)

effect of noise. The results when non-cooperative CCRS and $q = 0$ are also given for comparison purpose. OR rule outperforms the other fusion rules for different values of $q$. When $\bar{\gamma} = -5$dB, AND rule and $q$ increases from 0 to 0.05, the $Q_{de}$ decreases by 20.45% in the presence of Nakagami-$m$ fading (Fig. 7) and by 21.29% in the presence of Weibull fading (Fig. 8). Similarly, when $q = 0.05$, AND fusion rule and $\bar{\gamma}$ dB increases from $-5$ to 5 dB, the $Q_{de}$ increases by 97.26% in the presence of Nakagami-$m$ fading and increases by 34.77% in the presence of Weibull fading.

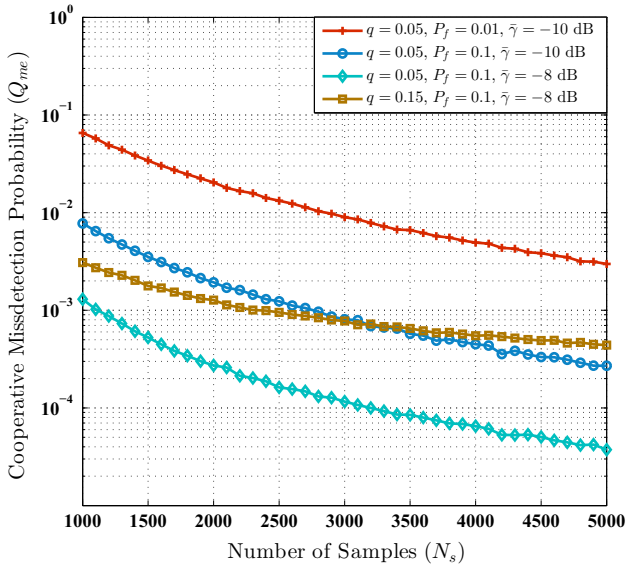



**Fig. 9** The $Q_{me}$ variation with respect to $N_s$, $P_f$, $\bar{\gamma}$, and $q$ is shown over Nakagami-$m$ fading channel ($m = 1.5$, $N = 5$ and OR rule)

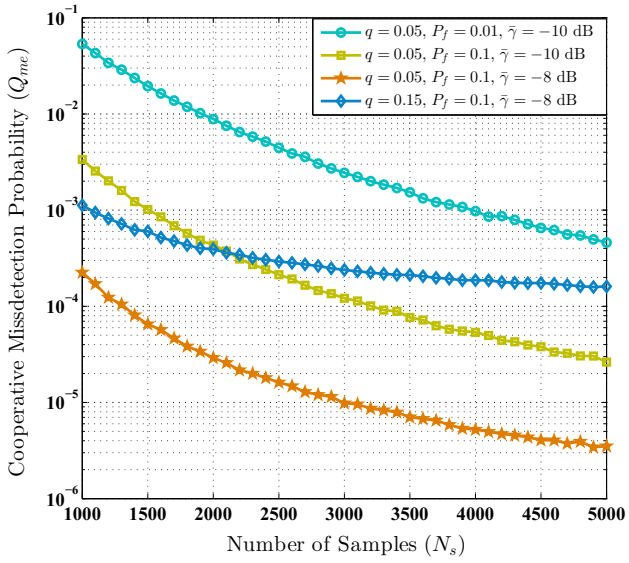



**Fig. 10** The $Q_{me}$ variation with respect to $N_s$, $P_f$, $\bar{\gamma}$, and $q$ is shown over Weibull fading channel ($v = 4$, $N = 5$ and OR rule)

In Figs. 9 and 10, the significant impact of $P_f$, SNR, and $q$ on $Q_{me}$ performance is depicted in the presence Nakagami-$m$ fading and Weibull fading, respectively. from these two figures it is observed that $Q_{me}$ decreases as any one of the $P_f$, SNR, and $N_s$ raises. It is also noted that there is a appreciable impact of $q$ on $Q_{me}$ performance. For example, when $N_s = 2500$ and $q$ goes up from 0.05 to 0.15, the $Q_{me}$ is decreases by 82.98% in the presence of Nakagami-$m$ fading (Fig. 9) and by 92.42% in the presence of Weibull fading (Fig. 10).

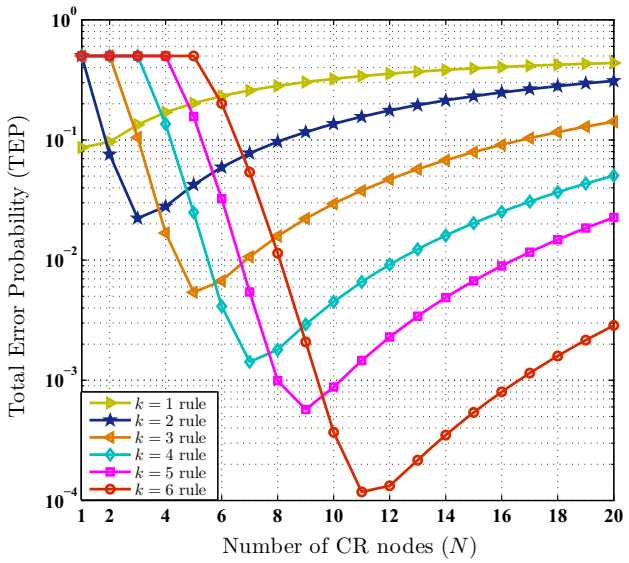



**Fig. 11** TEP versus $N$ is shown over Nakagami-$m$ fading channel for various fusion rules ($p(\mathcal{H}_0) = p(\mathcal{H}_1) = 0.5$, $m = 1.5$, $\bar{\gamma} = -10$dB, $N_s = 1000$, and $P_f = 0.05$)

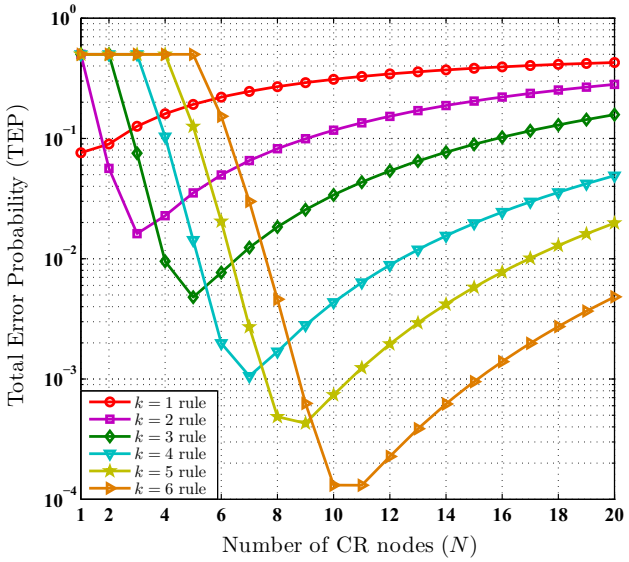



**Fig. 12** TEP versus $N$ is shown over Weibull fading channel for various fusion rules ($p(\mathcal{H}_0) = p(\mathcal{H}_1) = 0.5$, $q = 0.05$, $v = 4$, $\bar{\gamma} = -10$dB, $N_s = 1000$, and $P_f = 0.05$)

Figures 11 and 12 show the variation TEP with respect to $N$ under different fusion rules in Nakagami-$m$ and Weibull fading channels, respectively. It is seen from two figures that for $k = 1$ rule, as $N$ raises, initially TEP reduces and TEP raises when further raises of $N$. Optimal value of $N$ i.e., $N_{opt}$ can be found at which TEP is low. The similar behavior is also noted for higher order fusion rules (say $k = 5$ rule) and $N_{opt}$ shifts from left side to right side. When $N = 20$ and $k$ increases from 2 to 6, TEP decreases by

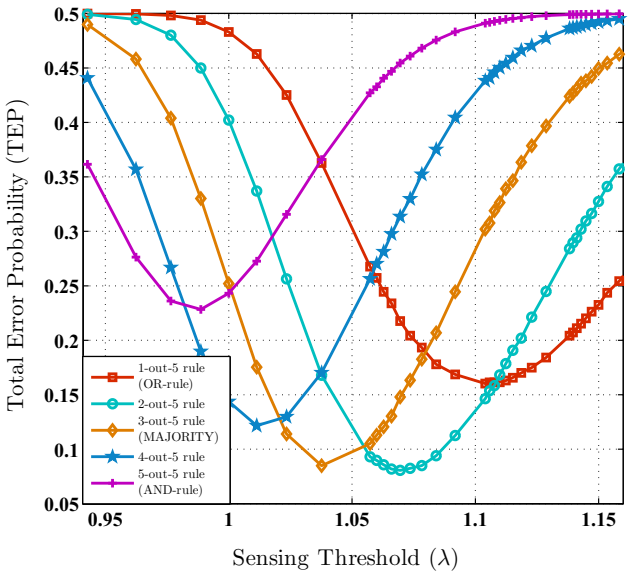



**Fig. 13** TEP versus $\lambda$ is shown over Weibull fading channel for various fusion rules ($p(\mathcal{H}_0) = p(\mathcal{H}_1) = 0.5$, $N = 5$, $q = 0.05$, $v = 4$, $\bar{\gamma} = -10$dB, and $N_s = 1000$)

107.45% in the presence of Nakagami-$m$ fading (Fig. 11) and by 57.34% in the presence of Weibull fading (Fig. 12).

In Fig. 13, the variation of TEP as a function of $\lambda$ under different fusion rules in Weibull fading environment is shown. It is noted that for $k = 1$ rule, as $\lambda$ value raises, initially TEP reduces and furthermore, TEP raises when $\lambda$ enhances. Hence, an optimal value of $\lambda$ i.e., $\lambda_{opt}$ can be found at which TEP is low. The similar behavior is also observed for higher order fusion rules (say $k = 5$ rule) and $\lambda_{opt}$ shifts from left side to right side. When $\lambda$ is equal to

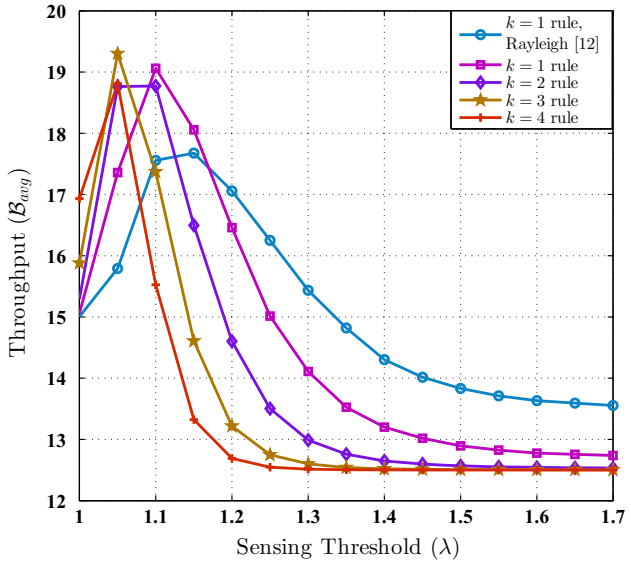



**Fig. 14** Throughput versus $\lambda$ is shown over Weibull fading channel for various fusion rules ($p(\mathcal{H}_0) = p(\mathcal{H}_1) = 0.5$, $N = 10$, $q = 0.05$, $v = 4$, $\bar{\gamma} = -10$dB, and $N_s = 1000$)

1.1, OR rule performs well, as sensing threshold value increases, where TEP decreases by 87.48% when compared to $k = 5$ rule.

Figure 14 depicts the impact of $\lambda$ on the throughput performance under various fusion rules in the presence of Weibull fading. It is observed that $\mathcal{B}_{avg}$ increases initially up to a certain value until it reaches a maximum value and decreases gradually as long as $\lambda$ increases. The maximum $\mathcal{B}_{avg}$ can be obtained at a particular value of $\lambda$. In particular, at $\lambda = 1.15$ when $k = 1$ rule is performed, the maximum $\mathcal{B}_{avg}$ is obtained as 17.68. Similarly, at $\lambda = 1.05$ and $k = 5$ rule is considered, the maximum $\mathcal{B}_{avg}$ is obtained as 18.81. It can also be noted that the suboptimal fusion rule for which minimum threshold and maximum $\mathcal{B}_{avg}$ can be obtained. The $k = 5$ rule is the suboptimal fusion rule which gives maximum $\mathcal{B}_{avg}$ as 19.3 and minimum threshold as 1.05. Throughput is enhanced when CCRS is operated over Weibull fading environment as compared to Rayleigh fading environment [12].

Figures 15 and 16 show the energy efficiency as a function $\lambda$ and $T_s$ is shown, respectively. In both the figures, $\eta$ for various values of $p(\mathcal{H}_0)$ is investigated in Nakagami-$m$ and Weibull fading channels. From Figs. 15 and 16, it is noticed that an optimal $\lambda$ and $T_s$ values can be found where $\eta$ is low. The $\eta$ raises for both the cases of raising $\lambda$ as well as $T_s$ initially,. The $\eta$ reaches to a maximum value, thereafter it decreases with further raise in $\lambda$ or $T_s$. Thus, there exists an optimal $\lambda$ as well as $T_s$ for which $\eta$ is maximum. It is found that $\eta$ is high when $p(\mathcal{H}_0)$ increases in both fading channels. It is found from these figures that an optimal $T_s$ is to be $T_s = 0.27$ ms and optimal

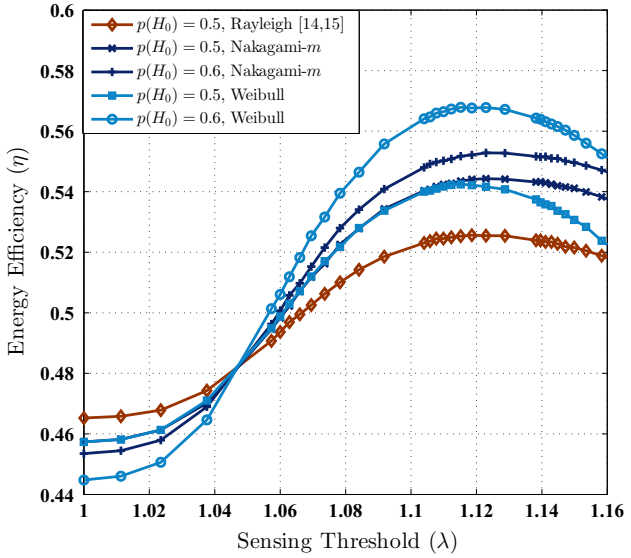



**Fig. 15** Energy efficiency as a function of $\lambda$ is shown over Nakagami-$m$ and Weibull fading channels for various values of $p(\mathcal{H}_0)$ ($N = 10$, $q = 0.05$, $v = 4$, $m = 1.5$, $\bar{\gamma} = -10$dB, and $N_s = 1000$)

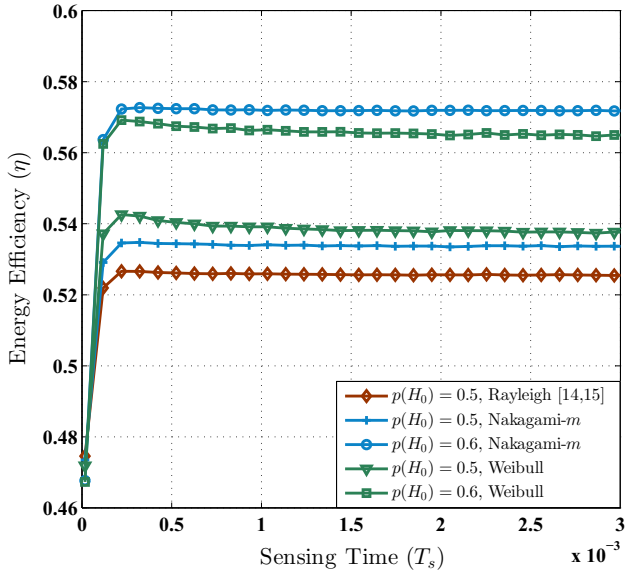



**Fig. 16** Energy efficiency as a function of sensing time $T_s$ is shown in Nakagami-$m$ and Weibull fading channels for various values of $p(\mathcal{H}_0)$ ($N = 10$, $q = 0.05$, $v = 4$, $m = 1.5$ $\lambda = 1.125$, $\bar{\gamma} = -10$dB, and $N_s = 1000$)

$\lambda$ is to be $\lambda = 1.12$. At these optimal values when $p(\mathcal{H}_0)$ increases from 0.5 to 0.7, maximum $\eta$ increases from 52.2 to 54.7% in case of Nakagami-$m$ fading and from 52.1 to 56.7% in case of Weibull fading are obtained. It is observed from both Figs. 15 and 16 that energy efficiency is enhanced by 19% when CCRS is operated over Nakagami-$m$ fading environment as compared to Rayleigh fading environment [14, 15].

Table 2 shows the performance metrics of significant parameters obtained under both Nakagami-$m$ and Weibull fading. The $Q_{de}$ have been evaluated for different values of SNR (−5 to 10 dB) with ($q = 5$) and without error ($q = 0$) under different fusion rules. OR rule performs better among other fusion rules even for low SNR values in the presence of fading.

Tables 3 and 4 show a comparison of TEP for various values of $p(\mathcal{H}_0)$ and $p(\mathcal{H}_1)$. Several values of $p(\mathcal{H}_0)$ and $p(\mathcal{H}_1)$ are considered to evaluate TEP. The parameters considered for these tables are $N_s = 1000$, $N = 5$, $P_f = 0.05$, and $\bar{\gamma} = -10$dB. When $q = 0.15$, OR-rule, and $p(\mathcal{H}_0)$ changes from 0.25 to 0.75 and $p(\mathcal{H}_1)$ changes from 0.75 to 0.25, TEP reduces by 63.65% in the presence of Nakagami-$m$ fading and by 64.87% in the presence of Weibull fading. When $p(\mathcal{H}_0) = 0.25$, $p(\mathcal{H}_1) = 0.75$, OR-rule, and $q$ changes from 0 to 0.15, TEP increases by 166.15% in the presence of Nakagami-$m$ fading and by 174.37% in the presence of Weibull fading. It is observed from table that MAJORITY rule performs better than AND-rule and OR-rule in the presence of fading.

**Table 2** The $Q_{de}$ performance comparison of different fading environments under SNR and channel error constraints, $p(\mathcal{H}_0) = p(\mathcal{H}_1) = 0.5$, $N_s = 1000$, $N = 5$, $P_f = 0.05$, $m = 1.5$ and $v = 4$

| Fusion type | SNR ($\bar{\gamma}$) (dB) | $q$ | Nakagami-$m$ fading | Weibull fading |
|---|---|---|---|---|
| Non-cooperation | − 5 | 0 | 0.863 | 0.9378 |
| | | 0.05 | 0.8249 | 0.8943 |
| | 5 | 0 | 0.9985 | 0.9971 |
| | | 0.05 | 0.9449 | 0.9494 |
| | 10 | 0 | 0.9954 | 0.9997 |
| | | 0.05 | 0.9491 | 0.9500 |
| OR rule | − 5 | 0 | 0.9998 | 0.9902 |
| | | 0.05 | 0.9998 | 0.9902 |
| | 5 | 0 | 0.9985 | 0.9971 |
| | | 0.05 | 0.9985 | 0.9971 |
| | 10 | 0 | 0.9954 | 0.9997 |
| | | 0.05 | 0.9954 | 0.9997 |
| Majority rule | − 5 | 0 | 0.9794 | 0.9902 |
| | | 0.05 | 0.9594 | 0.9902 |
| | 5 | 0 | 1.0000 | 0.9971 |
| | | 0.05 | 0.9985 | 0.9971 |
| | 10 | 0 | 0.9954 | 0.9997 |
| | | 0.05 | 0.9954 | 0.9997 |
| AND rule | − 5 | 0 | 0.4781 | 0.7274 |
| | | 0.05 | 0.3803 | 0.5725 |
| | 5 | 0 | 0.9689 | 0.9971 |
| | | 0.05 | 0.7502 | 0.7716 |
| | 10 | 0 | 0.9954 | 0.9997 |
| | | 0.05 | 0.7698 | 0.7736 |

**Table 3** The impact of $p(\mathcal{H}_0)$ and $p(\mathcal{H}_1)$ on TEP in Nakagami$m$ fading channel

| $q$ | Fusion rule | Nakagami-$m$ fading ($m = 1.5$) | | |
|---|---|---|---|---|
| | | $p(\mathcal{H}_0)$ | $p(\mathcal{H}_1)$ | TEP |
| 0 | OR-rule | 0.25 | 0.75 | 0.1835 |
| | MAJORITY-rule | 0.25 | 0.75 | 0.1011 |
| | AND-rule | 0.25 | 0.75 | 0.2369 |
| 0.15 | OR rule | 0.25 | 0.75 | 0.4884 2 |
| | MAJORITY-rule | 0.25 | 0.75 | 0.1441 |
| | AND-rule | 0.25 | 0.75 | 0.2392 |
| 0.15 | OR rule | 0.75 | 0.25 | 0.1775 |
| | MAJORITY-rule | 0.75 | 0.25 | 0.3341 |
| | AND-rule | 0.75 | 0.25 | 0.7163 |
| 0.15 | OR rule | 0.5 | 0.5 | 0.3332 |
| | MAJORITY-rule | 0.5 | 0.5 | 0.2386 |
| | AND-rule | 0.5 | 0.5 | 0.4779 |

**Table 4** The impact of $p(\mathcal{H}_0)$ and $p(\mathcal{H}_1)$ on TEP in Weibull fading channel

| $q$ | Fusion rule | Weibull fading ($v = 4$) | | |
|---|---|---|---|---|
| | | $p(\mathcal{H}_0)$ | $p(\mathcal{H}_1)$ | TEP |
| 0 | OR-rule | 0.25 | 0.75 | 0.1772 |
| | MAJORITY-rule | 0.25 | 0.75 | 0.0751 |
| | AND-rule | 0.25 | 0.75 | 0.2284 |
| 0.15 | OR rule | 0.25 | 0.75 | 0.4862 |
| | MAJORITY-rule | 0.25 | 0.75 | 0.1248 |
| | AND-rule | 0.25 | 0.75 | 0.2337 |
| 0.15 | OR rule | 0.75 | 0.25 | 0.1708 |
| | MAJORITY-rule | 0.75 | 0.25 | 0.2799 |
| | AND-rule | 0.75 | 0.25 | 0.7020 |
| 0.15 | OR rule | 0.5 | 0.5 | 0.3288 |
| | MAJORITY-rule | 0.5 | 0.5 | 0.2018 |
| | AND-rule | 0.5 | 0.5 | 0.4677 |

In Tables 5 and 6, the optimal of $N$ and $\lambda$ are investigated in Nakagami-$m$, and Weibull fading channels, respectively. The parameters considered for these tables are $p(\mathcal{H}_0) = p(\mathcal{H}_1) = 0.5$, $N_s = 1000$, $N = 20$, $P_f = 0.05$, and $\bar{\gamma} = -10$dB. It can be noted that when $N_{opt} = 7$ and $\lambda_{opt} = 1.104$ the minimum TEP cab be obtained as 0.0027 in the presence of Weibull fading.

**Table 5** The optimal values of $N$ and $\lambda$ for various fusion rules in Nakagami-$m$ fading channel

| Fusion rule | Nakagami-$m$ fading ($m = 1.5$) | | |
|---|---|---|---|
| | $N_{opt}$ | $\lambda_{opt}$ | TEP |
| $k = 1$ rule | 1 | 1.153 | 0.0864 |
| $k = 2$ rule | 3 | 1.138 | 0.0222 |
| $k = 3$ rule | 5 | 1.123 | 0.0053 |
| $k = 4$ rule | 8 | 1.104 | $9.91 \times 10^{-4}$ |
| $k = 5$ rule | 9 | 1.084 | $5.71 \times 10^{-4}$ |
| $k = 6$ rule | 11 | 1.074 | $1.17 \times 10^{-4}$ |

**Table 6** The optimal values of $N$ and $\lambda$ for various fusion rules in Weibull fading channel

| Fusion rule | Weibull fading ($v = 4$) | | |
|---|---|---|---|
| | $N_{opt}$ | $\lambda_{opt}$ | TEP |
| $k = 1$ rule | 1 | 1.15 | 0.0761 |
| $k = 2$ rule | 3 | 1.138 | 0.0161 |
| $k = 3$ rule | 5 | 1.119 | 0.0047 |
| $k = 4$ rule | 7 | 1.104 | 0.0027 |
| $k = 5$ rule | 9 | 1.084 | 0.0042 |
| $k = 6$ rule | 11 | 1.078 | $1.31 \times 10^{-4}$ |

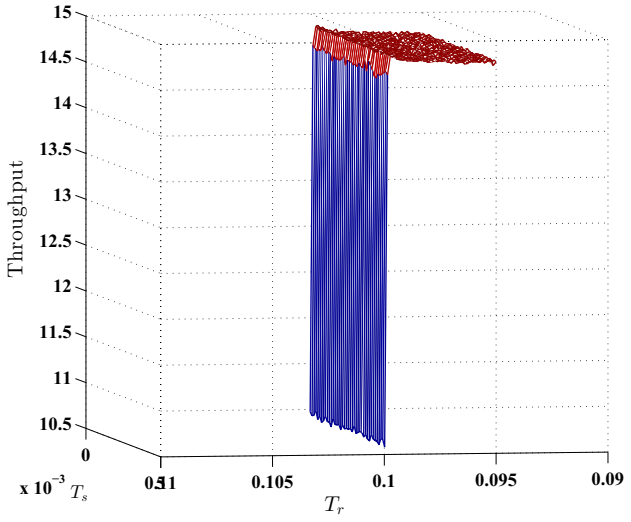



**Fig. 17** Energy efficiency as a function of $\lambda$ is shown in Nakagami-$m$ fading channel ($N = 10$, $q = 0.05$, $m = 1.5$, $\bar{\gamma} = -10$dB, and $N_s = 1000$)

From Figs. 17 and 18, Tables 7 and 8 are drawn where optimal values of $T_s$ and $T_r$ are given for various values of $q$ in Nakagami-$m$, and Weibull fading channels. We consider $p(\mathcal{H}_0) = 0.5$, $N = 10$, $\lambda = 1.125$, and $\bar{\gamma} = -10$dB for

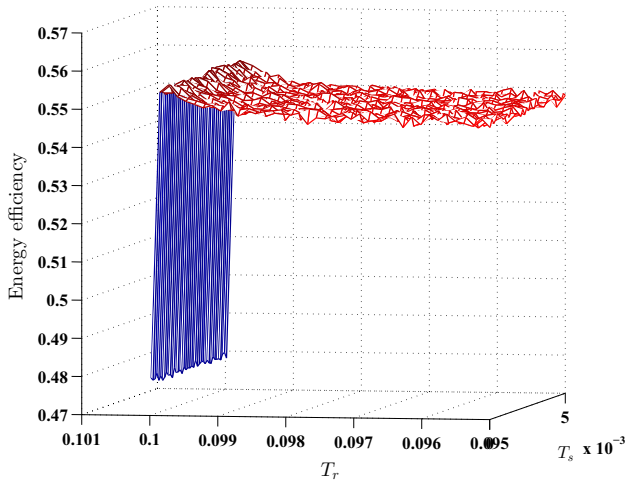



**Fig. 18** Energy efficiency as a function of $T_s/T_r$ is shown in Nakagami-$m$ fading channel ($N = 10$, $q = 0.05$, $m = 1.5$, $\bar{\gamma} = -10$dB, and $N_s = 1000$)

**Table 7** Trade-off between $\mathcal{B}_{avg}/\eta$ and $T_s/T_r$ for various values of $q$

| $q$ | $N_s$ | Nakagami-$m$ fading ($m = 1.5$) | | | |
|---|---|---|---|---|---|
| | | $T_s$ in ms | $T_r$ in ms | $\mathcal{B}_{avg}$ | $\eta$ |
| 0 | 2400 | 0.4 | 99.6 | 14.9339 | 0.5608 |
| | 7000 | 1.2 | 98.8 | 14.9242 | 0.5500 |
| | 10,067 | 1.7 | 98.3 | 14.9116 | 0.5428 |
| 0.05 | 2400 | 0.4 | 99.6 | 12.9437 | 0.5263 |
| | 7000 | 1.2 | 98.8 | 12.9355 | 0.5156 |
| | 10,067 | 1.7 | 98.3 | 12.9261 | 0.5057 |

**Table 8** Trade-off between $\mathcal{B}_{avg}/\eta$ and $T_s/T_r$ for various values of $q$

| $q$ | $N_s$ | Weibull fading ($v = 4$) | | | |
|---|---|---|---|---|---|
| | | $T_s$ in ms | $T_r$ in ms | $\mathcal{B}_{avg}$ | $\eta$ |
| 0 | 2400 | 0.4 | 99.6 | 14.8488 | 0.5541 |
| | 7000 | 1.2 | 98.8 | 14.7630 | 0.5404 |
| | 10,067 | 1.7 | 98.3 | 14.6918 | 0.5325 |
| 0.05 | 2400 | 0.4 | 99.6 | 12.8874 | 0.5216 |
| | 7000 | 1.2 | 98.8 | 12.7529 | 0.5138 |
| | 10,067 | 1.7 | 98.3 | 12.6491 | 0.5025 |

both the tables. The impact of $q$ on the $\mathcal{B}_{avg}$ and $\eta$ is investigated, It is observed that the erroneous faded channels degrades the $\mathcal{B}_{avg}$ and $\eta$ significantly. Also, it can be noted that when $T_s$ or $T_r$ increases both $\mathcal{B}_{avg}$ and $\eta$ decreases vice versa.

## 5 Conclusions

The performance of CCRS in the presence of noise plus Nakagami-$m$ or Weibull fading under channel error constraints has been investigated. The energy efficiency is enhanced by 19% when CCRS is operated over Nakagmi-$m$ fading environment as compared to Rayleigh fading environment for a fixed parameter values of the network. Various one-bit hard decision fusion rules have been performed at FC. The performance through cooperative missdetection, total error probability, and sensing time and throughput trade-off has been investigated for various parameters of the system such as $m$, $v$, $N$, $N_s$, $\bar{\gamma}$, $\lambda$, $T_s$, and $q$. Excellent performance improvement has been observed with $k = 1$ rule (OR fusion rule) among all fusion rules ($k > 1$). The analytical expressions have been formulated as well as validated through MATLAB based simulation model. High values of $\bar{\gamma}$ or $N_s$ or $N$, decreases both the missdetection and total error rate appreciably. The CCRS performance has been degraded when sensing and reporting channels are erroneous. Considerable performance improvement has been observed for bigger values of $m$ and $v$. The cooperative detection increases significantly with $\bar{\gamma}$ even for bigger values of $q$ present in sensing and reporting channels. Cooperative missdetection decreases as $N_s$ increases. As $N$ or $\lambda$ increases, initially TEP decreases and TEP increases when further increases of $N$ or $\lambda$. The optimum value of $N$ and $\lambda$ have been found at which minimum TEP for various 1-bit fusion rules has been obtained. The maximum throughput and energy efficiency have been obtained at a particular value of $\lambda$ under different fusion rules. It can be noted that when $T_s$ increases throughput and energy efficiency decrease vice versa, trade-off between sensing time and throughput/energy efficiency has been observed. The suboptimal fusion rule has been observed for different system parameters. The present work studied in this paper is useful for designing a CCRS in any faded environment under channel error constraints for the next generation wireless networks.

## Appendix

### Proof of (10)

The proof of (10) is provided here. From (7), $\mathcal{Q}$-function can be rewritten as $\mathcal{Q}(x) = \frac{1}{2} erfc\left(\frac{x}{\sqrt{2}}\right)$, where $erfc(.)$ represents complementary error function and can be expressed as [32]:

$$erfc(x) = 1 - \frac{2}{\sqrt{\pi}} \sum_{z=0}^{\infty} \frac{(-1)^z x^{2z+1}}{2^{z+\frac{1}{2}} z!(2z+1)} \tag{33}$$

Using (33), $\mathcal{Q}$-function can be obtained as [32]:

$$\mathcal{Q}(t) = \frac{1}{2} - \frac{1}{\sqrt{\pi}} \sum_{z=0}^{\infty} \frac{(-1)^z t^{2z+1}}{2^{z+\frac{1}{2}} z!(2z+1)}. \tag{34}$$

where $t = \left[\frac{(\lambda/\sigma_u^2)}{1+\gamma} - 1\right]\sqrt{\frac{\tau_s f_s}{2}}$. Further simplification of (8) after inserting (9) and (34) in (8), $\bar{P}_d$ can be expressed as

$$\begin{aligned} \bar{P}_{d,i}^{Naka} = & \frac{1}{2}\left(\frac{m}{\bar{\gamma}}\right)^m \frac{1}{\Gamma(m)} \underbrace{\int_0^{\infty} \gamma^{m-1} \exp\left(-\frac{m\gamma}{\bar{\gamma}}\right) d\gamma}_{I_1} \\ & - \frac{1}{\sqrt{\pi}}\left(\frac{m}{\bar{\gamma}}\right)^m \frac{1}{\Gamma(m)} \sum_{z=0}^{\infty} \frac{(-1)^z (\tau_s f_s)^{z+\frac{1}{2}}}{2^{2z+1}(2z+1)z!} \\ & \underbrace{\times \int_0^{\infty} \left(\frac{\lambda/\sigma^2}{1+\gamma} - 1\right)^{2z+1} \gamma^{m-1} \exp\left(-\frac{m\gamma}{\bar{\gamma}}\right) d\gamma}_{I_2} \end{aligned} \tag{35}$$

The solution to the first integral part, $I_1$ in (35) can be obtained using [33, (3.351.3)] as

$$\begin{aligned} I_1 &= \frac{1}{2}\left(\frac{m}{\bar{\gamma}}\right)^m \frac{1}{\Gamma(m)} \int_0^{\infty} \gamma^{m-1} \exp\left(-\frac{m\gamma}{\bar{\gamma}}\right) d\gamma, \\ &= \frac{1}{2}. \end{aligned} \tag{36}$$

The solution to second integral part, $I_2$ in (35) can be obtained as follows, from $I_2$

$$\left(\frac{\lambda/\sigma_u^2}{1+\gamma} - 1\right)^{2z+1} = \sum_{m=0}^{2z+1} (-1)^{2z+m+1} \binom{2z+1}{m} \left(\frac{\lambda/\sigma_u^2}{1+\gamma}\right)^m. \tag{37}$$

Then $I_2$ can be rewritten as

$$\begin{aligned} I_2 = & \sum_{m=0}^{2z+1} (-1)^{2z+m+1} \binom{2z+1}{m} \\ & \times \int_0^{\infty} \left(\frac{\lambda/\sigma_u^2}{1+\gamma}\right)^m \gamma^{m-1} \exp\left(-\frac{m\gamma}{\bar{\gamma}}\right) d\gamma \end{aligned} \tag{38}$$

The resultant integral in (38) is difficult to solve because of the complexity involved and an alternate approach (variable substitution method [34, 35]) is adopted here by substituting $\gamma$ with $\tan\theta$ and $d\gamma$ with $\sec^2\theta d\theta$ which gives an integral with definite limits from the integral with infinite limits and can go for numerical evaluation. Based on this method, (38) can be rewritten as

$$I_2 = \sum_{m=0}^{2z+1} (-1)^{2z+m+1} \binom{2z+1}{m} \times \int_0^{\pi/2} \left(\frac{\lambda/\sigma_u^2}{1+\tan\theta}\right)^m \tan\theta^{m-1} \exp\left(-\frac{m\tan\theta}{\bar{\gamma}}\right) \sec^2\theta d\theta. \quad (39)$$

## Proof of (12)

The proof of (12) is provided here. The expression for $\bar{P}_{d,i}^{Wieb}$ at $i$th CR node can be derived by substituting (7) and (11) in (8):

$$\bar{P}_{d,i}^{Wieb} = \frac{v}{2}\left[\frac{\Gamma(p)}{\bar{\gamma}}\right]^{\frac{v}{2}} \underbrace{\int_0^{\infty} \gamma^{\frac{v}{2}-1} \exp\left[-\left\{\frac{\gamma\Gamma(p)}{\bar{\gamma}}\right\}^{\frac{v}{2}}\right] d\gamma}_{I_1} - \frac{v}{2}\left[\frac{\Gamma(p)}{\bar{\gamma}}\right]^{\frac{v}{2}} \sum_{z=0}^{\infty} \frac{(-1)^z (\tau_s f_s)^{z+\frac{1}{2}}}{2^{2z+1}(2z+1)z!} \underbrace{\times \int_0^{\infty} \left(\frac{\lambda/\sigma_u^2}{1+\gamma}\right)^m \gamma^{\frac{v}{2}-1} \exp\left[-\left\{\frac{\gamma\Gamma(p)}{\bar{\gamma}}\right\}^{\frac{v}{2}}\right] d\gamma}_{I_2} \quad (40)$$

The solution to the first integral part, $I_1$ in (40) can be obtained using [33, (3.326.2)] as

$$I_1 = \frac{v}{2}\left[\frac{\Gamma(p)}{\bar{\gamma}}\right]^{\frac{v}{2}} \int_0^{\infty} \gamma^{\frac{v}{2}-1} \exp\left[-\left\{\frac{\gamma\Gamma(p)}{\bar{\gamma}}\right\}^{\frac{v}{2}}\right] d\gamma, = \frac{1}{2}. \quad (41)$$

The $I_2$ part can be modified as

$$I_2 = \sum_{m=0}^{2z+1} (-1)^{2z+m+1} \binom{2z+1}{m} \times \int_0^{\infty} \left(\frac{\lambda/\sigma_u^2}{1+\gamma}\right)^m \gamma^{\frac{v}{2}-1} \exp\left[-\left\{\frac{\gamma\Gamma(p)}{\bar{\gamma}}\right\}^{\frac{v}{2}}\right] d\gamma \quad (42)$$

The integral part in (42) is not easy to solve, following the same variable substitution approach presented in Sect. 2.1, an alternative expression for (42) can be obtained as

$$I_2 = \sum_{m=0}^{2z+1} (-1)^{2z+m+1} \binom{2z+1}{m} \times \int_0^{\pi/2} \left(\frac{\lambda/\sigma_u^2}{1+\tan\theta}\right)^m \tan\theta^{\frac{v}{2}-1} \exp\left[-\left\{\frac{\tan\theta\Gamma(p)}{\bar{\gamma}}\right\}^{\frac{v}{2}}\right] \sec^2\theta d\theta. \quad (43)$$

## References


1. Ball, S., Ferguson, A., & Rondeau, T. W. (November 2005). Consumer applications of cognitive radio defined networks. In *Proceedings of IEEE IEEE symposium on new frontiers in dynamic spectrum access networks* (pp. 518–525). Baltimore.
2. Gorcin, A., & Arslan, H. (October 2008). Public safety and emergency case communications: Opportunities from the aspect of cognitive radio. In: *Proceedings of IEEE symposium on new frontiers in dynamic spectrum access networks* (pp. 1–10). Chicago.
3. Sasipriya, S., & Vigneshram, R. (December 2016). An overview of cognitive radio in 5G wireless communications. In: *Proceedings of IEEE Computational Intelligence and Computing Research (ICCIC)* (pp. 1–5). Chennai.
4. Zhang, Z., Zhang, W., Zeadally, S., Wang, Y., & Liu, Y. (2015). Cognitive radio spectrum sensing framework based on multi-agent arc hitecture for 5G networks. *IEEE Communications Letters*, *22*(6), 34–39.
5. Ghasemi, A., & Sousa, E. S. (2007). Opportunistic spectrum access in fading channels through collaborative sensing? *IEEE Transactions on Wireless Communications*, *2*(2), 71–82.
6. Digham, F. F., Alouini, M.-S., & Simon, M. K. (2007). On the energy detection of unknown signals over fading channels. *IEEE Transactions on Communications*, *55*(1), 21–24.
7. Zhang, W., Mallik, R., & Letaief, K. B. (2009). Optimization of cooperative spectrum sensing with energy detection in cognitive radio networks. *IEEE Transactions on Wireless Communications*, *8*(12), 5761–5766.
8. Nallagonda, S., Chandra, A., Roy, S. D., Kundu, S., Kukolev, P., & Prokes, A. (2016). Detection performance of cooperative spectrum sensing with hard decision fusion in fading channels. *International Journal of Electronics (Taylor and Francis)*, *103*(2), 297–321.
9. Nallagonda, S., Chandra, A., Roy, S. D., & Kundu, S. (2017). Analytical performance of soft data fusion-aided spectrum sensing in hybrid terrestrial-satellite networks. *International Journal of Satellite Communications and Networking*, *35*(5), 461–480.
10. Nallagonda, S., Roy, S. D., Kundu, S., Ferrari, G., & Raheli, R. (2018). Censoring-based cooperative spectrum sensing with improved energy detectors and multiple antennas in fading channels. *IEEE Transactions on Aerospace and Electronic Systems*, *54*(2), 537–553.
11. Angueira, P., Fadda, M., Morgade, J., Murroni, M., & Popescu, V. (2016). Field measurements for practical unlicensed communication in the UHF band. *Telecommunication Systems*, *61*(3), 443–449.
12. Banavathu, N. R., & Khan, M. Z. A. (March 2016). On the throughput maximization of cognitive radio using cooperative spectrum sensing over erroneous control channel. In: *Proceedings of IEEE National Conference on Communication (NCC)* (pp. 1–6), IIT Guwahati, India.
13. Liang, Y. C., Zeng, Y., Peh, E. C. Y., & Hoang, A. T. (2008). Sensing-throughput tradeoff for cognitive radio networks. *IEEE Transactions on Wireless Communications*, *7*(4), 1326–1337.
14. Shi, Z., Teh, K. C., & Li, K. H. (2013). Energy efficient joint design of sensing and transmission durations for protection of primary user in cognitive radio systems. *IEEE Communications Letters*, *17*(3), 565–568.
15. Wu, Y., & Tsang, D. H. K. (2011). Energy efficient spectrum sensing and transmission for cognitive radio system. *IEEE Communications Letters*, *15*(5), 545–547.
16. Balam, S. K., Siddaiah, P., & Nallagonda, S. (May 2018). Throughput analysis of cooperative cognitive radio network over generalized $\kappa-\mu$ and $\eta-\mu$ fading channels. In: *Wireless*

*Networks* (pp. 1–14). Springer. https://doi.org/10.1007/s11276-018-1758-4.
17. Kiran Kumar, G., Mahendar, G., & Nallagonda, S. (February 2018). Cooperative spectrum sensing in log-normal shadowing environment with erroneous sensing and reporting channels. In: *Proceedings of IEEE International Conference on on Advances in Electronics, Computer and Communications (ICAECC)* (pp. 1–5). Bangalore, India.
18. Nallagonda, S., Chandra, A., Roy, S. D., & Kundu, S. (February 2018). The effects of channel knowledge on cooperative spectrum sensing in Nakagami-$n/q$ fading channels. In: *Wireless Networks* (pp. 1–13). https://doi.org/10.1007/s11276-018-1685-4.
19. Bhowmick, A., Roy, S. D., & Kundu, S. (2016). Sensing throughput trade-off for an energy efficient cognitive radio network under faded sensing and reporting channel. *International Journal of Communication Systems*, *29*, 1208–1218.
20. -Juarez, M. C., & Ghogho, M. (2011). Spectrum sensing and throughput trade-off in cognitive radio under outage constraints over nakagami fading. *IEEE Communications Letters*, *15*(10), 1110–1113.
21. Quang, P. M., Duy, T. T., & Bao, V. N. Q. (December 2016). Performance evaluation of underlay cognitive radio networks over Nakagami-*m* fading channels with energy harvesting. In: *Proceedings of IEEE International Conference on Advanced Technologies for Communications (ATC)* (pp. 108–113). Vietnam.
22. Van Nguyen, B., Jung, H., Har, D., & Kim, K. (2018). Performance analysis of a cognitive radio network with an energy harvesting secondary transmitter under Nakagami-*m* fading. *IEEE Access*, *6*, 4135–4144.
23. Musavian, L., & Aissa, S. (2010). Effective capacity of delay-constrained cognitive radio in Nakagami fading channels. *IEEE Transactions on Wireless Communications*, *9*(3), 1054–1062.
24. Han, B., Zeng, M., Guo, Q., Jiang, H., Zhang, Q., & Feng, L. (2018). Energy-efficient sensing and transmission for multi-hop relay cognitive radio sensor networks. *China Communications*, *15*(9), 106–117.
25. Haider, F., Wang, C. X., Haas, H., Hepsaydir, E., Ge, X., & Yuan, D. (2015). Spectral and energy efficiency analysis for cognitive radio networks. *IEEE Transactions on Wireless Communications*, *14*(6), 2969–2980.
26. Amjad, M., Rehmani, M. H., & Mao, S. (2018). Wireless multimedia cognitive radio networks: A comprehensive survey. *IEEE Communications Surveys and Tutorials*, *20*(2), 1056–1103.
27. Tang, L., Chen, Y., Hines, E. L., & Alouini, M.-S. (2011). Effect of primary user traffic on sensing-throughput tradeoff for cognitive radios. *IEEE Transactions on Wireless Communications*, *10*(4), 1063–1068.
28. Cao, H., & Yang, S. (April 2017). Sensing-throughput tradeoff in cognitive radio network based on high activity of primary user. In: *Proceedings of IEEE International Conference on Computing Intelligence and Information System* (pp. 121–125). Nanjing, China.
29. Tandra, R., & Sahai, A. (2008). SNR walls for signal detection. *IEEE Journal of Selected Topics in Signal Processing*, *2*(1), 4–17.
30. Banavathu, N. R., & Khan, M. Z. A. (2017). Optimal number of cognitive users in *K*-out-of-*M* rule. *IEEE Wireless Communications Letters*, *6*(5), 606–609.
31. Hu, H., Zhang, H., Yu, H., Chen, Y., & Jafarian, J. (2015). Energy-efficient design of channel sensing in cognitive radio networks. *Computers and Electrical Engineering*, *42*, 207–220.
32. Gurugopinath, S., & Shobitha, S. (January 2017). Energy-based bayesian spectrum sensing over $\alpha - \kappa - \mu$ fading channels. In: *Proceedings of IEEE International Conference on Communication Systems and Networks (COMSNETS)* (pp. 95–100). Bengaluru, India.
33. Gradshteyn, I. S., & Ryzhik, I. M. (2007). *Table of integrals, series and products* (7th ed.). San Diego: Elsevier.
34. Chandra, A., Bose, C., & Bose, M. K. (2013). Performance of non-coherent MFSK with selection and switched diversity over Hoyt fading channel. *Wireless Personal Communications*, *68*(2), 379–399.
35. Haghani, S., & Beaulieu, N. C. (2006). *M*-ary NCFSK with S + N selection combining in Rician fading. *IEEE Transactions on Communications*, *54*(3), 491–498.



**Kiran Kumar Godugu** received the B.Tech. degree in Electronics and Communication Engineering from JNTU Hyderabad, India, in 2008 and the M.Tech. degree in Digital Communication and Networking from SRM University, TamilNadu, India, 2011. He served as an Assistant Professor from 2011 to 2015 in Nagole Institute of Technology and Science. From 2015, he is presently working as an Assistant Professor in the Department of ECE Marri Laxman Reddy Institute of Technology and Management, Dundigal, India. His research interests include spectrum sensing issues in cognitive radio networks in different and fading environments. As of today, he has published a good number of journals and conferences.



**Ashok kumar Nallagonda** received his B.Tech. degree in Electronics and Communication Engineering in 2005 from Acharya Nagarjuna University, Andhra Pradesh, India. He received M. Tech. degree in Radar and Microwave Engineering in 2008 and Ph.D. in CMOS image processing in 2017 from Delhi Technological University, Delhi, India. Dr. A. K. Nallagonda served as an Assistant Professor from 2008 to 2017 in Avanthi group of institutions and presently, working as an Associate Professor in the Department of ECE at Kasireddy Narayan Reddy College of Engineering and Research, affiliated to JNTU Hyderabad, India. His research interests include image processing, antennas, radar systems, and microwave engineering. As of today, he has published a good number of research papers in various international conferences and journals. Dr. Ashok is a member of IEEE and also served as a Reviewer for several IEEE conferences.



**Srinivas Nallagonda** received the B.E. degree in Electronics and Communication Engineering from Osmania University, Hyderabad, India, in 2006, the M.Tech. degree in telecommunication engineering, and the Ph.D. degree in wireless communications from National Institute of Technology (NIT), Durgapur, India, in 2009 and 2014, respectively. He served as an Assistant Professor from 2009 to 2010 in Swami Ramananda Tirtha Institute of Science and Technology, Nalgonda and from 2014 to 2017 in Maturi Venkata Subba Rao Engineering College, Hyderabad. He is currently working as an Associate Professor in the Department of ECE, Marri Laxman Reddy Institute of Technology and Management, Dundigal, affiliated to JNTU Hyderabad, India. His research interests include fading models, diversity techniques, spectrum sensing issues in cognitive radio networks, energy harvesting, and green communications. As of today, he has published around 50 research papers in various international conferences (IWCMC, WPMC, NCC, INDICON, and ANTS etc.), journals (IEEE, Wiley, Springer, Taylor and Francis, IEICE, and Inderscience etc.), and book chapters (Springer and IGI Global). Dr. Nallagonda is a member of IEEE and also served as a Reviewer for several IEEE conferences and journals.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

International Journal of Environmental Science and Technology
https://doi.org/10.1007/s13762-019-02414-3

REVIEW

# Polycyclic aromatic hydrocarbons: soil pollution and remediation

**Sakshi**[1] · **S. K. Singh**[1] · **A. K. Haritash**[1]

Received: 7 January 2019 / Revised: 27 April 2019 / Accepted: 13 May 2019
© Islamic Azad University (IAU) 2019

**Abstract**
Soil is an important environmental matrix to support the life of all organisms directly or indirectly. Despite being the ultimate sink for all pollutants, it has been neglected for long, which has negatively affected the quality of the soil. Disposal of pollutants has resulted in changes in properties of soils and introduction of toxicity into it. The presence of heavy metals, pesticides, polychlorinated biphenyls and polycyclic aromatic hydrocarbons (PAHs) affects all forms of life since these chemicals have associated toxicity, mutagenicity, and carcinogenicity. PAHs are typical pollutants of soil which result in alteration in grain size, porosity and water-holding capacity of soil and affect diversity/population of microbes adversely. Significant changes in permeability, volume, plasticity, etc., are also brought about resulting in poor quality of contaminated soils. Considering the toxicity and global prevalence of PAHs, remediation of contaminated soils has become a challenge. Therefore, it is important to understand the detailed mechanism of physical, chemical or biological changes in soil. Simultaneously, it becomes pertinent to identify the environmentally sustainable treatment options for remediation of contaminated sites. Whereas physical and chemical treatment methods are either cost, chemical, or energy prohibitive, the biological treatment is emerging as an efficient and effective option which employs microorganisms for mitigation. Microorganisms are known for their enzyme-catalyzed catabolic activity when degradation/mineralization of a pollutant is aimed at and can prove useful in degradation of PAHs. Therefore, the present study reviews the effects of PAHs on soil properties, different remediation techniques and the role of microorganisms in remediating contaminated sites.

**Keywords** Advanced oxidation process · Bioremediation · Geotechnical properties · Toxicity · Treatment

## Introduction

With increasing awareness of the adverse effects of hazardous chemicals on human health and environment, the remediation of such chemicals has received more attention internationally. Environmental pollution caused by dumping of solid waste, untreated industrial effluents, persistent organic pollutants (POPs) like PAHs, PCBs, and pesticides, is a critical problem. These compounds, directly or indirectly released in the environment, are found to be very toxic and lead to contamination of water, soil, and air. Soils contaminated with organic pollutants and inorganic contaminants like heavy metals have high potential health risk because such contaminants can enter food chain (Xiaojun et al. 2007). Contamination due to leakage/spillage of crude oil can adversely affect exposed organisms since it is an intricate mixture of toxic aliphatic and complex aromatic hydrocarbons. Studies have established that toxicity increases with an increase in complexity of chemical structure (Patnaik 1999). Moreover, since complex/branched structure results in chemical stability, such pollutants are stable, resistant to biodegradation and as a result accumulate in environment which leads to accumulation of toxicity. When hazardous chemicals are introduced into soil, it may lead to soil contamination and can affect agricultural production too (Ibrahim 2004). Whereas contamination of soil can directly affect human health, the presence of PAHs in marine ecosystem may affect humans indirectly (Pathak et al. 2011). Soil contaminated with hazardous petroleum hydrocarbons can adversely affect human health, as it gets absorbed through direct dermal contact. Therefore, it is high time to adopt effective measures to reclaim contaminated sites and control other sites from being getting contaminated. If no efforts

Editorial responsibility: Ta Yeong Wu.

✉ A. K. Haritash
akharitash@dce.ac.in

[1] Department of Environmental Engineering, Delhi Technological University, Shahbad Daulatpur 110042, Delhi, India

Published online: 25 May 2019

are made to clean the contaminated sites, then the harmful effects of such sites will persist in environment, and add up to the Brownfield list.

Organic pollutants associated with petroleum such as PAHs are considered as the hazardous pollutants due to their toxic, mutagenic, and carcinogenic effects. They have varying physicochemical, and toxicological characteristics according to their molecular weight (Table 1). PAHs are composed of fused benzene rings consisting carbon and hydrogen arranged in a simple to complex structural configuration, i.e., in linear, angular, or cluster arrangements. PAHs are classified according to their aromatic benzene ring number, i.e., consisting two or three benzene rings are classified as low molecular weight (LMW) PAHs such as naphthalene, anthracene, and phenanthrene, and PAHs consisting more than three benzene rings are classified as high molecular weight (HMW) PAHs (pyrene, chrysene, benzo(a) pyrene, coronene, etc.). Usually, these are produced from anthropogenic as well as natural activities (Table 2). Forest fires, oil seeps, volcanic eruptions and exudates from trees are some natural sources; and burning of fossil fuel, solid biomass fuel (SBF) such as wood, crop residue or dung cake, coal tar, crude oil or petroleum spill (oil spillage and leakage), high-temperature industrial processes, and petroleum refinery effluent are some anthropogenic sources (Haritash and Kaushik 2009; Sharma and Jain 2019). There are many other natural as well as anthropogenic sources of PAHs and in addition viz. smoke from wood-burning stoves, automotive emissions, creosote waste materials and manufactured gas plants (coal gasification) are some other anthropogenic sources of PAHs (Abdel-Shafy and Mansour 2016). More than 100 unique PAHs are widely distributed and ubiquitously found in the environment, 16 PAHs among them have been classified and identified as priority pollutants, and some of them are classified as possible human carcinogen by U.S. EPA and European Union (Potin et al. 2004). PAHs are ubiquitously present around the globe and have contaminated the soil in different geographical regions of the world (Table 3), and they are hydrophobic with low water solubility so they have tendency to bind with organic matter present in soil/sediments. PAHs are relatively stable contaminants showing recalcitrant nature in soils, and it is difficult to degrade them as compared to many other organic contaminants. Considering the toxic properties and persistence of PAHs in soil, reclamation of PAH-contaminated soils become imperative. There are a number of physical (solvent extraction, air sparging, thermal desorption, microwave heating, vitrification, electrokinetic treatments, etc.); chemical (oxidation using Fenton's reagent, ozone, etc., photocatalytic degradation); and biological (microbial, landfarming, composting, phytoremediation, etc.) methods being used to remediate the contaminated site. The selection of treatment method is subject to type of contaminant, soil type, interferences present, and risk analysis associated with the techniques. It has been studied that during PAH degradation some epoxides and dihydrodiols are formed which are found to be more toxic than the parent PAHs (Cerniglia and Sutherland 2010). So, identification of intermediate PAH metabolites is also essential during its degradation. Although physical and chemical methods are effective and efficient, most of them are energy, cost, or chemical intensive. Moreover, these methods are responsible for generation of secondary pollutants. Keeping in view, the limitations associated with physicochemical treatment, the biological methods are gaining popularity owing to their ecofriendly nature and bioconversion of toxic pollutants into harmless/nontoxic chemical species.

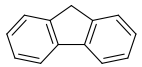
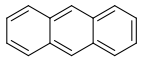
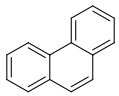
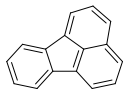
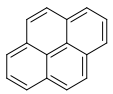
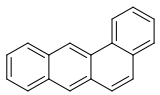
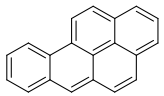
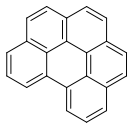
## Properties of contaminated soils

### Physicochemical properties

Soil is the major sink for a broad range of harmful contaminants, and these contaminants may affect the physical, chemical, and biological properties of soil. Crude oil is the major source of PAHs which may easily diffuse and get absorbed inside soil particles. Soil is extremely sensitive to different contaminants, and its behavior changes under different environmental conditions. Properties of soil are affected by different activities occurring on land and different types of contaminant, crude oil accidental spillage, and leakage, etc. Longer contamination may affect geotechnical properties such as Atterberg limits, permeability, hydraulic conductivity, strength parameters, consolidation, compaction and shear strength as well as the biological properties (biomass and enzyme activity) of soils. Since microbes have the ability to degrade hazardous contaminants, they are involved in remediation and soil self-purification processes. PAHs are toxic even to microbes and contaminated soil has lower self-purification capacity due to a decrease in microbial activity of soil (Hreniuc et al. 2015). Effects on properties of soil may be due to partial or full replacement of the soil pore liquid with PAH, and chocking of soil pores with contaminant due to which soil aeration and water infiltration could reduce.

The degree of effects of PAH is regulated by certain important parameters of soil viz. grain size of soil particles, organic carbon associated, and to some extent pH of the soil. A number of studies have confirmed that binding of PAHs takes place preferentially on finer grain size, i.e., silt and clay (Magi et al. 2002). The clay particles have more surface area, and as a result, more bonding sites resulting in adsorption of PAHs firmly on finer fraction of soil. Fine particles also result in less porosity and hence lesser movement of adsorbed contaminants over the period of time which results in persistent toxicity and long-term effects.

**Table 1** Physicochemical properties of major PAHs

| S. no. | Name | Mol. formula | Chemical structure | CAS registry no.[a] | B.Pt. (°C)[a] | M.Pt. (°C)[a] | V.P. (Pa at 25 °C)[a] | Aq. Sol. (mg/l)[b] | Log $K_{ow}^{c}$ | IARC group[a] | Half-life (days)[d] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. | Naphthalene | $C_{10}H_8$ |  | 91-20-3 | 218 | 80.2 | 11 | 31.7 ± 0.26 | 3.35 | NE | 5.66 |
| 2. | Fluorene | $C_{13}H_{10}$ |  | 86-73-7 | 295 | 116.7 | 0.08 | 1.98 ± 0.04 | 4.18 | 3 | 15.14 |
| 3. | Anthracene | $C_{14}H_{10}$ |  | 120-12-7 | 342 | 216.4 | $1\times10^{-3}$ | 0.073 ± 0.005 | 4.54 | 3 | 123 |
| 4. | Phenanthrene | $C_{14}H_{10}$ |  | 85-01-8 | 340 | 100.5 | $2\times10^{-2}$ | 1.29 ± 0.07 | 4.57 | 3 | 14.97 |
| 5. | Fluoranthene | $C_{16}H_{10}$ |  | 206-44-0 | 375 | 108.8 | $1.2\times10^{-3}$ | 0.26 ± 0.002 | 5.22 | 3 | 191.4 |
| 6. | Pyrene | $C_{16}H_{10}$ |  | 129-00-0 | 150.4 | 393 | $6.0\times10^{-4}$ | 0.135 ± 0.005 | 5.18 | 3 | 283.4 |
| 7. | Benzo[a]anthracene | $C_{18}H_{12}$ |  | 56-55-3 | 438 | 158 | 0.548 | 0.014 ± 0.0002 | 5.91 | 2A | 343.8 |
| 8. | Benzo[a]pyrene | $C_{20}H_{12}$ |  | 50-32-8 | 495 | 179 | $4\times10^{-7}$ | 0.0038 ± 0.00031 | 5.98 | 2A | 421.6 |
| 9. | Benzo[ghi]perylene | $C_{22}H_{12}$ |  | 191-24-2 | 500 | 277 | $6\times10^{-8}$ | 0.00026 ± 0.00001 | 7.10 | 3 | 517.1 |

Mol. formula, molecular formula; CAS, chemical abstracts service; B.Pt., boiling point; M.Pt., melting point; V.P., vapor pressure; Aq. Sol., aqueous solubility; Log $K_{ow}$, octanol/water partition coefficients; IARC, International Agency for Research on Cancer

[a]WGPAH (2001)

[b]Mackay and Shiu (1977)

[c]Michele et al. (1985)

[d]Howard et al. (2005)

**Table 2** Sources of PAHs in contaminated soil

| S. no. | Type of sources | Activities | Process | References |
|---|---|---|---|---|
| 1. | Pyrogenic (associated with combustion of wood, petroleum product, coal) | Industrial (anthropogenic) activities | Wood burning | Wilson and Jones (1993) |
| | | | Burning of tires | Downard et al. (2015) |
| | | | Burning of fossil fuel | McRae et al. (2000) |
| | | | Burning of tobacco | Haussmann (2012) |
| | | | Burning of agricultural waste | Lai et al. (2009) |
| | | | Combustion of oil, diesel, coal and oil products | Kaushik et al. (2012) |
| | | Natural activities | Volcanic eruption | Kozak et al. (2017) |
| | | | Forest fire | Denis et al. (2012) |
| 2. | Petrogenic (associated with substances originate from crude oil/petroleum) | Industrial (anthropogenic) activities | Oil spill | Soriano et al. (2006) |
| | | | Outlets from oil refinery | Pettersen et al. (1997) |
| | | | Petroleum handling facilities like kerosene tank, generating plant, petrol stations, mechanic workshops, leaking pipeline, and airport fuel dump | Nganje et al. (2007) |
| | | | Petroleum oil industries and industrial activities | Varjani et al. (2017) |
| | | | Creosote, asphalt production | Abdel-Shafy and Mansour (2016) |
| | | | Used engine oil, jet fuel, kerosene | Kaushik and Haritash (2006), Kaushik et al. (2012) |
| | | | Manufactured gas plants | Abdel-Shafy and Mansour (2016) |
| | | Natural activities | Oil seeps | Pampanin and Sydnes (2013) |
| 3. | Biogenic | Natural activities | Surface waxes of leaves and insect cuticles | Oleszczuk and Baran (2005) |
| | | | | Krauss et al. (2005) |
| | | | Wood of tropical forests | Krauss et al. (2005) |
| | | | PAH synthesis in termite organisms | Ratola et al. (2006) |
| | | | Pine needles | |

Some of the studies have also confirmed that HMW-PAHs bind firmly to the soil compared to LMW-PAHs (García-Alonso et al. 2008). Since the inter-particle spaces in fine soil are less, the porosity-mediated effects are minimum resulting in nontransfer of HMW-PAHs from one location to another due to their hydrophobic nature. Some of the studies have also reported that within-the-particle movement of PAHs from the surface to core also takes place (Bogan and Trbovic 2003). The intra-particle movement is ascribed to be the function of organic carbon associated with the soil particles. The two types of domains which interact with contaminant are surface layer of humic acid and fulvic acid (HA–FA), and core formed of hard layer of humin. The HA–FA layer is porous, flexible, and lipophilic, and humin layer is rigid, hard, and glassy (Xing and Pignatello 1997). The humin bound PAHs are firmly bonded to the soil which have been subjected to diagenetic alteration (derived from weathering of sedimentary rocks) (Lueking et al. 2000). On the other hand, soils which have not been subjected to diagenetic alterations have organic matter with soft carbon fraction (HA–FA). Therefore, the effect as well as treatment strategy of contaminated soil significantly depends upon the grain size, its paedogenetic profile and the type/fraction of organic carbon associated with the soil. The pH of soil also, to some extent, regulates the degree of contamination. The HA–FA-rich acidic soils bind preferably with PAHs resulting in higher and persistent contamination, whereas high/basic pH results in frequent dissociation of soil particles which does not favor the binding of PAHs (Saba et al. 2010). Thus, organic carbon-rich fine-grained acidic soil is more prone to PAH contamination and its effect and is difficult to remediate.

## Biological properties

The biological activities, i.e., microbial biomass and enzymatic activities of soil, are highly sensitive to environmental pressure/parameters (Labud et al. 2007). Soil contamination may affect the microbial community/population and microbial activity/enzymatic activities of the soil. Various experiments have been done under controlled environments to investigate the adverse effect of petroleum hydrocarbon (PAHs) contamination on soil microorganisms and metabolic activities. Microbial activity may be inhibited due to the presence of high concentration of organic contaminants. Soil contamination with crude oil may develop anaerobic

**Table 3** Polyaromatic hydrocarbons in soils and sediments of different geographical regions of world

| Country | Location | PAHs[a] | Total PAH range (ng $g^{-1}$) | Soil pH; TOC/OC (%) | References |
|---|---|---|---|---|---|
| India (Asia) | Agra (urban soils) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)F, CHRY, B(b)F, B(k)F, B(a)P, and B(ghi)P | 3190–28,540 | NM; NM | Masih and Taneja (2006) |
| | Delhi (vehicular traffic soil) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 1062–9652 | NM; NM | Agarwal (2009) |
| | Assam and Arunachal Pradesh (soil from Indian Himalayan Region) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 15.3–4762 | NM; NM | Devi et al. (2016) |
| China (Asia) | Beijing (urban soils, < 10 m from roads) | NAP, ACEY, ACE, FLU, PYR, B(a)A, CHRY, B(b)F, B(k) F, B(a)P, IPYR, D(ah)A, and B(ghi)P, | 1470–6610 | NM; NM | Chu et al. (2003) |
| | Macao (coastal sediments) | PHE, ANT, PYR, MPHE, FLA, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, B(e)P, IPYR, D(ah)A, and B(ghi)P, | 294 – 12 741 | NM; 0.42–1.7 | Mai et al. (2003) |
| | Hong Kong (urban soils) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 7.0–410 | 4.62–7.52; 0.532–2.92 | Zhang et al. (2006) |
| | Hong Kong (Mangrove swamps sediments) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 356–11,098 | NM; 1.23–7.50 | Tam et al. (2001) |
| | North and tropical forests in South China (forest soils) | ACEY, ACE, FLU, PHE, ANT, FLA, PYR, CHRY, B(a)A, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 24.81–4539 | NM; 2.007–52.375 | Syed et al. (2017) |
| Russia (Asia) | Moscow and the south Taiga near Moscow (grassland and forest soils) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(j)F, B(k)F, B(a)P, B(e)P, PERY, D(ah)A, B(ghi)P and COR | 59–1350 | 3.8–5.9; 1–6.8 | Wilcke et al. (2005) |
| Thailand (Asia) | Bangkok (urban soils) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(j)F, B(k)F, B(a)P, B(e)P, PERY, D(ah)A, B(ghi)P, and IPYR | 12–380 | 3.6–7.4; $1.04 \pm 0.69$–$2.03 \pm 1.94$ | Wilcke et al. (1999b) |

**Table 3** (continued)

| Country | Location | PAHs[a] | Total PAH range (ng $g^{-1}$) | Soil pH; TOC/OC (%) | References |
|---|---|---|---|---|---|
| South Korea (Asia) | Korean peninsula (paddy soils and upland soils) | NAP, ACEY, ACE, FLU, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 23.3–2834 | NM; NM | Nam et al. (2003) |
| Korea (Asia) | Kyeonggi Bay (sediments) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A,CHRY, B(b)F, B(k)F, B(a)P, PERY, D(ah)A, B(ghi)P, and IPYR | 9.1–1400 | NM; 0.08–1.2 | Kim et al. (1999) |
| Taiwan (Asia) | Gao-ping River (sediments) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 8 to 356 | NM; NM | Doong and Lin (2003) |
| Germany (Europe) | Stephanskirchen (urban soils) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, PERY, and B(e)P | 400–470,000 | 4.82–12.21; 1.2–44.5 | Wilcke et al. (1997) |
| Switzerland (Europe) | Swiss soil monitoring network (NABO) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 50–619 | 3.9–7.2; 1.0–22.7 | Bucheli et al. (2004) |
| Poland (Europe) | Upper Silesia (agricultural soil) | FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 28–2447 | 3.8–6.3; 1.45–17.20 | Maliszewska-Kordybach (1996) |
| Sweden (Europe) | Karlstad (gasworks plant soil), and Stockholm (wood tar production site) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, and B(ghi)P | 270–2,601,000 | 4.7–8.63; 2.0 ± 0.3 to 49.1–0.4 | Arp et al. (2014) |
| France (Europe) | Coke oven plant and metallurgy site | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 78,000–2,651,000 | 7.5–8.35; 7.7 ± 2.1 to 33 ± 0.6 | Arp et al. (2014) |
| Belgium (Europe) | Gasworks site | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 296,880 | 7.57; 2.5 | Arp et al. (2014) |
| Slovak Republic (Europe) | Ziar/Central Slovakia (forest soils) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(j)F, B(k)F, B(a)P, B(e)P, PERY, D(ah)A, B(ghi)P, and IPYR | 40,000–200,000 | 3.85–4.30; 0.81–2.46 | Wilcke et al. (1996) |

**Table 3** (continued)

| Country | Location | PAHs[a] | Total PAH range (ng $g^{-1}$) | Soil pH; TOC/OC (%) | References |
|---|---|---|---|---|---|
| Czech Republic (Europe) | Bohemian mountains (forest soils) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(j)F, B(k)F, B(a)P, B(e)P, PERY, B(ghi)P, and IPYR | 2000–30,000 | NM; 4.3–46.5 | Wilcke and Zech (1997) |
| | Valasske Mezirici (agricultural and forest soil) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 860–10,840 (agricultural) 7660–79,390 (forest soil) | 4.62–7.12; 1.78–3.82 3.24–4.03; 3.45–19.20 | Plachá et al. (2009) |
| Spain (Europe) | Tarragona County, Catalonia (soil and vegetation samples) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a) P, PERY, D(ah)A, B(ghi)P, and IPYR | 112–1002 | 7.08–8.22; 2.6–9.5 | Nadal et al. (2004) |
| Italy (Europe) | Chioggia and Ancona (sediments of Adriatic Sea) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 24.1–501.1 | NM; NM | Magi et al. (2002) |
| Tunisia (Africa) | Sfax (Sfax–Kerkennah coastal zone sediments) | NAP, PHE, ANT, FLA, PYR, B(k)F, CHRY, B(a)P, PERY, B(a)A, and B(ghi)P | 110–10,720 | NM; NM | Zaghden et al. (2007) |
| South Africa (Africa) | South Africa (soils from industrial, residential, and agricultural areas) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(j)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 44–39,000 | NM; 0.0000001–0.00000049 | Nieuwoudt et al. (2011) |
| Nigeria (Africa) | Okobo (soil from the vicinity of coal mine) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 0–10 | NM; 4–8 | Ugwu and Ukoha (2016) |
| USA (North America) | New Orleans, Louisiana (urban soils) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(j)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 731–2927 | NM; NM | Mielke et al. (2004) |
| Canada (North America) | Kitimat fjord (sediments) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 1000–10,000,000 | NM; 1.4–30 | Simpson et al. (1996) |
| Cuba (North America) | Havana (soils of Mayabeque) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 20–106 | 5.3–8.2; 1.26–4.70 | Sosa et al. (2017) |

**Table 3** (continued)

| Country | Location | PAHs[a] | Total PAH range (ng $g^{-1}$) | Soil pH; TOC/OC (%) | References |
|---|---|---|---|---|---|
| Mexico (North America) | San Nicolás (soils from brick manufacturing site) | NAP, ACEY, FLU, ANT, PHE, B(a)A, CHRY, FLA, PYR, B(b)F, B(k)F, B(a)P, and D(ah)A | 7–1384 | NM; NM | Barrán-Berdón et al. (2012) |
| | Uberlândia (urban soils) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(j)F, B(k)F, B(a)P,PERY, D(ah)A, B(ghi)P, and IPYR | 7–390 | 4.6–8.4; 0.37–2.34 | Wilcke et al. (1999a) |
| Brazil (South America) | Argentinean Pampas to the Patagonian Steppe (uppermost topsoil and the first subsoil horizons) | NAP, 2-MNAP, 1-MNAP, 1,3-DMNAP,ACEY, ACE, FLU, PHE,ANT, 1-MPHE, 3,6-DMPHE, FLA, PYR, B(a)A, CHRY, B(b)F, B(j)F, B(k)F, B(a)P, B(e)P, PERY, D(ah)A, B(ghi)P, COR, and IPYR | 2.4–38 | NM; NM | Wilcke et al. (2014) |
| Argentina (South America) | South Central Chile (Laja River Basin soil) | NAP, FLU, PHE,ANT, FLA, PYR, B(a)A, and CHRY | 600–4243 | NM; NM | Barra et al. (2005) |
| Chile (South America) | Sydney Harbor (surface sediment) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, PERY, D(ah)A, B(ghi)P, and IPYR | 100–380,000 | NM; NM | McCready et al. (2000) |
| Australia | Kogarah, Sydney (surface natural soils of Urban area)<br>(NS-surface natural soil, RDS-road deposited sediments, and WS-water sediments) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, PERY, D(ah)A, B(ghi)P, and IPYR | NS-400–7490: RDS-1650–4000; WS-490–5190 | NM; 1.8–3.8<br>NM; 0.8–13.0<br>NM; 0.3–3.0 | Nguyen et al. (2014) |
| | McMurdo Station (surface soil) | NAP, ACEY, ACE, PHE, FLU, ANT, FLA, PYR,FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(e)P, D(ah)P, D(ac)P, B(b)C, and COR | 1000,000–27,000,000 | NM; NM | Mazzera et al. (1999) |
| Ross Island (Antarctica) | Scott Base, Marble Point and the Wright Valley (surface and subsurface soil) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 41–8105 | NM; NM | Aislabie et al. (1999) |
| McMurdo Sound-Dry Valley Region (Antarctica) | James Ross Island (soil) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 34.9–171 | NM; 0.09–0.34 | Klánová et al. (2008) |

**Table 3** (continued)

| Country | Location | PAHs[a] | Total PAH range (ng $g^{-1}$) | Soil pH; TOC/OC (%) | References |
|---|---|---|---|---|---|
| Antarctic Peninsula (Antarctica) | James Ross Island (sediments) | NAP, ACEY, ACE, FLU, PHE, ANT, FLA, PYR, B(a)A, CHRY, B(b)F, B(k)F, B(a)P, D(ah)A, B(ghi)P, and IPYR | 1.4–205 | NM; 0.04–1.71 | Klánová et al. (2008) |

[a]NAP-Naphthalene, MNAP-Methylnaphthalene, 2-MNAP-2-methylnaphthalene, 1-MNAP-1-methylnaphthalene, 1,3-DMNAP-1,3-dimethylnaphthalene, ACEY-Acenaphthylene, ACE-Acenaphthene, FLU-Fluorene, PHE-Phenanthrene, 1-MPHE-1-methylphenanthrene, 3,6-DMPHE-3,6-dimethylphenanthrene, ANT-Anthracene, FLA-Fluoranthene, B(j)F-benzo(j)fluoranthene, PYR-Pyrene, B(a)A-Benzo(a)anthracene, CHRY-Chrysene, B(b)C-Benzo(b)chrysene, B(a)F-Benzo(a)fluoranthene, B(b)F-Benzo(b)fluoranthene, B(k)F-Benzo(k)flouranthene, B(a)P-Benzo(a)pyrene, D(ah)A-Dibenz(ah)anthracene, D(ah)P-Dibenzo(ah)pyrene, D(ac)P-Dibenzo(ah)pyrene B(ghi)P-Benzo(g,h,i)perylene, IPYR-Indeno(1,2,3-cd)pyrene, B(e)P-Benzo(e)pyrene, PERY-Perylene, and COR-Coronene

*NM* not mentioned

conditions in soil by blocking soil pore with consequent effects on microbial communities of soil (Sutton et al. 2013). By long-term contamination of soil by petroleum hydrocarbon and at high concentration of total petroleum hydrocarbons, there is decline in microbial biomass and soil enzyme activity owing to toxicity induced by high molecular weight hydrocarbons, i.e., PAHs (soil enzymes—dehydrogenase and urease), whereas at low petroleum hydrocarbon concentration, there is no effect on microbial biomass and enzymes (Verrhiest et al. 2002; Lipińska et al. 2014). Soil microbial diversity may be affected by PAHs contamination as PAHs may have toxic effects toward microorganisms present in soil. In an investigation, it was observed that PAH contamination has a significant effect on soil bacterial community structure (Khomarbaghi et al. 2019). In certain cases, PAH contamination may lead to complete loss of specific microbial species and if any particular species from contaminated soil have been damaged then certain essential soil function may be lost (Muckian et al. 2009). A study carried out to investigate the effect of pyrene on bacterial richness and microbial diversity in soil. It has been reported that after pyrene contamination microorganism's population belonging to phyla *Chlorflexi, Alphaproteobacteria, Actinobacteria*, *Deltaproteobacteria,* and *Crenarchaeota* were extensively reduced (Ren et al. 2015). The effect of three PAH mixture (phenanthrene, fluoranthene, and benzo(k)fluoranthene) on the indigenous microbial species of natural freshwater sediment was investigated. It was reported that due to heavy dose of PAHs, microbial activity was adversely affected. At low PAH concentration, i.e., 30 mg PAH/kg no negative effect on microbes was observed but at 300 mg PAH/kg harmful effects of PAHs, i.e., reduction in bacterial density and partial inhibition of the enzyme activity were observed as compared to the control 0 mg PAH/kg (Verrhiest et al. 2002). It has been studied that there is a significant effect of oil contamination on biological activity of soil, a strong decline in dehydrogenase and urease enzyme activity was observed in chainsaw oil (containing a complex mixture of PAHs, highly toxic.)-contaminated soil. It was also found that there is major influence of chainsaw oil on earthworm biomass and density (Klamerus-Iwan et al. 2015).

## Remediation of PAH-contaminated soils

Today, remediation of polyaromatic hydrocarbon (diesel/crude oil)-contaminated soils is a global concern due to adverse risk to public health. Reclamation/remediation of PAH-polluted soils is essential and it can be done by different methods which involve removal/isolation or alteration of the contaminant. Various physical, chemical, thermal and biological remediation techniques (ex situ and in situ) have been developed for soil reclamation (Fig. 1). In situ

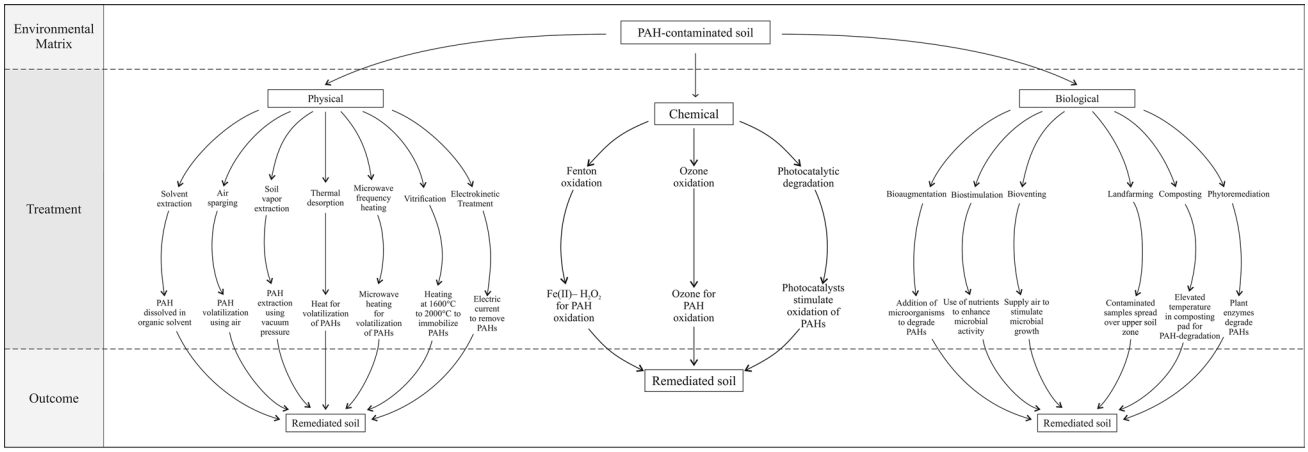



**Fig. 1** Physical, chemical, and biological methods for treatment of hydrocarbon-contaminated soil

remediation process takes place at the contaminated site, whereas ex situ remediation is an alternative to in situ remediation where contaminants are treated off-site. There are some advantages of in situ remediation. It has lower cost, lower risk factor, limited involvement of human, and environmental surroundings can help in remediation process to transform the contaminants. For ex situ remediation, contaminated soil is excavated (dig and haul) from the site and transfer to another location for treatment so it requires mechanical as well as civil work. It is costly, time-consuming, more human involvement, direct exposure to contaminant, and due to digging it can harm underground utilities. There are some regulatory constrains also for ex situ remediation. Some of these remediation techniques are solvent extraction, UV oxidation, photochemical or photocatalytic degradation, bioremediation, and phytoremediation. The selection of suitable remediation technique for contaminated soils depends on several factors such as type of contaminant, future use of contaminated soil, type and properties of soil, budget, etc.

## Physical treatment

### Solvent extraction/soil washing

PAHs have high tendency to get absorb on organic matter present in soil due to their hydrophobic nature. Solvent extraction/soil washing is a separation or cleanup technique which is used to separate compounds based on their solubility. This technique can be used for PAH removal from contaminated soils. Different organic solvents (individual solvent or mixture of solvents), cyclodextrins (β-cyclodextrin (BCD), hydroxypropyl-β-cyclodextrin (HPCD) and methyl-β-cyclodextrin (MCD)), and vegetable oils can be used for extraction of PAHs from soils. The extraction of PAHs from soils is a two-step process. The first step is desorption of compound from soil and the second step involves leaching (elution of compound into extraction fluid) of desorbed compound. Efficiency of solvent extraction is influenced by the nature of solvent used in extraction process and the ratio of the mass of contaminated soil to the volume of solvent (Silva et al. 2005; Viglianti et al. 2006). PAHs from extremely contaminated soils obtained from manufactured gas site can be extracted effectively using a solvent mixture containing ethanol or 2-propanol with 1-pentanol and water. The solvent mixture 1-pentanol (5%), water (10%), and ethanol (85%) is highly efficient and more effective than single solvent in removal of extractable PAH (19 PAH having two to six aromatic rings, e.g., naphthalene, 2-methyl naphthalene, acenaphthene, fluorene, phenanthrene, fluoranthene, pyrene, benzo(e)pyrene, benzo(a)pyrene, benzo(ghi) perylene, etc.) from soil. The extraction efficiency of this mixture is up to 95% when 4 ml solvent mixture is used for 1 g soil for one hour in three crosscurrent wash (extraction) stages (Khodadoust et al. 2000). Organic solvent mixture of Cyclohexane and ethanol (3:1) is suitable for fluoranthene extraction from contaminated soil. Its extraction efficiency is approximately 93% (Rababah and Matsuzawa 2002a). Similarly, it has been documented that solvent mixture of ethyl acetate (50%), acetone (40%), and water (10%) is suitable for extraction of hydrocarbons from soil (Silva et al. 2005). A major concern in this technique is the toxic nature of solvents and the liquid phase, i.e., the solvent containing desorbed PAHs. Therefore, selection of solvents is crucial and solvent containing desorbed PAHs exposed to other treatment for complete degradation of extracted PAHs.

### Organic and mineral soil amendments

The organic and mineral soil amendment is an important process for reclamation of contaminated soils. These amendments in the contaminated soils help in attaining high

remediation efficiency as many biological, chemical and physical processes start after these amendments. Compost, manures, organic by-products, etc., are organic amendments and foundry sand, gypsum, coal combustion products, volcanic ashes, etc., are mineral amendments which are found to be useful for pollutant degradation (Fernández-Luqueño et al. 2017). In an investigation, activated carbon and olive mill waste compost were used as amendments for PAH-contaminated soil and enhanced degradation of total PAHs was observed (García-Delgado et al. 2019). Similarly, it was studied that sand can be used as an efficient amendment as sand amendment increases both oxygen and proton passage which increase soil porosity, reduce Ohmic resistance, and increase charge output. Sand amendment was found to be an effective method to accelerate degradation of PAHs by bio-electrochemical treatment (Li et al. 2015). In order to allow soil amendments effectively improve remediation of contaminated soils, it is essential to find different amendments which can improve soil porosity, increase microbial activity, and increase pollutant mineralization in different soils while retaining soil functions.

## Thermal treatment

Thermal treatment for soil remediation uses heat to destroy contaminants. Organic chemicals such as PAHs can be destroyed or volatilize by heat, these contaminants changes into gases which results in increased mobility and these gases can be collected in wells for ex situ treatment. Different techniques have been used under thermal remediation such as thermal desorption, microwave frequency heating and vitrification.

### Thermal desorption

Thermal desorption is the process in which heat is applied to increase vapor pressure of organic contaminants results in volatilization of contaminants and release of them from contaminated sites such as soil (Rushton et al. 2007). In this process, volatilized contaminants carried away or sweep by gas for secondary treatment or removal. High PAHs removal efficiencies could be obtained using thermal desorption. Soil contaminated with different PAHs from a manufacturing gas plant treated by thermal desorption process at laboratory scale. After thermal treatment at maximum temperatures above 450 °C, the concentration of different PAHs is reduced to below 0.05 mg/kg dry weight. For dibenzo(a,h) anthracene, efficiency is around 87% at temperature 250 °C, and for fluoranthene and pyrene, efficiency is nearly 100% with temperature above 350 °C (Renoldi et al. 2003). The subsurface soil of a wood treatment plant contaminated with benzo(a)pyrene equivalents (B(a)P-E) was treated by thermal desorption at field scale. Approximately 12,385 $m^3$ (16,200 cubic yards) of predominantly silty soil containing 30.6 mg/Kg B(a)P-E was treated to a maximum depth of 32 m. After 130 days of post-treatment the remaining B(a)P-E concentration was 0.059 mg/Kg, which is equivalent to 99.8% PAHs removal. Therefore, thermal desorption is highly efficient at field scale for PAHs remediation (Baker et al. 2007). Temperature increase from 100 to 250 °C, results in an increase in removal efficiency of diesel from 47 to 100% from sand (Falciglia et al. 2011). Similar laboratory studies also showed that the efficiency of diesel removal from soil is increased during thermal desorption and removal efficiency depends on soil composition, temperature, temperature time, concentration of contaminants (Piña et al. 2002; Tatàno et al. 2013).

### Microwave frequency heating

Microwave frequency heating is an effective thermal remediation technique in which microwave energy converted into thermal energy to eliminate contaminants through volatilization via heating (Rushton et al. 2007). This technique has been effectively proposed for soil remediation (Falciglia et al. 2013). 99% diesel oil could be removed within 10 min from contaminated soil using microwave induced thermal treatment in a modified domestic microwave oven with power 800 W and frequency 2.45 GHz (Li et al. 2008). A domestic microwave oven with a power 700 W and frequency of 2.45 GHz was used to treat heavy oil (diesel fuels and marine fuels)-contaminated soil, and it was found that up to 92.5% diesel and 89.5% marine fuel removed from soil in 20–150 s (Chang et al. 2011). 75.6–98.4% petroleum removal was achieved within 3.5 h at field-scale remediation of petroleum hydrocarbon-contaminated soil using on-site microwave heating system using antenna of 4 m, with power 2 kW and frequency of 2450 MHz (Chien 2012). On the other hand, using a modified domestic microwave with a power 1000 W and frequency of 2450 MHz, maximum of 95% diesel removed from soil in 5–60 min (Falciglia et al. 2013). Other similar studies also investigated the efficiency of this technique for oil removal from soil. Using this technique up to 100% PAHs could be removed in 60 min from soil artificially contaminated with different PAHs (Bph, Flu, Phe, Ant, Flt, Pyr, BaA, Chr and Per) when microwave frequency 2.45 GHz applied at 1000 W for remediation process. Total removal was achieved for biphenyl and fluorene whereas up to 90% removal was achieved for phenanthrene and anthracene and 50–80% removal was achieved for other PAHs (Falciglia et al. 2017).

### Vitrification

Vitrification is a thermal technique which can be used for in situ soil remediation. This technique uses very high

temperatures, i.e., 1600–2000 °C to melt and immobilize contaminants in soil. Heat is delivered to soil through electric current via molybdenum electrodes. In this process, high-temperature melts the contaminants as well as the soil. After melting process, electrodes are turned off, allowing fusion of contaminants with soil and then after cooling both convert into a glass-like solid (Shearer 1991). Vitrification product, i.e., glass-like solid is a chemically stable, leach-resistant and crystalline material. Vitrification can be used to treat various organic contaminants such as petroleum products (Hinchee and Smith 1992). Fly ash is a major source of PAHs, after vitrification process in a coke bed furnace total 21 PAHs (naphthalene(NaP), acenaphthylene (AcPy), acenaphthene (Ant), fluoranthene (FL), pyrene (Pyr), cyclopenta(c,d)pyrene (CYC), benz(a)anthracene (BaA), chrysene (CHR), benzo(b)fluoranthene (BbF), benzo(k) fluoranthene (BkF), benzo(e)pyrene (BeP), benzo(a)pyrene (BaP), perylene (PER), indeno(l,2,3,-cd)pyrene (IND), dibenz-(a,h)anthracene (DBA), benzo(b)chrysene (BbC), benzo(ghi)perylene (BghiP) and coronene (COR)) present in fly ash significantly dropped from 69.6 µg to 3.46 ng, which results in high efficiency of PAH depletion. The lowest percent removal of individual PAH is higher than 99.9%. Therefore, PAHs in fly ash fully destroyed in vitrification process due to very high temperature (Kuo et al. 2003).

The thermal methods as mentioned above, i.e., thermal desorption, microwave frequency heating, and vitrification are essentially taking place under anaerobic conditions since oxygen in subsurface is absent. The added advantage of thermal processes is the absence of formation of secondary toxic pollutants that are formed sometimes during PAH-oxidation. Some toxic oxidized products such as epoxides and dihydrodiols are found to be formed and have higher toxicity than the parent PAHs (Cerniglia and Sutherland 2010).

## Electrokinetic treatment

Electrokinetic technique is an in situ remediation technique where direct electric current is used to remove organic and other contaminants (inorganic and heavy metal) from the soil (Huang et al. 2012; Karaca et al. 2016). During this method, low-voltage direct current electric potential is applied through electrodes (anode and cathode). Low-voltage electric current causes mobilization of contaminants and their transportation toward electrodes placed inside contaminated soil matrix. Contaminants collected on these electrodes are pumped out for further treatment. Different transport mechanisms are induced by electric current such as Electromigration, electro-osmosis, electrophoresis, and diffusion (Acar et al. 1995). It has been found that using upward electrokinetic remediation process 67% of phenanthrene could be removed after 6 days from contaminated soil (Wang et al. 2007). Electrokinetic treatment combined with different surfactants and complexing agents has been used to increase desorption and solubility of contaminant. It is observed that 70% phenanthrene was removed after 30 days from kaolin clay using electrokinetic treatment combined with1% Tween 80 and 0.1 M EDTA (Alcántara et al. 2012).

## Chemical treatment

### Chemical oxidation

Chemical oxidation treatment involves redox (oxidation/reduction) reactions that involve electron transfer from one chemical to other chemicals. This treatment converts hazardous contaminants into less toxic or nonhazardous compounds (Sharma et al. 2016; Verma and Haritash 2019). In recent days, advanced oxidation processes (AOPs) involving different oxidants have been used for treatment of PAH-contaminated soils. Different types of oxidants such as hydrogen peroxide, persulfate, ozone, Fenton's reagent, persulfate, peroxymonosulfate, and potassium permanganate have been used for the remediation of oil-contaminated soil via chemical oxidation reactions (Goi et al. 2006; Do et al. 2009, 2010; Yen et al. 2011). Chemical oxidation reactions have been used for treatment of oil or PAH-contaminated soils, different oxidants can be added to soil to oxidize contaminants (Tsai and Kao 2009; Rivas 2006).

### Fenton's reagent

Fenton's reagent ($Fe(II)–H_2O_2$) for Chemical oxidation remediation uses hydrogen peroxide as oxidant in the presence of ferrous sulfate to generate free radicals, i.e., hydroxyl radicals (OH·).

$$Fe(II) + H_2O_2 \rightarrow Fe(III)OH^{2+} + HO\cdot$$

These radicals are powerful oxidants (Fenton 1894). The use of the Fenton's reagent for chemical oxidation process has been found effective for the remediation of soil contaminated by polycyclic aromatic hydrocarbons (PAHs). Fenton oxidation treatment efficiency on PAH-contaminated soils has been found in different laboratory scale experiments. Soil samples contaminated with fluoranthene, benzo(b)fluoranthene, and benzo(a)pyrene were treated using Fenton oxidation process. Soil samples were mixed with water to form suspension (2 g soil sample in 10 ml water) in which 0.01 M Fe(II) was added with successive addition of $H_2O_2$. The suspension was then magnetically stirred for 24 h process to allow Fenton oxidation at room temperature with no pH adjustment. After 24 h oxidation process, high efficiency for PAH removal was observed, with removal of 85.7% fluoranthene, 87.4% benzo(b)fluoranthene, and 88.6% benzo(a)pyrene (Flotron et al. 2005). 24 different PAHs (2–6 ring) in different nine

contaminated soil samples were effectively degraded using Fenton's reagent. PAH degradation efficiency was 40–86% (20 g soil in 10 ml water) in the presence of 30% $H_2O_2$ and 4 mM Fe(II) at 70 °C and 3 pH (Jonsson et al. 2006). So, Fenton oxidation process could be used as an efficient remediation technique for PAH-contaminated sites.

### Ozone

Ozone is one of the stronger oxidizing agents for chemical oxidation technique, which can be used for remediation of PAH-contaminated soil. During ozonation, the ozone molecule may directly attack double bonds and can form reactive hydroxyl radicals. Different intermediates such as quinone and biphenyl-type products are formed in ring cleavage radical oxidation process (Yao et al. 1998). In situ ozone treatment for soils can be done by injecting gaseous ozone which is the most advanced method and by injecting aqueous ozone because gaseous ozone is more effective than aqueous ozone. It was found that 20% PAHs remain in soil after gaseous ozone treatment and 40% PAHs remain in soil after aqueous ozone treatment (Masten and Davies 1997). This is due to easy and more diffusivity of gaseous ozone, resulting in more contact between contaminants and ozone (oxidizing agent). In an investigation, it has been shown that ozone could be used to remove phenanthrene from phenanthrene-spiked farm soils. At least 50% of phenanthrene level reduction in air-dried soils was achieved when soil samples were exposed to ozone at 20 ppm for 6 h (O'Mahony et al. 2006). From PAH-contaminated soil 95% of phenanthrene removal was achieved with ozonation for 2.3 h at an ozone flux of 250 mg $h^{-1}$, 91% of pyrene and 50% of chrysene were removed using ozone flux of 600 mg $h^{-1}$ for 4 h (Masten and Davies 1997). Other similar studies also find the efficiency of ozone around 90–95% in removal of total petroleum hydrocarbons or diesel fuel from sand or unsaturated soils (Shin et al. 2005; Yu et al. 2007).

### Other oxidants

Several studies have investigated that other alternative oxidants like Persulfate/Fe(II), Peroxymonosulfate (PMS), Persulfate, $H_2O_2$, and permanganate can also be used for chemical oxidation treatment for diesel and fuel oil (Do et al. 2009, 2010; Yen et al. 2011). The feasibility of $KMnO_4$ (potassium permanganate) as an oxidant for removal of PAH from contaminated soil has been investigated and it has been found that there is significant reduction in PAHs (benzo(a) pyrene-72.1%, pyrene-64.2%, phenanthrene-56.2% and anthracene53.8%) present in soil (Brown et al. 2003).

## Photocatalytic degradation

Photocatalytic degradation is a process in which photocatalysts are used to stimulate oxidizing reactions, i.e., photoreaction. This process is used for treatment of petroleum-contaminated soils to destroy organic contaminants in teh presence of the light radiation. To enhance the degradation rate of fluoranthene, a photocatalytic solar reactor was developed. During photocatalysis process, both $TiO_2$ and $H_2O_2$ were used for fluoranthene degradation. Fluoranthene degradation efficiency was found 99% in the presence of both $TiO_2$ and $H_2O_2$. However, lower fluoranthene degradation (83%) was observed when only $TiO_2$ was present as catalyst (Rababah and Matsuzawa 2002b). The photocatalytic degradation of phenanthrene, pyrene, and benzo(a)pyrene on soil surfaces using titanium dioxide $TiO_2$ in the presence of ultraviolet (UV) light was investigated in a photodegradation chamber at 30 °C. Photocatalyst $TiO_2$ (0.5%) accelerated the photodegradation process of phenanthrene, pyrene and benzo(a)pyrene significantly as compared to degradation in absence of $TiO_2$ catalyst, their half-lives being reduced from 533.15 to 130.77 h, 630.09 to 192.53 h and 363.22 to 103.26 h, respectively. There is synergistic effect of ultraviolet light and $TiO_2$ catalyst for PAHs degradation in soil (Zhang et al. 2008). Phenanthrene degradation on soil surface using photocatalysis under UV-irradiation was investigated where nanometer anatase $TiO_2$ was used as photocatalyst. Soil samples spiked with phenanthrene, loaded with $TiO_2$ (0 wt%, 1 wt%, 2 wt%, 3 wt%, and 4 wt%) exposed to UV light for 25 h. It has been observed that $TiO_2$ significantly increase the degradation rate of phenanthrene with half-life reduced from 45.90 to 31.36 h for 0 wt% and 4 wt%, respectively (Gu et al. 2012). So, $TiO_2$ is an efficient photocatalyst which has been used for oxidation of PAHs through photocatalytic degradation for treatment of oil polluted soils.

## Biological treatment

### Bioremediation: role of microbes in PAH degradation

Over the last two decades, another method in which microorganisms are used to degrade PAHs (bioremediation) is more accepted. Bioremediation is considered an ecofriendly and sustainable remediation technique and has recently gained considerable interest all around the globe (Sharma and Reddy 2004; Reddy and Adams 2015). The natural degradative potential of microorganisms, i.e., bacteria, yeasts, fungi, and algae is used in bioremediation of contaminants to convert them into less toxic compounds or into water and carbon dioxide (Alexander 1994). The microorganisms can produce a number of enzymes to detoxify and mineralize PAHs eventually leading to its degradation. The microorganisms

isolated from contaminated sites show more degradation ability as these microorganisms have been adapted to the polluted environment and can survive in the presence of the pollutant. A very challenging task in bioremediation processes is the confirmation of the first step, i.e., which microorganism should be used for degradation and endpoint, i.e., the end products. For efficient contaminant degradation the microbe must be present in favorable environmental conditions, there are various factors affecting biodegradation are pH, temperature, nutrients, and metabolites (Boopathy 2000). Degradation of PAHs could be done by using various naturally occurring soil microorganisms, i.e., aerobic and anaerobic microorganisms belong to genera *Pseudomonas, Alcanivorax, Microbulbifer, Mycobacteria*, *Sphingomonas, Micrococcus, Alcaligenes*, *Ralstonia*, *Paenibacillus, Bacillus, Aeromonas, Xanthomonas, Arthrobacter, Acinetobacter, Corynebacterium Enterobacter* and others (Bayoumi 2009; Haritash and Kaushik 2016). Basidiomycetes *Pleurotus ostreatus* and *Irpex lacteus* were used for fungal bioremediation of soil contaminated with creosote (50–200 mg $kg^{-1}$ PAH) from a wood-preserving plant. 55–67% PAH removal was observed in *P. ostreatus* treatments (86–96% of 2-rings PAHs, 63–72% of 3-rings PAHs, 32–49% of 4-rings PAHs and 31–38% of 5–6-rings PAHs) and 27–36% PAH removal was observed in *I. lacteus* treatments (47–59%of 2-rings PAHs, 33–45% of 3-rings PAHs, 9–14% of 4-rings PAHs and 11–13% of 5–6-rings PAHs) in 120 days (Byss et al. 2008). There are some other fungal species, mainly *Aspergillus* and White Rot Fungi, such as *Phanerochaete chrysosporium, Bjerkandera adusta, and Pleurotus ostreatus*, have been documented for efficient biodegradation of PAHs (Haritash and Kaushik 2009, 2016). From various studies, several PAH catabolic genes encoding PAH catabolic enzymes have been characterized. Mono-oxygenases and PAH ring-hydroxylating dioxygenase enzymes were identified in *Novosphingobium pentaromativorans* proteome analysis. These enzymes are involved in the degradation of PAHs including phenanthrene, pyrene, and benzo(a)pyrene (Lyu et al. 2014). *Serratia marcesencs* L-11 strain has an ability to produce lipases and aromatic ring cleavage enzyme (catechol 1, 2-dioxygenase). It can degrade phenanthrene, anthracene, fluorene, and pyrene (Pandey et al. 2012). Similarly, *Pseudomonas aeruginosa* PSA5 and *Rhodococcus* sp. NJ2 isolated from petroleum sludge were investigated for B(a)P degradation. It was observed that various PAH catabolic enzymes such as salicylate hydroxylase, 2-carboxybenzaldehyde dehydrogenase, catechol 1,2-dioxygenase, and catechol 2,3-dioxygenase were differentially expressed in both bacterial species involved in B(a)P degradation (88% and 47% by *Pseudomonas aeruginosa* PSA5 and *Rhodococcus* sp. NJ2, respectively) (Mishra and Singh 2014). Various genes are involved in PAHs degradation and these were upregulated after exposing the microbe to PAHs (Lyu et al. 2014). In order to develop an effective sustainable strategy of bioremediation of PAHs, the characteristics, as well as the metabolic potential of the microbes need to be better understood in PAH-contaminated areas.

### In situ bioremediation

**Bioaugmentation** Bioaugmentation is a technique which involves the addition of microorganisms (indigenous or exogenous) to the contaminated sites/soils, these microorganisms degrade the contaminants. This bioremediation technique is used where natural degrading microbes are present in low number or absent. Microbes are selected on the basis of their metabolic capacity to degrade contaminant (Boopathy 2000). Various researches have been done to study the effects of microorganism addition into contaminated soil (bioaugmentation) for in situ PAH degradation. It has been studied that microbial communities are efficient for PAH degradation under aerobic as well as anaerobic conditions (Kiamarsi et al. 2018). Pollutant bioavailability, survival of microorganism and their enzymatic catabolic activities are important for bioaugmentation (Heinaru et al. 2005). PAH degradation in soil by augmentation with specific isolated bacteria or fungus has been documented in various studies. *Scopulariopsis brevicaulis* PZ-4 isolated from an aged polycyclic aromatic hydrocarbon (PAH)-contaminated soil was found to have the ability to degrade PAHs. In a PAH-contaminated soil, *Scopulariopsis brevicaulis* PZ-4 removed 77% of total PAHs and the highest removal of PAHs occurred for phenanthrene (89%) and benzo(a)pyrene (75%) after incubation for 28 days (Mao and Guan 2016). Fungal isolate, *Penicillium* sp. 06, was effective at oxidizing a range of PAH in petroleum-contaminated soils. After 28 days of incubation, 89% of the phenanthrene presents in oily waste residues from the petrochemical refining industry in Singapore oxidized by *Penicillium* sp. 06. This isolate could also oxidize more than 75% of the acenaphthene, fluorene, and fluoranthene after 30 days of incubation (Zheng and Obbard 2003). Enhanced fluorene degradation in soil slurry system by augmenting with *Absidia Cylindrospora* a fungal isolate has been reported. In soil slurry system augmented with *A. Cylindrospora* more than 90% of the fluorene was removed after 288 h, while in nonfungal soil slurry 576 h contact time required for 90% removal of the fluorene (Garon et al. 2004). This is an efficient bioremediation technique which can also be used in ex situ remediation processes.

**Biostimulation** The process of environmental modification through addition of nutrients (nitrogen, phosphorous, and carbon, organic biostimulants), and oxygen (electron acceptor) to stimulate the activity of contaminant/oil-degrading indigenous microorganisms is known as biostimulation. These nutrients are the building blocks of life, therefore

these nutrients allow microbe to synthesize necessary enzymes to degrade the contaminant. It is one of the important strategies for increasing the efficiency of bioremediation of crude oil/PAHs in soil (Garon et al. 2004).

Addition of nutrients to PAH-contaminated soil has been shown increased microbial biomass and activity, therefore results in enhanced degradation efficiency in soils. Inorganic nitrogen and phosphorous were added for in situ bioremediation of PAHs in creosote-contaminated soil of wood-preserving plant in Norway. The addition of nutrients stimulated PAH degradation rate in the topsoil and the aquifer sand (Breedveld and Sparrevik 2000). The influence of addition or amendments of nutrients in crude oil-contaminated soil has been investigated where crude oil biodegradation in soil was observed for fertilized (soil added with nitrogen, phosphorus, and potassium) and unfertilized soil for 150 days. It was observed that fertilized soil has high biodegradation efficiency (62%) as compared to unfertilized soil (47%) (Chaineau et al. 2005). Biostimulation of hydrocarbon-degrading bacterial community with mineral nutrient and surfactant solution in crude oil-contaminated soil has been investigated and, found 39.5% reduction of total hydrocarbon content (Zucchi et al. 2003). Similarly, Abed et al. (2015) investigated that addition of $NH_4CL$ and $NaH_2PO_4$ as nitrogen and phosphorus sources for biostimulation of oil-contaminated desert soil. They observed that after addition of nutrients oil removal efficiency has been increased by 20%. Organic biostimulants such as phycocyanin (a proteic emulsifier extracted from the *Spirulina platensis* biomass) and inactive biomass of *S. platensis*, or ammonium sulfate are effective in bioremediation application. The soil contaminated with 4% of diesel or biodiesel was biostimulated for 60 days with these organic biostimulants. The biomass of *S. platensis* was found most effective biostimulant for diesel removal as after 60 days of biostimulation 63.89% of 4% diesel was degraded and the extracted phycocyanin of *Spirulina platensis* was found most effective biostimulant for biodiesel removal as after 60 days of biostimulation a biodegradation value of 88.75% for biodiesel was obtained (Decesaro et al. 2017). Similar to bioaugmentation, this technique can also be used in in situ or ex situ remediation processes.

**Bioventing** It is the most common in situ remediation treatment that involves supplying air or oxygen through wells to contaminated soil (in the unsaturated zone) to stimulate growth of the indigenous microorganisms. Soils contaminated by petroleum hydrocarbons have been successfully remediated using this technique (Hinchee 1993). A pilot-scale bioventing for remediation treatment of polycyclic aromatic hydrocarbons including high-ring PAHs such as pyrene, fluoranthene, and benzo(a)anthracene) was done to treat a 15.2 $m^2$ area by The Reilly Tar and Chemical Corporation site in St. Louis Park, Minnesota. Preliminary results of this study showed 62% reduction in 2-ring PAHs; 50% reduction in 3-ring PAHs; 31% reduction in 4-ring PAHs; 20% reduction in 5-ring PAHs; and 24% reduction in 6-ring PAHs (Alleman et al. 1995). 93% of phenanthrene was removed after 7 months of bioventing treatment of artificially phenanthrene-contaminated soil (1000 mg/kg soil) under Optimum conditions of mineralization (humidity = 60% WHC; C/N/P = 100:20:1) (García Frutos et al. 2010). Maximum degradation effectiveness of 85% was observed for remediation of 4 wt% of B20 (blend of diesel and biodiesel fuel)-contaminated soil (clay) through bioventing (Thomé et al. 2014).

### Ex situ bioremediation

**Landfarming** Landfarming is a simple bioremediation technique in which contaminated soil is excavated, transported to the landfarming site and spread over a prepared bed and periodically tilled (turned over that provides aeration) until pollutants are degraded. Contaminants are degraded and transformed by microbiological metabolic processes and by oxidation (Riser-Roberts 1998). After 3 months of landfarming treatment, there is 63% reduction in total PAH concentration in a field contaminated mainly due to high concentrations of PAHs (1140 mg/Kg dry weight). There is up to 79% reduction in PAHs concentration with 2, 3 and 4 rings (Picado et al. 2001). Similarly, landfarrming was used for bioremediation of creosote-contaminated soil in South Africa. After 6 months of treatment low molecular weight PAHs, i.e., 2–3 ring PAHs (naphthalene, anthracene, phenanthrene, and fluorene) were removed from soil but high molecular weight PAHs were still found in soil after 6 months treatment. At the end of treatment, i.e., after 10 months 76–87% of 4–5 ring PAHs (86.8% pyrene, 78.64% B(a)P, 76% chrysene and 85.5% fluoranthene)were removed (Atagana 2004). Effective remediation of diesel-contaminated soils at former military base at Resolution Island, Nunavut was done using landfarming for 3-year period with rototilling. 80% diesel fuel level was reduced after 3 years of landfarming treatment (Paudyn et al. 2008). Apart from PAHs, other petroleum hydrocarbons such as diesel-range organics, trimethylbenzenes, gasoline-range organics and BTEX compounds can also be degraded using landfarming (McCarthy et al. 2004).

**Composting** The process in which microorganisms (mesophilic and thermophilic) degrade organic contaminants at elevated temperature, i.e., 55–66 °C, is known as composting. During this process, microorganisms release heat results in an increase in temperature which further results in more solubility of contaminants and higher microbial activity in compost. For composting, contaminated soil is transferred to the composting pad (Namkoong et al. 2002).

Various studies have been done to investigate the efficiency of composting to degrade PAHs. An investigation on spent mushroom compost was done for bioremediation of soil contaminated with PAHs and it was observed that naphthalene, phenanthrene, benzo(a)pyrene and benzo(g,h,i) perylene was completely degraded after 48 h at 80 °C (Lau et al. 2003). Remediation of soil contaminated with PAHs through composting has been investigated in thermally insulated composting chamber. Mushroom compost, which consists wheat straw, chicken manure, and gypsum, has been used for this investigation. Substantial PAH removal or degradation has been observed during composting. At the end of 54 days of composting 20–60%, PAHs were removed, and after another 100 days 37–80%, PAHs were removed (Sasek et al. 2003). Diesel-, coal tar- and coal ash-contaminated soil mixed with compost results in enhancement of PAHs bioavailability and its increased removal rate up to 90% (Wu et al. 2013).

**Phytoremediation** Phytoremediation or plant-assisted bioremediation can be defined as an in situ technique that uses green plants and associated microorganisms to remove (extract, degrade, or immobilize) contaminants (PAHs) from environment (soil). It is an ecofriendly and cheaper alternative to other physical and chemical treatments, which prevents excavation of soil from contaminated sites. Different phytoremediation techniques such as phytoextraction, phytotransformation, phytostabilization, phytodegradation rhizodegradation can be used for treatment of oil-contaminated soils (Germida et al. 2002; McCutcheon, and Schnoor 2003). Plant species secrete various enzymes such as monooxygenase, dioxygenase, dehydrogenase, hydrolase, peroxidases, and dehalogenase into the soil which have ability to transform or degrade aromatic contaminants (Campos et al. 2008). Various factors such as contaminant nature, soil properties, bioavailability of contaminant and type of plant affect phytoremediation efficiency (Sreelal and Jayanthi 2017). However, few studies have investigated efficiency of plants for PAH remediation in soils, different types of plants and grasses such as *Festuca arundinacea,* ryegrass (*Lolium multiflorum*, *Lolium perenne*), *Dactylis glomerata, Festuca rubra,* yellow sweet clover (*Melilotus officinalis*), *Lotus corniculatus* (birdsfoot-trefoil), *Trifolium pratense* (red clover) and *Trifolium repensm* (white clover) are found to degrade different PAHs like naphthalene, acenaphthylene, acenaphthene, phenanthrene, anthracene, fluoranthene, pyrene, benzo(a)anthracene, chrysene, benzo(b)fluoranthene, benzo(k)fluoranthene, benzo(a)pyrene, indeno(1,2,3-c,d) pyrene, benzo(g,h,i)perylene, dibenz(a,h)anthracene (Parrish et al. 2005; Smith et al. 2006; Rezek et al. 2008). It has been observed that ryegrasses could effectively reduce the amount of mixture of hydrocarbons included n-alkanes, pristine, hexadecane, phenanthrene, anthracene, fluoranthene and pyrene (Gunther et al. 1996). In another study, it has been found that after 6 months of phytoremediation using three plant species *Medicago sativa* (alfalfa), *Panicum virgatum* (switch grass), and *Schizachyrium scoparium* (little bluestem grass) there is 57% reduction of total PAHs in soils collected from a former MGP (manufactured gas plant) in Newark, New Jersey (Pradhan et al.1998). Similarly, phytoremediation efficiency of Arctared red fescue and annual ryegrass was examined after planted together in crude oil- or diesel-contaminated soil. Results showed lower concentrations of TPH in contaminated soils planted with both plants (Reynolds and Wolf 1999). A pot-culture experiment for reduction in petroleum hydrocarbons was done using five different plant species, *Echinacea purpurea,* Fawn (*Festuca arundinacea*), Fire Phoenix, *Gaillardia aristata* and Alfalfa (*Medicago sativa*). These plants significantly remove TPH including saturated hydrocarbon, aromatic hydrocarbon, asphaltene, and polar compound (Liu et al. 2012). Phytoremediation of contaminated soils can be enhanced along with different technologies such as bioremediation or electrokinetic treatment. It was observed that it is possible to enhance removal of phenanthrene and anthracene in soils by electrophytoremediation with *Brassica rapa* (Cameselle and Gouveia 2019). 47% TPH removal was observed using *Medicago sativa* and the TPH removal rate was found to increase up to 68% for bioaugmentation assisted phytoremediation where *Psedomonas aeruginosa* was used for bioaugmentation (Agnello et al. 2016). There is synergistic effect of bioremediation and phytoremediation for TPH removal (Chaudry et al. 2005).

The remediation technologies as discussed above, i.e., physical, chemical, and biological treatment are effective for PAH degradation at contaminated sites. In order to achieve high efficiency, one has to choose the best remediation technologies as efficiency of these technologies depend on various factors such as type of soil, toxicity associated with contaminant, and environmental conditions. Each of these remediation technologies is associated with their advantage and disadvantage (Lim et al. 2016) with some opportunities and challenges and discussed in Table 4.

## Future outlook and perspective

The contamination of soils with PAHs is inescapable in some sectors such as petroleum industry. A number of treatment methods for PAH-contaminated soils have been developed to remediate such sites. Some of these treatment methods have high efficiency in field-scale application. The understanding of contaminated site, i.e., site-specific conditions such as contaminant nature, soil properties, and weather is important to choose better remediation treatment method as some of them are not efficient for sites

**Table 4** Analysis of strengths, weaknesses, opportunities, and challenges (SWOC) of techniques for soil remediation

| Classification | Treatment method | Strength | Weakness | Opportunities | Challenges | Cost |
|---|---|---|---|---|---|---|
| Physical treatment | Solvent extraction/soil washing | Volume reduction or minimization of contaminants, implement easily and quick | Contaminant transfer to another phase, large amount of solvent required, high cost | Investigate effective and environmental friendly solvent, recycling/regeneration of solvent | May not be effective in treatment of clayey or silty soils, adsorption of solvent into soil particles | Moderate |
| Thermal treatment | Thermal desorption | Volatile contaminants treated over a broad range of moisture contents, high removal efficiency, no excavation required for in situ treatment | Cost-ineffective for large volumes of contaminated soils, volatization and release of contaminants from the soil into the air at high temperatures, less efficiency at low temperature | Investigate different factors such as temperature, soil matrix, contaminant nature and concentration, and moisture content, which effect thermal desorption | Limited to volatile contaminants | Moderate |
| | Microwave frequency heating | High efficiency, quick, treat significant volumes of contaminated soil | High costs, affected with high moisture content | Investigate geochemical changes after treatment | Microwave exposure may cause health problem | High |
| | Vitrification | Easily applied for remediation of highly contaminated soils, destroys contaminants quickly | High cost, transforms treated soil into glassy composition | Investigate vitrified end products which may be recycled or used as other reusable materials | Collection and treatment of off-gases, utility of end product | High |
| | Electrokinetic treatment | Minimal requirement of equipment, safe, suitable for clays with low permeability and heterogeneous soil, used for in situ or ex situ treatment, cost-effective, low electrical energy | Time-consuming, microbial activity may be affected as a result of soil nutrients mobilization | Investigate environmental friendly electrolytes, interaction between electrokinetic parameters and contaminants | Not efficient in low contaminant concentration | Moderate |
| Chemical treatment | Chemical oxidation (AOPs oxidants, i.e., Fenton's reagent, ozone) | Oxidizes organic contaminants simply, efficiency not modified due to toxicity of contaminant | High treatment cost, treatments of extracts, excavation, pH dependency, natural microorganisms destroyed | Investigate other oxidants with in situ application and optimization of parameters such as pH and oxidant concentration. | Restricted by low soil permeability, oxidant high reactivity with soil components | High |
| | Photocatalytic degradation | Pollutant removal at very low concentration, complete degradation of organic contaminant within few hours, efficiency not modified due to toxicity of contaminant | High cost, highly dependent on various parameters such as pH, temperature, amount and nature of photocatalyst, intensity and time of light irradiation, | Investigate active and cheap photocatalyst, and optimization of parameters to achieve high efficiency | Not suitable for commercial application for volatile contaminants removal | High |
| Biological treatment | Bioaugmentation | Natural process and perceived by public, cheaper cost, environmental friendly, contaminants transform into nonhazardous substance, safe | Time-consuming, contaminant toxicity and bioavailability, exogenous microbes decline later addition to a site, i.e., failure of bioaugmentation | Investigate conditions which are required to make bioaugmentation an effective bioremediation tool | Competition between added exogenous microbes and indigenous microbes for limited nutrients | Low |

**Table 4** (continued)

| Classification | Treatment method | Strength | Weakness | Opportunities | Challenges | Cost |
|---|---|---|---|---|---|---|
| | Biostimulation | Natural process and perceived by public, environmental friendly, contaminants transform into nonhazardous substance, safe, native microbes well adapted to environment | Time-consuming, contaminant toxicity and bioavailability, additives for biostimulation not readily available to native microbes | Investigate effects of biostimulation on the native microbial community structure and metabolic activities | Additives might promote growth of other microbes which are not involve in degradation and create competition with microbes involve in degradation | Low |
| | Bioventing | Natural process and perceived by public, cost-effective, environmental friendly, high efficiency, safe, contaminants transform into nonhazardous substance | Not suitable for volatile contaminants, cannot be used for surficial soils, time-consuming, contaminant toxicity and bioavailability | Off-gases treatment in case of volatile contaminant | Suitable only for less volatile contaminant | Moderate |
| | Landfarming | Natural process and perceived by public, environmental friendly, extensive primary assessment of contaminated site prior to treatment not required, contaminants transform into nonhazardous substance, safe | Require large operating space, unfavorable environmental conditions for microbial activities, additional cost due to excavation, time-consuming, contaminant toxicity and bioavailability, low efficiency as compared to other bioremediation treatment methods | Investigate the capability in complex contaminant degradation | Contaminated soil and dust particles spread during excavation and transportation, weather sensitive, ground water contamination | Moderate |
| | Composting | Natural process and perceived by public, low cost, simple operation, high microbial diversity high efficiency, environmental friendly, parameters such as moisture content, nutrients, and pH can be controlled, no surface mixing, contaminants transform into nonhazardous substance, safe | Require large area, time-consuming, contaminant toxicity and bioavailability, capture off-gasses is difficult | Field studies with optimized environmental conditions and microbial consortia for contaminant removal | Controlling moisture, nutrients and pH during treatment | Low |
| | Phytoremediation | Low cost, environmental friendly, large-scale operation, low installation and maintenance cost, soil structure preservation, safe, plants can be easily monitored | Time-consuming, slow plant growth rate, contaminant toxicity and bioavailability | Treatment of sites contaminated with complex mixture of different contaminants | Plant growth requirement (climate, altitude and temperature) | Low |

comprising soils with low permeability or having mixed contaminants. In addition to these site-specific conditions, the selection of remediation technology also depends on their advantages, limitations, cost of alternative remediation method, implementation applicability, and probable environmental impact. So, selection of the best remediation method for field-scale application is very crucial step for PAH removal from contaminated sites. However, each remediation method has its own advantages and constrains, and any single treatment method cannot be used universally for PAH removal. Therefore, two or more remediation methods could be integrated/combined and investigated in future research for different type of soils as enhanced PAH degradation/removal efficiency could be achieved with integrated treatment technologies/methods. Further, comprehensive research is essential to find precisely the existing condition of PAH-contaminated sites.

## Conclusion

Soil pollution arises from different industrial activities, open burning of solid waste, combustion of fossil fuels, etc., and is a critical challenge due to emission of PAHs and its harmful effects on human health and soil ecosystem. This review is an attempt to overview the effects of PAHs on physicochemical and microbial properties of soil, as well as the remediation techniques for reclamation of soils contaminated with PAH. The PAH-contaminated soils may hold the persistent pollutants for long period of time depending on its physicochemical properties (pH, grain size, porosity and organic carbon) and its microbial diversity. Considering the toxic, mutagenic, and carcinogenic effects of PAHs, it is imperative to device a sustainable, cost-effective, and universally adopted treatment method. Whereas the physical and chemical methods are cost and energy intensive, biological methods for reclamation of PAH-contaminated soils are gaining popularity. Landfarming, bioventing, composting, phytoremediation are some of the biotreatment technologies in addition to microbial degradation. Factors like availability of nutrients, favorable environment, microbial adaptation, and their catabolic activity are important regulators of PAH biodegradation. Depending on the chemical structure of PAH, its binding affinity to soil, toxicity profile, characteristics of soil substrate, possible future use of soil, availability of infrastructure, and risk analysis of different treatment options, a suitable method or an integrated approach of two or more methods may be adopted for reclamation of PAH-contaminated soils sustainably. Screening of genetic makeup of native flora and fauna for stimulated microbial activity may be foreseen as a potential tool for remediation of PAH contamination.

**Acknowledgements** The authors acknowledge Mr. Mohnish Sapra, undergraduate student of Department of Environmental Engineering, Delhi Technological University, for his help in various ways.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

Abdel-Shafy HI, Mansour MSM (2016) A review on polycyclic aromatic hydrocarbons: source, environmental impact, effect on human health and remediation. Egypt J Pet 25:107–123

Abed RMM, Al-Kharusi S, Al-Hinai M (2015) Effect of biostimulation, temperature and salinity on respiration activities and bacterial community composition in an oil polluted desert soil. Int Biodeterior Biodegrad 98:43–52

Acar YB, Gale RJ, Alshawabkeh AN, Marks RE, Puppala S, Bricka M, Parker R (1995) Electrokinetic remediation: basics and technology status. J Hazard Mater 40(2):117–137

Agarwal T (2009) Concentration level, pattern and toxic potential of PAHs in traffic soil of Delhi, India. J Hazard Mater 171(1–3):894–900

Agnello AC, Bagard M, Van Hullebusch ED, Esposito G, Huguenot D (2016) Comparative bioremediation of heavy metals and petroleum hydrocarbons co-contaminated soil by natural attenuation, phytoremediation, bioaugmentation and bioaugmentation-assisted phytoremediation. Sci Total Environ 563–564:693–703

Aislabie J, Balks M, Astori N, Stevenson G, Symons R (1999) Polycyclic aromatic hydrocarbons in fuel-oil contaminated soils, Antarctica. Chemosphere 39(13):2201–2207

Alcántara MT, Gómez J, Pazos M, Sanromán MA (2012) Electrokinetic remediation of lead and phenanthrene polluted soils. Geoderma 173–174:128–133

Alexander M (1994) Biodegradation and bioremediation. Academic Press, San Diego

Alleman BC, Hinchee RE, Brenner RC, McCauley PT (1995) Bioventing PAH contamination at the Reilly Tar site. In: Hinchee RE, Miller RN, Johnson PC (eds) In-situ aeration: air sparging, bioventing, and related remediation processes. Battelle Press, Columbus, pp 473–482

Arp HPH, Lundstedt S, Josefsson S, Cornelissen G, Enell A, Allard AS, Kleja DB (2014) Native Oxy-PAHs, N-PACs, and PAHs in Historically contaminated soils from Sweden, Belgium, and France: their soil-porewater partitioning behavior, bioaccumulation in *Enchytraeus crypticus*, and bioavailability. Environ Sci Technol 48(19):11187–11195

Atagana HI (2004) Bioremediation of creosote-contaminated soil in South Africa by landfarming. J Appl Microbiol 96:510–520

Baker RS, Tarmasiewicz D, Bierschenk JM, King J, Landler T, Sheppard D (2007) Completion of in situ thermal remediation of PAHs, PCP and dioxins at a former wood treatment facility. In: International conference on incineration and thermal treatment technologies, Orlando, Florida, US

Barra R, Peter Popp P, Roberto Quiroz R, Coretta Bauer C, Hernan Cid H, Tümpling W (2005) Persistent toxic substances in soils and waters along an altitudinal gradient in the Laja River Basin, Central Southern Chile. Chemosphere 58:905–915

Barrán-Berdón AL, González VG, Aboytes GP, Rodea-Palomares I, Carrillo-Chávez A, Gómez Ruiz H, Cuéllar BV (2012) Polycyclic aromatic hydrocarbons in soils from a brick manufacturing location in central Mexico. Rev Int Contam Ambie 28(4):277–288

Bayoumi RA (2009) Bacterial bioremediation of polycyclic aromatic hydrocarbons in heavy oil contaminated soil. J Appl Sci Res 5(2):197–211
Bogan BW, Trbovic V (2003) Effect of sequestration on PAH degradability with Fenton's reagent: roles of total organic carbon, humin, and soil porosity. J Hazard Mater 100:285–300
Boopathy R (2000) Factors limiting bioremediation technologies. Bioresour Technol 74:63–67
Breedveld GD, Sparrevik M (2000) Nutrient-limited biodegradation of PAH in various soil strata at a creosote contaminated site. Biodegradation 11:391–399
Brown GS, Barton LL, Thomson BM (2003) Permanganate oxidation of sorbed polycyclic aromatic hydrocarbons. Waste Manag 23:737–740
Bucheli TD, Blum F, Desaules A, Gustafsson Ö (2004) Polycyclic aromatic hydrocarbons, black carbon, and molecular markers in soils of Switzerland. Chemosphere 56:1061–1076
Byss M, Elhottová D, Tříska J, Baldrian P (2008) Fungal bioremediation of the creosote-contaminated soil: influence of *Pleurotus ostreatus* and *Irpex lacteus* on polycyclic aromatic hydrocarbons removal and soil microbial community composition in the laboratory-scale study. Chemosphere 73(9):1518–1523
Cameselle C, Gouveia S (2019) Phytoremediation of mixed contaminated soil enhanced with electric current. J Hazard Mater 361:95–102
Campos VM, Merino I, Casado R, Pacios LF, Gómez L (2008) Review: Phytoremediation of organic pollutants. Span J Agric Res 6:38–47
Cerniglia CE, Sutherland JB (2010) Degradation of polycyclic aromatic hydrocarbons by fungi. In: Timmis KN, McGenity TJ, van der Meer JR, de Lorenzo V (eds) Handbook of hydrocarbon and lipid microbiology. Springer, Berlin, pp 2080–2110
Chaineau C, Rougeux G, Yepremian C, Oudot J (2005) Effects of nutrient concentration on the biodegradation of crude oil and associated microbial populations in the soil. Soil Biol Biochem 37:1490–1497
Chang HJ, Jou CJG, Lee CL (2011) Treatment of heavy oil contaminated sand by microwave energy. Environ Eng Sci 28:869–873
Chaudry Q, Blom-Zandstra M, Gupta S, Joner EJ (2005) Utilising the synergy between plants and rhizosphere microorganisms to enhance breakdown of organic pollutants in the environment. Environ Sci Pollut Res 12(1):34–48
Chien YC (2012) Field study of in situ remediation of petroleum hydrocarbon contaminated soil on site using microwave energy. J Hazard Mater 199:457–461
Chu SG, Liu H, Ma LL, Xu XB (2003) Polycyclic aromatic hydrocarbons in soil adjacent to highways in Beijing, People's Republic of China. Bull Environ Contam Toxicol 70:972–977
Decesaro A, Rampel A, Machado TS, Thomé A, Reddy K, Margarites AC, Colla LM (2017) Bioremediation of soil contaminated with diesel and biodiesel fuel using biostimulation with microalgae biomass. J Environ Eng 143(4):04016091
Denis EH, Toney JL, Tarozo R, Anderson RS, Roach LD, Huang Y (2012) Polycyclic aromatic hydrocarbons (PAHs) in lake sediments record historic fire events: validation using HPLC-fluorescence detection. Org Geochem 45:7–17
Devi NL, Yadav IC, Shihua Q, Dan Y, Zhang G, Raha P (2016) Environmental carcinogenic polycyclic aromatic hydrocarbons in soil from Himalayas, India: implications for spatial distribution, sources apportionment and risk assessment. Chemosphere 144:493–502
Do SH, Jo JH, Jo YH, Lee HK, Kong SH (2009) Application of a peroxymonosulfate/cobalt (PMS/Co (II)) system to treat diesel-contaminated soil. Chemosphere 77:1127–1131
Do SH, Kwon YJ, Kong SH (2010) Effect of metal oxides on the reactivity of persulfate/Fe(II) in the remediation of diesel-contaminated soil and sand. J Hazard Mater 182:933–936
Doong R, Lin Y (2003) Characterization and distribution of polycyclic aromatic hydrocarbon contaminations in surface sediment and water from Gao-ping River, Taiwan. Water Res 38(7):1733–1744
Downard J, Singh A, Bullard R, Jayarathne T, Rathnayake C, Simmons DL, Stone EA (2015) Uncontrolled combustion of shredded tires in a landfill—part 1: characterization of gaseous and particulate emissions. Atmos Environ 104:195–204
Falciglia P, Giustra M, Vagliasindi FG (2011) Low-temperature thermal desorption of diesel polluted soil: influence of temperature and soil texture on contaminant removal kinetics. J Hazard Mater 185:392–400
Falciglia PP, Urso G, Vagliasindi FGA (2013) Microwave heating remediation of soils contaminated with diesel fuel. J Soils Sediments 13:1396–1407
Falciglia PP, Guidib GD, Catalfob A, Vagliasindia FGA (2017) Contaminant removal mechanisms in microwave heating remediation of PAH-contaminated soils. Chem Eng Trans 57:361–366
Fenton HJH (1894) Oxidation of tartaric acid in presence of iron. J Chem Soc Trans 65:899–910
Fernández-Luqueño F, López-Valdez F, Pérez-Morales C, García-Mayagoitia S, Sarabia-Castillo CR, Pérez-Ríos SR (2017) Enhancing decontamination of PAHs-polluted soils: role of organic and mineral amendments. In: Anjum N, Gill S, Tuteja N (eds) Enhancing cleanup of environmental pollutants. Springer, Cham
Flotron V, Delteil C, Padellec Y, Camel V (2005) Removal of sorbed polycyclic aromatic hydrocarbons from soil, sludge and sediment samples using the Fenton's reagent process. Chemosphere 59:1427–1437
García Frutos FJ, Escolano O, García S, Babín M, Fernández MD (2010) Bioventing remediation and ecotoxicity evaluation of phenanthrene-contaminated soil. J Hazard Mater 183:806–813
García-Alonso S, Pérez-Pastor RM, Sevillano-Castaño ML, Escolano O, García-Frutos FJ (2008) Influence of particle size on the quality of PAH concentration measurements in a contaminated soil. Polycycl Aromat Compd 28(1):67–83. https://doi.org/10.1080/10406630701815253
García-Delgado C, Fresno T, Rodríguez-Santamaría JJ, Diaz E, Mohedano AF, Moreno-Jimenez E (2019) Co-application of activated carbon and compost to contaminated soils: toxic elements mobility and PAH degradation and availability. Int J Environ Sci Technol 16:1057–1068
Garon D, Sage L, Wouessidjewe D, Seigle-Murandi F (2004) Enhanced degradation of fluorine in soil slurry by *Absidia cylindrospora* and maltosyl-cyclodextrin. Chemosphere 56:159–166
Germida J, Frick C, Farrell R (2002) Phytoremediation of oil-contaminated soils. Dev Soil Sci 28:169–186
Goi A, Trapido M, Kulik N, Palmroth M, Tuhkanen T (2006) Ozonation and Fenton treatment for remediation of diesel fuel contaminated soil. Ozone Sci Eng 28:37–46
Gu J, Dong D, Kong L, Zheng Y, Li X (2012) Photocatalytic degradation of phenanthrene on soil surfaces in the presence of nanometer anatase $TiO_2$ under UV-light. J Environ Sci 24(12):2122–2126
Gunther T, Dornberger U, Fritsche W (1996) Effects of ryegrass on biodegradation of hydrocarbons in soil. Chemosphere 33(2):203–215
Haritash AK, Kaushik CP (2009) Biodegradation aspects of polycyclic aromatic hydrocarbons: a review. J Hazard Mater 169:1–15
Haritash AK, Kaushik CP (2016) Degradation of low molecular weight polycyclic aromatic hydrocarbons by microorganisms isolated from contaminated soil. Int J Environ Sci 6:0976–4402

Haussmann HJ (2012) Use of hazard indices for a theoretical evaluation of cigarette smoke composition. Chem Res Toxicol 25(4):794–810
Heinaru E, Merimaa M, Viggor S, Lehiste M, Leito I, Truu J, Heinaru A (2005) Biodegradation efficiency of functionally important populations selected for bioaugmentation in phenol- and oil-polluted area. FEMS Microb Ecol 51:363–373
Hinchee RE (1993) Bioventing of petroleum hydrocarbons. Handbook of bioremediation. CRC Press, Boca Raton
Hinchee RE, Smith LA (1992) In situ thermal technologies for site remediation. CRC Press, Boca Raton
Howard P, Meylan W, Aronson D, Stiteler W, Tunkel J, Comber M, Parkerton TF (2005) A new biodegradation prediction model specific to petroleum hydrocarbons. Environ Toxicol Chem 24:1847–1860
Hreniuc M, Coman M, Cioruţa B (2015) Consideration regarding the soil pollution with oil products in Sacel-Maramures. In: International conference of scientific paper AFASES, Brasov, 28–30
Huang D, Xu Q, Cheng J, Lu X, Zhang H (2012) Electrokinetic remediation and its combined technologies for removal of organic pollutants from contaminated soils. Int J Electrochem Sci 7:4528–4544
Ibrahim AM (2004) Soil pollution: origin, monitoring and remediation. Springer, Berlin
Jonsson S, Persson Y, Frankki S, Lundstedt S, Bavel BV, Haglund P, Tysklind M (2006) Comparison of Fenton reagent and ozone oxidation of polycyclic aromatic hydrocarbon in aged contaminated soils. J Soil Sediment 6:208–214
Karaca O, Cameselle C, Reddy KR (2016) Electrokinetic removal of heavy metals from mine tailings and acid lake sediments from Can Basin, Turkey. Geotechnical special publication (273 GSP), pp 225–234
Kaushik CP, Haritash AK (2006) Polycyclic aromatic hydrocarbons (PAHs) and environmental health. Our Earth 3(3):1–7
Kaushik CP, Sangwan P, Haritash AK (2012) Association of polycyclic aromatic hydrocarbons (PAHs) with different sizes of atmospheric particulate in Hisar City and its health aspects. Polycycl Aromat Compd 32(5):626–642
Khodadoust AP, Bagchi R, Suidan MT, Brenner RC, Sellers NG (2000) Removal of PAHs from highly contaminated soils found at prior manufactured gas operations. J Hazard Mater 80(1–3):159–174
Khomarbaghi Z, Shavandi M, Amoozegar MA, Dastgheib SMM (2019) Bacterial community dynamics during bioremediation of alkane- and PAHs-contaminated soil of Siri Island, Persian Gulf: a microcosm study. Int J Environ Sci Technol. https://doi.org/10.1007/s13762-018-02198-y
Kiamarsi Z, Soleimani M, Nezami A, Kafi M (2018) Biodegradation of n-alkanes and polycyclic aromatic hydrocarbons using novel indigenous bacteria isolated from contaminated soils. Int J Environ Sci Technol 1(1):1–12
Kim GB, Maruya KA, Lee RF, Lee JH, Koh CH, Tanabe S (1999) Distribution and sources of polycyclic aromatic hydrocarbons in sediments from Kyeonggi Bay, Korea. Mar Pollut Bull 38(1):7–15
Klamerus-Iwan A, Błońska E, Lasota J, Kalandyk A, Waligórski P (2015) Influence of oil contamination on physical and biological properties of forest soil after chainsaw use. Water Air Soil Pollut 226(11):389
Klánová J, Matykiewiczová N, Máčka Z, Prošek P, Láska K, Klán P (2008) Persistent organic pollutants in soils and sediments from James Ross Island, Antarctica. Environ Pollut 152:416–423
Kozak K, Ruman M, Kosek K, Karasiński G, Stachnik Ł, Polkowska Z (2017) Impact of volcanic eruptions on the occurrence of PAHs compounds in the aquatic ecosystem of the southern part of west Spitsbergen (Hornsund Fjord, Svalbard). Water 9:42
Krauss M, Wilcke W, Martius C, Bandeira A, Garcia M, Amelung W (2005) Atmospheric versus biological sources of polycyclic aromatic hydrocarbons (PAHs) in a tropical rain forest environment. Environ Pollut 135:143–154
Kuo YM, Lin TC, Tsai PJ, Lee WJ, Lin HY (2003) Fate of polycyclic aromatic hydrocarbons during vitrification of incinerator ash in a coke bed furnace. Chemosphere 51(4):313–319
Labud V, Garcia C, Hernandez T (2007) Effect of hydrocarbon pollution on the microbial properties of a sandy and a clay soil. Chemosphere 66:1863–1871
Lai CH, Li HC, Chen KS (2009) Source characterization and environment impact of open burning of rice straw residues on polycyclic aromatic hydrocarbons in agricultural county, Taiwan. J Environ Eng Manag 19(2):79–88
Lau KL, Tsang YY, Chiu SW (2003) Use of spent mushroom compost to bioremediate PAH-contaminated samples. Chemosphere 52:1539–1546
Li D, Quan X, Zhang Y, Zhao Y (2008) Microwave-induced thermal treatment of petroleum hydrocarbon-contaminated soil. Soil Sediment Contam 17:486–496
Li X, Wang X, Ren ZJ, Zhang Y, Li N, Zhou Q (2015) Sand amendment enhances bioelectrochemical remediation of petroleum hydrocarbon contaminated soil. Chemosphere 141:62–70
Lim MW, Lau EV, Poh PE (2016) A comprehensive guide of remediation technologies for oil contaminated soil—present works and future directions. Mar Pollut Bull 109(1):14–45
Lipińska A, Kucharski J, Wyszkowska J (2014) The effect of polycyclic aromatic hydrocarbons on the structure of organotrophic bacteria and dehydrogenase activity in soil. Polycycl Aromat Compd 34(1):35–53
Liu R, Jadeja RN, Zhou Q, Liu Z (2012) Treatment and remediation of petroleum contaminated soils using selective ornamental plants. Environ Eng Sci 29:494–501
Lueking AD, Huang WL, Soderstrom-Schwarz S, Kim M, Weber WJ (2000) Relationship of soil organic matter characteristics to organic contaminant sequestration and bioavailability. J Environ Qual 29:317–323
Lyu Y, Zheng W, Zheng T, Tian Y (2014) Biodegradation of Polycyclic aromatic hydrocarbons by *Novosphingobium pentaromativorans* US6-1. PLoS ONE 9(7):e101438
Mackay D, Shiu WY (1977) Aqueous solubility of polynuclear aromatic hydrocarbons. J Chem Eng 22:399–402
Magi E, Bianco R, Ianni C, Carro MD (2002) Distribution of polycyclic aromatic hydrocarbons in the sediments of the Adriatic Sea. Environ Pollut 119:91–98
Mai B, Qi S, Zeng E, Yang Q, Hang G, Fu J, Sheng G, Peng P, Wang Z (2003) Distribution of polycyclic aromatic hydrocarbons in the coastal region off Macao, China: assessment of input sources and transport pathways using compositional analysis. Environ Sci Technol 37:4855–4863
Maliszewska-Kordybach B (1996) Polycyclic aromatic hydrocarbons in agricultural soils in Poland: preliminary proposals for criteria to evaluate the level of soil contamination. Appl Geochem 11(1–2):121–127
Mao J, Guan W (2016) Fungal degradation of polycyclic aromatic hydrocarbons (PAHs) by *Scopulariopsis brevicaulis* and its application in bioremediation of PAH-contaminated soil. Acta Agric Scand B Soil Plant Sci 66(5):399–405
Masih A, Taneja A (2006) Polycyclic aromatic hydrocarbons (PAHs) concentrations and related carcinogenic potencies in soil at a semi-arid region of India. Chemosphere 65:449–456
Masten SJ, Davies SHR (1997) Efficacy of in situ ozonation for the remediation of PAH contaminated soils. J Contam Hydrol 28:327–335
Mazzera D, Hayes T, Lowenthal D, Zielinska B (1999) Quantification of polycyclic aromatic hydrocarbons in soil at McMurdo Station, Antarctica. Sci Total Environ 229:65–71

McCarthy K, Walker L, Vigoren L, Bartel J (2004) Remediation of spilled petroleum hydrocarbons by in situ landfarming at an artic site. Cold Reg Sci Technol 40:31–39

McCready S, Slee DJ, Birch JF, Taylor SE (2000) The distribution of polycyclic aromatic hydrocarbons in surficial sediments of Sydney Harbour, Australia. Mar Pollut Bull 40:999–1006

McCutcheon SC, Schnoor JL (2003) Phytoremediation: transformation and control of contaminants. Wiley, Hoboken

McRae C, Sun C, McMillan CF, Snape CE, Fallick AE (2000) Sourcing of fossil fuel-derived PAH in the environment. Polycycl Aromat Compd 20(1–4):97–109

Michele MM, Stanley PW, Guo LH, Wan YS, Donald M (1985) Relationships between octanol-water partition coefficient and aqueous solubility. Environ Sci Technol 19(6):522–529

Mielke HW, Wang G, Gonzales CR, Powel ET, Le B, Quach VN (2004) PAHs and metals in the soils of inner-city and suburban New Orleans, Louisiana, USA. Environ Toxicol Pharmacol 18:243–247

Mishra S, Singh SN (2014) Biodegradation of benzo(a)pyrene mediated by catabolic enzymes of bacteria. Int J Environ Sci Technol 11:1571

Muckian L, Grant R, Clipson N, Doyle E (2009) Bacterial community dynamics during bioremediation of phenanthrene and fluoranthene amended soil. Int Biodeterior Biodegrad 63:52–56

Nadal M, Schuhmacher M, Domingo JL (2004) Levels of PAHs in soil and vegetation samples from Tarragona County, Spain. Environ Pollut 132:1–11

Nam JJ, Song BH, Eom KC, Lee SH, Smith A (2003) Distribution of polycyclic aromatic hydrocarbons in agricultural soils in South Korea. Chemosphere 50:1281–1289

Namkoong W, Hwang E, Park J, Choi J (2002) Bioremediation of diesel contaminated soil with composting. Environ Pollut 119:23–31

Nganje TN, Edet AE, Ekwere SJ (2007) Distribution of PAHs in surface soils from petroleum handling facilities in Calabar. Environ Monit Assess 130(1–3):27–34

Nguyen TC, Loganathan P, Nguyen TV, Vigneswaran S, Kandasamy J, Slee D, Stevenson G, Naidu R (2014) Polycyclic aromatic hydrocarbons in road-deposited sediments, water sediments, and soils in Sydney, Australia: comparisons of concentration distribution, sources and potential toxicity. Ecotoxicol Environ Saf 104:339–348

Nieuwoudt C, Pieters R, Quinn LP, Kylin H, Borgen AR, Bouwman H (2011) Polycyclic aromatic hydrocarbons (PAHs) in soil and sediment from industrial, residential, and agricultural areas in central south Africa: an initial assessment. Soil Sediment Contam 20:188–204

O'Mahony MM, Dobson ADW, Barnes JD, Singleton I (2006) The use of ozone in the remediation of polycyclic aromatic hydrocarbon contaminated soil. Chemosphere 63:307–314

Oleszczuk P, Baran S (2005) Leaching of individual PAHs in soil varies with the amounts of sewage sludge applied and total organic carbon content. Pol J Environ Stud 14(4):491–500

Pampanin DM, Sydnes MO (2013) Polycyclic aromatic hydrocarbons a constituent of petroleum: presence and influence in the aquatic environment. In: Kutcherov V, Kolesnikov A (eds) Hydrocarbon. Intech Open Ltd., London, UK. ISBN 978-953-51-0927-3

Pandey AK, Chaudhary P, Singh SB, Arora A, Kumar K, Chaudhry S, Nain L (2012) Deciphering the traits associated with PAH degradation by a novel *Serratia marcesencs* L-11 strain. J Environ Sci Health A Tox Hazard Subst Environ Eng 47(5):755–765

Parrish ZD, Banks MK, Schwab AP (2005) Assessment of contaminant liability during phytoremediation of polycyclic aromatic hydrocarbon impacted soil. Environ Pollut 137:187–197

Pathak H, Bhatnagar K, Jaroli DP (2011) Physico-chemical properties of petroleum polluted soil collected from Transport Nagar (Jaipur). Indian J Fundam Appl Life Sci 1(3):2231–6345

Patnaik P (1999) A comprehensive guide to the properties of hazardous chemical substances, 2nd edn. Wiley, Hoboken

Paudyn K, Rutter A, Rowe RK, Poland JS (2008) Remediation of hydrocarbon contaminated soils in the Canadian Arctic by landfarming. Cold Reg Sci Technol 53:102–114

Pettersen H, Näf C, Broman D (1997) Impact of PAH outlets from an oil refinery on the receiving water area—sediment trap fluxes and multivariate statistical analysis. Mar Pollut Bull 34(2):85–95

Picado A, Nogueira A, Baeta-Hall L, Mendonça E, de Fátima Rodrigues M, do Céu Sàágua M, Martins A, Anselmo AM (2001) Landfarming in a PAH-contaminated soil. J Environ Sci Health A Tox Hazard Subst Environ Eng 36(9):1579–1588

Piña J, Merino J, Errazu AF, Bucalá V (2002) Thermal treatment of soils contaminated with gas oil: influence of soil composition and treatment temperature. J Hazard Mater 94:273–290

Plachá D, Raclavská H, Matýsek D, Rümmeli MH (2009) The polycyclic aromatic hydrocarbon concentrations in soils in the Region of Valasske Mezirici, the Czech Republic. Geochem Trans 10:12

Potin O, Veignie E, Rafin C (2004) Biodegradation of polycyclic aromatic hydrocarbon by *Cladosporium sphaerospermum* isolated from an aged (PAHs) contaminated soil. FEMS Microbiol Eco 51(1):71–78

Pradhan SP, Conrad JR, Paterek JR, Srivastava VJ (1998) Potential of phytoremediation for treatment of PAHs in soil at MGP sites. J Soil Contam 7(4):467–480

Rababah A, Matsuzawa S (2002a) Treatment system for solid matrix contaminated with fluoranthene I. Modified extraction technique. Chemosphere 46(1):39–47

Rababah A, Matsuzawa S (2002b) Treatment system for solid matrix contaminated with fluoranthene. II—Recirculating photodegradation technique. Chemosphere 46:49–57

Ratola N, Lacorte S, Alves A, Barceló D (2006) Analysis of polycyclic aromatic hydrocarbons in pine needles by gas chromatography-mass spectrometry: comparison of different extraction and clean-up procedures. J Chromatogr A 1114(2):198–204

Reddy KR, Adams JA (2015) Sustainable remediation of contaminated sites. Momentum Press, New York

Ren G, Ren W, Teng Y, Li Z (2015) Evident bacterial community changes but only slight degradation when polluted with pyrene in a red soil. Front Microbiol 6:22–33

Renoldi F, Lietti L, Saponaro S, Bonomo L, Forzatti P (2003) Thermal desorption of a PAH-contaminated soil: a case study. In: Brebbia CA (ed) Ecosystems and sustainable development, vol 2. WIT Press, Ashurst, pp 1123–1132

Reynolds CM, Wolf DC (1999) Microbial based strategies for assessing rhizosphere enhanced phytoremediation. In: Proceedings of the phytoremediation technical seminar, Calgary, AB, Environment Canada, Ottawa, pp 125–135

Rezek J, Wiesche C, Mackova M, Zadrazil F, Macek T (2008) The effect of ryegrass (*Lolium perenne*) on decrease of PAH content in long term contaminated soil. Chemosphere 70:1603–1608

Riser-Roberts E (1998) Bioremediation of petroleum contaminated soils. Lewis Publishers, Boca Raton

Rivas FJ (2006) Polycyclic aromatic hydrocarbons sorbed on soils: a short review of chemical oxidation based treatments. J Hazard Mater 138:234–251

Rushton D, Ghaly AE, Martinell K (2007) Assessment of Canadian regulations and remediation methods for diesel oil contaminated soils. Am J Appl Sci 4:465

Saba B, Rafique U, Hashmi I (2010) Adsorption kinetics of anthracene and phenanthrene in different soils of Attock Refinery Limited (ARL) Rawalpindi, Pakistan. Desalin Water Treat 30(1–3):333–338

Sasek V, Bhatt M, Cajthaml T, Malachová K, Lednická D (2003) Compost-mediated removal of polycyclic aromatic hydrocarbons from contaminated soil. Arch Environ Contam Toxicol 44:336–342

Sharma D, Jain S (2019) Impact of intervention of biomass cookstove technologies and kitchen characteristics on indoor air quality and human exposure in rural settings of India. Environ Int 123:240–255

Sharma HD, Reddy KR (2004) Geoenvironmental engineering: site remediation, waste containment, and emerging waste management technologies. Wiley, Hoboken

Sharma A, Verma M, Haritash AK (2016) Degradation of toxic azo dye (AO7) using Fenton's process. Adv Environ Res 5(3):189–200

Shearer TL (1991) A comparison of In situ vitrification and rotary kiln incineration for soils treatment. J Air Waste Manag Assoc 41(9):1259–1264

Shin KH, Jung H, Chang P, Choi H, Kim KW (2005) Earthworm toxicity during chemical oxidation of diesel-contaminated sand. Environ Toxicol Chem 24:1924–1929

Silva A, Delerue-Matos C, Fiuza A (2005) Use of solvent extraction to remediate soils contaminated with hydrocarbons. J Hazard Mater B 124:224–229

Simpson CD, Mosi AA, Cullen WR, Reimer KJ (1996) Composition and distribution of polycyclic aromatic hydrocarbon contamination in surficial marine sediments from Kitimat Harbor, Canada. Sci Total Environ 181(3):265–278

Smith MJ, Flowers TH, Duncan HJ, Alder J (2006) Effects of polycyclic aromatic hydrocarbons on germination and subsequent growth of grasses and legumes in freshly contaminated soil and soil with aged PAHs residues. Environ Pollut 141:519–525

Soriano JA, Viñas L, Franco MA, González JJ, Ortiz L, Bayona JM, Albaigés J (2006) Spatial and temporal trends of petroleum hydrocarbons in wild mussels from the Galician coast (NW Spain) affected by the Prestige oil spill. Sci Total Environ 370:80–90

Sosa D, Hilber I, Faure R, Bartolomé N, Fonseca O, Keller A, Schwab P, Escobar A, Bucheli TD (2017) Polycyclic aromatic hydrocarbons and polychlorinated biphenyls in soils of Mayabeque, Cuba. Environ Sci Pollut Res Int 24(14):12860–12870

Sreelal G, Jayanthi R (2017) Review on phytoremediation technology for removal of soil contaminant. Indian J Sci Res 14(1):127–130

Sutton NB, Maphosa F, Morillo JA, Abu Al-Soud W, Langenhoff AA, Grotenhuis T, Rijnaarts H, Smidt H (2013) Impact of long-term diesel contamination on soil microbial community structure. Appl Environ Microbiol 79(2):619–630

Syed JH, Iqbal M, Zhong G, Katsoyiannis A, Yadav IC, Li J, Zhang G (2017) Polycyclic aromatic hydrocarbons (PAHs) in Chinese forest soils: profile composition, spatial variations and source apportionment. Sci Rep 7:2692

Tam NFY, Ke L, Wang XH, Wong YS (2001) Contamination of polycyclic aromatic hydrocarbons in surface sediments of mangrove swamps. Environ Pollut 114(2):255–263

Tatàno F, Felici F, Mangani F (2013) Lab-scale treatability tests for the thermal desorption of hydrocarbon-contaminated soils. Soil Sediment Contam Int J 22:433–456

Thomé A, Reginatto C, Cecchin I, Colla LM (2014) Bioventing in a residual clayey soil contaminated with a blend of biodiesel and diesel oil. J Environ Eng 140(11):06014005

Tsai T, Kao C (2009) Treatment of petroleum-hydrocarbon contaminated soils using hydrogen peroxide oxidation catalyzed by waste basic oxygen furnace slag. J Hazard Mater 170:466–472

Ugwu KE, Ukoha PO (2016) Analysis and sources of polycyclic aromatic hydrocarbons in soil and plant samples of a coal mining area in Nigeria. Bull Environ Contam Toxicol 96(3):383–387

Varjani SJ, Joshi RR, Kumar PS, Srivastava VK, Kumar V, Banerjee C, Kumar RP (2017) Polycyclic aromatic hydrocarbons from petroleum oil industry activities: effect on human health and their biodegradation. In: Waste bioremediation, energy, environment, and sustainability. Springer, Singapore, pp 185–199

Verma M, Haritash AK (2019) Degradation of amoxicillin by Fenton and Fenton-integrated hybrid oxidation processes. J Environ Chem Eng 7(1):102886. https://doi.org/10.1016/j.jece.2019.102886

Verrhiest GJ, Clement B, Volat B, Montuelle B, Perrodin Y (2002) Interactions between a polycyclic aromatic hydrocarbon mixture and the microbial communities in a natural freshwater sediment. Chemosphere 46(2):187–196

Viglianti C, Hanna K, Brauer C, Germain P (2006) Removal of polycyclic aromatic hydrocarbons from aged-contaminated soil using cyclodextrins: experimental study. Environ Pollut 140:427–435

Wang JY, Huang XJ, Kao JCM, Stabnikova O (2007) Simultaneous removal of organic contaminants and heavy metals from kaolin using an upward electrokinetic soil remediation process. J Hazard Mater 144:292–299

Wilcke W, Zech W (1997) Polycyclic aromatic hydrocarbons (PAHs) in forest floors of the northern Czech mountains. Z Pflanzenernähr Bodenkd 160:369–378

Wilcke W, Zech W, Kobza J (1996) PAH pools in soils along a PAH deposition gradient. Environ Pollut 92:307–313

Wilcke W, Amelung W, Zech W (1997) Heavy metals and polycyclic aromatic hydrocarbons (PAHs) in a rural community leewards of a waste incineration plant. Z Pflanzenernähr Bodenkd 160:369–378

Wilcke W, Lilienfein J, Lima SDC, Zech W (1999a) Contamination of highly weathered urban soils in Uberlândia, Brazil. J Plant Nutr Soil Sci 162:539–548

Wilcke W, Müller S, Kanchanakool N, Niamskul C, Zech W (1999b) Polycyclic aromatic hydrocarbons (PAHs) in hydromorphic soils of the tropical metropolis Bangkok. Geoderma 91:297–309

Wilcke W, Krauss M, Safronov G, Fokin AD, Kaupenjohann M (2005) Polycyclic aromatic hydrocarbons in soils of the Moscow region: concentrations, temporal trends, and small-scale distribution. Environ Pollut 141:327–335

Wilcke W, Bandowe BAM, Lueso MG, Ruppenthal M, H-d Valle, Oelmann Y (2014) Polycyclic aromatic hydrocarbons (PAHs) and their polar derivatives (oxygenated PAHs, azaarenes) in soils along a climosequence in Argentina. Sci Total Environ 473–474:317–325

Wilson SC, Jones KC (1993) Bioremediation of soil contaminated with polynuclear aromatic hydrocarbons (PAHs): a review. Environ Pollut 81:229–249

Working Group on Polycyclic Aromatic Hydrocarbons (WGPAH) (2001) Ambient air pollution by polycyclic aromatic hydrocarbons: position paper, 2001 Annexes. https://ec.europa.eu/environment/air/pdf/pp_pah.pdf. Accessed 18 Oct 2016

Wu G, Kechavarzi C, Li X, Sui H, Pollard SJ, Coulon F (2013) Influence of mature compost amendment on total and bioavailable polycyclic aromatic hydrocarbons in contaminated soils. Chemosphere 90:2240–2246

Xiaojun L, Peijun L, Xin L, Chungui Z, Qi L, Zongqiang G (2007) Biodegradation of aged polycyclic aromatic hydrocarbons by microbial consortia in soil and slurry phases. J Hazard Mater 1016:6–12

Xing B, Pignatello JJ (1997) Dual-mode sorption of low-polarity compounds in glassy poly(vinyl chloride) and soil organic matter. Environ Sci Technol 31(3):792–799

Yao JJ, Huang ZH, Masten SJ (1998) The ozonation of pyrene: pathway and product identification. Water Res 32:3001–3012

Yen CH, Chen KF, Kao CM, Liang SH, Chen TY (2011) Application of persulfate to remediate petroleum hydrocarbon-contaminated soil: feasibility and comparison with common oxidants. J Hazard Mater 186:2097–2102

Yu DY, Kang N, Bae W, Banks MK (2007) Characteristics in oxidative degradation by ozone for saturated hydrocarbons in soil contaminated with diesel fuel. Chemosphere 66:799–807

Zaghden H, Kallel M, Elleuch B, Oudot J, Saliot A (2007) Sources and distribution of aliphatic and polyaromatic hydrocarbons in sediments of Sfax, Tunisia, Mediterranean Sea. Mar Chem 105:70–89

Zhang HB, Luo YM, Wong MH, Zhao QG, Zhang GL (2006) Distribution and concentrations of PAHs in Hong Kong soils. Environ Pollut 141:107–114

Zhang L, Li P, Gong Z, Li X (2008) Photocatalytic degradation of polycyclic aromatic hydrocarbons on soil surfaces using $TiO_2$ under UV light. J Hazard Mater 158:478–484

Zheng ZM, Obbard JP (2003) Oxidation of polycyclic aromatic hydrocarbons by fungal isolates from oil contaminated refinery soil. Environ Sci Pollut Res 10:173–176

Zucchi M, Angiolini L, Borin S, Brusetti L, Dietrich N, Gigliotti C, Barbieri P, Sorlini C, Daffonchio D (2003) Response of bacterial community during bioremediation of an oil-polluted soil. J Appl Microbiol 94:248–257

2019 6th International Conference on Signal Processing and Integrated Networks (SPIN)

# Realisation of CBTA Based Current Mode Frequency Agile Filter

Kshitij Chhabra
*Department of Electronics and Communication Engineering*
*Delhi Technological University*
Delhi, India
kchhabra.1998@gmail.com

Sanjhi Singhal
*Department of Electrical Engineering*
*Delhi Technological University*
Delhi, India
sanjhisinghal007@gmail.com

Neeta Pandey
*Department of Electronics and Communication Engineering*
*Delhi Technological University*
Delhi, India
n66pandey@rediffmail.com

***Abstract* - In this paper, current mode frequency agile filter based on Current Backwards Trans-Conductance Amplifier (CBTA) is proposed. The proposed filter provides high tunability and agility. The internal structure of CBTA is modified to produce two low pass responses simultaneously. Using the modified CBTA schematic, Low pass controlled low pass, Low pass controlled high pass and High pass controlled low pass agile filters are proposed. These responses are functionally verified using PSPICE simulations. The simulations are performed using the PR100N (PNP) and NR100N (NPN) transistors model of the bipolar arrays ALA400 and filter parameters with variation in external gain is examined.**

***Keywords*— Current backward trans-conductance amplifier (CBTA), Frequency agile filter (FAF), biquads, current mode circuits.**

## I. INTRODUCTION

The demand of front end analog solutions is the consequence of rapid development in the field of wireless communications[1]. These newly introduced services have therefore contributed to different corresponding standards[1]. This in turn necessitated the development of multi-standard transceivers owing to its small size and low power consumption. A vital component of multi-standard transceivers is reconfigurable filter. The electronically reconfigurable filters are of great importance in different domains of analog electronics. The Radio frequency (RF) filters used widely in telecommunications, are the electronically tunable second order filters. A new class of second order filters, named as agile filters as explained in [1]-[10], which can bring a revolution in the telecommunications by virtue of its re-configurability, and agility are discussed in this paper.

Out of various current mode blocks available [11], current backwards trans-conductance amplifier (CBTA) [12] has proven to be beneficial owing to its independence from input capacitance [13]. Various application of CBTA like integrators, amplifiers have been cited in the literatures [14]-[20]. Considering these facts, a CBTA based current mode FAF is proposed in this paper. Section II gives a brief implementation scheme of FAF in general, and CBTA based FAF is proposed thereafter. Section III outlays the CBTA implementation Class 1 FAF. The proposal is verified through SPICE simulations and corresponding results are arranged in the Section IV. The conclusions are made in Section V.

## II. IMPLEMENTATION SCHEME OF FAF

The FAF as proposed in [2] is a newly developed method for tuning filter parameters of the second order filter. Here, an external amplifier is added in the feedback loop and by controlling its gain (A), the parameters of the resulting filter can then be tuned. The concept is illustrated by a second order current mode filter exhibiting two current outputs, namely $I_{OUT1}$ and $I_{OUT2}$ is shown in Fig.1

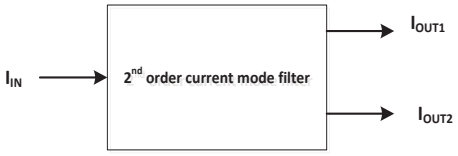



Fig. 1. Second Order Filter

A current mode FAF(Class 1) [2] is illustrated in Fig. 2 where $I_{OUT2}$ is amplified through variable gain A and is fed back to the input of the filter. The resulting characteristics of the filter are found to be dependent on the external gain A.

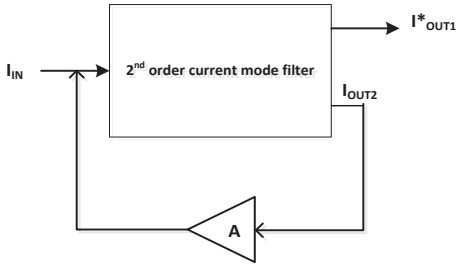



Fig 2. FAF (Class 1)

Considering $I_{OUT1}$ and $I_{OUT2}$ (Fig. 1) as low pass response given by:

$$\frac{I_{OUT1}}{I_{IN}} = \frac{I_{OUT2}}{I_{IN}} = \frac{1}{as^2+bs+c} \quad (1)$$

where coefficients *a, b* and *c* are related to pole frequency ($\omega_0$) and the quality factor ($Q_O$) of filter through (2)

$$\omega_o = \sqrt{\frac{c}{a}}\ , Q_O = \frac{a}{b}\sqrt{\frac{c}{a}} \quad (2)$$

The transfer function for Fig. 2 is obtained as:

$$\frac{I^*_{OUT1}}{I_{IN}} = \frac{1}{as^2+bs+c+A} \quad (3)$$

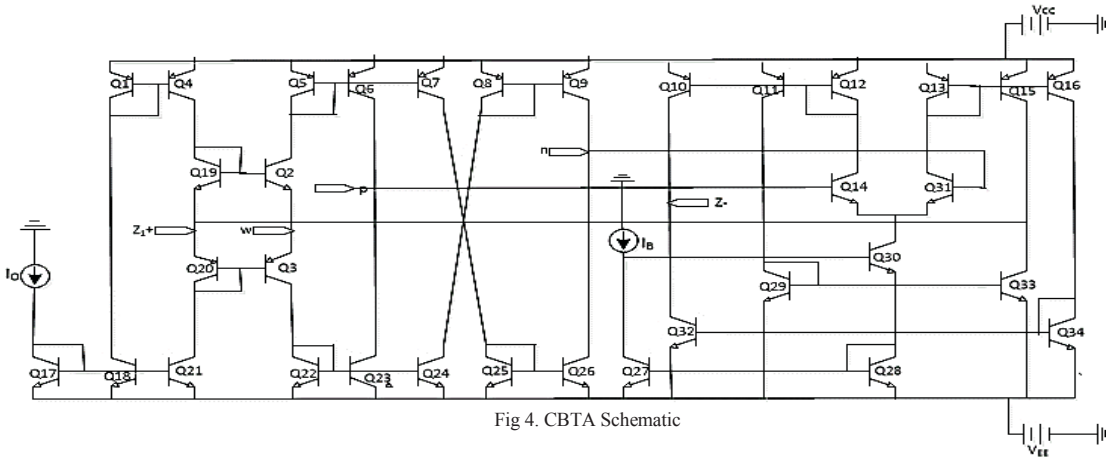



Fig 4. CBTA Schematic

which is a low pass filter with modified pole frequency ($\omega_{OLPCLP}$) and the Q factor ($Q_{LPCLP}$) as given in (4).

$$\omega_{0LPCLP} = \omega_o\sqrt{1+\frac{A}{c}},\ \mathrm{Q_{LPCLP}} = \mathrm{Q}\sqrt{1+\frac{\mathrm{A}}{\mathrm{c}}} \tag{4}$$

It is clear from (4) that both the parameters can be controlled by gain of amplifier. The subscript LPCLP is indicative of controlling of Low Pass response through Low Pass response.

Similarly, if $I_{OUT2}$ is considered high pass and $I_{OUT1}$ is low pass output in Fig 2, the values of modified pole frequency ($\omega_{OHPCLP}$) and the Q factor ($Q_{HPCLP}$) of the High Pass Controlled Low Pass FAF are obtained as:

$$\omega_{0\mathrm{HPCLP}} = \frac{\omega_o}{\sqrt{1+\frac{\mathrm{A}}{\mathrm{c}}}},\ \mathrm{Q_{HPCLP}} = \mathrm{Q}\sqrt{1+\frac{\mathrm{A}}{\mathrm{c}}} \tag{5}$$

Also, for the case when $I_{OUT1}$ is a high pass output and $I_{OUT2}$ is a low pass function in Fig 2, the modified pole frequency ($\omega_{OLPCLP}$) and the Q factor ($Q_{LPCHP}$) of the Low Pass Controlled High Pass FAF are expressed as:

$$\omega_{0LPCHP} = \omega_o\sqrt{1+\frac{A}{c}},\ Q_{LPCHP} = Q\sqrt{1+\frac{A}{c}} \tag{6}$$

Thus, it is clear that FAF concept helps in tuning of filter parameters through gain of external amplifier. Other combinations like low pass controlled band pass, band pass controlled low pass, band pass controlled band pass etc. are also possible. As the proposed structure addresses LPCLP, HPCLP and LPCHP therefore the discussion on FAF is restricted to the modification of filter parameters related to these filters only.

## III. Circuit Description

The section outlays the port relationship of the CBTA block along with its schematic diagram. A CBTA based current mode second order filter is described next around which the proposed Agile filters are built. A CBTA based current amplifier is also presented.

### A. Port Relationship

The CBTA circuit symbol [12]-[18] is depicted in Fig. 3. The terminals p, n, z+ and z- have high input impedance while terminal w has low input impedance.

Following are the terminal characteristics of CBTA

$$I_{z+} = g(V_P - V_N), I_{z-} = -g(V_P - V_N), I_P = I_W, I_N = -I_W, V_W = V_{Z+} \tag{7}$$

where $g$ is transconductance.

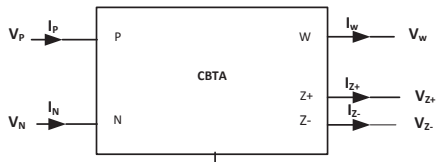



Fig 3. CBTA Block

### B. Schematic Diagram

The BJT based realization of CBTA is depicted in in Fig. 4 [12], [20]. It consists of a trans-conductance section and a current-controlled current conveyor. The trans-conductance (g) of the CBTA is controlled by the bias current ($I_{bias}$) and is given as $\mathrm{g} = \frac{\mathrm{I_{bias}}}{2\mathrm{V_T}}$, where $V_T$ is the thermal voltage and its value at room temperature is 26 mV

### C. Second Order Filter (Biquad)

The CBTA based biquad [21] (Fig. 5) is used for developing FAF and is presented in this section for ready reference. It employs two CBTA and two capacitors as the only passive components is given in Fig. 5 [21]. The expression for the transfer functions for low pass, band pass and high pass responses are given in (8) – (10) respectively.

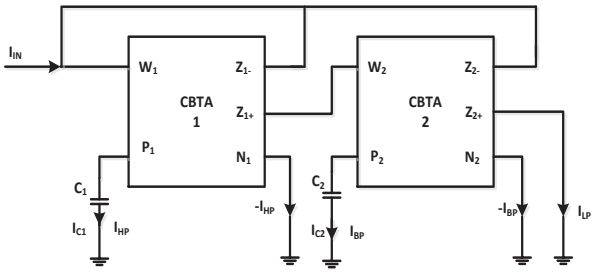



Fig 5. CBTA Based Biquad

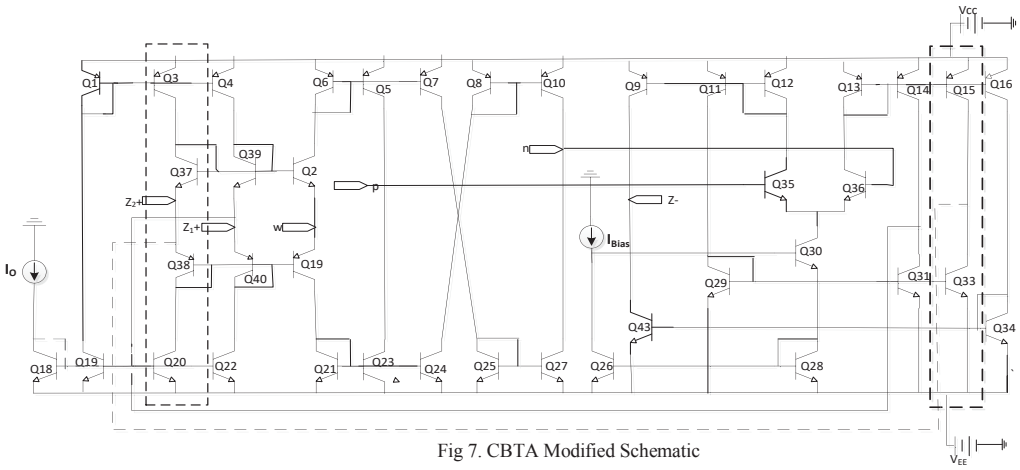



Fig 7. CBTA Modified Schematic

$$H_{LP} = \frac{1}{c_1c_2s^2+c_2s+1} \quad (8)$$

$$H_{BP} = \frac{c_2s}{c_1c_2s^2+c_2s+1} \quad (9)$$

$$H_{HP} = \frac{c_1c_2s^2}{c_1c_2s^2+c_2s+1} \quad (10)$$

The pole frequency ($\omega_o$) and quality factor ($Q_o$) of the biquad are computed as.

$$\omega_o = \frac{1}{\sqrt{c_1\,c_2}} = \frac{g}{\sqrt{C_1\,C_2}} \quad (11)$$

$$Q_o = \sqrt{\frac{c_1}{c_2}} = \sqrt{\frac{C_1}{C_2}} \quad (12)$$

respectively, where $c_1 = \frac{C_1}{g}, c_2 = \frac{C_2}{g}$.

### D. Proposed Agile Filters

For the realization of the FAF a current amplifier with tunable gain A is needed. Therefore, a CBTA based current amplifier is designed and is shown in Fig 6.

The gain of CBTA based amplifier (A) is given by:

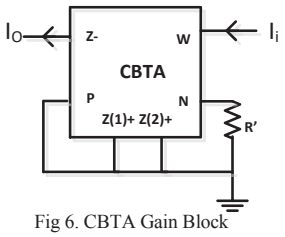



Fig 6. CBTA Gain Block

$$A = \frac{I_o}{I_i} = gR' \quad (13)$$

The Gain (A) can be adjusted either by varying $I_{bias}$ making it an electronically adjustable FAF or by varying the value of R

As the presented FAF is current mode, therefore additional current terminal namely a z+ terminal is needed for amplification purpose. Figure 7 shows the modified schematic wherein the dotted block shows additional transistors (Q3, Q37, Q38, Q20, Q15, Q33).

### 1. Low Pass Controlled Low Pass Filter

The schematic of LPCLP FAF is placed as Fig. 8. With additional z+ terminal the low pass responses are available through two terminals. One of these responses is applied to amplifier input and is fed back to the input while the other is used as the output of the FAF. The routine analysis of the circuit (Fig. 8) yields in the following transfer function ($TF_{LPCLP}$):

$$TF_{LPCLP} = \frac{1}{c_1c_2s^2+c_2s+(1+A)} \quad (14)$$

The value of $\omega_o$ and $Q_o$ for this filter given in Fig. 8 are termed as $\omega_{OLPCLP}$ and $Q_{OLPCLP}$ respectively:

$$\omega_{OLPCLP} = \sqrt{\frac{1+A}{c_1c_2}} = \omega_o\sqrt{1+A} \quad (15)$$

$$Q_{OLPCLP} = \sqrt{\frac{c_1}{c_2}}\sqrt{1+A} = Q\sqrt{1+A} \quad (16)$$

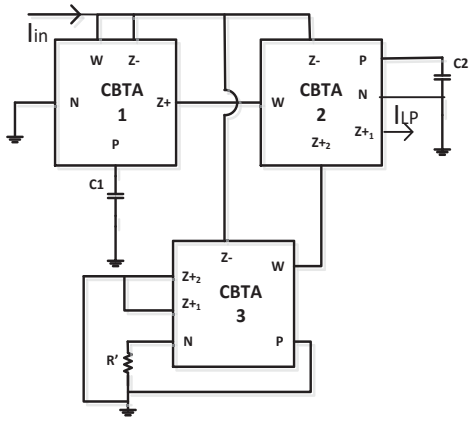



Fig 8. Low Pass Controlled Low Pass FAF

### 2. Low Pass Controlled High Pass Filter

The Low Pass controlled High Pass FAF is designed as given in Fig 9. The routine analysis of the circuit yields in the following transfer function ($TF_{LPCHP}$):

$$TF_{LPCHP} = \frac{c_1c_2s^2}{c_1c_2s^2+c_2s+(1+A)} \quad (17)$$

The value of $\omega_o$ and $Q_o$ for this filter given in Fig. 9 are termed as $\omega_{OLPHLP}$ and $Q_{OLPCHP}$ respectively:

$$\omega_{OLPCHP} = \sqrt{\frac{1+A}{c_1c_2}} = \omega_o\sqrt{1+A} \quad (18)$$

$$Q_{OLPCHP} = \sqrt{\frac{c_1}{c_2}}\sqrt{1+A} = Q\sqrt{1+A} \quad (19)$$

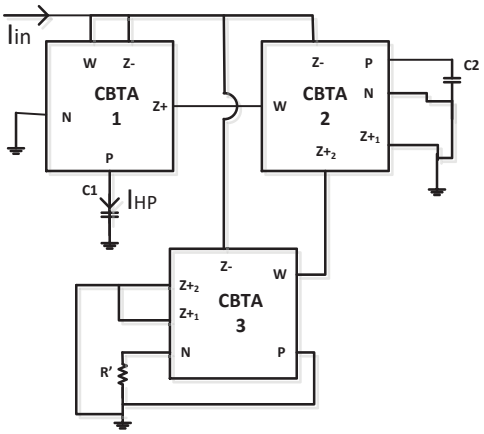



Fig 9. Low Pass Controlled High Pass FAF

### 3. *High Pass Controlled Low Pass Filter*

The High Pass controlled Low Pass FAF is designed as depicted in Fig 10. The routine analysis of the circuit yields in the following transfer function ($TF_{HPCLP}$):

$$TF_{HPCLP} = \frac{1}{c_1c_2(1+A)s^2+c_2s+1} \quad (20)$$

The value of $\omega_o$ and $Q_o$ for this filter given in Fig. 10 are termed as $\omega_{OHPCLP}$ and $Q_{OHPCLP}$ respectively:

$$\omega_{OHPCLP} = \sqrt{\frac{1}{c_1c_2\,(1+A)}} = \frac{\omega_o}{\sqrt{1+A}} \quad (21)$$

$$Q_{OHPCLP} = \sqrt{\frac{c_1(1+A)}{c_2}} = Q\sqrt{1+A} \quad (22)$$

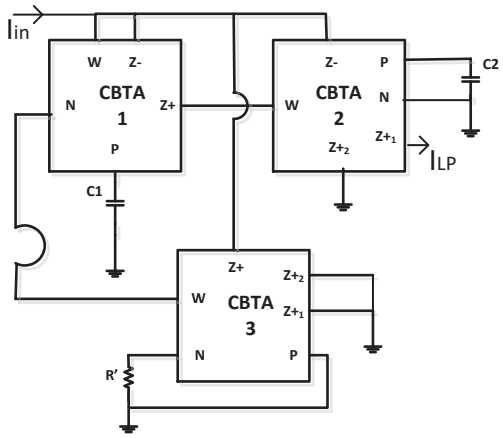


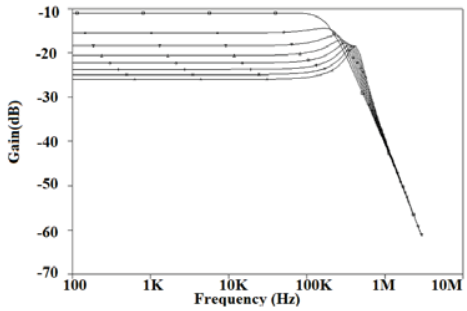

Fig 10. High Pass Controlled Low Pass FAF

## IV. FUNCTIONAL VERIFICATION

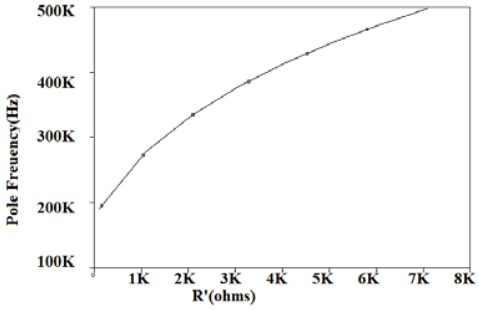
The second order filters in Figs. 8 – 10 are verified through SPICE simulations using PR100N (PNP) and NR100N (NPN) transistors model of the bipolar arrays ALA400 from AT&T21 [22]. The DC voltages of ± 2.5V is considered. For the simulation purpose bias currents ($I_o$, $I_{bias}$) of all CBTAs are taken to be 52µA which gives g of 1mS. The value of capacitances C1 and C2 are chosen as 1nF. The frequency response is studied by varying the resistance R' from 100 to 7.1K ohms in step of 1K ohm. This setting results in gain ranging from 0.1 to 7.1.

### A. *Simulation of Low Pass Controlled Low Pass Filter*

The simulated frequency response of proposed LPCLP FAF (Fig. 8) is placed in Fig 11 (a). The variation of pole frequency with respect to R' is plotted in Fig. 11(b). It may be observed that pole frequency of the proposed FAF shows an upward trend with increasing the value of the R'.




(a)

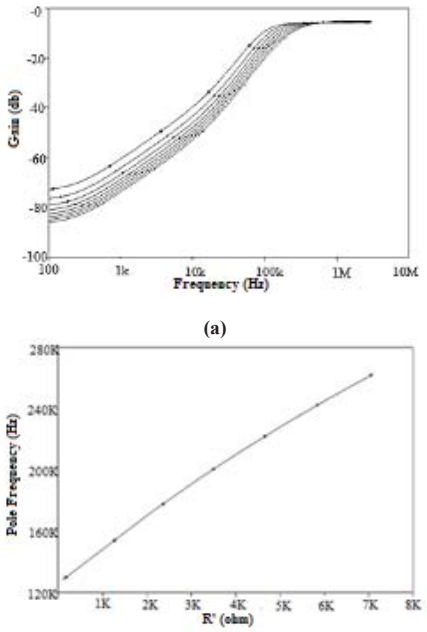



(b)

Fig. 11. Simulated outputs for proposed LPCLP FAF (a) magnitude response (b) Pole frequency variation

### B. *Simulation of Low Pass Controlled High Pass Filter*

The simulated frequency response of proposed LPCHP FAF (Fig. 9) is placed in Fig 12 (a). The variation of pole frequency with respect to R' is plotted in Fig. 12(b). It may be observed that pole frequency of the proposed FAF shows an upward trend with increasing the value of the R'



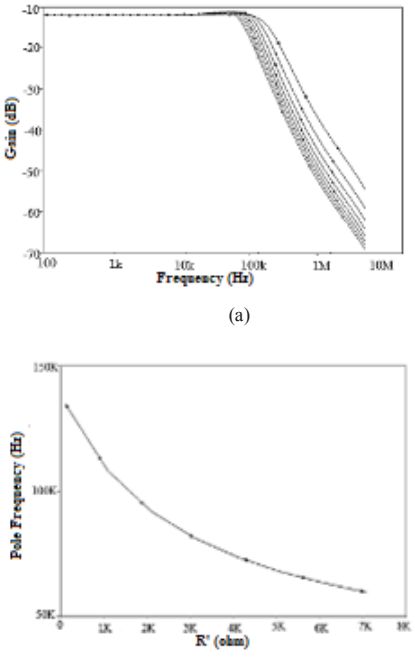

(a)

(b)

Fig. 12. Simulated outputs for proposed LPCLP FAF (a) magnitude response (b) Pole frequency variation

### C. Simulation of High Pass Controlled Low Pass Filter

The simulated frequency response of proposed HPCLP FAF (Fig. 10) is placed in Fig 13 (a). The variation of pole frequency with respect to R' is plotted in Fig. 13(b). It may be observed that pole frequency of the proposed FAF shows a downward trend with increasing the value of the R'




(a)

(b)

Fig. 13. Simulated outputs for proposed LPCLP FAF (a) magnitude response (b) Pole frequency variation

## V. CONCLUSION

A new concept of Current Backwards Trans-conductance Amplifier (CBTA) based Current Mode Frequency Agile Filter (FAF) is presented in this paper. Three configurations of FAFs (low pass controlled low pass, low pass controlled high pass and high pass controlled low pass) have been analyzed for their frequency response. The proposed concept is validated by plotting the 3db frequency against the external gain A of the FAF through PSPICE simulations using PR100N (PNP) and NR100N (NPN) transistors model of the bipolar arrays ALA400 from AT&T21 [22]. The simulation results are included to demonstrate the workability of the circuit.

## REFERENCES

[1] P. I. Mak, B. U. Seng-Pan and R. P. Martins, "Analog – baseband architecture and circuits for multistandard and low voltage wireless transceivers", Springer, 2007.

[2] Y. Lakys, A. Fabre, "Multistandard transceivers state of the art and a new versatile implementation for fully active frequency agile filters", Analog Integrated Circuits and Signal Processing, vol. 74, no. 1, pp. 63 - 78, 2013.

[3] H. H. Sigmarsson, J. Lee, D. Peroulis and W. J. Chappell, "Reconfigurable-order bandpass filter for frequency agile systems", International Microwave Symposium Digest, pp. 1756-1759, 2010.

[4] G. Subramanyam, F.W.V. Keuls and F. A. Miranda, "A k-band-frequency agile microstrip bandpass filter using a thin-film hts/ferroelectric/dielectric multilayer configuration", IEEE Transactions on Microwave Theory and Techniques, vol. 48, no. 4, pp. 525 – 530, 2000.

[5] H. Chandrahalim, S.A. Bhave, R.G. Polcawich, J.Pulskamp and R. Kaul, "A Pb (Zr0.55TiO.45) O3-transduced fully differential mechanically coupled frequency agile filter", IEEE Electron Device Letters, vol. 30, no. 12,2009.

[6] S. Kaehlert, D. Bormann, T.D. Werth, M.D. Wei, L. Liao and S. Heinen, "Design of frequency agile filters in RF frontend circuits", Radio and Wireless Symposium (RWS), pp. 13 - 16, 2012.

[7] A.J.X. Chen, Y. Wu, J. Hodiak and P.K.L. Yu, "Frequency agile digitally tunable microwave photonic filter", International Topical Meeting on Microwave Photonics, pp. 89-92,2003.

[8] Y. Lakys, B. Godara and A. Fabre, "Cognitive and encrypted communications, part 2: A new approach to active frequency-agile filters and validation results for an agile bandpass topology in SiBiCMOS", International Conference on Electrical and Electronics Engineering, pp. II 16-II 29,2009.

[9] Y. Lakys and A. Fabre, "A fully active frequency agile filter for multistandard transceivers", International Conference on Applied Electronics, pp. 1 - 7, 2011.

[10] M.W. Wyville, R.C. Smiley and J.S. Wight, "Frequency agile rf filter for interference attenuation", Radio and Wireless Symposium (RWS), pp. 399 - 402, 2012.

[11] D. Biolek, R. Senani, V. Biokova and Z. Kolka, "Active elements for analog signal processing; classification, review and new proposals", Radioengineering, vol. 17, no. 4, pp. 15 – 32, 2008.

[12] U.E. Ayten, M.Sagbas and H. Sedef, "Current mode leapfrog ladder filters using a new active block", AEU Int. J. Electron. Commun. 64 pp 503 – 511, 2010.

[13] M. Sagbas, U.E. Ayten and H. Sedef, "Current and voltage transfer function filter using as single active device", IET Circuits, Devices and Systems, vol. 4, no. 1, pp. 78-86, 2010.

[14] U.E. Ayten, M. Sagbas, N. Herencsar and J. Koton, "Novel floating general element simulators using CBTA", Radioengineering, vol. 21, pp 11-19, 2012.

[15] M. Sagbas, "Component reduced floating ±L, ±C and ±R simulators with grounded passive components", International Journal of Electronics and Communication (AEU), vol. 65, no. 10, pp. 794-798, 2011.

[16] M. Koksal, "Realisation of a general all-pole current transfer function by using CBTA", International Journal of Circuit Theory and Applications, vol. 41, no. 3, 2013.

[17] U.E. Ayten, M. Sagbas, N. Herencsar and J. Koton, "Novel floating FDNR, inductor and capacitor simulator using CBTA", In proceedings of the 2011 34th International Conference on Telecommunications And Signal Processing, Budapest (Hungary), 2011, pp. 312-316.

[18] U. E. Ayten, M. Sagbas, N. Herencsar and J. Koton, "A new compact CMOS realization of sinusoidal oscillator using a single modified CBTA", In Proceedings of the 21st International Conference Radio elektronika, Bmo (Czech Republic), 2011, pp. 41-44.

[19] U. E. Ayten, M. Sagbas and S. Minaei, "Realization of low-voltage modified CBTA and design of cascadable current-mode all-pass filter", Radioengineering, vol. 23, pp. 523-531, 2014.

[20] A. Fabre, O. Saaid, F. Wiest and C. Boucheron, "High frequency applications based on a new current controlled conveyor", IEEE Trans Circuits Syst. I, Fundam. Theory Appl, vol. 43, pp. 82 – 91, 1996.

[21] M. Sagbas and M. Koksal, "Realization of a general resistorless active biquad using CBTA", Jounal of Circuits, Systems and Computers, 2012, vol. 21, no. 1, 1250013.

[22] D.R. Frey, "Log domain filtering: An approach to current mode filtering", IEE Proc.-G: Circuits, Dev. Syst., vol.140, pp. 406 – 416, 1993.

Computational and Applied Mathematics (2019) 38:130
https://doi.org/10.1007/s40314-019-0866-9

# Stability of a delayed SIR epidemic model by introducing two explicit treatment classes along with nonlinear incidence rate and Holling type treatment

**Abhishek Kumar**[1] · **Nilam**[1]

Received: 8 August 2017 / Revised: 3 July 2018 / Accepted: 13 May 2019
© SBMAC - Sociedade Brasileira de Matemática Aplicada e Computacional 2019

**Abstract**
In this article, we analyze the stability of a time-delayed susceptible–infected–recovered (S–I–R) epidemic model by introducing two explicit treatment classes (or compartments) along with nonlinear incidence rate. The treatment classes are named as a pre-treated class ($T_1$) and post-treated class ($T_2$). The pre-treatment and post-treatment rates are being considered as Holling type I and Holling type III, respectively. Long-term qualitative analysis has been carried out after incorporating incubation time delay ($\tau$) into the incidence rate. The model analysis shows that the model has two equilibrium points, named as disease-free equilibrium (DFE) and endemic equilibrium (EE). The disease-free equilibrium is locally asymptotically stable when the basic reproduction number ($R_0$) is less than one and unstable when $R_0$ is greater than one for time lag $\tau \geq 0$, and when $R_0 = 1$ by Castillo-Chavez and Song theorem, the disease-free equilibrium changes its stability from stable to unstable and the model exhibits transcritical bifurcation. Furthermore, some conditions for stability of the endemic equilibrium are obtained. Finally, numerical simulations are presented to exemplify the analytical studies.

**Keywords** Epidemic · Delayed SIR model · Holling type treatment rates · Stability · Center manifold theory · Transcritical bifurcation

**Mathematics Subject Classification** 34D20 · 92B05 · 37M05

Communicated by Geraldo Diniz.

✉ Nilam
rathi.nilam@gmail.com

Abhishek Kumar
abhishek.dtu14@gmail.com

[1] Department of Applied Mathematics, Delhi Technological University, Delhi 110042, India

Published online: 28 May 2019

## 1 Introduction

Control of an epidemic in society is a major problem for governments, health agencies, and doctors. Every year, government spends the money to control the spread of the disease in the societies and through doctors and health agencies they provide the treatment to the individuals by various control strategies. In mathematical epidemiology literature, many epidemic models like SIS (Anderson and May 1991), SIR (Zhang and Suo 2010; Karim and Razali 2011; Zhou and Fan 2012; Adebimpe et al. 2015; Kumar and Nilam 2018a, b), SEIR (Dubey et al. 2013), SIRS (Xu and Ma 2009), SVEIR (Gumel et al. 2006) (where S, I, E, V, and R denote the population of susceptible, infected, exposed, vaccinated and recovered individuals, respectively), and many more have been proposed for the disease transmission and control. To contribute the mathematical modeling of the epidemic, in this article, we proposed a time-delayed SIR model by introducing two explicit treatment classes along with nonlinear incidence rate and Holling type treatment rates. We know that incidence rate and treatment rate play an important role in modeling dynamics and control of an epidemic. Several authors have suggested that the disease transmission process may have a nonlinear incidence rate. In many epidemic models (Gumel et al. 2006; Michael et al. 1999), the bilinear incidence rate $\beta SI$ and the standard incidence rate $\frac{\beta SI}{N}$ are frequently used. In the literature, the bilinear incidence rate is based on the law of mass action. Standard incidence rate may be a good approximation and it is practically feasible if the number of available partners is large enough and everybody could not make more contacts. In fact, the infection probability per contact is likely influenced by the number of infective individuals, because more infective individuals can increase the infection risk (Xu and Ma 2009). Capasso and Serio (1978) introduced a saturated incidence rate $g(I)S$ into epidemic models, where $g(I)$ tends to a saturation level when $I$ gets large, that is

$$g(I) = \frac{\beta I}{1+\alpha I},$$

where $\beta I$ measures the infection force of the disease and $\frac{1}{1+\alpha I}$ measures the inhibition effect from the behavioral change of the susceptible individuals when their number increases or from the crowding effect of the infective individuals. This incidence rate seems more reasonable than the bilinear incidence rate $\beta IS$, because it includes the behavioral change and crowding effect of the infective individuals and prevents the unboundedness of the contact rate by choosing suitable parameters (Xu and Ma 2009). Furthermore, the incidence rate in SIR model has been considered as saturated nonlinear by Adebimpe et al. (2015).

Proper and timely treatment methodology can substantially reduce the effect of disease on the society. In classical epidemic models, the treatment rate of infected individuals is assumed to be either constant or proportional to the number of the infected individuals. However, we know that there are limited treatment resources available in the community. Therefore, this is very important to choose a suitable treatment rate of a disease. In the absence of effective therapeutic treatments and vaccines, the epidemical control strategies are based on taking appropriate preventive measures. Wang and Ruan (2004) considered an SIR epidemic model with constant treatment rate (i.e., the recovery from infected subpopulation per unit time) as given by the following:

$$h(I) = \begin{cases} b, & I > 0 \\ 0, & I = 0 \end{cases},$$

where $b$ is a positive constant and $I$ is the number of infected individuals. They studied stability analysis and showed that this model exhibits various bifurcations. Furthermore, Zhou and Fan (2012) modified the treatment rate to Holling type II:

$$h(I) = \frac{aI}{1+bI}, \quad I \geq 0,\ a > 0,\ b > 0.$$

They have shown that, with varying amount of medical resources and their supply efficiency, the target model admits both backward and Hopf bifurcation. Dubey et al. (2013) changed it into Holling type III and type IV treatment rates.

In the present article, we have analyzed the effect of delay on SIR model by introducing two explicitly treatment classes along with nonlinear incidence rate and Holling type treatment rates for the better understanding of disease dynamics. Furthermore, we evaluate the basic reproduction number $R_0$, analyze the dynamical behavior of the model, and also discuss the stability of the model. The stability analysis of the model has been done by Descartes's rule (Wang 2004) along with Routh–Hurwitz criterion.

The rest of paper is organized as follows: After the abstract and introduction, the mathematical model is formulated in Sect. 2. Basic properties of solutions are discussed in Sect. 3. Equilibrium and stability analysis of the model are in Sect. 4. Numerical simulation and discussion are in Sect. 5. Finally, the conclusion is summarized in Sect. 6.

## 2 Mathematical model

In this section, we proposed a mathematical model for epidemic through compartments (or classes). For this, we consider the total population $N(t)$ at time $t$, with the immigration of susceptible individuals with a constant rate $A$. All susceptible population who are in effective contact of infected individuals does not become infected. Some of them enter into infected compartment, but some of them already vaccinated before showing the symptoms of the disease. On the other hand, individuals who have already entered into infected compartment will be given treatment according to the severity of the disease. Therefore, to take into account these pre-treated and post-treated individuals, two compartments have been incorporated in the SIR model namely pre-treated compartment and post-treated compartment, respectively. Hence, we divided the total population into five compartments: susceptible ($S$) compartment, pre-treated ($T_1$) compartment, infected ($I$) compartment, post-treated ($T_2$) compartment, and recovered ($R$) compartment. Susceptible are those who can get a disease under appropriate conditions, pre-treated are those who are vaccinated before the infection and can get a disease under appropriate conditions, infected are those who have been contracted with the disease and can spread the disease to susceptible and pre-treated individuals under appropriate conditions, post-treated are those who are infected and taking the medical treatments, and recovered are those who are immunized after the medical treatment and are free from the effect of disease. The pre-treatment and post-treatment to individuals are given by the Holling type I and Holling type III treatment rates, respectively. It has been considered that treatment before getting infected will be given in a linear form; therefore, Holling type I functional response has been considered for the treatment rate in the pre-treated class. The treatment given to infected individuals will be more than linearly increasing function because of the limitation to the treatment availability which is well explained by Holling type III functional response, and therefore, Holling type III treatment rate will be considered for the treatment of infected individuals. Incubation time delay has been incorporated as a time delay in force of infection and inhibition effect which form nonlinear incidence rate. This time delay ($\tau$)

will be a fixed time during which the infectious agents develop in the vector and it is only after that time that the infected vector can infect a susceptible individual. Let $\mu$ is the natural death rate of the population, $d$ is the death rate due to the disease, and $\theta$ is the recovery rate of post-treated individuals. The progression of the epidemic in different compartments is shown in the block diagram below (Fig. 1).

The rate of change in each compartment is given by the following system of nonlinear delay differential equations:

$$\frac{ds(t)}{dt} = A - (\mu + \delta)S(t) - \frac{\beta S(t)I(t-\tau)}{1+\alpha I(t-\tau)}, \tag{1}$$

$$\frac{dT_1(t)}{dt} = \delta S(t) - \mu T_1(t) - \gamma I(t)T_1(t), \tag{2}$$

$$\frac{dI(t)}{dt} = \frac{\beta S(t)I(t-\tau)}{1+\alpha I(t-\tau)} + \gamma I(t)T_1(t) - (\mu + d)I(t) - \frac{aI^2(t)}{1+bI^2(t)}, \tag{3}$$

$$\frac{dT_2(t)}{dt} = \frac{aI^2(t)}{1+bI^2(t)} - (\mu + \theta)T_2(t), \tag{4}$$

$$\frac{dR(t)}{dt} = \theta T_2(t) - \mu R(t), \tag{5}$$

where $\tau > 0$ is a fixed time during which the infectious agents develop in the vector and it is only after that time that the infected vector can infect a susceptible individual.

The initial conditions of the model are given by the following:

$$S(\theta) = \varphi_1(\theta), \quad T_1(\theta) = \varphi_2(\theta), \quad I(\theta) = \varphi_3(\theta), \quad T_2(\theta) = \varphi_4(\theta), \quad R(\theta) = \varphi_5(\theta),$$

$$\varphi_i(\theta) \geq 0, \theta \in [-\tau, 0], \quad \varphi_i(0) > 0 \ (i = 1, 2, 3, 4, 5), \tag{6}$$

where $(\varphi_1(\theta), \varphi_2(\theta), \varphi_3(\theta), \varphi_4(\theta), \varphi_5(\theta)) \in C([-\tau, 0], \mathbb{R}^5_+)$. Here, $C$ denotes the Banach space of continuous functions mapping the interval $[-\tau, 0]$ into $\mathbb{R}^5_+$.

In the model (1)–(5), the derivative $, \frac{d}{dt}$, represents the rate of change in corresponding compartment or class. The term $\delta S$ in Eqs. (1) and (2) represents the Holling type I treatment rate, where $\delta$ is named as pre-treatment rate and the term $\frac{aI^2}{1+bI^2}$ in the Eqs. (4)–(5) represents the Holling type III treatment rate, where $a$ and $b$ both are positive constants and named as treatment rate given to the infected individuals and limitation rate in treatment availability

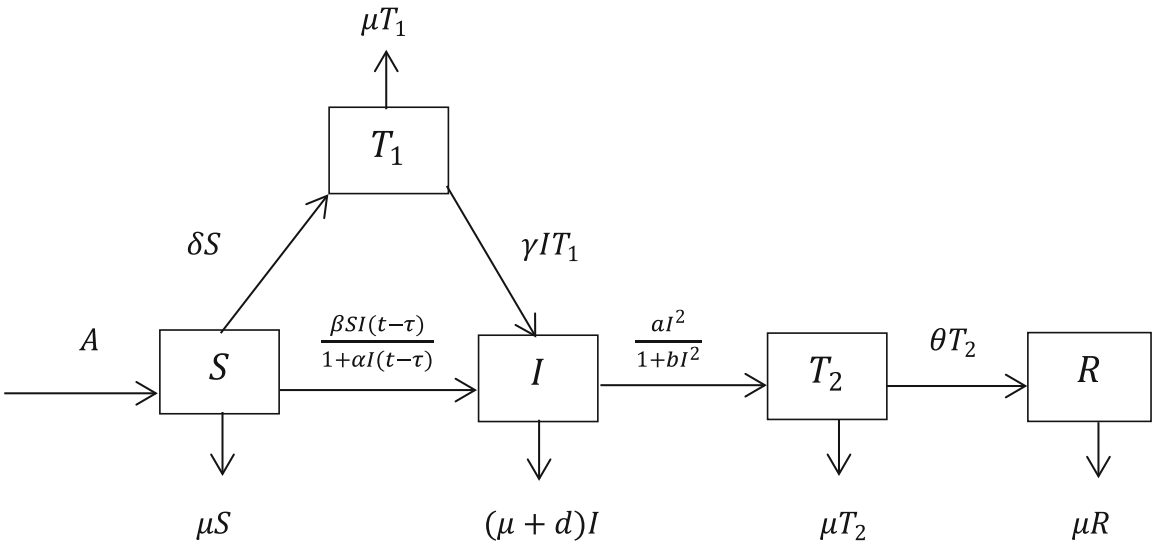



**Fig. 1** Progression of infection from susceptible ($S$) and pre-treated ($T_1$) compartments through infected ($I$), post-treated ($T_2$), and recovered ($R$) compartments for the model

(Dubey et al. 2016), respectively. The term $\frac{\beta S(t)I(t-\tau)}{1+\alpha I(t-\tau)}$ in the model (1)–(5) represents the nonlinear incidence rate, where $\beta$ is named as transmission rate of infection in susceptible individuals, and $\alpha$ is the measure of inhibition effect due to the infected individuals. The term $\gamma I(t)T_1(t)$ in the model represents the bilinear incidence rate of infection between pre-treated and infected individuals, where $\gamma$ is named as transmission rate of infection in pre-treated individuals. The incidence rate $\frac{\beta S(t)I(t-\tau)}{1+\alpha I(t-\tau)}$ represents the rate at time $t-\tau$ at which susceptible individuals leave the susceptible class and enter in the infectious class at time $t$. Furthermore, it is assumed that all other parameters of the model are positive which is also biologically true. The description of the model parameters given in Table 1 are as follows.

**Table 1** Parameters and description

| Parameter | Description |
|---|---|
| $A$ | Recruitment rate per day |
| $\alpha$ | Measure of inhibition effect |
| $\beta$ | Transmission rate of infection in susceptible individuals |
| $\mu$ | Natural mortality rate |
| $d$ | Disease-induced mortality rate |
| $\theta$ | Recovery rate of post-treated individuals |
| $\gamma$ | Transmission rate of infection in pre-treated individuals |
| $\delta$ | Pre-treatment rate |
| $a$ | Treatment rate of infected |
| $b$ | Limitation rate in treatment availability |

It should be mentioned that although the recovered population continues to make contacts with other individuals of the population, it does not contribute to the transmission dynamics of disease. Since the recovered population $R$ does not feature in the first four equations of the model, the rest paper will consider only Eqs. (1)–(4).

## 3 Basic properties of the model

The model (1)–(4) monitors the populations. It is assumed that all state variables of the model are positive, i.e., $(S, T_1, I, T_2) \in \mathbb{R}^4_+$. It follows that nonnegative cone $\mathbb{R}^4_+$ is invariant, as the disease-free plane ($I = T_2 = 0$).

It is convenient to define $N_0 = \frac{A}{\mu}$, $S_0 = \frac{A}{(\mu+\delta)}$ and $T_{10} = \frac{\delta A}{\mu(\mu+\delta)}$.

The rate of change of the total population, obtained by adding the all Eqs. (1)–(5), is given by the following:

$$\frac{\mathrm{d}N}{\mathrm{d}t} = A - \mu N - dI \leq A - \mu N, \tag{7}$$

with $N = S + T_1 + I + T_2 + R$.

Since as $t \to \infty$, disease will disappear; therefore, $\limsup_{t\to\infty} I(t) = 0$

$$\text{equation (7)} \Rightarrow \frac{\mathrm{d}N}{\mathrm{d}t} = A - \mu N,$$
$$\Rightarrow N(t) = \frac{A}{\mu} + \left(N(0) - \frac{A}{\mu}\right)e^{-\mu t},$$
$$\Rightarrow \lim_{t\to\infty} N(t) = \frac{A}{\mu} = N_0.$$

This follows that

$$0 < \limsup_{t\to\infty} N(t) \le N_0 \quad \text{if and only if} \quad \limsup_{t\to\infty} I(t) = 0.$$

From the first equation of the model, it follows that

$$0 < \limsup_{t\to\infty} S(t) \le S_0, \tag{8}$$

and then from Eq. (2),

$$0 < \limsup_{t\to\infty} T_1(t) \le T_{10}. \tag{9}$$

From Eq. (7), it follows that if $N > N_0$, then $\frac{dN}{dt} < 0$. This establishes the following lemma.

**Lemma 1** *The set*

$$D = \{(S, T_1, I, T_2) \in \mathbb{R}^4_+ : S + T_1 + I + T_2 \le N_0, S \le S_0, T_1 \le T_{10}\}$$

*is a positively invariant and attracting region for the disease transmission model given by the Eqs.* (1)–(4) *with the initial conditions in* $\mathbb{R}^4_+$.

Consequently, in the absence of disease ($I = 0$), the total population $N$ approaches the carrying capacity $N_0$ asymptotically, and in the presence of disease, the total population is less than or equals to $N_0$.

Thus, every solution of the system (1)–(4) with initial conditions in $\mathbb{R}^4_+$ tends towards $D$ as $t \to \infty$, and every solution with an initial condition in $D$ remains there for $t > 0$. Furthermore, in $D$, the usual existence, uniqueness, and continuation results hold for the system so that the system (1)–(4) is well posed mathematically and biologically.

## 4 Stability analysis

The equilibrium points of the system (1)–(4) are obtained by setting the right-hand side of the equations equal to zero. Let $Q^* = (S^*, T_1^*, I^*, T_2^*)$ be any equilibrium point of the model. Naturally, a stable equilibrium point $Q^*$ is the one for which all the nearby trajectories approach it asymptotically as $t \to \infty$. Hence, the equilibrium solutions of a system with time delays are the same as those of the corresponding system with zero delay (Tipsri and Chinviriyasit 2014).

### 4.1 Disease-free equilibrium (DFE)

The system (1)–(4) has a unique disease-free equilibrium of the form $Q = Q\left(\frac{A}{\mu+\delta}, \frac{\delta A}{\mu(\mu+\delta)}, 0, 0\right)$, obtained by setting the right-hand side of the equation of the system (1)–(4) equal to zero.

The characteristic equation at $Q$ of the system (1)–(4) is given by the following:

$$(\lambda+\mu)(\lambda+\mu+\theta)(\lambda+\mu+\delta)\left(\lambda+(\mu+d)\left(1-\frac{A(\gamma\delta+\beta\mu e^{-\lambda\tau})}{\mu(\mu+\delta)(\mu+d)}\right)\right) = 0. \tag{10}$$

At $\tau = 0$, the term $\frac{A(\gamma\delta+\beta\mu e^{-\lambda\tau})}{\mu(\mu+\delta)(\mu+d)}$ is known as a basic reproduction number $R_0$. Therefore, we define the basic reproduction number $R_0$ of our model by $\frac{A(\gamma\delta+\beta\mu)}{\mu(\mu+\delta)(\mu+d)}$. The basic reproduction number $R_0$ is defined as "the expected number of secondary cases produced, in a

completely susceptible population, by a typical infective individual" (van den Driessche and Watmough 2002).

### 4.1.1 Analysis for $R_0 \neq 1$

Clearly, Eq. (10) always has three negative roots: $\lambda_1 = -\mu$, $\lambda_2 = -(\mu+\theta)$, $\lambda_3 = -(\mu+\delta)$ and one other root is determined by the solution of equation:

$$\lambda + \mu + d - \frac{A(\gamma\delta + \beta\mu e^{-\lambda\tau})}{\mu(\mu+\delta)} = 0.$$

Let

$$f(\lambda) = \lambda + \mu + d - \frac{A(\gamma\delta + \beta\mu e^{-\lambda\tau})}{\mu(\mu+\delta)}.$$

If $R_0 > 1$, it is readily seen that, for $\lambda$ real:

$$f(0) = \mu + d - \frac{A(\gamma\delta + \beta\mu)}{\mu(\mu+\delta)} < 0, \quad \lim_{\lambda\to\infty} f(\lambda) = +\infty.$$

Hence, $f(\lambda) = 0$ has a positive real root if $R_0 > 1$.
If $R_0 < 1$, we assume that Re $\lambda \geq 0$.
We notice that

$$\text{Re}\,\lambda = \frac{A(\gamma\delta + \beta\mu e^{-\text{Re}\lambda\tau}\cos \text{Im}\,\lambda\tau)}{\mu(\mu+\delta)} - (\mu+d) < \frac{A(\gamma\delta+\beta\mu)}{\mu(\mu+\delta)} - (\mu+d) < 0,$$

a contradiction to our assumption. Hence, if $R_0 < 1$, then $\lambda$ is a root of Eq. (10) with negative real part.

Hence, we state the following theorem (Dubey et al. 2013):

**Theorem 1** *The disease-free equilibrium* $Q\left(\frac{A}{\mu+\delta}, \frac{\delta A}{\mu(\mu+\delta)}, 0, 0\right)$ *of the system* (1)–(4) *is locally asymptotically stable if* $R_0 < 1$ *and unstable if* $R_0 > 1$ *for* $\tau \geq 0$.

### 4.1.2 Analysis at $R_0 = 1$

This section is to analyze the behavior of the system (1)–(4) when basic reproduction number $R_0$ is equal to one. We notice that the Jacobian matrix of the system (1)–(4) evaluated at $R_0 = 1$ and bifurcation parameter $\beta = \beta^* = \left(\frac{\mu(\delta+\mu)(d+\mu)-\gamma\delta A}{A\mu}\right)$ has one of the eigenvalue as zero and the remaining eigenvalues are negative. The stability behavior of equilibrium points at $R_0 = 1$ cannot be determined using linearization, so we apply center manifold theory (Sastry 1999). To apply central manifold theorem to the system (1)–(4), we consider that $S = x_1$, $T_1 = x_2$, $I = x_3$ and $T_2 = x_4$; then, the system (1)–(4) can be rewritten as follows:

$$\frac{dx_1}{dt} = A - \delta x_1 - \mu x_1 - \frac{\beta x_1 x_3}{1+\alpha x_3}, \tag{11}$$

$$\frac{dx_2}{dt} = \delta x_1 - \mu x_2 - \gamma x_2 x_3, \tag{12}$$

$$\frac{dx_3}{dt} = \frac{\beta x_1 x_3}{1+\alpha x_3} + \gamma x_2 x_3 - (\mu+d)x_3 - \frac{a x_3^2}{1+b x_3^2}, \tag{13}$$

$$\frac{dx_4}{dt} = \frac{ax_3^2}{1+bx_3^2} - \theta x_4. \tag{14}$$

Let $J^*$ be the Jacobian matrix evaluated at $R_0 = 1$ and $\beta = \beta^*$. In addition, let $u = [u_1, u_2, u_3, u_4]$ and $w = [w_1, w_2, w_3, w_4]^T$ be the left eigenvector and right eigenvector of $J^*$ corresponding to the zero eigenvalue:

$$J^* = \begin{pmatrix} -\mu-\delta & 0 & -\beta^* A/(\mu+\delta) & 0 \\ \delta & -\mu & -\gamma\delta A/\mu(\mu+\delta) & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\mu-\theta \end{pmatrix}.$$

Then, we have the following:

$$u_1 = 0, u_2 = 0, u_3 = 1, u_4 = 0 \text{ and } w_1 = -\frac{\beta^* A}{(\delta+\mu)^2}, w_2 = -\frac{(\gamma\delta A\mu + \gamma A\delta^2 + \mu\delta A\beta^*)}{\mu^2(\delta+\mu)^2}, w_3 = 1, w_4 = 0.$$

The non-zero partial derivatives associated with the functions of the system (11)–(14) evaluated at $R_0 = 1$ and $\beta = \beta^*$ are as follows:

$$\left(\frac{\partial^2 f_3}{\partial x_1 \partial x_3}\right)_Q = \beta^*, \left(\frac{\partial^2 f_3}{\partial x_2 \partial x_3}\right)_Q = \gamma, \left(\frac{\partial^2 f_3}{\partial x_3 \partial x_1}\right)_Q = \beta^*, \left(\frac{\partial^2 f_3}{\partial x_3 \partial x_2}\right)_Q = \gamma, \left(\frac{\partial^2 f_3}{\partial x_3^2}\right)_Q = -2\left(\frac{A\alpha\beta^*}{\mu+\delta} + a\right) \text{ and } \left(\frac{\partial^2 f_3}{\partial x_3 \partial \beta^*}\right)_Q = \frac{A}{(\delta+\mu)}.$$

Then, from the Theorem 4.1 of Castillo-Chavez and Song (2004), the bifurcation constants $a_1$and$b_1$ are as follows:

$$\begin{aligned} a_1 &= \sum_{k,i,j=1}^{4} u_k w_i w_j \left(\frac{\partial^2 f_k}{\partial x_i \partial x_j}\right)_Q \\ &= u_3\left(w_1 w_3 \beta^* + w_2 w_3 \gamma + w_3 w_1 \beta^* + w_3 w_2 \gamma - w_3^2\left(\frac{A\alpha\beta^*}{\mu+\delta} + a\right)\right) \\ &= -\frac{2}{\mu+\delta}\left(A\beta^{*2} + \frac{\gamma(\mu\delta\beta^* A + \gamma A\delta\mu + \gamma\delta^2 A)}{\mu(\mu+\delta)} + \alpha A\beta^* + a(\mu+\delta)\right) < 0, \end{aligned}$$

$$\begin{aligned} b_1 &= \sum_{k,i=1}^{4} u_k w_i \left(\frac{\partial^2 f_k}{\partial x_i \partial \beta^*}\right)_Q \\ &= w_3\left(\frac{A}{(\delta+\mu)}\right) \\ &= \frac{A}{(\delta+\mu)} > 0. \end{aligned}$$

Thus, from Theorem 4.1(iv) (Castillo-Chavez and Song 2004), we state the following Theorem:

**Theorem 2** *The disease-free equilibrium changes its stability from stable to unstable at $R_0 = 1$ and there exists a positive equilibrium as $R_0$ crosses one. Hence, the system (11)–(14) exhibits transcritical bifurcation with bifurcation parameter $\beta^*$ at $R_0 = 1$.*

### 4.2 Endemic equilibrium

To find conditions for the existence of an equilibrium $Q^* = (S^*, T_1^*, I^*, T_2^*)$ for which the disease is endemic in the population, the systems (1)–(4) are rearranged to get $S^*, T_1^*, I^*$and$T_2^*$ which give the following:

(i) $S^* = \frac{A(1+\alpha I^*)}{(\delta+\mu)(1+\alpha I^*)+\beta I^*}$,
(ii) $T_1^* = \frac{A\delta(1+\alpha I^*)}{(\mu+\gamma I^*)((\delta+\mu)(1+\alpha I^*)+\beta I^*)}$,
(iii) $T_2^* = \frac{aI^{*2}}{(\mu+\theta)(1+bI^{*2})}$,
(iv) $I^*$.

where $I^*$ is given by the following equation:

$$C_1 I^{*4} + C_2 I^{*3} + C_3 I^{*2} + C_4 I^* + C_5 = 0, \tag{15}$$

where

$$C_1 = \{b\gamma(\mu+d)(\beta+\alpha(\delta+\mu))\},$$

$$C_2 = \{b(\mu+d)(\gamma(\mu+\delta)+\mu(\beta+\alpha(\delta+\mu))) + a\gamma\beta + a\gamma\alpha(\delta+\mu) - Ab(\beta\gamma+\gamma\delta\alpha)\},$$

$$C_3 = \{\mu b(\mu+d)(\mu+\delta) + a\mu(\beta+\alpha(\delta+\mu)) + \gamma^2(\mu+d)(\beta+\alpha(\delta+\mu)) - Ab(\beta\mu+\gamma\delta)\},$$

$$C_4 = \{a\mu(\mu+\delta) + (\mu+d)(\gamma(\delta+\mu)+\mu(\beta+\alpha(\delta+\mu))) - A(\beta\gamma+\gamma\delta\alpha)\},$$

$$C_5 = \{\mu(\mu+d)(\mu+\delta) - A(\beta\mu+\gamma\delta)\}.$$

Now, using Descartes' rule of sign (Wang 2004), the biquadratic Eq. (15) has unique positive real root $I^*$ if any one of the following conditions hold:

(i) $C_1 > 0, C_2 < 0, C_3 < 0, C_4 < 0$ and $C_5 < 0$.
(ii) $C_1 > 0, C_2 > 0, C_3 < 0, C_4 < 0$ and $C_5 < 0$.
(iii) $C_1 > 0, C_2 > 0, C_3 > 0, C_4 < 0$ and $C_5 < 0$.
(iv) $C_1 > 0, C_2 > 0, C_3 > 0, C_4 > 0$ and $C_5 < 0$..

After finding the value of $I^*$, we can find the value of $S^*, T_1^*$, and $T_2^*$. This implies that there exists a unique positive endemic equilibrium $Q^*(S^*, T_1^*, I^*, T_2^*)$ if one of the above conditions holds true.

The local stability of $Q^*$ is explored as follows: The characteristic equation of the system (1)–(4) evaluated at endemic equilibrium point $Q^*$ is given by the following:

$$(\lambda+\mu+\theta)(\lambda^3 + p_0\lambda^2 + q_0\lambda + r_0 - (p_1\lambda^2 + q_1\lambda + r_1)e^{-\lambda\tau}) = 0, \tag{16}$$

where

$$p_0 = \left(3\mu + \delta + \frac{\beta I^*}{1+\alpha I^*} + \gamma I^* + d + \frac{2aI^*}{(1+bI^{*2})^2}\right),$$

$$q_0 = \left(\left(2\mu + \gamma I^* + d + \frac{2aI^*}{(1+bI^{*2})^2}\right)\left(\mu + \delta + \frac{\beta I^*}{1+\alpha I^*}\right)\right.$$

$$\left.+(\mu+\gamma I^*)\left(\mu + d + \frac{2aI^*}{(1+bI^{*2})^2}\right) + \gamma^2 T_1^* I^*\right),$$

$$r_0 = \left(\mu + \delta + \frac{\beta I^*}{1+\alpha I^*}\right)\left((\mu+\gamma I^*)\left(\mu + d + \frac{2aI^*}{(1+bI^{*2})^2}\right) + \gamma^2 T_1^* I^*\right),$$

$$p_1 = \frac{\beta S^*}{1+\alpha I^*},$$

$$q_1 = (\delta + \gamma I^*)\frac{\beta S^*}{1+\alpha I^*},$$
$$r_1 = (\mu^2 + \mu\gamma I^* + \delta\mu)\frac{\beta S^*}{1+\alpha I^*}.$$

Clearly, $\lambda_1 = -(\mu+\theta)$ is a root of the characteristic Eq. (16) for $\tau \geq 0$; others' roots $\lambda_2, \lambda_3 \text{and} \lambda_4$ can be obtained by the following equation:

$$\lambda^3 + p_0\lambda^2 + q_0\lambda + r_0 - (p_1\lambda^2 + q_1\lambda + r_1)\mathrm{e}^{-\lambda\tau} = 0. \tag{17}$$

**Theorem 3** *For $\tau = 0$, $Q^*$ is locally asymptotically stable if all three $\frac{S^*}{I^*} \leq 1$, $\frac{S^*}{I^*} \leq \frac{(2\mu+d)}{(\delta+\gamma I^*)}$ and $\frac{S^*}{I^*} \leq \frac{(\mu+\gamma I^*)(\mu+d)}{(\mu^2+\mu\gamma I^*+\delta\mu)}$ are satisfied simultaneously.*

***Proof*** At $Q^*$, the roots of the system (1)–(4) are $\lambda_1 = -(\mu+\theta)$, and other roots $\lambda_2, \lambda_3 \text{and} \lambda_4$ of the system are given by the equation $\lambda^3 + p_0\lambda^2 + q_0\lambda + r_0 - (p_1\lambda^2 + q_1\lambda + r_1) = 0$ for $\tau = 0$.

It is easy to show that if $\frac{S^*}{I^*} \leq 1$, $\frac{S^*}{I^*} \leq \frac{(2\mu+d)}{(\delta+\gamma I^*)}$ and $\frac{S^*}{I^*} \leq \frac{(\mu+\gamma I^*)(\mu+d)}{(\mu^2+\mu\gamma I^*+\delta\mu)}$ are satisfied simultaneously, then

$$\begin{aligned}
p_0 - p_1 &= 3\mu + \delta + \frac{\beta I^*}{1+\alpha I^*} + \gamma I^* + d + \frac{2aI^*}{(1+bI^{*2})^2} - \frac{\beta S^*}{1+\alpha I^*} \\
&= 3\mu + \delta + \gamma I^* + d + \frac{2aI^*}{(1+bI^{*2})^2} + \frac{\beta(I^*-S^*)}{1+\alpha I^*} > 0, \\
q_0 - q_1 &= \left(2\mu + \gamma I^* + d + \frac{2aI^*}{(1+bI^{*2})^2}\right)\left(\mu + \delta + \frac{\beta I^*}{1+\alpha I^*}\right) \\
&\quad + (\mu+\gamma I^*)\left(\mu + d + \frac{2aI^*}{(1+bI^{*2})^2}\right) + \gamma^2 T_1^* I^* - (\delta+\gamma I^*)\frac{\beta S^*}{1+\alpha I^*} \\
&= (\mu+\gamma I^*)\left(\mu + d + \frac{2aI^*}{(1+bI^{*2})^2}\right) + \gamma^2 T_1^* I^* + \left(2\mu + \gamma I^* + d + \frac{2aI^*}{(1+bI^{*2})^2}\right) \\
&\quad \times (\mu+\delta) + \left(\frac{2aI^*}{(1+bI^{*2})^2} + \gamma I^*\right) \\
&\quad \times \left(\frac{\beta I^*}{1+\alpha I^*}\right) + \frac{\beta}{1+\alpha I^*}((2\mu+d)I^* - (\delta+\gamma I^*)S^*) > 0, \\
r_0 - r_1 &= \left(\mu + \delta + \frac{\beta I^*}{1+\alpha I^*}\right)\left((\mu+\gamma I^*)(\mu+d) + \frac{2aI^*(\mu+\gamma I^*)}{(1+bI^{*2})^2} + \gamma^2 T_1^* I^*\right) \\
&\quad - (\mu^2 + \mu\gamma I^* + \delta\mu)\frac{\beta S^*}{1+\alpha I^*} \\
&= \left((\mu+\gamma I^*)(\mu+d) + \frac{2aI^*(\mu+\gamma I^*)}{(1+bI^{*2})^2} + \gamma^2 T_1^* I^*\right)(\mu+\delta) \\
&\quad + \left(\frac{2aI^*(\mu+\gamma I^*)}{(1+bI^{*2})^2} + \gamma^2 T_1^* I^*\right)\frac{\beta I^*}{1+\alpha I^*} + \frac{\beta}{1+\alpha I^*}((\mu+\gamma I^*)(\mu+d)I^* \\
&\quad - (\mu^2 + \mu\gamma I^* + \delta\mu)S^*) > 0.
\end{aligned}$$

Hence, by the definition of Descartes rule of sign and Routh–Hurwitz criterion, the endemic equilibrium $Q^*$ is locally asymptotically stable when $\tau = 0$.

**Theorem 4** *For $\tau > 0$, $Q^*$ is locally asymptotically stable if all three $L_1 > L_2, L_3 > L_4$ and $\frac{S^*}{I^*} \leq \frac{(\mu+\gamma I^*)(\mu+d)}{(\mu^2+\mu\gamma I^*+\delta\mu)}$ are satisfied simultaneously,*
*where*

$$L_1 = \left(3\mu + \delta + \gamma I^* + d + \frac{2aI^*}{(1+bI^{*2})^2} + \frac{\beta(I^*+S^*)}{1+\alpha I^*}\right) \times \left(3\mu + \delta + \gamma I^* + d + \frac{2aI^*}{(1+bI^{*2})^2} + \frac{\beta(I^*-S^*)}{1+\alpha I^*}\right),$$

$$L_2 = 2\left(\left(2\mu + \gamma I^* + d + \frac{2aI^*}{(1+bI^{*2})^2}\right)\left(\mu + \delta + \frac{\beta I^*}{1+\alpha I^*}\right) + (\mu+\gamma I^*)\left(\mu + d + \frac{2aI^*}{(1+bI^{*2})^2}\right) + \gamma^2 T_1^* I^*\right),$$

$$L_3 = \left(2\left((\mu^2+\mu\gamma I^*+\delta\mu)\frac{\beta S^*}{1+\alpha I^*}\right)\left(\frac{\beta S^*}{1+\alpha I^*}\right) + \left(\left(2\mu + \gamma I^* + d + \frac{2aI^*}{(1+bI^{*2})^2}\right) \times \left(\mu + \delta + \frac{\beta I^*}{1+\alpha I^*}\right) + (\mu+\gamma I^*)\left(\mu + d + \frac{2aI^*}{(1+bI^{*2})^2}\right) + \gamma^2 T_1^* I^*\right)^2\right),$$

$$L_4 = \left(2\left(\mu + \delta + \frac{\beta I^*}{1+\alpha I^*}\right)\left((\mu+\gamma I^*)\left(\mu + d + \frac{2aI^*}{(1+bI^{*2})^2}\right) + \gamma^2 T_1^* I^*\right) \times \left(3\mu + \delta + \frac{\beta I^*}{1+\alpha I^*} + \gamma I^* + d + \frac{2aI^*}{(1+bI^{*2})^2}\right) + \left((\delta+\gamma I^*)\frac{\beta S^*}{1+\alpha I^*}\right)^2\right).$$

***Proof*** At $Q^*$, the roots of the Eq. (16) are $\lambda_1 = -(\mu+\theta)$, and other roots $\lambda_2, \lambda_3$and$\lambda_4$ are given by the equation $\lambda^3 + p_0\lambda^2 + q_0\lambda + r_0 - (p_1\lambda^2 + q_1\lambda + r_1)e^{-\lambda\tau} = 0$ for $\tau > 0$.

Assume that $\lambda = i\omega, \omega > 0$ is the root of the characteristic Eq. (17).
Substituting $\lambda = i\omega$ in Eq. (17), we get the following:

$$(r_0 - p_0\omega^2) - i\omega(\omega^2 - q_0) = (r_1 - p_1\omega^2)\cos\omega\tau + q_1\omega\sin\omega\tau - i((r_1 - p_1\omega^2)\sin\omega\tau - q_1\omega\cos\omega\tau). \tag{18}$$

On separating real and imaginary parts of Eq. (18):

$$(r_0 - p_0\omega^2) = (r_1 - p_1\omega^2)\cos\omega\tau + q_1\omega\sin\omega\tau \tag{19}$$

$$\omega(\omega^2 - q_0) = ((r_1 - p_1\omega^2)\sin\omega\tau - q_1\omega\cos\omega\tau). \tag{20}$$

Squaring and adding the both side of Eqs. (19) and (20) yield the following:

$$(r_0 - p_0\omega^2)^2 + \omega^2(\omega^2 - q_0)^2 = (r_1 - p_1\omega^2)^2 + q_1^2\omega^2. \tag{21}$$

Letting $\omega^2 = z_1$; Eq. (21) becomes the following:

$$z_1^3 + (p_0^2 - 2q_0 - p_1^2)z_1^2 + (q_0^2 + 2r_1p_1 - 2r_0p_0 - q_1^2)z_1 + (r_0^2 - r_1^2) = 0. \tag{22}$$

It is easy to show that if $L_1 > L_2, L_3 > L_4$ and $\frac{S^*}{I^*} \leq \frac{(\mu+\gamma I^*)(\mu+d)}{(\mu^2+\mu\gamma I^*+\delta\mu)}$ are satisfied simultaneously, then

$$p_0^2 - 2q_0 - p_1^2 = (p_0+p_1)(p_0-p_1) - 2q_0 = \left(3\mu + \delta + \gamma I^* + d + \frac{2aI^*}{(1+bI^{*2})^2} + \frac{\beta(I^*+S^*)}{1+\alpha I^*}\right)\left(3\mu + \delta + \gamma I^* + d + \frac{2aI^*}{(1+bI^{*2})^2} + \frac{\beta(I^*-S^*)}{1+\alpha I^*}\right)$$

$$-\left(\left(2\mu+\gamma I^*+d+\frac{2aI^*}{(1+bI^{*2})^2}\right)\left(\mu+\delta+\frac{\beta I^*}{1+\alpha I^*}\right)+(\mu+\gamma I^*)\left(\mu+d+\frac{2aI^*}{(1+bI^{*2})^2}\right)+\gamma^2 T_1^* I^*\right)$$

$$-2\left(\left(2\mu+\gamma I^*+d+\frac{2aI^*}{(1+bI^{*2})^2}\right)\left(\mu+\delta+\frac{\beta I^*}{1+\alpha I^*}\right)+(\mu+\gamma I^*)\left(\mu+d+\frac{2aI^*}{(1+bI^{*2})^2}\right)+\gamma^2 T_1^* I^*\right)$$

$$= L_1 - L_2 > 0,$$

$$2r_1 p_1 + q_0^2 - 2r_0 p_0 - q_1^2 = 2\left((\mu^2+\mu\gamma I^*+\delta\mu)\frac{\beta S^*}{1+\alpha I^*}\right)\left(\frac{\beta S^*}{1+\alpha I^*}\right)$$

$$+\left(\left(2\mu+\gamma I^*+d+\frac{2aI^*}{(1+bI^{*2})^2}\right)\left(\mu+\delta+\frac{\beta I^*}{1+\alpha I^*}\right)+(\mu+\gamma I^*)\left(\mu+d+\frac{2aI^*}{(1+bI^{*2})^2}\right)+\gamma^2 T_1^* I^*\right)^2$$

$$-\left(2\left(\mu+\delta+\frac{\beta I^*}{1+\alpha I^*}\right)\left((\mu+\gamma I^*)\left(\mu+d+\frac{2aI^*}{(1+bI^{*2})^2}\right)+\gamma^2 T_1^* I^*\right)\right.$$

$$\left.\left(3\mu+\delta+\frac{\beta I^*}{1+\alpha I^*}+\gamma I^*+d+\frac{2aI^*}{(1+bI^{*2})^2}\right)+\left((\delta+\gamma I^*)\frac{\beta S^*}{1+\alpha I^*}\right)^2\right) = L_3 - L_4 > 0,$$

$$r_0^2 - r_1^2 = (r_0+r_1)(r_0-r_1) = \left(\left((\mu+\gamma I^*)(\mu+d)+\frac{2aI^*(\mu+\gamma I^*)}{(1+bI^{*2})^2}+\gamma^2 T_1^* I^*\right)(\mu+\delta)\right.$$

$$+\left(\frac{2aI^*(\mu+\gamma I^*)}{(1+bI^{*2})^2}+\gamma^2 T_1^* I^*\right)\frac{\beta I^*}{1+\alpha I^*}+\frac{\beta}{1+\alpha I^*}\left(((\mu+\gamma I^*)(\mu+d)I^*+(\mu^2+\mu\gamma I^*+\delta\mu)S^*)\right)\Big)$$

$$\left(\left((\mu+\gamma I^*)(\mu+d)+\frac{2aI^*(\mu+\gamma I^*)}{(1+bI^{*2})^2}+\gamma^2 T_1^* I^*\right)(\mu+\delta)+\left(\frac{2aI^*(\mu+\gamma I^*)}{(1+bI^{*2})^2}+\gamma^2 T_1^* I^*\right)\frac{\beta I^*}{1+\alpha I^*}\right.$$

$$\left.+\frac{\beta}{1+\alpha I^*}((\mu+\gamma I^*)(\mu+d)I^*-(\mu^2+\mu\gamma I^*+\delta\mu)S^*)\right) > 0.$$

Hence, by the definition of Descartes rule of sign (Wang 2004) and Routh–Hurwitz criterion, the endemic equilibrium $Q^*$ is locally asymptotically stable for $\tau > 0$.

## 5 Numerical simulation and discussion

In this section, we perform the computer simulation using MATLAB 2012b to illustrate the results. We choose the following set of tested values of parameters from literature (Dubey et al. 2015, 2016) (Table 2).

With initial values:

$$S(0) = 400.0, \quad T_1(0) = 10.0, \quad I(0) = 30.0, \quad T_2(0) = 2.0.$$

**Table 2** List of parameters

| Parameter | Value | Unit |
|---|---|---|
| $A$ | 9 | person $(\text{day})^{-1}$ |
| $\alpha$ | 0.05 | $(\text{person})^{-1}$ |
| $\beta$ | 0.003 | $(\text{person})^{-1}$ $(\text{day})^{-1}$ |
| $\mu$ | 0.02 | $(\text{day})^{-1}$ |
| $d$ | 0.05 | $(\text{day})^{-1}$ |
| $\gamma$ | 0.001 | $(\text{person})^{-1}$ $(\text{day})^{-1}$ |
| $\delta$ | 0.01 | $(\text{day})^{-1}$ |
| $a$ | 0.2 | $(\text{day})^{-1}$ |
| $b$ | 0.02 | $(\text{person})^{-1}$ |
| $\theta$ | 0.002 | $(\text{day})^{-1}$ |

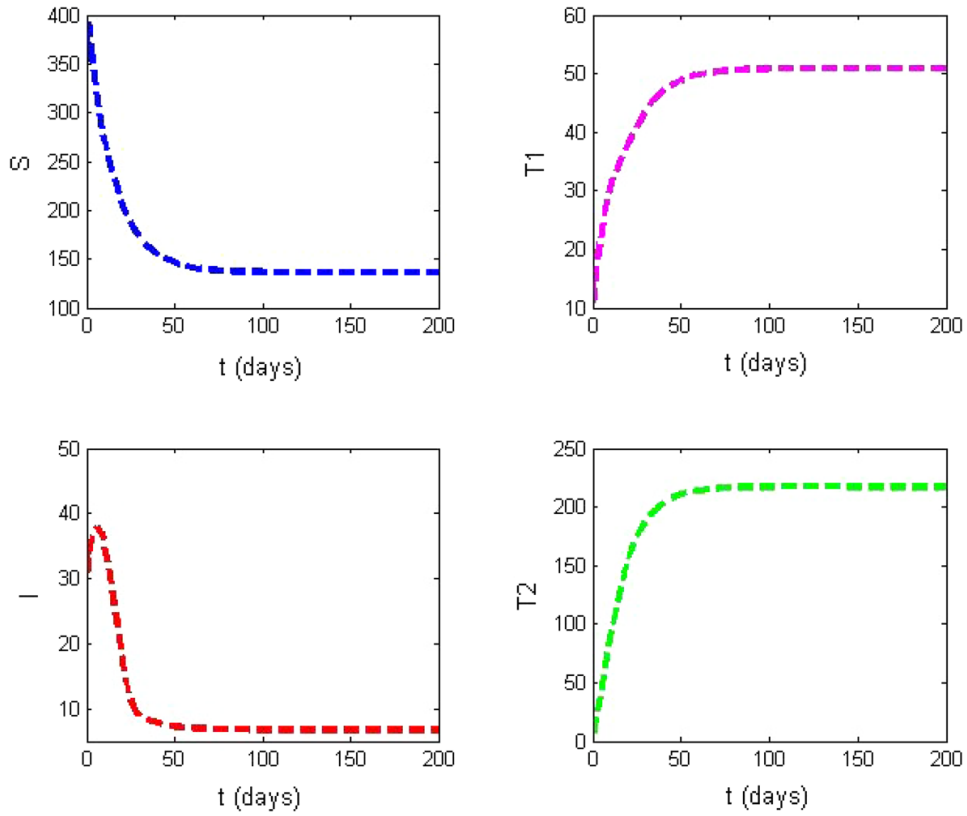



**Fig. 2** Population at time lag $\tau = 1$ day

Figures 2 and 3 show the population in various compartments at two different values of incubation time delay $\tau = 1, 2$ days approaches to endemic equilibrium $Q^*(136.36, 50.94, 6.77, 217.32)$. The number of infected (I) individuals is initially increasing, and as time passes, they are decreasing due to treatment. The susceptible (S) population decreases to attain a steady state. In addition, pre-treated ($T_1$) and post-treated ($T_2$) populations increase to attain a steady state for both the time delays, as shown in Figs. 2 and 3. The increment in the pre-treated population is because of the vaccination of susceptible population as a result of which susceptible population decreases. The post-treated population increases, because the treatment is given to infected population, and hence, infected population decreases.

Figure 4 shows the effect of delay on infected population. Clearly, as the value of $\tau$ increases, the number of infected individuals also increases in comparison to decreasing values of time lag $\tau$. Delay clearly indicates that as much time we take to initiate the preventive measures, the more number of individuals will be infected.

Figure 5 shows the effect of delay on pre-treated population. Clearly, as the value of $\tau$ increases, the pre- treated individual decreases. This decrease may be due to time taken by health agencies to provide the precautionary treatment to the susceptible individuals. It can be understood that time will be consumed to provide vaccination or any other precautionary treatment to susceptible will result the increment of infected population.

The effect of delay on post-treated population is shown in Fig. 6. Clearly, as $\tau$ increases, the number of post-treated individuals increases. Delay clearly indicates that as much time we take to initiate the post-treatment to infected, we have to provide the treatment to more infected individuals. The graphs in Fig. 6 are closer than the graphs in

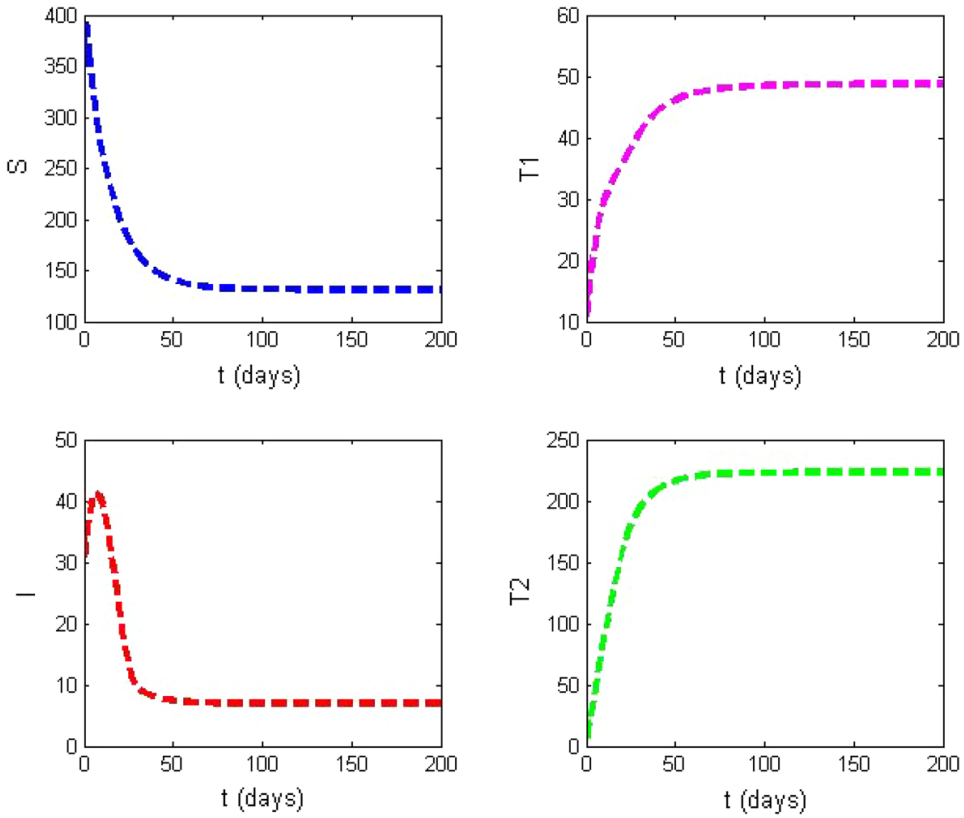



**Fig. 3** Population at time lag $\tau = 2$ days

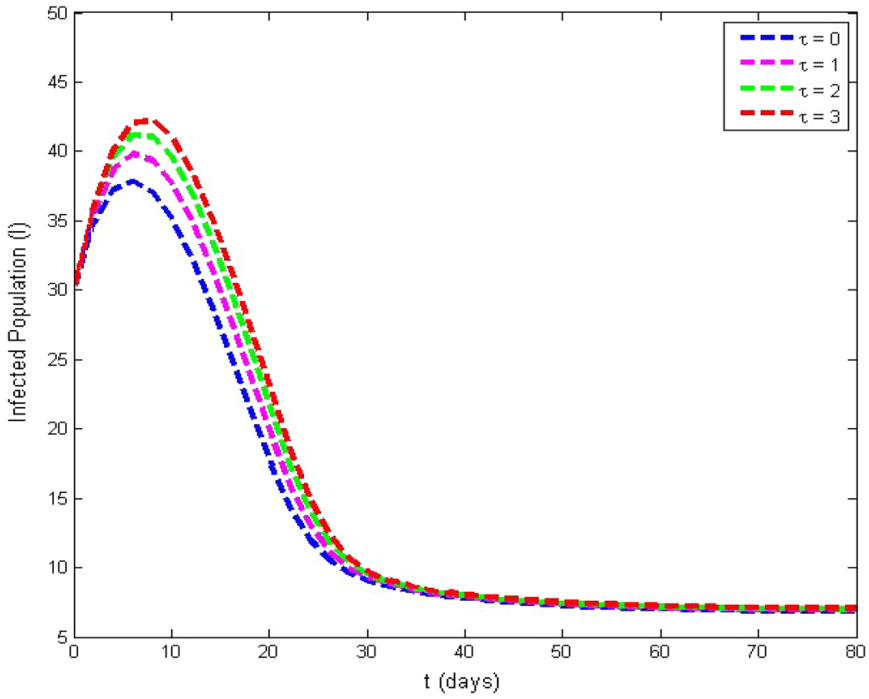



**Fig. 4** Infected population for various values of time lag $\tau$

Fig. 5 because of different type of treatment rate in $T_1$ and $T_2$ compartments. The treatment rate is linear (Holling type I) in $T_1$, while it is nonlinear, i.e., Holling type III in $T_2$.

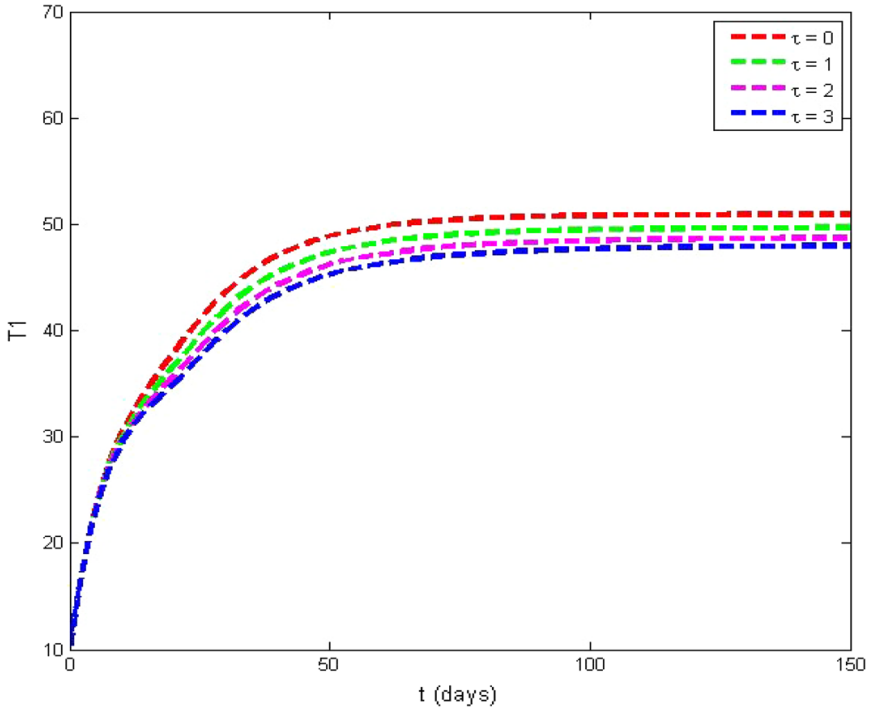



**Fig. 5** Pre-treated population at various values of time lag $\tau$

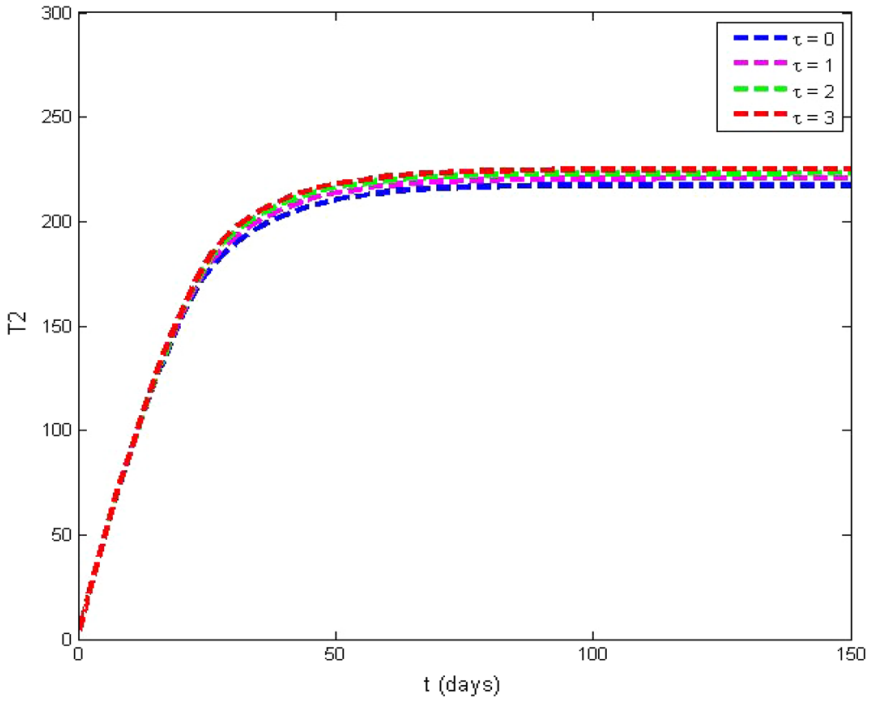



**Fig. 6** Post-treated population at various values of time lag $\tau$

Figures 7 and 8 show the infected population at various values of $\beta$ and $\alpha$, respectively. The graphs show the decrease in infected population with a decline in transmission rate $\beta$ and an enhancement of inhibition rate $\alpha$.

Figure 9 shows the pre-treated population at various values of pre-treatment rate $\delta$. Clearly, as $\delta$ increases, the number of pre-treated individuals also increases.

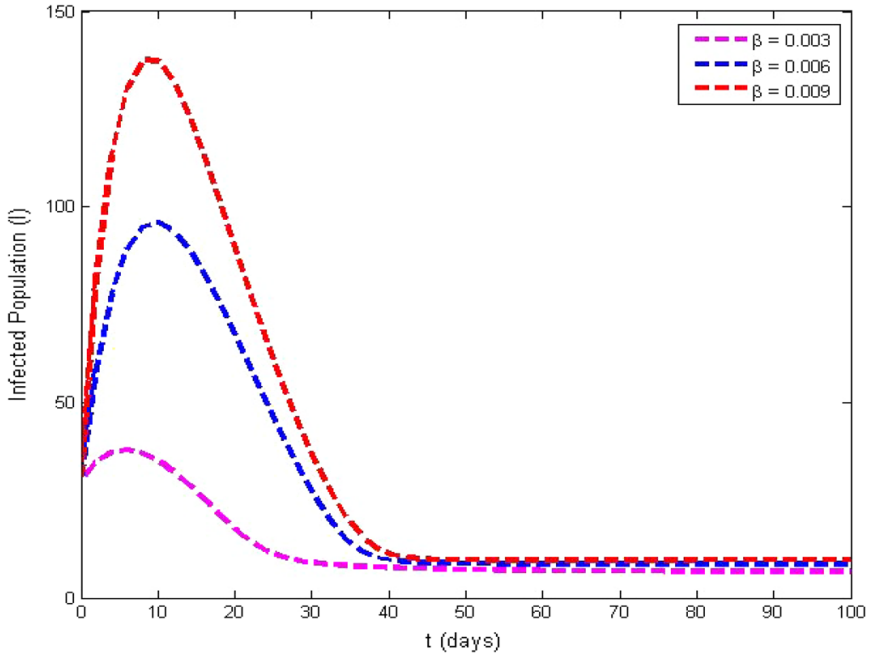



**Fig. 7** Infected population at various values of $\beta$ for $\tau = 0$

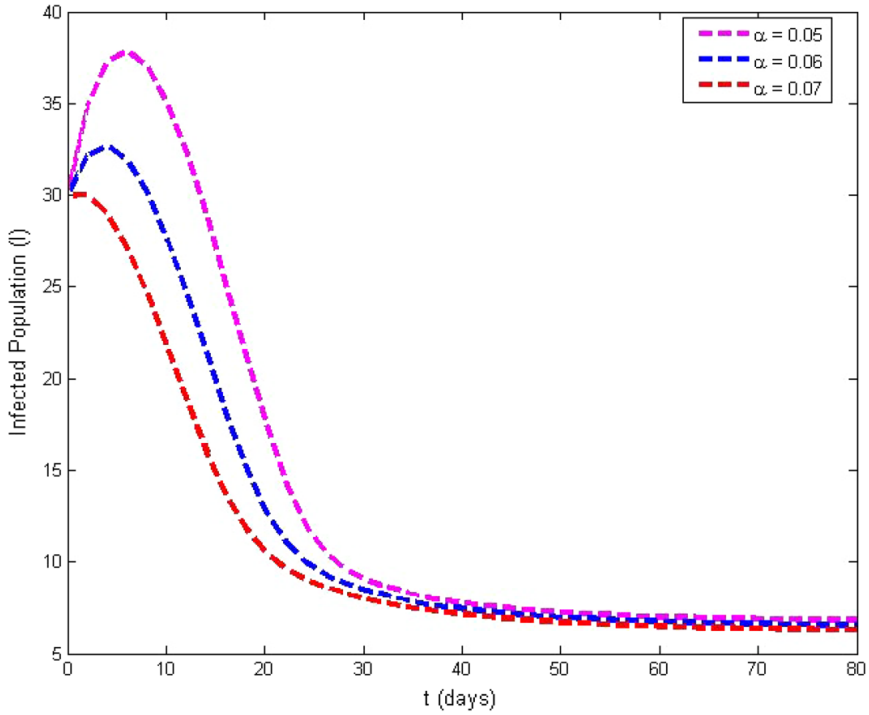



**Fig. 8** Infected population at various values of $\alpha$ for $\tau = 0$

Figures 10 and 11 show the infected population at various values of '$a$' and '$b$', respectively. Figure 10 shows the increment in an infected population with decrement in '$a$' and it settles down at its steady state, but the disease is not getting totally eradicated as it will persist at a much lower level. Figure 11 shows the increment in an infected population with increment in '$b$' which is due to limited availability of resources in the society.

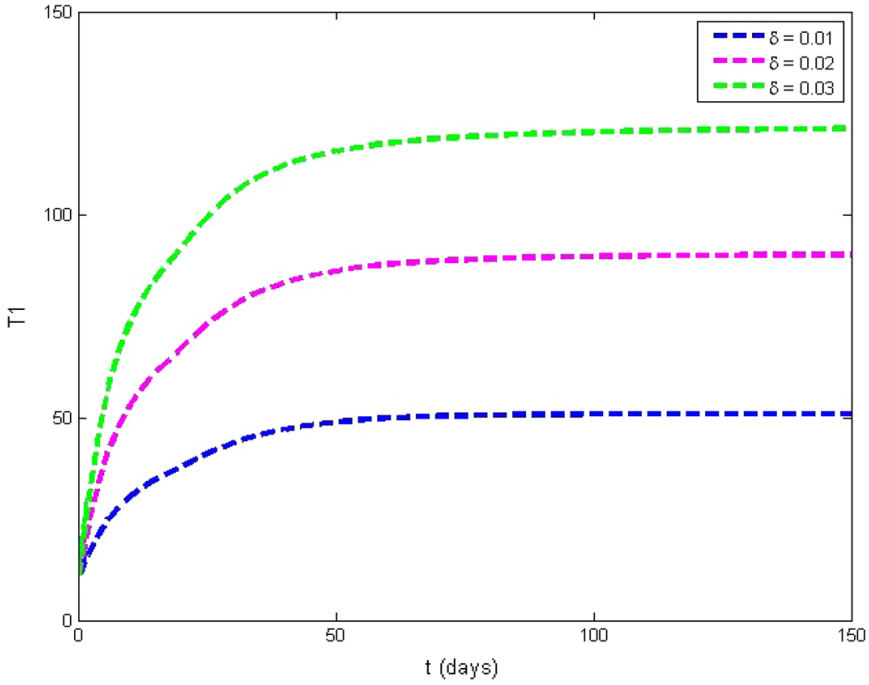



**Fig. 9** Pre-treated population at various values of $\delta$ for $\tau = 0$

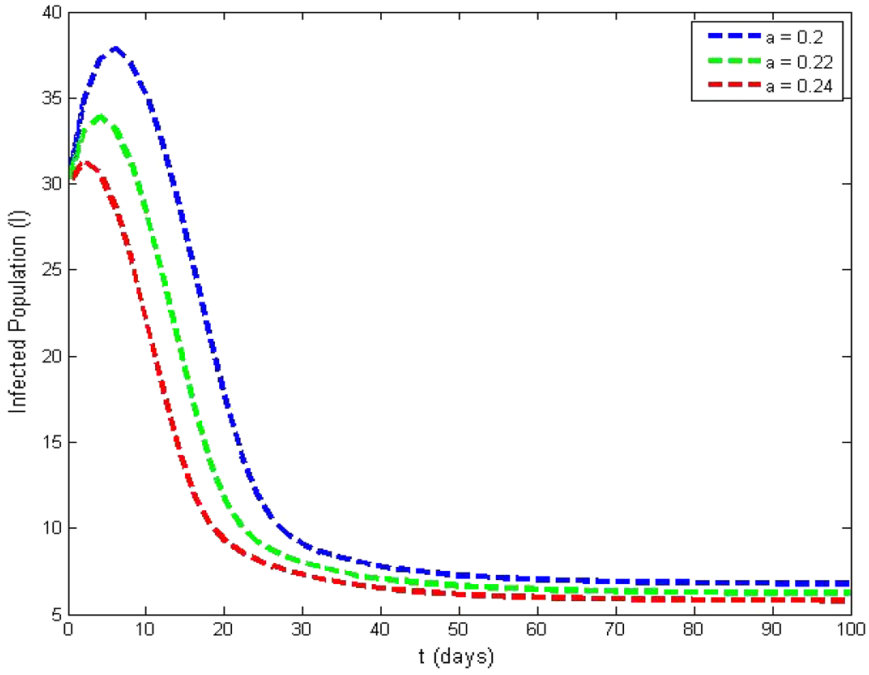



**Fig. 10** Infected population at various values of '$a$' for $\tau = 0$

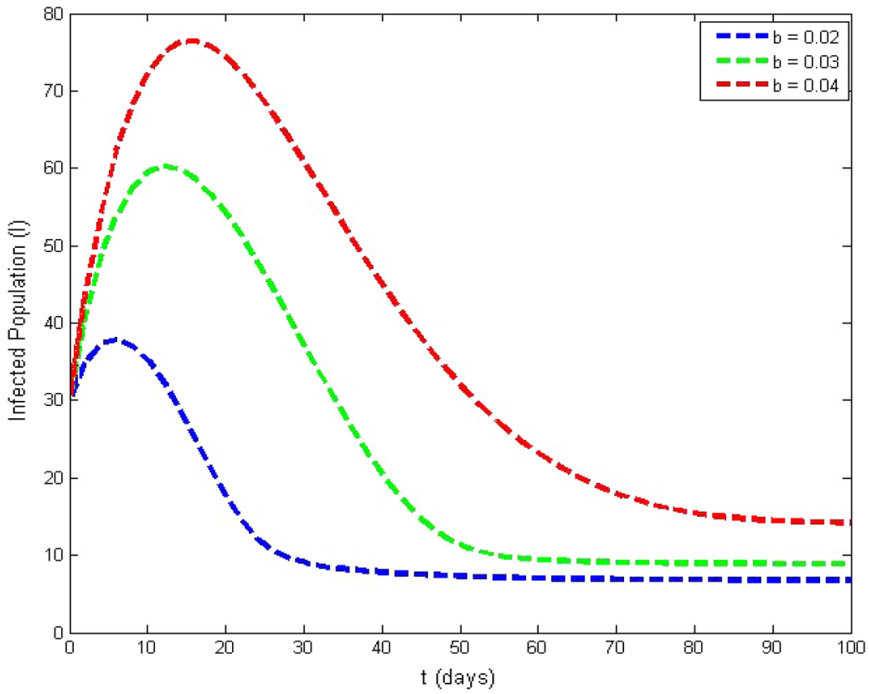



**Fig. 11** Infected population at various values of '$b$' for $\tau = 0$

## 6 Conclusion

In this article, we proposed a delayed SIR model by introducing two explicit treatment classes along with the saturated incidence rate: Holling type I treatment to susceptible population and Holling type III treatment to infected population. The model analysis shows that there exist only two types of equilibrium points: disease-free equilibrium point, i.e., when there is no infection in the society, and endemic equilibrium point, i.e., when infection is present in the society. The disease-free equilibrium is locally asymptotically stable when $R_0 < 1$ and unstable when $R_0 > 1$ for $\tau \geq 0$. We have also shown that the system (1)–(4) undergoes transcritical bifurcation at $R_0 = 1$. We have found that endemic equilibrium of the system (1)–(4) at $\tau = 0$ and $\tau > 0$ is locally asymptotically stable if inequalities as stated in theorem 3 and theorem 4, respectively, hold true. A numerical study presented the effect of incubation time delay on various classes of population and we found that delay can play a very crucial role to control the disease and in providing the preventive measures. The model also shows a decrease in an infected population with a decline in the transmission rate $\beta$ and an enhancement of inhibition rate $\alpha$.

**Acknowledgements** The authors would like to acknowledge Delhi Technological University, Delhi, India, for providing the financial support to carry out this research work.

## References

Adebimpe O, Bashiru KA, Ojurongbe TA (2015) Stability analysis of an SIR epidemic model with non-linear incidence rate and treatment. Open J Model Simul 3:104–110

Anderson RM, May RM (1991) Infectious disease of humans. Oxford University Press Inc, New York

Capasso V, Serio G (1978) A generalization of the Kermack–Mckendrick deterministic epidemic model. Math Biosci 42:41–61

Castillo-Chavez C, Song B (2004) Dynamical models of tuberculosis and their applications. Math Biosci Eng 1:361–404

Dubey B, Patra A, Srivastava PK, Dubey US (2013) Modeling and analysis of a SEIR model with different types of nonlinear treatment rates. J Biol Syst 21(3):1350023
Dubey B, Dubey P, Dubey US (2015) Dynamics of n SIR model with nonlinear incidence and treatment rate. Appl Appl Math 10(2):718–737
Dubey P, Dubey B, Dubey US (2016) An SIR model with nonlinear incidence rate and Holling type III treatment rate. In: Applied analysis in biological and physical sciences, Springer Proceedings in mathematics and statistics, vol 186, pp 63–81
Gumel AB, Mccluskey CC, Watmough J (2006) An SVEIR model for assessing the potential impact of an imperfect anti-SARS vaccine. Math Biosci Eng 3:485–494
Karim SAA, Razali R (2011) A proposed mathematical model of influenza A, H1N1 for Malaysia. J Appl Sci 11(8):1457–1460
Kumar A, Nilam (2018a) Stability of a time delayed SIR epidemic model along with nonlinear incidence rate and Holling type II treatment rat. Int J Comput Methods 15(6):1850055
Kumar A, Nilam (2018b) Dynamical model of epidemic along with time delay; Holling type II incidence rate and monod—Haldane treatment rate. Differ Equ Dyn Syst. https://doi.org/10.1007/s12591-018-0424-8
Michael YL, Graef JR, Wang L, Karsai J (1999) Global dynamics of a SEIR model with varying total population size. Math Biosci 160:191–213
Sastry S (1999) Analysis, stability and control. Springer, New York
Tipsri S, Chinviriyasit W (2014) Stability analysis of SEIR model with saturated incidence and time delay. Int J Appl Phys Math 4(1):42–45
van den Driessche P, Watmough J (2002) Reproduction numbers and sub-threshold endemic equilibria for compartment models of disease transmission. Math Biosci 180:29–48
Wang X (2004) A simple proof of Descartes's rule of signs. Am Math Mon. https://doi.org/10.2307/4145072
Wang W, Ruan S (2004) Bifurcation in an epidemic model with constant removal rates of the infective. J Math Anal Appl 21:775–793
Xu R, Ma Z (2009) Stability of a delayed SIRS epidemic model with a nonlinear incidence rate. Chaos Solutions Fract 41:2319–2325
Zhang Z, Suo S (2010) Qualitative analysis of a SIR epidemic model with saturated treatment rate. J Appl Math Comput 34:177–194
Zhou L, Fan M (2012) Dynamics of a SIR epidemic model with limited medical resources revisited. Nonlinear Anal RWA 13:312–324

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations

2019 6th International Conference on Signal Processing and Integrated Networks (SPIN)

# State-of-the-Arts Person Re-Identification using Deep Learning

Shradha Jaiswal[1], Dinesh Kumar Vishwakarma[2]
Department of Information Technology
Delhi Technological University
Delhi, India
jaiswalshraddha1994@gmail.com[1], dvishwakarma@gmail.com[2]

*Abstract*— **Person Re-Identification has become prominent because of various reasons majorly due to its high-performance methods based on deep-learning. It is the process of person recognition from various images captured by different cameras. Provided two set of images the purpose is to find that the given set of images are identical or not. Person Re-Id is often a challenging task due to the similarity in nature like people with identical features, color or clothes. Images are taken from various angles and distances of a given subject in order to achieve high accuracy, so it identifies correctly. Re-Identification has broadly two major categories: i) Image Re-ID and ii) Video Re-ID. Based on the category Re-ID has numerous applications like robotics, automated video surveillance, forensics and multimedia that are deployed using various public datasets like Market1501, VIPeR, MARS, CUHK01, CUHK02, CUHK03, DukeMTMC-reID, MSMT17 etc. In this paper we aim to briefly discuss the process, datasets, recent work on Re-ID, challenges, its approaches and techniques that has been implemented using deep learning systems.**

***Keywords—Person Re-Identification, deeply learned systems, Convolutional Neural Networks, Deep Learning.***

## I. INTRODUCTION

Person re-identification is the process of comparing the set of two input images captured from several cameras. Person re-id has emerged from the past several years and has gained popularity due to the need of public safety in streets, airports, university campuses, railway stations, parks, office buildings. It can effectively identify or track a person on different camera view points. Video surveillance, pedestrian counting, including multiple camera tracking, and multiple camera activity detection and analysis are some of the majorly used re-Id applications. Person re-identification video surveillance systems comprises of person detection, person tracking and person retrieval. Person Re-ID is a highly complex task as individual's look changes because of change in viewpoint, pose, different camera angles, luminescence, similar clothing etc. Humans can easily identify objects and people as our comprehension units are trained in this way, but for a machine it is difficult task to identify objects without human interruption. A general re-identification system consists of two simple factors: 1) capturing the unique description/model of a person and 2) matching those two models for concluding whether they are matching or not [2]. To automatize each component, various work should be performed, that has its issues and challenges which adds to the efficiency and challenges of Re-ID. Hand-craft features are used by conventional methods for deriving low-level features such as colour, shape, edge, etc. In Recent years, various researchers have applied the CNN for deriving high-level features which has shown a substantial enhancement as compared to hand-craft features.

*Image Re-ID*. Provided the test image taken by a camera, person re-ID system tries to recognize the person from a gallery of candidate images taken from different cameras. The viewable traces of people after the spatio-temporal segmentation method of forepart analysis was employed in [17]. Visual comparison is related to colour and edge histograms which are implemented by an enunciated person model or the Hessian-Affine interest point operator. This work [17] is the beginning of independent computer vision task and point the separation of person re-identification by tracking with multiple cameras [10], [11].

*Video Re-ID*. Maximum of the re-identification previous works were based on image matching. Two methods [18], [19] based on tracking of videos for multi-shot re-identification were designed by Farenzena et al. [19] in 2010 which selects frames randomly. Both the methods have colour as the common feature and a segmentation model for detecting the foreground features. Performance and accuracy of re-ID system increases as the increase in number of selected frames [18].

*Deep Re-ID*. The key objective of a deep re-ID model is learning discriminative feature representation. These features need to be evaluated at different scales [3]. There are two types of DL models that have been developed in this research field: 1) a classification method for the person re-ID work, and 2) a pair or triplet comparison based Siamese model [12]. However, deep learning models in person re-ID task still suffers from lack of training data samples and because of overfitting, these models fails during testing. A Siamese Neural network was implemented by Yi et al. [20] and Li et al. [21] to identify whether a set of two input images belong to same Id or not, which contributed to the success of deep learning models in classification of images. Deep learning approaches has then become prominent choice in re-identification. Various CNN based approaches focus on extracting global features of

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE
978-1-7281-1380-7/19/$31.00 ©2019 IEEE

pedestrian images [15]. Deep Learning methods have become a popular choice in person re-identification systems due to its higher efficiency, performance and accuracy over conventional person re-ID methods.

## II. PERSON RE-IDENTIFICATION PROCESS

Person re-ID process will begin with an automated identification of person from videos captured from several cameras by eliminating background noise [16]. Observing pedestrians in a network of cameras will enhance the system's performance. The process of person re-ID is shown in Fig. 1. There are following three types of functions on which these observations are based upon:

*1) Person Detection:* Its goal is the detection of persons in the image. There is only one image, no physical evaluation or treatment effects.

*2) Person tracking:* Its purpose is to provide time period, position of an image, moving targets, basically people in a crowded scene. Unlike detection, the difficulty here is time.

*3) Person re-identification*: Recognizing a person in a group of persons over different places unprovided with a common field of view in a camera network.

At the re-ID level, the main problem is observed which can be carried out in two levels: a) A learning level for detection and tracking of people in an order, and for building the model, object descriptors are needed b) The recognition level which will help to analyse whether the same person is on another camera using the models from the learning steps [4], [14].

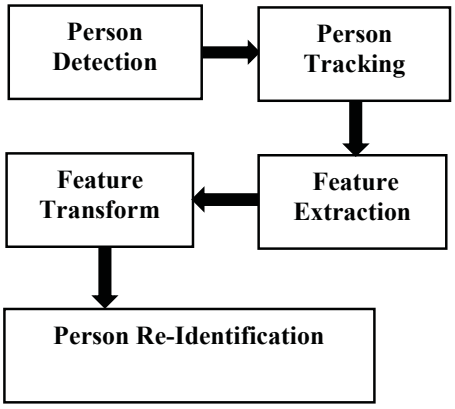



Fig. 1. Process of Person Re-Identification.

## III. DEEP LEARNING METHODS

This section presents an outline of recent works on person re-ID problem solved using DL models. To improve the performance of person re-ID systems, several deep learning methods have been proposed which has achieved greater performance compared to hand-craft feature methods. The state-of-the-art results on person re-ID system are achieved by CNN-based methods. Based on the feature representation [30], these are categorized into three feature methods: local feature methods, global feature methods and their combined features. Filter Pairing Neural Network (FPNN) is designed by Li et al. [21] to resolve the erroneous detection, background clutter and strong occlusion. To build a CNN model by joining hand craft features, in [23] a fusion model based on deep neural network feature was proposed. Deep learning features illustrations by multiple domains using CNN is given in [24]. They produced a strong model that works on multiple datasets by accumulating the data and labels from various re-id datasets and trained the CNN model with a SoftMax loss simultaneously [5], [6]. Extraction of global features using identification and verification methods was proposed by [25]. In [38], Part-based Convolutional Baseline has achieved considerable increase in performance by dividing feature maps in six horizontal sections. In [22], [34], a pairwise Siamese model was designed by adding a metric learning method on top of the network for the gain of spatio-temporal features. Multi-region bilinear subnetwork for training every part of the pedestrian image for the local feature extraction is employed by Ustinova et al. [35]. In [39], in order to combine local and global features of persons image, a multiple channel-based CNN method was designed. To learn similarity between set of two images, a Deep Hybrid Similarity Learning (DHSL) [37] like Siamese networks is designed in [37] based on CNN model.

To increase the efficiency of person re-identification systems, a Multiple scale Context-Aware Network model was proposed by Li et al. [40] for integration of global features by capturing information from local context of an image. In [45], an architecture based on Long-Short Term Memory was designed to increase local features performance by levering contextual dependencies. Two types of networks were analysed in [41] for training the networks and combining local and global features in it. To generate feature representation from an image of pedestrian, a framework based on deep neural network was designed by Ding et al. [43]. A multiple scale triplet network comprises of two shallow neural networks and a deep CNN with shared parameters was proposed by Liu et al. [44]. A deep CNN model for the generation of features based on body parts and global features from a pedestrian's image was designed by Bai et al. [42].

In [36], a distinct and novel approach that seeks global optimal matching over the cameras and covers all the camera networks was designed. Comparison of various deep learning model [12] is mentioned in Table I.

## IV. TASK AND CHALLENGES

Extraction of features and generating descriptors are carried out by re-ID system that consists of images or videos as an input. For re-ID system Building feature representation can turn out to be a complex task [9]. Features that are to be extracted must be adaptive, robust in case of variations, barriers, perceptual image quality and background consternation. Given multiple pictures, in-depth comprehensive understanding is required between detected objects across the frames in order to assure that the extracted features are from expected person. This process attach a persistent label to individual objects in multiple frames which is referred as tracking. Therefore, multiple instances of an individual object can be used for extracting features and consequently generating descriptors which could be further in re-ID [13].

In re-identification, person detection and multiple person tracking are challenging tasks. Person detection techniques

Table I. Comparison of existing Deep Learning models based on Rank-1 recognition rates Person Re-ID [12].

| S.No. | Methods | Model Type | Datasets (Accuracy %) | Architecture Used | Publication Year |
|---|---|---|---|---|---|
| 1 | Li et al. [21] | Pairwise | CUHK01 (20.65) | FPNN | 2014 |
| 2 | Zhang et al. [46] | Pairwise | ViPER (12.50), CAVIER (7.20) | Deep CNN with SVM | 2014 |
| 3 | Ahmed et al. [8] | Pairwise | ViPER (34.81), CUHK01 (65.0), CUHK03 (57.74) | Deep CNN | 2015 |
| 4 | Ding et al. [43] | Triplet | ViPER (40.50), i-LIDS (52.10) | DNN | 2015 |
| 5 | Zhang et al. [47] | Triplet | CUHK03 (18.74) | Deep Hashing | 2015 |
| 6 | Cheng et al. [39] | Triplet | ViPER (47.80), CUHK01 (53.70), PRID 2011 (22.0) | CNN based on Multi-channel | 2016 |
| 7 | Xiao et al. [24] | Classification | ViPER (38.60), CUHK01 (66.60), CUHK03 (75.30), i-LIDS (64.60), PRID 2011 (64.0) | CNN with multiple domains | 2016 |
| 8 | Wang et al. [34] | Pairwise | CUHK01 (38.28), CUHK03 (27.92) | Deep CNN based on Spatio-temporal features | 2016 |
| 9 | Liu et al. [44] | Triplet | Market 1501 (55.40) | CNN based on multi-scale triplet | 2016 |
| 10 | Iodice et al. [48] | Pairwise | ViPER (18.04) | CNN based 3D pyramidal NN | 2016 |
| 11 | Su et al. [49] | Classification | ViPER (51.27), CUHK03 (78.29), Market 1501 (63.14) | PDC | 2017 |
| 12 | Bai et al. [42] | Triplet | CUHK03 (91.50), Market 1501 (92.31) | LSTM | 2017 |
| 13 | Cheng et al. [50] | Triplet | CUHK01 (70.09), CUHK03 (84.70), Market 1501 (83.6) | CNN based structured Laplacian embedding | 2017 |
| 14 | Li et al. [51] | Classification | ViPER (38.08), CUHK03 (74.21), MARS (71.77), Market 1501 (80.31) | MSCAN | 2017 |
| 15 | Tao et al. [52] | Pairwise | ViPER (46.0) | DMVFL | 2017 |

have been used for so many years and there have been impressive amount of progress towards it. Uninterrupted tracking with varying environment has been an issue despite of broad researches which took place in Multiple Object Tracking as a part of a single camera and development of new methods over past several years.

In special cases, when people may not be cooperative, robust and critical visual descriptors require to extract the data captured in deprived and out of control environments. Thus, extraction of reliable descriptor depends on availability of superior amount of observation. Inaccurate detections and false trajectory evaluation could introduce errors during extraction and generation process of descriptor which may harm the quality of re-ID [9].

## V. PUBLIC DATASETS

There are various datasets that are available to test re-ID models. The mostly used datasets for person re-ID models are Market1501, i-LIDS, MSMT, ETHZ, CUHK, CAVIER, ViPER, and MARS, etc., shown in Table II.

*A. ViPER:* This [27] dataset includes a pair of images of 632 pedestrians with illumination and pose changes captured from two camera views. There is change in viewpoints and deviations in illumination exists in the cameras. Every image is adjusted and cropped by 128 × 48 pixels [7].

*B. ETHZ*: This [26] dataset consists of sequences of three videos from two movable cameras in a crowded street. There are 83, 35 and 28 persons in sequence 1, 2 and 3 respectively. Images exhibit strong occlusions, changes in illumination and variations in scale.

*C. MARKET 1501:* Images in this [29] dataset is collected by six manually configured cameras, in which there are 32,668 pedestrian images of 1,501 persons. There are two subsets of this database: a test set containing 19,732 images from 750 different identities and a training set which consists of 12,936 persons image of 750 identities.

*D. CUHK:* This dataset consists of three parts setups individually. CUHK01 [31] have 1,942 images of 971 persons. There are 1,816 pedestrians in CUHK02 [20] dataset captured by 10 (5 pairs) cameras. Images in CUHK03 [21] dataset are captured by 6 view cameras, contains 14,097 images of 1,467 different persons. Every image is manually cropped and shows variations in missing body-part, occlusions and illumination [8].

*E. i-LIDS:* Images in this [28] dataset is captured by two cameras which are non-overlapping. It has 476 images of 119 people. Both the cameras show changes in illumination and occlusions, and database contains exactly 4 images of every person and on average there are 2 images per person. Every image is adjusted to 128 × 64 pixels.

*F. MARS:* This [1] dataset consists of 1,191,003 bounding boxes with 20,478 tracklets. It has video sequences of 1,261 different persons captured by two or more cameras.

*G. CAVIAR:* There are 72 pedestrians in this [32] dataset, which consists of two views. 22 pedestrians have appeared in only one view while 50 pedestrians appeared in both the views. The camera contains differing variations, pose, occlusions and light conditions with 5 images per view of every person.

*H. MSMT*: This [33] dataset consists of 126,441 pedestrian images of 4,101 different persons. Images are captured by 3

Table II. Illustration of some currently available datasets for image-based and video-based re-ID [2].

| | Datasets | Year | #Camera | #Person | Label |
|---|---|---|---|---|---|
| Image Re-ID | ViPER | 2007 | 2 | 632 | Hand |
| | GRID | 2009 | 8 | 250 | Hand |
| | iLIDS | 2009 | 2 | 119 | Hand |
| | PRID 2011 | 2011 | 2 | 200 | Hand |
| | CAVIER | 2011 | 2 | 72 | Hand |
| | WARD | 2012 | 3 | 70 | Hand |
| | CUHK01 | 2012 | 2 | 971 | Hand |
| | CUHK02 | 2013 | 10 | 1,816 | Hand |
| | CUHK03 | 2014 | 2 | 1,467 | Hand/DPM |
| | PRID 450S | 2014 | 2 | 450 | Hand |
| | RAiD | 2014 | 4 | 43 | Hand |
| | Market 1501 | 2015 | 6 | 32,668 | Hand/DPM |
| Video Re-ID | ETHZ | 2007 | 1 | 148 | Hand |
| | PRID 2011 | 2011 | 2 | 200 | Hand |
| | 3DPES | 2011 | 8 | 200 | Hand |
| | iLIDS-VID | 2014 | 2 | 300 | Hand |
| | MARS | 2016 | 6 | 1,261 | DPM & GMMCP |

outdoor and 12 indoor cameras with variation in pose, illumination and scale.

## VI. CONCLUSION

The paper presents the fundamental process of person re-identification, an outline of person re-ID with respect to Deep learning methods and the current progress towards it. We have also discussed the ongoing researches, task and challenges of person re-identification. Available Datasets that are images-based as well as video-based for person re-ID reviewed. As applications starts to grow in wide variety of range, it turns out to be a complex task. Person re-Identification catching eyes in not only academics but in industries also due to its various applications. Though models and methods have been utilized but there is still a strong need of in-depth exploration in real time problems.

## REFERENCES


[1] L. Zheng, Z. Bie, Y. Sun, J. Wang, C. Su, S. Wang and Q. Tian, "Mars: A video benchmark for large-scale person re-identification," *European Conference on Computer Vision*, pp. 868-884, 2016.

[2] L. Zheng, Y. Yang and A. G. Hauptmann, "Person re identification: past, present and future," *arXiv preprint arXiv:1610.02984*, 2016.

[3] X. Qian, Y. Fu, Y. Jiang, T. Xiang and X. Xue, "Multi-scale Deep Learning Architectures for Person Re-identification," *IEEE International Conference on Computer Vision (ICCV)*, pp. 5409-5418, 2017.

[4] B. Hadjkacem, W. Ayedi and M. Abid, "A comparison between person re-identification approaches," *Third International Conference on Artificial Intelligence and Pattern Recognition (AIPR)*, pp. 1-4, 2016.

[5] W. Chen, X. Chen, J. Zhang and K. Huang, "A Multi-Task Deep Network for Person Re-Identification," *AAAI*, pp. 3988-3994, 2017.

[6] Z. Zhang, T. Si and S. Liu, "Integration Convolutional Neural Network for Person Re-Identification in Camera Networks," *in IEEE Access*, vol. 6, pp. 36887-36896, 2018.

[7] M. Xu, Z. Tang, Y. Yao, L. Yao, H. Liu and J. Xu, "Deep Learning for Person Reidentification Using Support Vector Machines," *Advances in Multimedia*, pp. 1-12, 2017.

[8] E. Ahmed, M. Jones and T. K. Marks, "An improved deep learning architecture for person re-identification," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3908-3916, 2015.

[9] A. Bedagkar-Gala and S. K. Shah, "A survey of approaches and trends in person re-identification," *Image and Vision Computing*, pp. 270-286, 2014.

[10] A. Schumann and R. Stiefelhagen, "Person Re-identification by Deep Learning Attribute-Complementary Information," *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1435-1443, 2017.

[11] Y. Zhou, W. Zheng, A. Wu, H. Yu, H. Wan and J. Lai, "Learning a Semantically Discriminative Joint Space for Attribute Based Person Re-identification," *arXiv preprint arXiv:1712.01493*, 2017.

[12] B. Lavi, M. F. Serj and I. Ullah, "Survey on Deep Learning Techniques for Person Re-Identification Task," *arXiv preprint arXiv:1807.05284*, 2018.

[13] S. Gong, M. Cristani, C. C. Loy and T. M. Hospedales, "The re-identification challenge," *in Person re-identification*, pp. 1-20, 2014.

[14] F. M. Khan and F. Bremond, "Person re-identification for real-world surveillance systems," *arXiv preprint arXiv:1607.05975*, 2016.

[15] S. Liao, Z. Mo, J. Zhu, Y. Hu and S. Z. Li, "Open-set person re-identification," *arXiv preprint arXiv:1408.0872*, 2014.

[16] X. Wang and R. Zhao, "Person re-identification: System design and evaluation overview," *in Person Re-Identification*, pp. 351-370, 2014.

[17] N. Gheissari, T. B. Sebastian and R. Hartley, "Person reidentification using spatio-temporal appearance," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, vol. 2, pp. 1528–1535, 2006.

[18] L. Bazzani, M. Cristani, A. Perina, M. Farenzena and V. Murino, "Multiple-shot person re-identification by hpe signature," *International Conference on Pattern Recognition (ICPR)*, pp. 1413–1416, 2010.

[19] M. Farenzena, L. Bazzani, A. Perina, V. Murino and M. Cristani, "Person re-identification by symmetry-driven accumulation of local features," *in Computer Vision and Pattern Recognition (CVPR)*, pp. 2360–2367, 2010.

[20] D. Yi, Z. Lei, S. Liao and S. Z. Li, "Deep Metric Learning for Person Re-identification," *International Conference on Pattern Recognition*, pp. 34-39, 2014.

[21] W. Li, R. Zhao, T. Xiao and X. Wang, "Deepreid: Deep filter pairing neural network for person re-identification," *in IEEE Conference on Computer Vision and Pattern Recognition*, pp. 152–159, 2014.

[22] J. Bromley, J. W. Bentz, L. Bottou, I. Guyon, Y. LeCun, C. Moore, E. Sackinger and R. Shah, "Signature verification using a Siamese time delay neural network," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 7, no. 04, pp. 669–688, 1993.

[23] S. Wu, Y. Chen, X. Li, A. Wu, J. You and W. Zheng, "An enhanced deep feature representation for person re-identification," *IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1-8, 2016.

[24] T. Xiao, H. Li, W. Ouyang and X. Wang, "Learning deep feature representations with domain guided dropout for person re-identification," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1249–1258, 2016.

[25] Z. Zheng, L. Zheng and Y. Yang, "A discriminatively learned CNN embedding for person re-identification," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 14, no. 1, pp. 13, 2017.

[26] A. Ess, B. Leibe and L. Van Gool, "Depth and appearance for mobile scene analysis," *in IEEE International Conference on Computer Vision*, pp. 1–8, 2007.

[27] D. Gray, S. Brennan and H. Tao, "Evaluating appearance models for recognition, reacquisition, and tracking," *IEEE International Workshop on Performance Evaluation for Tracking and Surveillance (PETS)*, vol. 3, no. 5, pp. 41-49, 2007.

[28] T. Wang, S. Gong, X. Zhu and S. Wang, "Person re-identification by video ranking," *in European Conference on Computer Vision*, pp. 688–703, 2014.

[29] L. Zheng, L. Shen, L. Tian, S.Wang, J.Wang and Q. Tian, "Scalable person re-identification: A benchmark," *IEEE International Conference on Computer Vision*, pp. 1116–1124, 2105.

[30] W. Li and X. Wang, “Locally aligned feature transforms across views,” *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3594–3601, 2013.

[31] W. Li, R. Zhao and X. Wang, “Human reidentification with transferred metric learning,” *in Asian Conference on Computer Vision*, pp. 31–44, 2012.

[32] D. S. Cheng, M. Cristani, M. Stoppa, L. Bazzani and V. Murino, "Custom pictorial structures for re-identification," *In Bmvc*, vol. 1, no. 2, pp. 1-11. 2011.

[33] L. Wei, S. Zhang, W. Gao and Q. Tian, "Person transfer gan to bridge domain gap for person re-identification," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 79-88, 2018.

[34] S. Wang, C. Zhang, L. Duan, L. Wang, S. Wu and L. Chen, "Person re-identification based on deep spatio-temporal features and transfer learning," *International Joint Conference on Neural Networks (IJCNN)*, pp. 1660-1665, 2016.

[35] E. Ustinova, Y. Ganin and V. Lempitsky, “Multi-region bilinear convolutional neural networks for person re-identification,” *IEEE Conference on Advanced Video and Signal-based Surveillance*, pp. 1–6, 2017.

[36] J. Lin, L. Ren, J. Lu, J. Feng and J. Zhou, "Consistent-aware deep learning for person re-identification in a camera network," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5771-5780, 2017.

[37] J. Zhu, H. Zeng, S. Liao, Z. Lei, C. Cai and L. Zheng, "Deep Hybrid Similarity Learning for Person Re-Identification," *in IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 11, pp. 3183-3193, 2018.

[38] Y. Sun, L. Zheng, Y. Yang, Q. Tian and S.Wang, “Beyond part models: person retrieval with refined part pooling,” *arXiv preprint arXiv:1711.09349*, 2017.

[39] D. Cheng, Y. Gong, S. Zhou, J. Wang and N. Zheng, “Person reidentification by multi-channel parts-based CNN with improved triplet loss function,” *International Conference on Pattern Recognition*, pp. 1335–1344, 2016.

[40] D. Li, X. Chen, Z. Zhang and K. Huang, “Learning deep context-aware features over body and latent parts for person re-identification,” *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 384–393, 2017.

[41] Z. Zhang and T. Si, “Learning deep features from body and parts for person re-identification in camera networks,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, pp. 52, 2018.

[42] X. Bai, M. Yang, T. Huang, Z. Dou, R. Yu and Y. Xu, "Deep-Person: Learning Discriminative Deep Features for Person Re-Identification," *arXiv preprint arXiv:1711.10658*, 2017.

[43] S. Ding, L. Lin, G. Wang and H. Chao, "Deep feature learning with relative distance comparison for person re-identification," *Pattern Recognition*, pp. 2993-3003, 2015.

[44] J. Liu, Z. Zha, Q. I. Tian, D. Liu, T. Yao, Q. Ling and T. Mei, "Multi-scale triplet cnn for person re-identification," *ACM on Multimedia Conference*, pp. 192-196, 2016.

[45] R. R. Varior, B. Shuai, J. Lu, D. Xu and G. Wang, "A siamese long short-term memory architecture for human re-identification," *in European Conference on Computer Vision*, pp. 135-153, 2016.

[46] G. Zhang, J. Kato, Y. Wang and K. Mase, "People re-identification using deep convolutional neural network," *International Conference on Computer Vision Theory and Applications (VISAPP)*, pp. 216-223, 2014.

[47] R. Zhang, L. Lin, R. Zhang, W. Zuo and L. Zhang, "Bit-Scalable Deep Hashing With Regularized Similarity Learning for Image Retrieval and Person Re-Identification," *in IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 4766-4779, 2015.

[48] S. Iodice, A. Petrosino and I. Ullah, "Strict pyramidal deep architectures for person re-identification," *in International Workshop on Neural Networks*, pp. 179-186, 2015.

[49] C. Su, J. Li, S. Zhang, J. Xing, W. Gao and Q. Tian, "Pose-driven deep convolutional model for person re-identification," *IEEE International Conference on Computer Vision (ICCV)*, pp. 3980-3989, 2017.

[50] D. Cheng, Y. Gong, X. Chang, W. Shi, A. Hauptmann and N. Zheng, "Deep Feature Learning via Structured Graph Laplacian Embedding for Person Re-Identification," *Pattern Recognition*, pp. 94-104, 2018.

[51] D. Li, X. Chen, Z. Zhang and K. Huang, "Learning deep context-aware features over body and latent parts for person re-identification," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 384-393, 2017.

[52] D. Tao, Y. Guo, B. Yu, J. Pang and Z. Yu, "Deep Multi-View Feature Learning for Person Re-Identification," *in IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 10, pp. 2657-2666, 2018.

EPMA Journal
https://doi.org/10.1007/s13167-019-00169-y

RESEARCH

# The burden of prostate cancer is associated with human development index: evidence from 87 countries, 1990–2016

Rajesh Sharma [1]

Received: 22 March 2019 / Accepted: 26 April 2019
© European Association for Predictive, Preventive and Personalised Medicine (EPMA) 2019

## Abstract

**Aim** To examine the temporal patterns of the prostate cancer burden and its association with human development.

**Subject and methods** The estimates of the incidence and mortality of prostate cancer for 87 countries were obtained from the Global Burden of Disease 2016 study for the period 1990 to 2016. The human development level of a country was measured using its human development index (HDI): a summary indicator of health, education, and income. The association between the burden of prostate cancer and the human development index (HDI) was measured using pairwise correlation and bivariate regression. Mortality-to-incidence ratio (MIR) was employed as a proxy for the survival rate of prostate cancer.

**Results** Globally, 1.4 million new cases of prostate cancer arose in 2016 claiming 380,916 lives which more than doubled from 579,457 incident cases and 191,687 deaths in 1990. In 2016, the age-standardised incidence rate (ASIR) was the highest in very high–HDI countries led by Australia with ASIR of 174.1/100,000 and showed a strong positive association with HDI ($r = 0.66$); the age-standardised mortality rate (ASMR), however, was higher in low-HDI countries led by Zimbabwe with ASMR of 78.2/100,000 in 2016. Global MIR decreased from 0.33 in 1990 to 0.26 in 2016. Mortality-to-incidence ratio (MIR) exhibited a negative gradient ($r = -0.91$) with human development index with tenfold variation globally with seven countries recording MIR in excess of 1 with the USA recording the minimum MIR of 0.10.

**Conclusion** The high mortality and lower survival rates in less-developed countries demand all-inclusive solutions ranging from cost-effective early screening and detection to cost-effective cancer treatment. In tackling the rising burden of prostate cancer predictive, preventive and personalised medicine (PPPM) can play a useful role through prevention strategies, predicting PCa more precisely and accurately using a multiomic approach and risk-stratifying patients to provide personalised medicine.

**Keywords** Prostate cancer · Incidence · Mortality · Mortality-to-incidence ratio · Human development index · Predictive preventive personalised medicine · Precision medicine

## Introduction

Prostate cancer (PCa) is one of the leading causes of death in males worldwide, and its incidence is increasing in both developed and developing countries alike in the last three decades [1]. While the incidence of prostate cancer is high in developed countries such as the USA and European countries, the survival rates, however, are lower in the less-developed countries [2, 3]. Prostate cancer is mostly diagnosed in older ages with a large percent of it being diagnosed as indolent and a significant portion of older patients die with prostate cancer rather than due to it [4]. There are wide disparities in survival rates in world regions which largely depend upon the stage of disease presentation, healthcare infrastructure, and availability of treatment options [5].

This paper examines the burden of prostate cancer (incidence and mortality) in 87 countries for the period 1990–2016 paying careful attention to the temporal patterns of prostate cancer burden and its association with human development. The estimates of prostate cancer burden were procured from the Global Burden of Disease 2016 study [1, 6], and the human development of a country was measured using its human development index (HDI) compiled by United Nations Development Program (UNDP). Against previous studies limited by the number of countries (or regions) or time period

✉ Rajesh Sharma
rajesh.sharma@dtu.ac.in

[1] University School of Management and Entrepreneurship, Delhi Technological University, East Delhi Campus, Room No. 305, Vivek Vihar Phase II, Delhi 110095, India

Published online: 08 May 2019

examined, we investigate prostate cancer burden in a large set of countries (87 countries) and pay careful attention to the temporal patterns of prostate cancer burden. In addition to investigating the incidence and mortality of prostate cancer, we also examine mortality-to-incidence ratio (MIR) of prostate cancer which has been demonstrated to be a proxy indicator of 5-year survival rates [7–10]. Although MIR is an imperfect proxy for survival rate, yet, we conjecture that it can reveal patterns of relative survival rates across countries as per their development status. In this study, we employed the human development index (HDI) as the indicator of development in the place of the socio-demographic index (SDI) employed by GBD. This is because SDI is a composite measure of income, education, and fertility whereas HDI is the indicator of income, education, and life expectancy, and as cancer burden has been found to have positive association with age, life expectancy seems to be a better indicator than fertility while describing the prostate cancer burden. Examination of past trends of incidence, mortality, and survival rate (proxied using MIR) and its association with HDI can reveal the rate of our progress in the fight against one of the diseases causing substantial mortality and morbidity in males worldwide. Moreover, PPPM approaches to cancer management will first require the data or the estimates as to the regions or countries in which cancer burden is rising the fastest and to identify the countries in which the survival rates are the worst. Identifying the most-burdened countries in terms of incidence, mortality, and survival can help in accelerating PPPM practices. Secondly, our analysis can illuminate the role of human development (and its components) in ameliorating the burden of prostate cancer. Thirdly, developed countries seem to be in a better position to offer precision medicine to tackle the cancer burden. If the survival rates proxied by MIR can be better predicted by human development measured using HDI, burdened countries must strive for broader human development rather than just boosting per capita income levels.

## Data and methods

The estimates of incidence and mortality of prostate cancer (both all-age numbers and age-standardised rates) were procured from the Global Burden of Disease (GBD) 2016 study for the period 1990–2016 [6, 11].[1] The Global Burden of Disease 2016 study provides age-, sex-, and year-wise estimates of incidence, prevalence, and mortality for 264 causes for 195 countries for the period 1990 to 2016 [6]. The Global Burden of Disease provides estimates of three additional population health metrics: years of life lost (YLLs) due to premature death, years lived with disability (YLDs), and disability-adjusted life years (DALYs).

[1] For more details on GBD methodology, the readers are referred to the GBD 2016 disease, injuries, and risk factor study [1, 6, 12].

### GBD estimation framework

To obtain estimates of incidence, mortality, and DALYs, GBD procured data from various sources such as vital registration, cancer registries, and verbal autopsies, which goes through multiple modelling steps detailed elsewhere [1, 6]; here, we briefly discuss how GBD arrives at the estimates of the incidence and mortality of prostate cancer. The GBD methodology starts with the processing of incidence and mortality data from cancer registries and the processed data was matched by cancer, age, sex, year, and location to generate crude mortality-to-incidence ratios (MIR). Second, the crude mortality-to-incidence ratios were used as input for a three-step modelling approach using the general spatio-temporal Gaussian process regression (ST-GPR) along with socio-demographic index (SDI) as a covariate in the linear-step mixed effects model using a logit link function to obtain the final MIR. In the third step, the incidence from cancer registries was multiplied with mortality-to-incidence ratios (MIR) obtained in the second step to generate mortality estimates; these mortality estimates were then used as input for cause-of-death ensemble model (CodeM) to obtain the final mortality estimates [6, 13]. The CodeM model was developed by Foreman et al. [13] to arrive at the cause-specific mortality estimates; it is an ensemble modelling approach in which a large number of models are tested for their predictive validity before arriving at the final mortality estimates [13]. In this modelling approach, only those covariates are chosen which have been found to have a plausible relationship with the particular cause of death. The relationship need not be causal but at least plausible. Few of the covariates used to model prostate cancer mortality in CodeM were as follows: education, lag-distributed income, socio-demographic index, and healthcare access and quality index [6]. The final mortality estimates were divided by the MIR estimates obtained in the second step to generate the final incidence estimates [6, 12].

### Mortality, incidence, and mortality-to-incidence ratio

In this study, we procured estimates of incidence and mortality of prostate cancer for the period 1990 to 2016. Both all-age prostate cancer numbers and age-standardised rates were employed to describe the PCa burden. We employed mortality-to-incidence ratio to proxy for the survival rate of prostate cancer which has been demonstrated to be a proxy of the survival rate of different neoplasia [7–10]. Although an imperfect proxy of survival rates, it can provide useful direction as to the relative survival rates of PCa across countries. Mortality-to-incidence ratio was calculated directly from the ratio of crude mortality rate and crude incidence rate obtained from GBD 2016. In this study, we examined prostate cancer for 87 countries which had more than 1000 incident cases in 2016. It was done to focus on heavily-burdened countries as

well as to prevent skewness in MIR values due to very low values of incidence or mortality, which could have resulted in too high or too low MIR values. Following the GBD convention, we reported the estimates of incidence and mortality with 95% uncertainty intervals (UI)[2] inside the square brackets and where percent change was reported alone, it was calculated from the mean estimates.

### Development level of a country

The development level of a country was measured using its human development index (HDI)—a composite measure of three indicators of development: health (life expectancy at birth), education (mean years of schooling and expected number of years of schooling),[3] and income (gross national income (GNI) per capita). For HDI calculation, each indicator gets scaled on a 0 to 1 scale with 0 corresponding to the lowest value and 1 corresponding to the highest value, and the final HDI value of a country was calculated by taking the geometric mean of the health, education, and income indices [14]. Data pertaining to the human development index (HDI) for the period 1990 to 2015 was procured from the UNDP database. For descriptive as well as analytical purposes, countries were categorised into four groups as per UNDP classification in 2015: very high (HDI > 0.800, 36 countries), high (0.700 < HDI < 0.799, 26 countries), medium (0.550 < HDI < 0.669, 15 countries), and low (HDI < 0.550, 9 countries) HDI categories.[4] In this paper, we employ the term "developed" for countries in high- and very high–HDI categories and "developing" for countries in low- and medium-HDI categories.

In addition to defining PCa burden as per development measured by HDI, we also examined the association between the mortality-to-incidence ratio (MIR), and income per capita and life expectancy of males. We procured data of GNI per capita (at 2010 USD) and life expectancy of males from World Development Indicators (WDI) database of the World Bank. The entire data analysis was conducted using MS Excel 2016, Python 3.6, and Stata 13.0.

## Results

### Global temporal patterns: 1990–2016

Globally, prostate cancer resulted in 380,916 [95% uncertainty interval (UI), 320,808–412,868] deaths in 2016 from 191,687 [168886–209,254] in 1990 (Fig. 1). The incidence of prostate cancer more than doubled from 579,457 [521564–616,107] in 1990 to 1.4 million [1.3–1.6 million] incident cases in 2016. Against the rising number of deaths due to prostate cancer, age-standardised mortality rate (ASMR) fallen from 16.7 [14.9–18.3] in 1990 to 14.9 [12.7–16.2] per 100,000 males in 2016, whereas the age-standardised incidence rate (ASIR) rose from 43.2 [38.8–46.1] per 100,000 males in 1990 to 49.9 [45.0–56.1] in 2016. Global mortality-to-incidence ratio (MIR) witnessed a decline from 0.33 in 1990 to 0.26 in 2016. Figure 1 illustrates that all-age incidence and deaths due to prostate cancer increased continuously between 1990 and 2016, whereas MIR exhibited a declining trend over the same period. ASIR increased sharply in the early 1990s before undergoing a flattened trajectory in the time period thereafter. ASMR too decreased in the study period before peaking around mid-1990s.

The sample of 87 countries in this study accounted for 1,402,465 incident cases and 358,056 deaths which contributed towards 98% of global incidence and deaths due to prostate cancer in 2016. Incidence increased in all the 87 countries, whereas ASIR increased in all countries except for two countries: Ghana and the USA. Deaths due to PCa also increased in all the sample countries, whereas ASMR increased in 52 of the 87 countries—21 of the 24 less-developed countries (low/medium HDI) and 31 of 62 developed countries (high/very high HDI) (Table 1; Fig. 2).

### Incidence and mortality in 2016

In terms of incidence, very high– and high-HDI countries recorded the highest all-age incidence and deaths with the USA recording the highest incidence of 371,672 [348,609–485,368] followed by China with 111,604 [88324–138,314] newly diagnosed cases in 2016 (Table 2). The age-standardised incidence rate (ASIR) was again the highest in developed countries (high/very high HDI) with all the top ten countries belonged to the very high–HDI category. With an ASIR of 174.1 [144.2–213.4] per 100,000 males, Australia recorded the largest ASIR in 2016 closely followed by the USA with ASIR of 172.5

[2] In this paper, we report UIs for the estimates directly available from GBD which were calculated on the basis of 1000 draws for each estimate and do not report UIs for MIR as we calculated MIR on the basis on mean estimates.

[3] UNDP defines mean number of years of schooling as the number of years of schooling received by people ages 25 and older, and expected years of schooling is defined as number of years of schooling that a child of school entrance age can expect to receive if prevailing patterns of age-specific enrolment rates persist throughout life.

[4] **Very High HDI:** Argentina, Australia, Austria, Belgium, Canada, Chile, Croatia, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Israel, Italy, Japan, Latvia, Lithuania, Netherlands, New Zeeland, Norway, Poland, Portugal, Romania, Russia, Saudi Arabia, Slovakia, Slovenia, South Korea, Spain, Sweden, Switzerland, UK, and USA.

**High HDI:** Algeria, Belarus, Brazil, Bulgaria, China, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, Iran, Jamaica, Kazakhstan, Lebanon, Malaysia, Mexico, Panama, Peru, Serbia, Sri Lanka, Thailand, Tunisia, Turkey, Ukraine, Uruguay, and Venezuela.

**Medium HDI**: Bangladesh, Bolivia, Egypt, Ghana, Guatemala, India, Indonesia, Kenya, Morocco, Myanmar, Pakistan, Paraguay, Philippines, South Africa, and Vietnam.

**Low HDI:** Cote d'Ivoire, Democratic Republic of Congo, Ethiopia, Haiti, Madagascar, Nigeria, Tanzania, Uganda, and Zimbabwe.

**No HDI category:** Puerto Rico

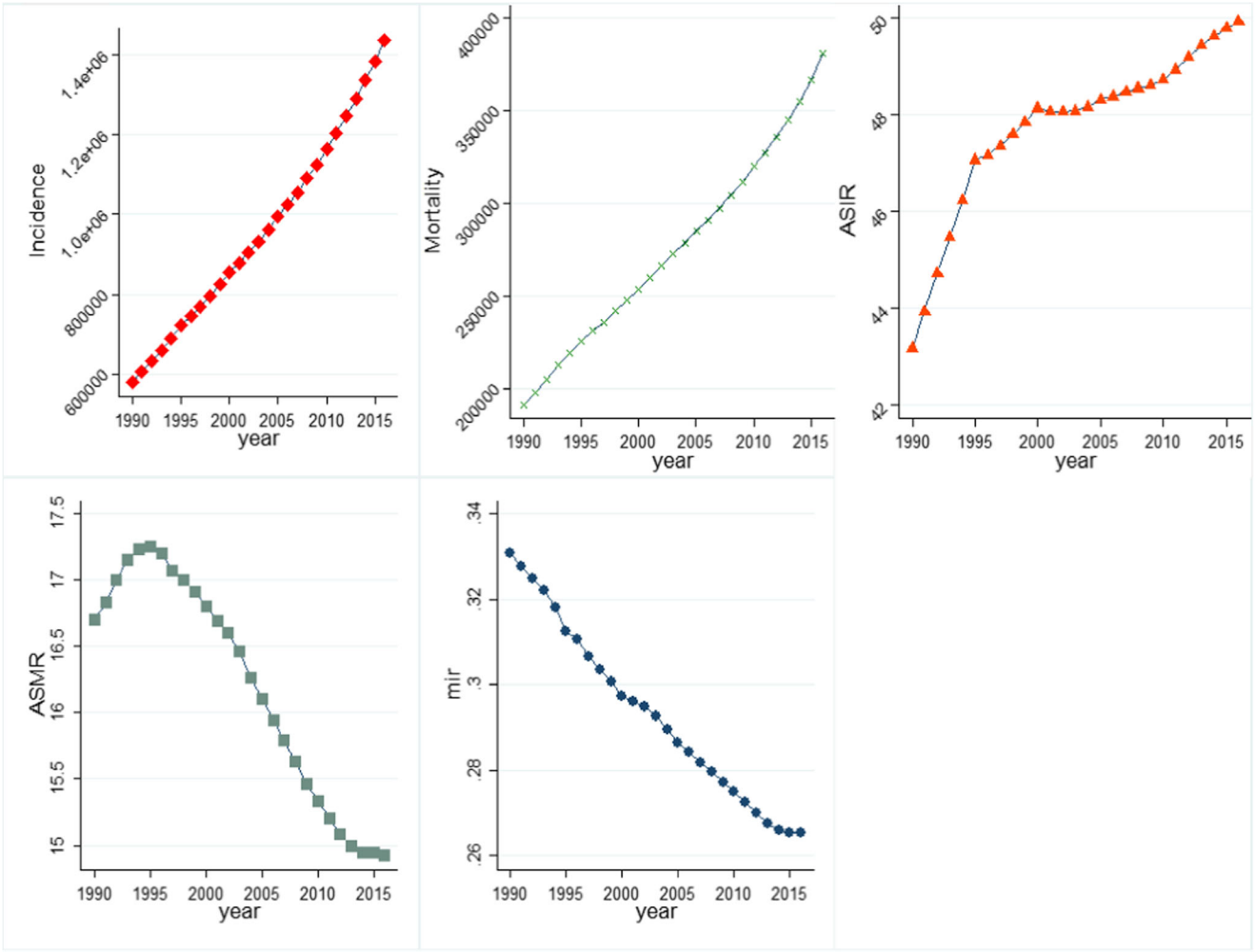



**Fig. 1** Temporal movement of global prostate cancer burden in terms of key indicators, 1990–2016: **a** incidence, **b** mortality, **c** ASIR, **d** ASMR, **e** MIR. Incidence, all-age incidence numbers; mortality, all-age deaths; ASIR, age-standardised incidence rate; ASMR, age-standardised mortality rate; MIR, mortality-to-incidence ratio. Data source: Global Burden of Disease 2016 study

[161.9–225.0] per 100,000 males. The populous countries of India and China recorded low ASIR of 16.1 [12.6–20.1] and 9.0 [7.1–10.8] per 100,000 males, respectively, in 2016.

In terms of mortality, the developed countries (high/very high HDI) occupied eight of the top ten positions with the USA ranked first with 35,468 [32,984–48,659] deaths followed by China with 35,327 [26,800–44,281] deaths in 2016. India (medium HDI) and Nigeria (low HDI) were two exceptions in top ten with prostate cancer claiming 20,954 [16,144–25,842] and 11,410 [95% UI, 4812–17,297] lives, respectively in 2016 (Table 2). The age-standardised mortality rate (ASMR) exhibited different patterns as per human development in comparison to ASIR with seven of the top ten countries belonged to the low-HDI category in 2016. High ASMR was concentrated in regions of Africa with Zimbabwe recording the highest ASMR of 78.2/100,000 [34.8–109.5] closely followed by Uganda with ASMR of 75.5 [57.6–95.4] per 100,000 in 2016. India and China were present at the bottom with ASMR of 6.7 [5.1–8.0] and 6.2 [4.7–7.6] per 100,000 males, respectively in 2016.

## Mortality-to-incidence ratio (MIR) in 1990 and 2016

MIR exhibited different patterns of association with HDI in comparison to the incidence and mortality with all nine low-HDI countries present at the top (Table 1). Seven low-HDI countries registered MIR in excess of 1 in 2016 with Cote d'Ivoire recording the highest MIR of 1.22 followed by Zimbabwe and Ethiopia with MIR of 1.17 and 1.13, respectively, in 2016. Nineteen of the bottom 20 countries belonged to the very high–HDI category with the USA recording the lowest MIR of 0.10 in 2016. Between 1990 and 2016, there was not much improvement recorded in low HDI-countries; a significant reduction, however, was recorded in other HDI categories (Fig. 3).

**Table 1** Country-wise prostate cancer burden in 2016

| Location | Incidence | Deaths | ASIR | ASMR | MIR |
|---|---|---|---|---|---|
| High HDI | | | | | |
| Algeria | 1434 [1099–1896] | 857 [638–1178] | 11.74 [9.04–15.62] | 7.57 [5.69–10.38] | 0.6 |
| Belarus | 2959 [2301–3477] | 869 [664–1086] | 59.32 [47.09–69.7] | 19.22 [14.92–23.9] | 0.29 |
| Brazil | 47,681 [45015–64,039] | 19,104 [17893–25,276] | 64.92 [61.36–85.82] | 30.17 [28.28–39.12] | 0.4 |
| Bulgaria | 2299 [1926–2798] | 1000 [799–1281] | 40.34 [33.83–48.35] | 17.74 [14.34–22.53] | 0.44 |
| China | 111,604 [88324–138,314] | 35,327 [26800–44,281] | 16.06 [12.61–20.05] | 6.15 [4.68–7.62] | 0.32 |
| Colombia | 9718 [8452–11,214] | 3359 [2788–3945] | 60.32 [53.09–69.23] | 24.01 [19.97–28.14] | 0.35 |
| Costa Rica | 1386 [1055–1523] | 506 [354–580] | 67.46 [49.78–74.33] | 26.65 [18.21–30.48] | 0.37 |
| Cuba | 7828 [6200–8602] | 3068 [2417–3494] | 102.78 [81.36–112.8] | 40.28 [31.71–46] | 0.39 |
| Dominican Republic | 3011 [2199–3518] | 1665 [1155–2008] | 85.47 [61.99–99.91] | 48.8 [33.86–58.72] | 0.55 |
| Ecuador | 2464 [2203–3167] | 1190 [1020–1492] | 47.02 [42.07–60.68] | 24.28 [20.84–30.16] | 0.48 |
| Iran | 7411 [4990–8640] | 2766 [1773–3559] | 31.21 [21.32–36.6] | 13.48 [8.78–17.19] | 0.37 |
| Jamaica | 1432 [789–1752] | 794 [416–1087] | 113.08 [62.31–138.58] | 61.98 [32.56–84.84] | 0.55 |
| Kazakhstan | 1172 [830–1352] | 443 [304–568] | 22.37 [15.63–25.76] | 10.15 [6.98–12.85] | 0.38 |
| Lebanon | 1810 [1143–2188] | 534 [324–698] | 73.53 [46.54–88.62] | 24.41 [14.92–31.56] | 0.29 |
| Malaysia | 2104 [1668–2553] | 721 [551–888] | 22.92 [17.83–27.29] | 9.32 [6.91–11.21] | 0.34 |
| Mexico | 25,183 [23725–31,533] | 7261 [6725–8997] | 60.34 [56.85–75.3] | 18.75 [17.36–23.2] | 0.29 |
| Panama | 1607 [1169–1790] | 403 [283–485] | 106.9 [77.51–119.29] | 27.77 [19.39–33.38] | 0.25 |
| Peru | 4712 [3835–6183] | 2194 [1633–3008] | 46.63 [38.02–61.17] | 23.49 [17.55–32.35] | 0.47 |
| Serbia | 2975 [2363–3874] | 1278 [1049–1709] | 47.33 [37.74–61.44] | 21.31 [17.56–28.64] | 0.43 |
| Sri Lanka | 1165 [948–1565] | 547 [392–798] | 13.54 [11.08–18.41] | 7.25 [5.25–10.6] | 0.47 |
| Thailand | 6057 [3882–7180] | 2597 [1703–3244] | 17.71 [11.21–21.02] | 8.4 [5.45–10.49] | 0.43 |
| Tunisia | 1019 [602–1234] | 558 [302–776] | 24.28 [14.24–29.6] | 14.55 [7.98–20.19] | 0.55 |
| Turkey | 9301 [7575–11,784] | 3694 [2742–4810] | 33.67 [27.24–42.43] | 14.91 [11.08–19.35] | 0.4 |
| Ukraine | 9872 [8555–12,185] | 3788 [2743–5297] | 38.12 [33.09–47.61] | 15.27 [11.19–21.97] | 0.38 |
| Uruguay | 1624 [1428–1836] | 750 [640–866] | 84.64 [74.63–95.31] | 38.79 [33.24–44.62] | 0.46 |
| Venezuela | 10,341 [8106–11,543] | 2821 [2097–3562] | 108.26 [83.95–121.22] | 35.6 [26.71–44.67] | 0.27 |
| Low HDI | | | | | |
| Cote d'Ivoire | 1564 [773–2025] | 1913 [909–2588] | 44.55 [20.86–57.91] | 59.24 [27.09–79.97] | 1.22 |
| Democratic Republic of the Congo | 2468 [1447–3093] | 2584 [1530–3429] | 26.17 [15.1–32.84] | 29.57 [17.55–39.85] | 1.05 |
| Ethiopia | 6186 [2488–8295] | 7020 [2501–10,347] | 37.35 [14.97–49.74] | 45.1 [16.2–65.27] | 1.13 |
| Haiti | 1357 | 1274 | 63.93 | 64.11 | 0.94 |

**Table 1** (continued)

| Location | Incidence | Deaths | ASIR | ASMR | MIR |
|---|---|---|---|---|---|
| | [996–1693] | [917–1694] | [46.65–79.18] | [47.08–84.21] | |
| Madagascar | 1432 | 1313 | 39.61 | 38.79 | 0.92 |
| | [584–1915] | [472–1975] | [16.07–52.79] | [14.08–56.92] | |
| Nigeria | 15,474 | 11,410 | 58.63 | 50.8 | 0.74 |
| | [6348–20,883] | [4812–17,297] | [26.27–77.54] | [23.98–72.99] | |
| Tanzania | 3439 | 3732 | 40.97 | 47.49 | 1.09 |
| | [1388–4655] | [1399–5393] | [16.43–55.38] | [17.99–68.17] | |
| Uganda | 2892 | 3165 | 64.17 | 75.5 | 1.09 |
| | [2255–3519] | [2402–3991] | [50.16–78.68] | [57.62–95.39] | |
| Zimbabwe | 1304 | 1533 | 62.87 | 78.19 | 1.17 |
| | [610–1671] | [666–2185] | [29.24–80.26] | [34.77–109.54] | |
| Medium HDI | | | | | |
| Bangladesh | 2897 | 1949 | 7.21 | 5.15 | 0.67 |
| | [2260–3999] | [1502–2756] | [5.75–9.9] | [4.02–7.2] | |
| Bolivia | 1378 | 920 | 41.69 | 28.65 | 0.67 |
| | [1042–1710] | [670–1186] | [31.44–51.74] | [20.97–36.82] | |
| Egypt | 2855 | 1496 | 12.9 | 8.07 | 0.52 |
| | [2322–4230] | [1090–2330] | [9.98–20.04] | [5.68–13.15] | |
| Ghana | 1054 | 1096 | 24.95 | 29.53 | 1.04 |
| | [862–1494] | [874–1565] | [20.62–34.78] | [23.77–40.74] | |
| Guatemala | 1247 | 823 | 34.3 | 23.95 | 0.66 |
| | [1098–1424] | [588–1077] | [30.06–39.34] | [17.33–31.02] | |
| India | 32,726 | 20,954 | 9.01 | 6.7 | 0.64 |
| | [26011–39,749] | [16143–25,842] | [7.1–10.84] | [5.1–8.05] | |
| Indonesia | 15,732 | 7336 | 24.29 | 14 | 0.47 |
| | [10337–18,450] | [4635–8854] | [15.87–28.48] | [8.92–16.87] | |
| Kenya | 1174 | 823 | 16.95 | 13.36 | 0.7 |
| | [695–1339] | [466–1050] | [10.02–19.31] | [7.74–16.9] | |
| Morocco | 1600 | 933 | 14.06 | 9.14 | 0.58 |
| | [938–1987] | [576–1160] | [8.83–17.01] | [6.1–11.13] | |
| Myanmar | 2572 | 1756 | 19.81 | 15.65 | 0.68 |
| | [1688–3075] | [1079–2176] | [12.82–23.77] | [9.6–19.34] | |
| Pakistan | 5794 | 3308 | 12.9 | 8.14 | 0.57 |
| | [3099–7364] | [1864–4315] | [7.33–16.11] | [4.93–10.27] | |
| Paraguay | 1004 | 526 | 50.58 | 29.23 | 0.52 |
| | [667–1160] | [346–653] | [33.7–58.48] | [19.34–36.09] | |
| Philippines | 6397 | 3534 | 32.17 | 21.97 | 0.55 |
| | [5430–7306] | [2804–4261] | [27.31–36.63] | [17.61–26.5] | |
| South Africa | 9736 | 3825 | 77.2 | 36.55 | 0.39 |
| | [7789–12,031] | [3012–4836] | [60.81–94.72] | [28.83–45.73] | |
| Vietnam | 2853 | 1820 | 10.32 | 7.03 | 0.64 |
| | [1964–3363] | [1220–2201] | [7.01–12.24] | [4.71–8.49] | |
| Very high HDI | | | | | |
| Argentina | 12,768 | 4936 | 66.07 | 26.97 | 0.39 |
| | [11391–15,740] | [4263–6176] | [58.92–81.93] | [23.34–33.92] | |
| Australia | 28,519 | 3987 | 174.1 | 23.8 | 0.14 |
| | [23426–35,056] | [3424–4921] | [144.16–213.37] | [20.49–29.2] | |
| Austria | 8304 | 1405 | 124.19 | 20.5 | 0.17 |
| | [7075–9671] | [1096–1629] | [106.9–143.6] | [15.99–23.63] | |
| Belgium | 10,182 | 2079 | 116.84 | 22.56 | 0.2 |
| | [8775–15,573] | [1718–3456] | [100.66–179.03] | [18.75–37.7] | |
| Canada | 28,020 | 5652 | 107.46 | 22.05 | 0.2 |
| | [22636–36,426] | [4834–7161] | [86.99–139.99] | [18.92–27.96] | |
| Chile | 6710 | 2428 | 77.16 | 29.41 | 0.36 |
| | [5093–7505] | [1639–3222] | [58.22–86.41] | [19.77–38.95] | |
| Croatia | 2364 | 819 | 74.19 | 26.31 | 0.35 |
| | [1769–2923] | [602–986] | [55.37–90.33] | [19.26–31.52] | |
| Czech Republic | 6913 | 1590 | 88.77 | 22.53 | 0.23 |
| | [4920–8627] | [1144–1855] | [63.12–109.72] | [15.92–26.1] | |
| Denmark | 5261 | 1404 | 112.82 | 31.4 | 0.27 |
| | [3535–5988] | [952–1691] | [76.18–127.8] | [21.26–37.61] | |

**Table 1** (continued)

| Location | Incidence | Deaths | ASIR | ASMR | MIR |
|---|---|---|---|---|---|
| Finland | 6363 [5081–7404] | 997 [767–1171] | 138.06 [109.94–159.54] | 22.37 [17.11–26.07] | 0.16 |
| France | 53,184 [45717–64,514] | 12,761 [10871–15,751] | 104.79 [90.34–126.65] | 22.86 [19.64–28.23] | 0.24 |
| Germany | 92,890 [79412–110,335] | 16,156 [13461–19,988] | 127.63 [109.7–149.48] | 21.37 [17.91–26.27] | 0.17 |
| Greece | 7354 [6419–11,160] | 2178 [1867–3215] | 73.24 [63.87–114.39] | 18.36 [15.7–27.53] | 0.3 |
| Hungary | 3865 [3156–5232] | 1368 [1137–1855] | 57.13 [47.09–76.45] | 21.72 [18.11–29.53] | 0.35 |
| Ireland | 3434 [2878–4002] | 683 [532–834] | 121.54 [101.74–141.59] | 25.51 [20.02–30.96] | 0.2 |
| Israel | 1966 [1730–2740] | 586 [452–903] | 47.75 [41.89–67.16] | 13.74 [10.52–21.36] | 0.3 |
| Italy | 57,726 [49449–69,184] | 10,179 [8550–12,144] | 107.99 [92.32–130.77] | 16.73 [14.22–20] | 0.18 |
| Japan | 73,795 [63205–80,381] | 12,995 [10452–14,202] | 53.77 [46.81–58.79] | 8.76 [7.13–9.56] | 0.18 |
| Latvia | 1048 [578–1241] | 367 [180–450] | 82.57 [45.67–97.64] | 29.64 [14.49–36.02] | 0.35 |
| Lithuania | 3038 [1622–3622] | 534 [263–626] | 163.55 [88.73–195.29] | 28.26 [13.99–32.98] | 0.18 |
| Netherlands | 18,465 [15016–21,482] | 3267 [2605–3828] | 136.71 [110.79–158.55] | 25.57 [20.38–29.76] | 0.18 |
| New Zealand | 4695 [3842–5705] | 701 [592–893] | 150.93 [124.3–182.46] | 22.88 [19.29–29.17] | 0.15 |
| Norway | 6024 [4568–6865] | 1158 [837–1362] | 158.42 [120.46–180.25] | 30.88 [22.36–36.26] | 0.19 |
| Poland | 13,554 [10720–16,192] | 5322 [4280–6322] | 54.79 [43.82–65.64] | 23.17 [18.82–27.4] | 0.39 |
| Portugal | 7784 [6816–10,990] | 2256 [1927–3129] | 90.52 [79.45–129.46] | 24.55 [21.08–34.15] | 0.29 |
| Romania | 8201 [6634–10,384] | 2181 [1838–2705] | 59.55 [48.36–75.21] | 15.98 [13.45–19.8] | 0.27 |
| Russian Federation | 41,994 [33488–48,671] | 11,836 [8062–16,085] | 56.34 [45.69–65.29] | 17.79 [12.36–23.9] | 0.28 |
| Saudi Arabia | 1352 [1071–1680] | 403 [296–494] | 21.01 [16.37–25.73] | 8.99 [6.57–10.87] | 0.3 |
| Slovakia | 2128 [1263–2756] | 626 [362–792] | 68.85 [41.04–88.76] | 22.5 [13.11–28.35] | 0.29 |
| Slovenia | 1160 [792–1401] | 449 [276–548] | 72.71 [49.2–87.72] | 29.57 [17.81–35.92] | 0.39 |
| South Korea | 10,943 [7083–13,259] | 2262 [1222–3232] | 38.08 [24.15–46.09] | 9.2 [4.88–12.98] | 0.21 |
| Spain | 24,184 [20966–34,860] | 7355 [6272–10,646] | 65.05 [56.29–94.21] | 17.8 [15.21–25.94] | 0.3 |
| Sweden | 12,792 [10009–14,418] | 2731 [2034–3216] | 153.79 [121.7–172.1] | 31.4 [23.38–36.94] | 0.21 |
| Switzerland | 8497 [6964–9972] | 1513 [1108–2045] | 132.35 [108.26–155.52] | 22.75 [16.71–30.67] | 0.18 |
| UK | 49,432 [45379–55,963] | 13,636 [12471–15,441] | 97.17 [89.54–109.82] | 25.5 [23.46–28.8] | 0.28 |
| USA | 371,672 [348609–485,368] | 35,468 [32984–48,659] | 172.51 [161.85–225.02] | 16.98 [15.8–23.23] | 0.1 |
| Global | 1,435,742 [1293395–1,618,655] | 380,916 [320808–412,868] | 49.93 [44.99–56.05] | 14.92 [12.7–16.15] | 0.27 |
| Puerto Rico | 2579 [2292–3248] | 653 [563–810] | 116.94 [104.05–147.45] | 28.86 [24.92–35.64] | 0.25 |

*Incidence*, all-age incidence numbers; *mortality*, all-age deaths; *ASIR*, age-standardised incidence rate; *ASMR*, age-standardised mortality rate; *MIR*, mortality-to-incidence ratio. Data Source: Global Burden of Disease 2016 study. Countries were categorised into four groups as per HDI value in 2015: very high (HDI > 0.800), high (0.700 < HDI < 0.799), medium (0.550 < HDI < 0.669), and low (HDI < 0.550)

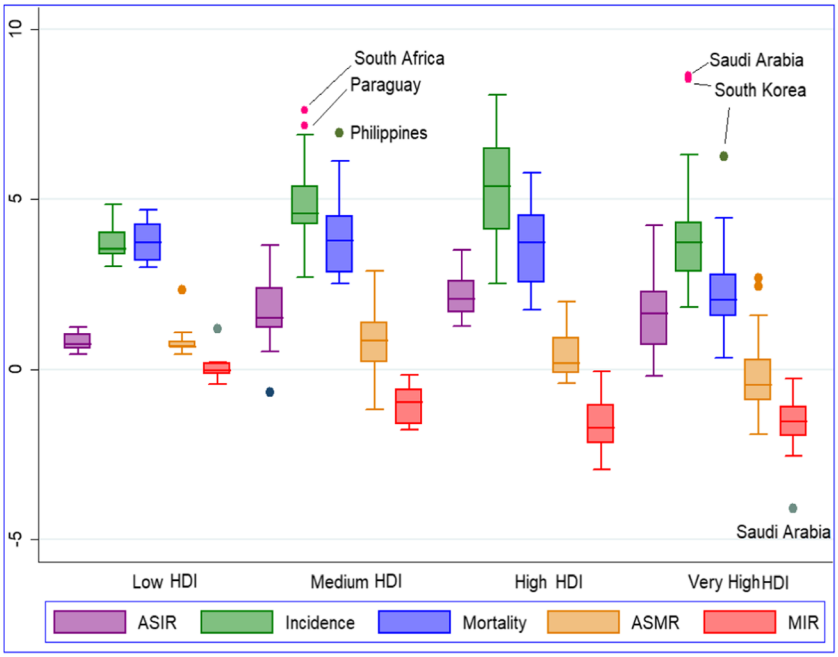



**Fig. 2** HDI-group wise annual percentage change of key indicators of prostate cancer burden over 1990–2016. Incidence, all-age incidence numbers; mortality, all-age deaths; ASIR, age-standardised incidence rate; ASMR, age-standardised mortality rate; MIR, mortality-to-incidence ratio. Data pertains to low HDI (15 countries), medium HDI (18 countries), high HDI (31 countries), and very high HDI (36 countries) for the period 1990 to 2016 and is procured from Global Burden of Disease study 2016. Countries were categorised into four groups as per HDI value in 2015: very high (HDI > 0.800), high (0.700 < HDI < 0.799), medium (0.550 < HDI < 0.669), and low (HDI < 0.550)

### Prostate cancer burden and factors affecting PCa burden

Prostate cancer burden exhibited a different pattern of association with HDI as per the metric used for measuring the cancer burden (Table 3). MIR had a stronger correlation with HDI ($r = -0.91$) than male life expectancy ($r = -0.80$) and income per capita ($r = -0.78$). ASIR was predicted better by income per capita ($R^2 = 0.59$, $p < 0.01$) than the all-age incidence number ($R^2 = 0.12$, $p < 0.01$). The age-standardised mortality rate (ASMR), however, was predicted better by life expectancy ($R^2 = 0.08$, $p < 0.01$) than by income per capita ($R^2 = 0.0001$, $p < 0.01$) and even proved to be a better predictor than the human development index ($R^2 = 0.02$, $p < 0.01$).

## Discussion

In this paper, we examined the burden of prostate cancer in 87 countries and their temporal patterns over 1990 to 2016. Using GBD 2016 data, we showed that the PCa burden increased considerably across the globe in the last three decades with wide geographical inequalities as per development levels. While incidence rates (both all-age numbers and rates) were the highest in developed countries, the mortality rates, however, were on the higher side in low-HDI countries. The developing countries now account for a significant share of incidence and mortality with the share of prostate cancer decreasing in the developed countries (high/very high HDI). Figure 4 illustrates this transition with less-developed countries (low/medium HDI) accounting for 9.9% of incident cases in 2016 versus 7.1% in 1990 and resulted in 23.8% of global cancer deaths against 18% in 1990 indicating higher incidence rates in developed countries but low survival rates in less-developed countries. The disparity between developing and developed countries becomes more noticeable in the case of mortality-to-incidence ratio (MIR), a proxy indicator of a 5-year survival rate [7] than that observed in the case of incidence or mortality with the USA, for instance, recorded mortality-to-incidence ratio of 0.10 and seven of the low-HDI countries registering MIR in excess of 1 (more than a tenfold difference) in 2016.

The incidence rate of prostate cancer was the highest in developed countries led by Australia with ASIR of 174.1/100,000 followed by the USA with ASIR of 172.5/100,000 in 2016. High incidence, both all-age and age-standardised in developed countries, is attributed to population ageing, increasing life spans, and screening of prostate cancer [15]. Screening using prostate-specific antigen (PSA) testing, in particular, has almost become a routine in developed countries and has been instrumental in detecting PCa in early stages, thereby, resulting in improved survival and reduced mortality rates [16–19]. Early screening using PSA testing or in combination with the digital rectal examination (DRE), however, is also not without limitations such as high false positive rates and overdiagnosis/

**Table 2** Annual percentage change of key indicators: 1990–2016

| Country | ASIR | ASMR | Incidence | Mortality | MIR |
|---|---|---|---|---|---|
| High HDI | | | | | |
| Algeria | 1.5370 | 0.0352 | 6.1951 | 4.6786 | − 1.4280 |
| Belarus | 3.0497 | 1.9444 | 4.0876 | 3.0730 | − 0.9747 |
| Brazil | 2.0030 | 0.1351 | 6.3735 | 4.5349 | − 1.7284 |
| Bulgaria | 2.0767 | 0.9355 | 2.6348 | 1.7637 | − 0.8488 |
| China | 2.2075 | − 0.2099 | 5.7682 | 3.1788 | − 2.4482 |
| Colombia | 1.6526 | − 0.1664 | 5.5033 | 3.5321 | − 1.8684 |
| Costa Rica | 2.0644 | 0.1097 | 6.7295 | 4.8350 | − 1.7750 |
| Cuba | 1.8735 | 0.7753 | 4.0909 | 2.9783 | − 1.0688 |
| Dominican Republic | 2.8552 | 1.6420 | 6.6670 | 5.7726 | − 0.8385 |
| Ecuador | 1.2697 | − 0.4062 | 5.2661 | 3.5576 | − 1.6230 |
| Iran | 3.5092 | 1.0339 | 7.6685 | 5.0377 | − 2.4434 |
| Jamaica | 2.5983 | 1.9901 | 4.5921 | 4.0185 | − 0.5485 |
| Kazakhstan | 1.8702 | 0.7167 | 3.3073 | 2.0032 | − 1.2624 |
| Lebanon | 3.1693 | 0.1119 | 8.0772 | 4.8950 | − 2.9444 |
| Malaysia | 2.6567 | − 0.0735 | 7.1554 | 4.1233 | − 2.8296 |
| Mexico | 2.4963 | 0.2146 | 6.2223 | 3.9017 | − 2.1846 |
| Panama | 2.5969 | 0.2532 | 6.5112 | 4.3381 | − 2.0403 |
| Peru | 1.3215 | − 0.0015 | 5.2190 | 4.1004 | − 1.0632 |
| Serbia | 1.3281 | 0.4903 | 3.1566 | 2.5563 | − 0.5820 |
| Sri Lanka | 1.7452 | − 0.3367 | 4.3141 | 2.0258 | − 2.1937 |
| Thailand | 1.9844 | − 0.1334 | 6.4189 | 4.3276 | − 1.9651 |
| Tunisia | 1.5165 | − 0.3082 | 5.0557 | 3.2425 | − 1.7259 |
| Turkey | 2.1336 | − 0.1153 | 5.0437 | 2.5085 | − 2.4135 |
| Ukraine | 1.6762 | 1.3357 | 2.5308 | 2.4640 | − 0.0652 |
| Uruguay | 2.1108 | 0.7237 | 3.2429 | 2.1144 | − 1.0931 |
| Venezuela | 3.2980 | 1.2585 | 7.4094 | 5.2107 | − 2.0470 |
| Low HDI | | | | | |
| Cote d'Ivoire | 0.6120 | 0.5772 | 3.7088 | 3.6024 | − 0.1026 |
| Dem. Rep of Congo | 0.4779 | 0.6493 | 4.0655 | 4.2807 | 0.2068 |
| Ethiopia | 0.4487 | 0.6486 | 3.5497 | 3.7347 | 0.1786 |
| Haiti | 0.7482 | 0.8132 | 3.0753 | 3.1202 | 0.0436 |
| Madagascar | 0.7876 | 0.6534 | 3.3821 | 3.1957 | − 0.1803 |
| Nigeria | 0.7063 | 0.4438 | 3.4433 | 2.9967 | − 0.4318 |
| Tanzania | 1.0269 | 0.8232 | 4.8371 | 4.6797 | − 0.1501 |
| Uganda | 1.1690 | 1.0749 | 4.0247 | 3.9821 | − 0.0409 |
| Zimbabwe | 1.2391 | 2.3363 | 3.0327 | 4.2616 | 1.1927 |
| Medium HDI | | | | | |
| Bangladesh | 0.5294 | − 1.1923 | 4.3538 | 2.6201 | − 1.6613 |
| Bolivia | 1.3083 | 0.1909 | 5.3770 | 4.4910 | − 0.8407 |
| Egypt | 1.5038 | − 0.0255 | 4.4732 | 2.6088 | − 1.7845 |
| Ghana | − 0.6882 | − 0.7540 | 2.7137 | 2.5305 | − 0.1784 |
| Guatemala | 3.6480 | 2.7836 | 7.6339 | 6.9690 | − 0.6177 |
| India | 1.1414 | 0.7761 | 4.5904 | 3.8666 | − 0.6921 |
| Indonesia | 2.0606 | 1.4347 | 5.1643 | 4.0762 | − 1.0347 |
| Kenya | 1.2254 | 0.8313 | 4.2493 | 3.6296 | − 0.5944 |
| Morocco | 1.4743 | 0.4662 | 5.2118 | 4.1119 | − 1.0455 |
| Myanmar | 1.6124 | 1.3752 | 4.1724 | 3.6429 | − 0.5083 |
| Pakistan | 1.5770 | 1.1039 | 4.3929 | 3.7858 | − 0.5816 |

**Table 2** (continued)

| Country | ASIR | ASMR | Incidence | Mortality | MIR |
|---|---|---|---|---|---|
| Paraguay | 2.8132 | 1.2018 | 6.9169 | 5.1893 | − 1.6159 |
| Philippines | 3.5621 | 2.8943 | 7.1606 | 6.1234 | − 0.9679 |
| South Africa | 2.3915 | 1.0385 | 4.7190 | 2.9565 | − 1.6831 |
| Vietnam | 1.4199 | 0.2543 | 4.0471 | 2.8474 | − 1.1530 |
| Very high HDI | | | | | |
| Argentina | 1.8956 | 0.3017 | 3.8181 | 2.3581 | − 1.4063 |
| Australia | 1.5229 | − 0.9585 | 4.5671 | 2.5899 | − 1.8909 |
| Austria | 0.5544 | − 0.7122 | 2.6404 | 1.7704 | − 0.8477 |
| Belgium | 0.3448 | − 1.8972 | 2.1041 | 0.3842 | − 1.6845 |
| Canada | 0.2374 | − 1.5985 | 3.0923 | 1.5422 | − 1.5036 |
| Chile | 2.6958 | 0.8928 | 6.2340 | 4.4388 | − 1.6898 |
| Croatia | 1.2041 | 0.2889 | 2.8559 | 2.2125 | − 0.6255 |
| Czech Republic | 1.7677 | 0.0055 | 3.7286 | 2.0422 | − 1.6258 |
| Denmark | 2.1633 | − 0.0293 | 3.7502 | 1.4404 | − 2.2263 |
| Finland | 1.5982 | − 1.1071 | 4.4030 | 1.9276 | − 2.3710 |
| France | 0.8855 | − 1.5033 | 2.9512 | 1.0339 | − 1.8624 |
| Germany | 1.6615 | − 0.7884 | 3.9866 | 1.8847 | − 2.0213 |
| Greece | 1.6684 | − 0.1156 | 3.6237 | 2.6293 | − 0.9595 |
| Hungary | 0.7166 | − 0.8255 | 1.8188 | 0.3243 | − 1.4677 |
| Ireland | 2.0496 | − 0.6039 | 4.2290 | 1.5951 | − 2.5271 |
| Israel | 0.3349 | − 1.4455 | 3.7059 | 2.1005 | − 1.5481 |
| Italy | 1.0676 | − 0.1256 | 2.8900 | 2.6072 | − 0.2748 |
| Japan | 2.8300 | 0.1138 | 6.3137 | 4.1201 | − 2.0634 |
| Latvia | 3.0982 | 2.4482 | 3.8680 | 3.4974 | − 0.3568 |
| Lithuania | 4.2389 | 2.6860 | 5.1480 | 3.9244 | − 1.1637 |
| Netherlands | 1.6428 | − 0.5229 | 4.0780 | 1.9963 | − 2.0001 |
| New Zealand | 0.6429 | − 0.8797 | 3.4024 | 2.4002 | − 0.9692 |
| Norway | 1.8829 | − 0.7809 | 3.3081 | 0.6631 | − 2.5603 |
| Poland | 2.7000 | 0.8459 | 4.8185 | 2.9692 | − 1.7643 |
| Portugal | 0.6925 | − 0.8812 | 2.5879 | 1.6610 | − 0.9035 |
| Romania | 2.5626 | 1.0268 | 3.5632 | 2.3499 | − 1.1715 |
| Russian Federation | 2.3975 | 1.6041 | 3.9786 | 3.4854 | − 0.4743 |
| Saudi Arabia | 3.5521 | 0.2771 | 8.6369 | 4.1770 | − 4.1054 |
| Slovakia | 2.0880 | 0.5503 | 3.7948 | 2.0754 | − 1.6565 |
| Slovenia | 1.7295 | 0.1148 | 4.4834 | 3.1370 | − 1.2887 |
| South Korea | 3.4242 | 0.9575 | 8.5278 | 6.2751 | − 2.0757 |
| Spain | 0.5609 | − 1.1851 | 2.8579 | 1.7760 | − 1.0519 |
| Sweden | 1.3622 | − 0.4195 | 2.6529 | 1.0865 | − 1.5260 |
| Switzerland | 0.3721 | − 1.8175 | 2.4766 | 0.5167 | − 1.9125 |
| UK | 1.3179 | − 0.5054 | 2.9965 | 1.5764 | − 1.3788 |
| USA | − 0.1869 | − 1.8532 | 2.0685 | 0.5382 | − 1.4993 |
| No HDI | | | | | |
| Puerto Rico | 1.1453 | − 0.4420 | 2.6613 | 1.2322 | − 1.3920 |

*Incidence*, all-age incidence numbers; *mortality*, all-age deaths; *ASIR*, age-standardised incidence rate; *ASMR*, age-standardised mortality rate; *MIR*, mortality-to-incidence ratio. Data pertaining to low HDI (15 countries), medium HDI (18 countries), high HDI (31 countries), and very high HDI (36 countries) for the period 1990 to 2016. Data source: Authors' calculation from Global Burden of Disease study 2016. Countries were categorised into four groups as per HDI value in 2015: very high (HDI > 0.800), high (0.700 < HDI < 0.799), medium (0.550 < HDI < 0.669), and low (HDI < 0.550)

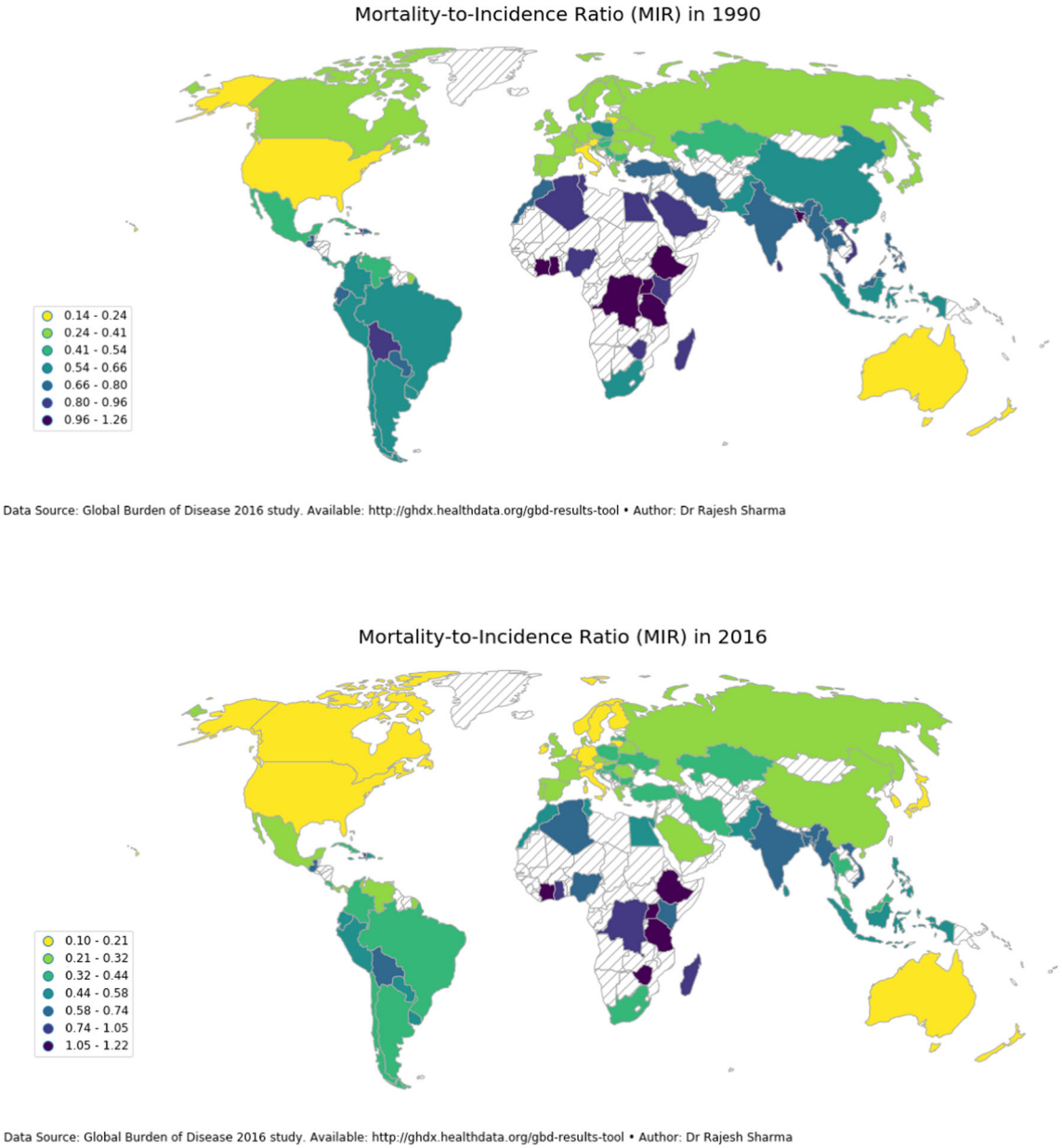



**Fig. 3** Geographical distribution of MIR in sample countries **a** 1990 and **b** 2016. MIR, mortality-to-incidence ratio. MIR is calculated as the ratio of all-age crude mortality rate to all-age crude incidence rate which was procured from Global Burden of Disease study 2016

overtreatment [20–22].[5] Cancer therapeutics is moving away from a reactive approach to a predictive, preventative, and personalised approach where attention is given to predict cancer before its onset, diagnose early or in pre-cancerous state, and risk-stratify patients to provide personalised treatment regimens [24–26]. In this approach, overtreatment and overdiagnosis can be tackled using a multiomic approach to cancer diagnosis [27, 28]. Omics technologies comprising of genomics, transcriptomics, metabolomics, and radiomics have wide-ranging applications in cancer detection, risk-stratification of patients, and improving the accuracy of diagnostic and prognostic procedures [28, 29]. TMPRSS2-ERG gene fusion, for instance, has been demonstrated to have prognostic value, and the fact that it can be measured in urine makes it an ideal biomarker supplementing the PSA test [30–32]. Cheng and Zhan [27] highlighted the limited applicability of use of a single molecule as a biomarker for prediction and personalised treatment for cancer and advocated for multi-parameter strategies in PPPM. In this regard, prostate health index (*phi*), which combines total, free, and (-2) proPSA and approved by the United States Food and Drug Administration (USFDA) for PCa testing, has been demonstrated to have greater specificity in determining clinically significant prostate cancer than PSA testing alone [33–36].

[5] Given high overdiagnosis and overtreatment, USPSTF in 2012 statement recommended omitting PSA screening from routine primary care for men [23].

**Table 3** Correlation of indicators of prostate cancer burden and HDI

| | Income | LE_male | HDI |
|---|---|---|---|
| Overall | | | |
| Incidence | 0.3473 | 0.1942 | 0.2536 |
| Mortality | 0.2330 | 0.1269 | 0.1604 |
| ASMR | 0.0097 | − 0.2910 | − 0.1313 |
| ASIR | 0.7684 | 0.4964 | 0.6567 |
| MIR | − 0.7812 | − 0.7974 | − 0.9092 |
| Very high HDI | | | |
| Incidence | 0.2912 | 0.1267 | 0.2594 |
| Mortality | 0.2376 | 0.1239 | 0.2470 |
| ASMR | 0.1736 | 0.1297 | 0.2491 |
| ASIR | 0.6295 | 0.4920 | 0.6815 |
| MIR | − 0.5431 | − 0.4827 | − 0.6481 |
| High HDI | | | |
| Incidence | − 0.0541 | 0.1230 | − 0.0813 |
| Mortality | − 0.1054 | 0.0812 | − 0.1637 |
| ASMR | − 0.0272 | 0.2337 | 0.1470 |
| ASIR | 0.1865 | 0.3764 | 0.3815 |
| MIR | − 0.4858 | − 0.2572 | − 0.6200 |
| Medium HDI | | | |
| Incidence | 0.1146 | 0.0206 | 0.1328 |
| Mortality | − 0.0069 | 0.0291 | 0.0546 |
| ASMR | 0.5255 | − 0.3623 | 0.3351 |
| ASIR | 0.6791 | − 0.2999 | 0.4842 |
| MIR | − 0.7573 | − 0.2272 | − 0.5751 |
| Low HDI | | | |
| Incidence | 0.8311 | − 0.0905 | 0.2830 |
| Mortality | 0.7966 | − 0.0401 | 0.2453 |
| ASMR | 0.1511 | − 0.0879 | 0.2783 |
| ASIR | 0.3577 | − 0.0430 | 0.4405 |
| MIR | − 0.4842 | − 0.0498 | − 0.3481 |

*Incidence*, all-age incidence numbers; *mortality*, all-age deaths; *ASIR*, age-standardised incidence rate; *ASMR*, age-standardised mortality rate; *MIR*, mortality-to-incidence ratio; *income*, gross national income (GNI) per capita; *LE_male*, life expectancy of males (at birth); *HDI*, human development index. Data pertains to low HDI (15 countries), medium HDI (18 countries), high HDI (31 countries), and very high HDI (36 countries) for the period 1990 to 2016 and is procured from Global Burden of Disease study 2016. Countries were categorised into four groups as per HDI value in 2015: very high (HDI > 0.800), high (0.700 < HDI < 0.799), medium (0.550 < HDI < 0.669), and low (HDI < 0.550)

A key component of PPPM is targeted therapy which seeks to offer individualised treatment options on the basis of grade and stage of the tumour along with patients' characteristics [24–26]. Treatment paradigms for prostate cancer range between active surveillance (for indolent tumour), radiation therapy (external-beam RT, brachytherapy, etc.), surgical therapy (radical prostatectomy, pelvic lymph node dissection, etc.), and androgen deprivation therapy (ADT) or a combination of these. Again, the role of PPPM becomes pertinent in this respect to provide personalised treatment out of the basket of treatment options on the basis of the stage of diagnosis, risk-category, estimated life expectancy of the patient, Gleason score,[6] PSA level and velocity, and familial history of prostate cancer along with other comorbid conditions [4].

Although the aetiology of prostate carcinoma is not yet completely understood, there are a few well-established risk factors of prostate cancer such as ageing, ethnicity (African-American), family history, and behavioural and environmental risk factors [15, 37–39]. Ageing is found to be the leading risk factor of prostate cancer with incidence rates being higher in countries with a high proportion of the elderly population, whereas MIR and age-standardised mortality rate were higher in demographically younger countries with lesser developed healthcare system (Fig. 5). While HDI was positively associated with ASIR, it had a strong negative effect on MIR. The benefit of human development in developed countries is that majority of the PCa detected in these countries is in early stages whereas advanced tumour is detected in developing countries resulting in poor prognosis and survival. The proactive approach of PPPM can further help in early detection and boosting survival rates of PCa patients. A developed country, therefore, is in a relatively better position to adopt PPPM practices of prevention, offer multiomic diagnostic and prognostic approaches and thus can offer tailor-made solutions for individualised treatments.

There was a high burden of prostate cancer in Africa exemplified by the highest ASMR of Zimbabwe (78.2/100,000) followed by Uganda with ASMR of 75.5/100,000 in 2016. The low survival rates in these countries were evidenced by high MIR (a proxy for survival) with many of these countries recording MIR greater than 1 in 2016. The high genetic susceptibility of prostate carcinoma in men of African origin (be it African-Americans or men in African countries), as well as high mortality and low survival rates, has been documented before [40]. There is a multitude of factors that govern such differences in the survival rates such as lack of screening, late disease presentation, limited healthcare access, and low awareness and low socioeconomic status [30]. Apart from that, there are few inherent weaknesses of the healthcare system of these countries. Figure 6 illustrates this point, showing that current health expenditures (% age of GDP) are lower, out-of-pocket expenditure (OOP) as a percentage of total health expenditure is higher, density of specialist surgeons is lower, and universal health coverage is lower in countries having high mortality-to-incidence ratio (MIR) which are evidently the less-developed countries.

[6] The WHO endorsed the Gleason grading system in the 2004 classification of prostate cancer.

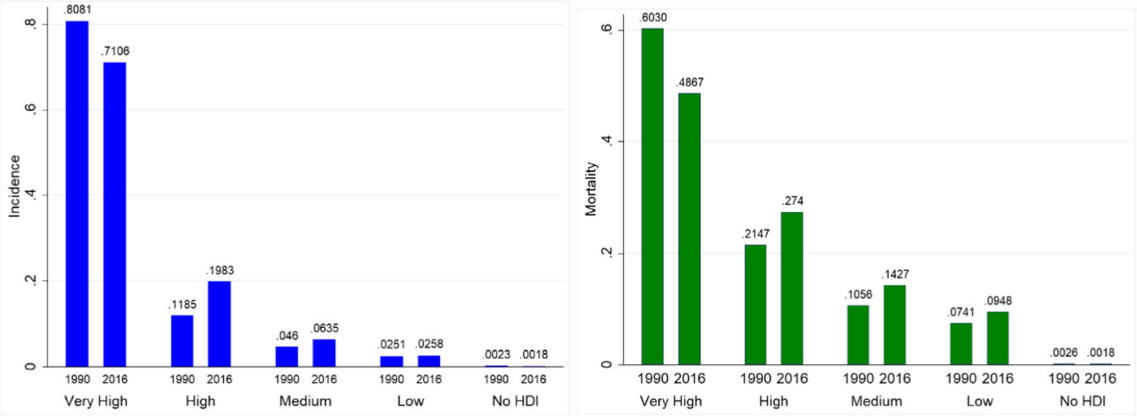



**Fig. 4** Percentage share of HDI categories in prostate cancer burden **a** incidence and **b** mortality in 2016. Incidence, all-age incidence numbers; mortality, all-age deaths. Data pertains to low HDI (15 countries), medium HDI (18 countries), high HDI (31 countries), and very high HDI (36 countries) for the period 1990 to 2016 and is procured from Global Burden of Disease study 2016. Countries were categorised into four groups as per HDI value in 2015: very high (HDI > 0.800), high (0.700 < HDI < 0.799), medium (0.550 < HDI < 0.669), and low (HDI < 0.550)

## Limitations

First, the estimates of prostate cancer burden were procured from the Global Burden of Disease 2016 study and the quality of estimates crucially depends upon data availability from cancer registries. As cancer registries provide incomplete coverage and lack completeness in many low-income countries, GBD employs data available spatially and temporally from neighbouring locations. Lack of cancer registry data yields wide uncertainty intervals which may be too wide to provide meaningful policy implications in these countries. Second, due to under-reporting or misclassification of cancer deaths in less-

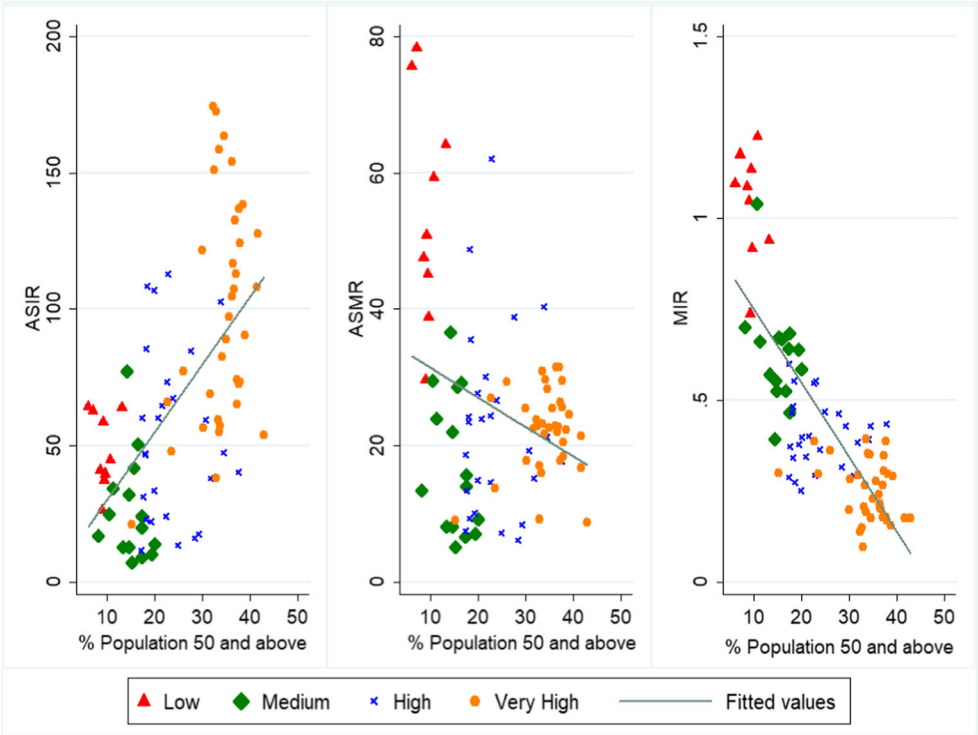



**Fig. 5** Prostate cancer burden versus the proportion of population aged 50 years and above in 2016. ASIR, age-standardised incidence rate; ASMR, age-standardised mortality rate; MIR, mortality-to-incidence ratio. Data pertains to low HDI (15 countries), medium HDI (18 countries), high HDI (31 countries), and very high HDI (36 countries) for the period 1990 to 2016. Countries were categorised into four groups as per HDI value in 2015: very high (HDI > 0.800), high (0.700 < HDI < 0.799), medium (0.550 < HDI < 0.669), and low (HDI < 0.550). Data source: ASIR, ASMR data is from Global Burden of Disease 2016 study; proportion of population aged 50 years and above is from World Development Indicators (WDI) database of World Bank

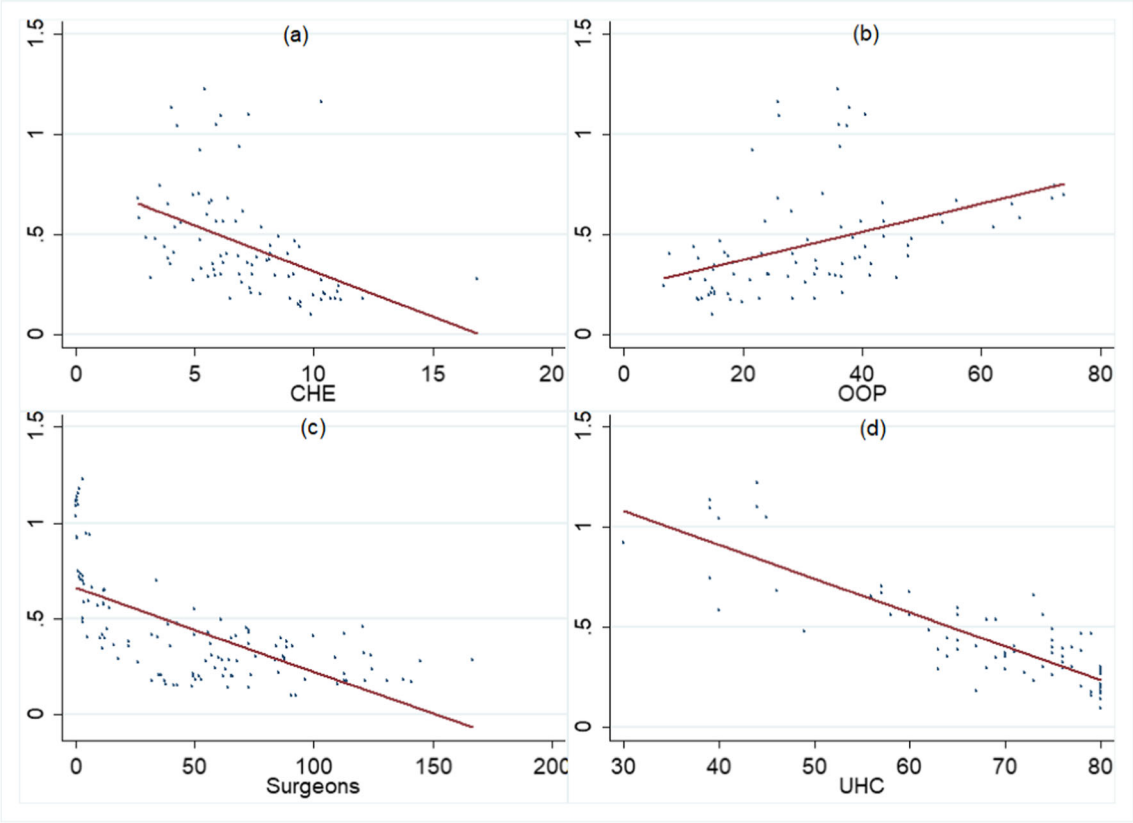



**Fig. 6** Relation between prostate cancer mortality-to-incidence ratio (MIR) and country-specific healthcare parameters **a** MIR vs CHE, **b** MIR vs OOP, **c** MIR vs surgeon workforce, **d** MIR vs UHC. MIR, mortality-to-incidence ratio; CHE, current health expenditure as %age of GDP; OOP, out-of-pocket expenditure as %age of total health expenditure; surgeon workforce, number of specialist surgeons per 100,000 population; UHC, universal health coverage service index. The data corresponding to MIR, CHE, UHC, and OOP pertains to the year 2015. The data of the surgeon workforce pertains to years ranging from 2010 to 2015 depending upon country-level availability of data. All the data is procured from health nutrition and population statistics database of World Bank

developed countries, the estimates utilised and reported here may be downward biased. Third, as is well known, cancer detected early (clinical stage 1) is less severe and life-threatening as compared to a case encountered later (stage 3 or 4), but there are no estimates of cases diagnosed in various clinical stages; therefore, we had to rely on the total number of cancer cases (irrespective of stage) to examine the burden of prostate cancer. Finally, we employed MIR as a proxy for survival rate which is not an exact measure of survival rate but a proxy of survival rate. The usefulness of MIR as a predictor of a 5-year survival rate crucially depends upon the quality of data and data completeness [32]. As this study contains estimates of many developing countries in which data from cancer registries is far from complete, therefore, MIR can only serve as a crude measure of survival rate and can only indicate the direction of relative survival in different countries.

## Conclusions and expert recommendations

The burden of prostate cancer is rising globally with the incidence being higher in developed parts of the world against high mortality rate and MIR (a proxy for survival) in less-developed countries particularly African regions. Greater human development leads to early detection of malignancy, boosts awareness levels, and offers better treatment paradigms in developed countries. On the basis of past trends, the burden of PCa is expected to escalate further in the future due to a sedentary lifestyle, eating habits and obesity, etc. However, due to the lack of early detection, late disease presentation, and the lack of active surveillance systems in less-developed countries, it is expected to put pressure on the resource-constrained healthcare systems. A patient afflicted in these countries, therefore, is in a relatively disadvantageous state due to late diagnosis, access to treatment options, and a high cost of treatment which is incurred mostly from own pocket. These factors result in low survival rates and great economic hardships for a cancer patient in these countries. As few individuals are more likely to get cancer than others and cancer treatment is more costly than its prevention, the focus of medicine must shift from a reactive to a proactive one in line with PPPM principles of prevention, diagnostic, and prognostic accuracy, and offering personalised treatments rather than offering one size fits all type of solutions. In order to prevent the

disease onset, risk-stratify patients, and provide personalised medicine, the capability of both healthcare system as well as patients has to be boosted. As survival rates are positively associated with the human development level of a country, the PPPM strategies are more likely to be adopted in developed countries but must be promoted in low-resource economies too, where survival rates are low.

**Acknowledgements** We thank the Institute of Health Metrics and Evaluation (IHME) for making Global Burden of Disease (GBD) data pertaining to prostate cancer available in the public domain.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical approval** The research is conducted using data available in the public domain and does not include any human participants and/or animals.

## References

1. Global Burden of Disease Cancer Collaborators. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 1990 to 2016 a systematic analysis for the Global Burden of Disease Study. JAMA Oncol. 2018;4:1553. https://doi.org/10.1001/jamaoncol.2018.2706.
2. Dorr M, Holzel D, Schubert-Fritschle G, Engel J, Schlesinger-Raab A. Changes in prognostic and therapeutic parameters in prostate cancer from an epidemiological view over 20 years. Oncol Res Treat. 2015;38:8–14.
3. Winter A, Sirri E, Jansen L, Wawroschek F, Kieschke J, Castro FA, et al. Comparison of prostate cancer survival in Germany and the USA: can differences be attributed to differences in stage distributions? BJU Int. 2017;119:550–9.
4. Mohler JL, Bahnson RR, Boston B, Busby JE, D'Amico A, Eastham JA, et al. Prostate cancer: clinical practice guidelines in oncology. J Natl Compr Cancer Netw. 2010;8:162–200.
5. Quinn M, Babb P. Patterns and trends in prostate cancer incidence, survival, prevalence and mortality. Part II: individual countries. BJU Int. 2002 Jul 1;90(2):174–84.
6. GBD 2016 Cause of Death Collaborators. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet. 2017;390:1151–210.
7. Vostakolaei F, Karim-Kos HE, Janssen-Heijnen ML, Visser O, Verbeek AL, Kiemeney LA. The validity of the mortality to incidence ratio as a proxy for site-specific cancer survival. Eur J Pub Health. 2010;21:573–7.
8. Sunkara V, Hebert JR. The colorectal cancer mortality-to-incidence ratio as an indicator of global cancer screening and care. Cancer 2015. 2015;121(10):1563–9.
9. Chen SL, Wang SC, Ho CJ, Kao YL, Hsieh TY, Chen WJ, et al. Prostate cancer mortality-to-incidence ratios are associated with cancer care disparities in 35 countries. Sci Rep. 2017 Jan 4;7:40003.
10. Sharma R. Breast cancer incidence, mortality and mortality-to-incidence ratio (MIR) are associated with human development, 1990–2016: evidence from Global Burden of Disease Study 2016. Breast Cancer. 2019:1–18. https://doi.org/10.1007/s12282-018-00941-4.
11. Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2016 (GBD 2016) results. Seattle, United States: Institute for Health Metrics and Evaluation (IHME), 2017. Available from http://ghdx.healthdata.org/gbd-results-tool (Accessed: August 14–16, 2018).
12. GBD 2016 Disease and Injury, Incidence and Prevalence collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet. 2017;390:1211–59.
13. Foreman KJ, Lozano R, Lopez AD, Murray CJ. Modeling causes of death: an integrated approach using CODEm. Popul Health Metrics. 2012;10(1).
14. Human Development Database: http://hdr.undp.org/en/data# (Accessed: 30.6.2018 and 1.7.2018).
15. Crawford ED. Epidemiology of prostate cancer. Urology. 2003;62(6):3–12.
16. Crawford ED, Robert Grubb AB III, Andriole GL Jr, Chen MH, Izmirlian G, Berg CD, et al. Comorbidity and mortality results from a randomized prostate cancer screening trial. J Clin Oncol. 2011;29:355–61.
17. Roobol MJ, Kranse R, Bangma CH, van Leenders AG, Blijenberg BG, van Schaik RH, et al. Screening for prostate cancer: results of the Rotterdam section of the European randomized study of screening for prostate cancer. Eur Urol. 2013;64:530–9.
18. Schröder FH, Hugosson J, Roobol MJ, Tammela TL, Zappa M, Nelen V, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. Lancet. 2014;384:2027–35.
19. Arnsrud RG, Holmberg E, Lilja SJ, Hugosson J. Opportunistic testing versus organized prostate-specific antigen screening: outcome after 18 years in the Göteborg randomized population-based prostate cancer screening trial. Eur Urol. 2015;68:354–60.
20. Smith RA, Cokkinides V, Brawley OW. Cancer screening in the United States, 2008: a review of current American Cancer Society guidelines and cancer screening issues. CA Cancer J Clin. 2008;58:161–79.
21. Barry MJ, Mulley AJ Jr. Why are a high overdiagnosis probability and a long lead time for prostate cancer screening so important? J Natl Cancer Inst. 2009;101:362–3.
22. Grossman DC, Curry SJ, Owens DK, Bibbins-Domingo K, Caughey AB, Davidson KW, et al. Screening for prostate cancer: US Preventive Services Task Force recommendation statement. Jama. 2018;19:1901–13.
23. Moyer VA. Screening for prostate cancer: US Preventive Services Task Force recommendation statement. Ann Intern Med. 2012;157:120–34.
24. Grech G, Zhan X, Yoo BC, Bubnov R, Hagan S, Danesi R, et al. EPMA position paper in cancer: current overview and future perspectives. EPMA J. 2015;6:9. https://doi.org/10.1186/s13167-015-0030-6.
25. Golubnitschaja O, Baban B, Boniolo G, Wang W, Bubnov R, Kapalla M, et al. Medicine in the early twenty-first century: paradigm and anticipation - EPMA position paper 2016. EPMA J. 2016;7:23.
26. Janssens JP, Schuster K, Voss A. Preventive, predictive, and personalized medicine for effective and affordable cancer care. EPMA J. 2018;9:113–23. https://doi.org/10.1007/s13167-018-0130-1.
27. Cheng T, Zhan X. Pattern recognition for predictive, preventive, and personalized medicine in cancer. EPMA J. 2017;8:51–60. https://doi.org/10.1007/s13167-017-0083-9.
28. Lu M, Zhan X. The crucial role of multiomic approach in cancer research and clinically relevant outcomes. EPMA J. 2018;9:77–102. https://doi.org/10.1007/s13167-018-0128-8.

29. Horgan RP, Kenny LC. 'Omic' technologies: genomics, transcriptomics, proteomics and metabolomics. Obstet Gynaecol. 2011;13:189–95.
30. Tomlins SA, Day JR, Lonigro RJ, Hovelson DH, Siddiqui J, Kunju LP, et al. Urine TMPRSS2: ERG plus PCA3 for individualized prostate cancer risk assessment. Eur Urol. 2016;70:45–53.
31. Leyten GH, Hessels D, Jannink SA, Smit FP, de Jong H, Cornel EB, et al. Prospective multicentre evaluation of PCA3 and TMPRSS2-ERG gene fusions as diagnostic and prognostic urinary biomarkers for prostate cancer. Eur Urol. 2014;65:534–42.
32. Salagierski M, Schalken JA. Molecular diagnosis of prostate cancer: PCA3 and TMPRSS2: ERG gene fusion. J Urol. 2012;187:795–801.
33. Catalona WJ, Partin AW, Sanda MG, Wei JT, Klee GG, Bangma CH, et al. A multi-center study of [−2] pro-prostate-specific antigen (PSA) in combination with PSA and free PSA for prostate cancer detection in the 2.0 to 10.0 ng/mL PSA range. J Urol. 2011;185:1650–5.
34. Huang YQ, Sun T, Zhong WD, Wu CL. Clinical performance of serum [-2] proPSA derivatives, % p2PSA and PHI, in the detection and management of prostate cancer. Am J Clin Exp Urol. 2014;2:343.
35. Loeb S, Sanda MG, Broyles DL, Shin SS, Bangma CH, Wei JT, et al. The prostate health index selectively identifies clinically significant prostate cancer. J Urol. 2015;193:1163–9.
36. Tosoian JJ, Druskin SC, Andreas D, Mullane P, Chappidi M, Joo S, et al. Use of the Prostate Health Index for detection of prostate cancer: results from a large academic practice. Prostate Cancer Prostatic Dis. 2017;20:228–33.
37. Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, et al. Environmental and heritable factors in the causation of cancer–analyses of cohorts of twins from Sweden, Denmark, and Finland. N Engl J Med. 2000;343:78–85.
38. Mucci LA, Hjelmborg JB, Harris JR, Czene K, Havelick DJ, Scheike T, et al. Familial risk and heritability of cancer among twins in Nordic countries. Jama. 2016;315:68–76.
39. McGinley KF, Tay KJ, Moul JW. Prostate cancer in men of African origin. 2016 Nat Rev Urol. 2016;13:99–107.
40. Parkin DM, Bray F. Evaluation of data quality in the cancer registry: principles and methods part II. Completeness. Eur J Cancer. 2009;45:756–64.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 8 (2019) pp. 1940-1946
© Research India Publications. http://www.ripublication.com

# Variation of flow parameters and Magnetic flux in Viscoelastic Magnetic Abrasive Finishing Process

**K. Srinivas[1], Q. Murtaza[2], A.K Aggarwal[3]**

[1] *Assistant Professor Delhi Technological University New Deli, India.*

[2, 3] *Professor Delhi Technological University New Delhi, India.*

**Abstract**

The advancement of manufacturing technology give rise to the development of various non-traditional machining processes which have their specific applications. Magnetic abrasive finishing is a non- traditional machining process mainly developed to accomplish the finishing processes by considering the magnetic effect into account. Various ferromagnetic particles and viscous fluids are being used for the purpose to provide a better finishing to the work piece. In this paper a serious effort is made to study the distribution of magnetic field intensity on the work piece along with the computation of force and torque value experienced by the ferromagnetic particle and abrasive particle also a plot of variation of magnetic field intensity against the input current is obtained using ANSYS Maxwell. The distribution of various flow parameters is also studied by using ANSYS Fluent 16. After going through the simulated result, it was concluded that the suggested viscous fluid is suitable for the process and an optimised value of parameters is required for controlling the machining process.

**Keyword:** Viscoelastic Magnetic Abrasive Finishing, Magnetic Abrasive Finishing, Magnetic Field

## INTRODUCTION

Magnetic abrasive finishing involves introduction of magnetic abrasive particle filled between the work piece and the magnetic poles. In this process work piece is made one of the pole with the magnet as another pole. This magnetic field results in the alignment of magnetic particle thus making a flexible magnetic abrasive brush which behaves like a multipoint cutting tool for finishing process. Shimamura et al. [1] , [2] modelled and tested that magnetic abrasive particles are subjected to the pressure for machining and that pressure is a function of magnetic flux density, number of abrasive particles and the permeability of abrasive medium. Fox et al [3] observed that the magnetic force (function of the volume and magnetic susceptibility of the ferrous particles in the magnetic field, magnetic field intensity and the gradients at the finishing area). Kim and Choi [4] mathematically modelled the magnetic field gradient which resulted in the attractive force between the machining pressure in air gap and abrasives. Magnetic potential was computed using Maxwell's equations and FEM by mathematical modelling by Gurvinder and Yadav [5] which computed magnetic potential distribution within the gap between tool bottom surface and work piece top surface. Mulik and Pandey [6] determined the effect of process variables on the surface finish obtained during the MAF process using unbounded Fe + SiC magnetic abrasive particles. Surface finish was evaluated the considering the electromagnet rotation in the range of 180-450 rpm and magnetic flux density in the range of 0.02 -0.2T. Srivastav and Komma [7] used thermal analysis of Magnetic abrasive finishing process on the work piece in both transient and steady condition. They took Magnetic stainless steel (AISI-420) work piece as their work piece and the abrasive particle used was chromium oxide abrasives of dimension 387 μm diameter they took magnetic flux density is of 1 Tesla and the rotation given to the tool of 6 mm diameter, was 5835 rpm. Srivastav and Katiyar [8] did thermal simulation on silicon nitride work piece and the temperature distribution of the work tool interface was computed. Thermal analysis was done for both steady and transient condition with the help of ANSYS 15. Singh et all [9] studied the surface temperature finishing surface of mild steel using abrasive finishing process. L9 orthogonal Array was used with current, Rotational speed, Abrasive weight and working gap as the parameters. Fuzzy logic was also used to model the surface temperature which was further validated. Li et all [10] established a mathematical model of interface debonding and proposed a surface modification method for filler particle using silane coupling agent. they analysed the mechanism between surface modified filler particle and the matrix and determine the finishing effect of matrix. They concluded that the surface treatments would improve the processing capability of the viscoelastic magnetic abrasive tools. Li et all [11] developed a finishing setup having new media for finishing the rotatory surface ANSYS Maxwell was used to determine the optimum angle between the poles and the mathematical model was made to compute the MRR in which pressure and velocity based on arced wear model. A MRR computation model was made which predicted MRR as a function of pole rotational speed, Magnetic flux density, cam rotational speed and the ferromagnetic and abrasive particles' diameter. The present scenario, abrasive flowing finishing, grinding, abrasive brush deburring, magnetorheological finishing, manual deburring etc. are broad name in term of finishing but still they possess some limitations in finishing quality and efficiency [12-13]. Li et all [14] proposed a new viscoelastic magnetic Abrasive based on analysis of field characteristic and existing finishing process. They also discussed the motion locus of the abrasives. Mishra et all [15] made a model of magnetic pressure and magnetic intensity

and proposed a magnetic circuit for magnetic abrasive finishing. The aim of the paper is to explore a viscoelastic fluid which could be used in the Viscoelastic magnetic abrasive finishing process so that the combination of magnetic effect and machining may take place effectively. Thus, to accomplish the purpose the authors decided to analyse the solid model on ANSYS Maxwell for the magnetic effect and ANSYS Fluent for the flow parameters.

## MODELLING AND ANALYSIS

The setup is designed in such a way that the hollow work piece is kept between the viscous fluid which is allowed to reciprocate between the work piece by means of pressure difference and the magnetic field is applied around the work piece so the result combination of the two results in the effective machining. The initial model is made on CREO 3.0 software which is exported to the Maxwell. The modal consists of Iron particle which are surrounded by four abrasive particles which are taken as alumina. The work piece taken is of aluminium which support magnetic properties.

## METHODOLOGY FOR VISCOELASTIC MAGNETIC ABRASIVE FINISHING PROCESS

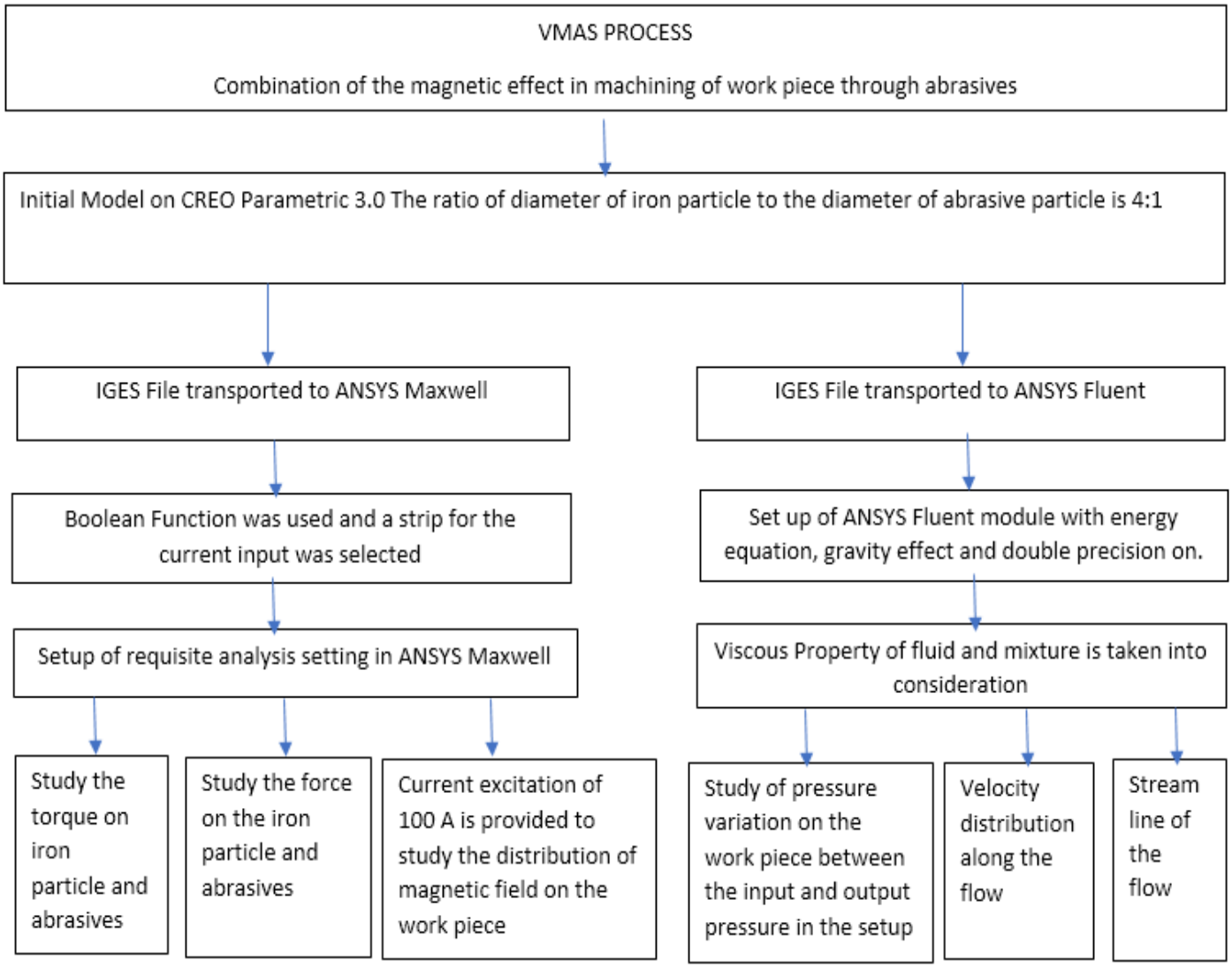



**Fig.1** Methodology of Viscoelastic Magnetic Abrasive Process

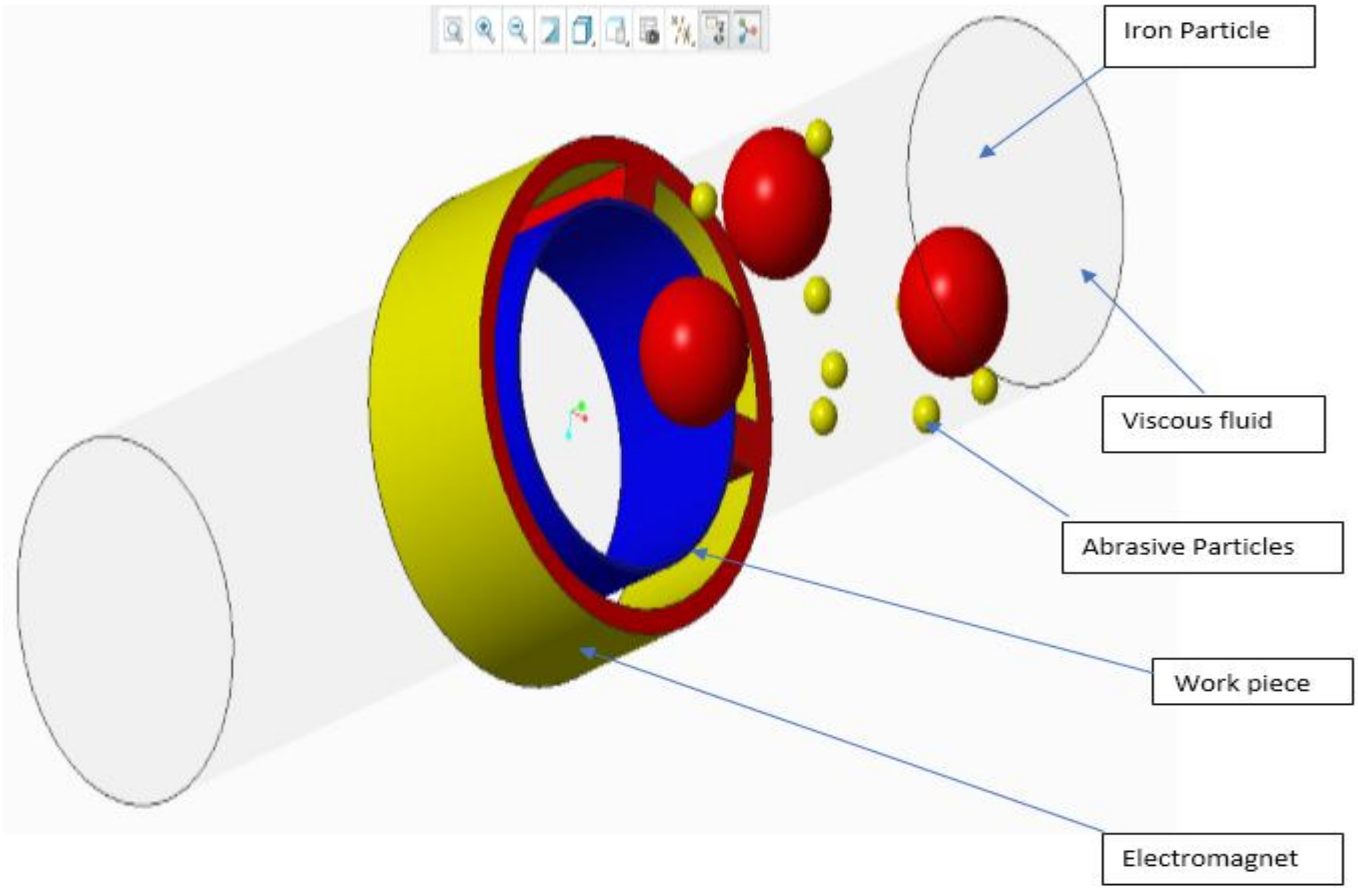



**Fig. 2** CREO Parametric Model Showing Exploded view of component of Set UP

The above figure shows the CREO model of the setup which consist of the iron and Abrasive particles in the ratio of 4 to 1 according to diameter and also 4 abrasive particles associated with each iron particle the transparent part represents the fluid which has specific properties. Blue part is aluminium work piece with magnet around the work piece touching the work piece at two positions and current is supplied through 1 of them. The viscoelastic fluid so chosen had the physical properties which are viscosity as 0.789 Kg/m-s, density 1219 Kg/m$^3$, Thermal conductivity 0.22 W/m-K and the specific heat as 20.25 J/Kg-K. The setup is based on the mechanism that the iron particle when comes in the influence of magnetic field would hit the abrasive particle that would come in its path and the abrasive would be directed towards the work piece which would do the machining based on pressure difference. For analysis of the magnetic field's effect on the work piece ANSYS Maxwell is used and the magnetic field is produced by inducing the current in a strip cut through the core which surround the work piece. This step was done because the current can be given through the surfaces only. Although the variation of current and the produced magnetic field is plotted in the results and the value of current for the analysis is 100 A.

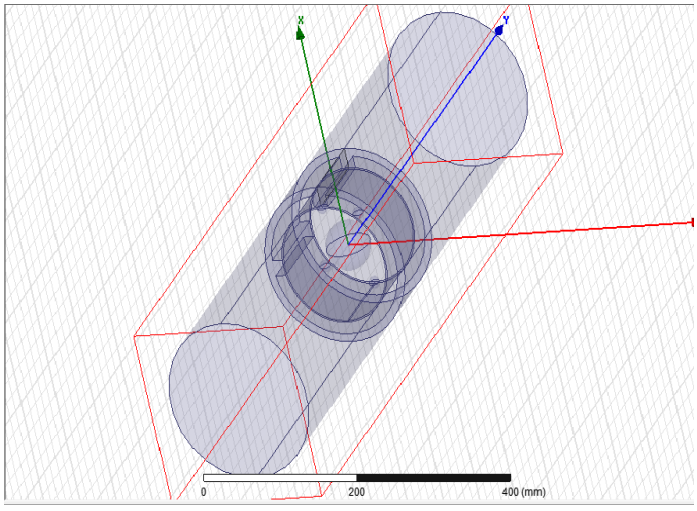



**Fig 3** Initial Setup Geometry in ANSYS Maxwell

The above figure represents the geometry of the modal used for analysis of magnetic field. It consists of a magnet wrapped around the hollow aluminium work piece which contains the viscous fluid having iron particle and the abrasive particle. Then a strip is cut using Boolean function in the magnet around the work piece and then the excitation of 100 amperes is applied on the strip. Since the current has the magnetic

effect, it induces the magnetic field on the contact surface of the work piece whose distribution would be discussed later in the results. The next part is to check the flow parameters of the Viscoelastic magnetic abrasive media for which the authors used ANSYS Fluent, since there is a limitation of software on simultaneous use of ANSYS Fluent and ANSYS Maxwell so the work was done separately. Once the Geometry is imported to the ANSYS Fluent and the pressure difference is given to the fluid at the inlet and the outlet condition. In this modal the number of abrasive particles is increased to 3 and the abrasive particle are made 12 with 4 abrasive particles surrounding the particular iron particle. The work piece is taken as aluminium and same viscous fluid is used to the analysis with similar physical properties as used in the ANSYS Maxwell. The shape of the mess was Hexagonal and the total number of elements including all domains was 27260, with the size of 3X10 $^{-003}$.The table for physical properties is as follows which contains the physical and magnetic property of various material used in the simulation bot in the ANSYS Maxwell and in ANSYS Fluent.

**Table 1.** Property of Various Materials Used in Simulation

| MATERIAL | TYPE | DENSITY (Kg/m$^3$) | VISCOSITY (Kg/ m-s) | THERMAL CONDUCTIVITY (W/ m-K) | SPECIFIC HEAT (J/Kg-K) | Relative permeability |
|---|---|---|---|---|---|---|
| Viscous Fluid | Fluid | 1219 | 0.789 | 0.22 | 20.25 | 1.00000004 |
| Aluminium | Solid | 2719 | | 202.4 | 871 | 1.000021 |
| Iron Particle | Solid | | | 16.27 | 502.48 | 4000 |
| Alumina abrasives | Solid | 3950 | | 12 | 451 | |

The basic generalised equations of Maxwell are [16]

$\vec{\nabla}.\vec{E} = \frac{\rho}{\varepsilon 0}$

$\vec{\nabla}.\vec{B} = 0$

$\vec{\nabla}.\vec{E} = -(\frac{\partial \vec{B}}{\partial \vec{t}})$

$\vec{\nabla}.\vec{B} = \mu 0\vec{J} + \mu 0\varepsilon 0(\frac{\partial \vec{E}}{\partial t})$

For magneto statics problem

$\vec{\nabla}.\vec{B} = 0$

$\vec{\nabla}.\vec{B} = \mu o\vec{J}$

The calculation for Force for magneto statics problem was done by Singh and Jain and proposed following results in x and y directions

$Fx = V\chi\mu_0\frac{\partial H}{\partial x}$

$Fy = V\chi\mu_0\frac{\partial H}{\partial y}$

Where

ε0 = permittivity of vacuum (F/m)

ρ = density (A/m$^2$)

$\vec{E}$ is electric field intensity (V/m)

$\vec{B}$ = magnetic flux density (T)

$\vec{J}$ = electric current density (A/m²)

μ0 = relative magnetic permeability in vacuum (Wb/(A-m))

$\chi$ = susceptibility of Particle

V = volume of particle

H = magnetic field strength at point

∂H/∂x = gradients of magnetic field strength in the x directions

∂H/∂y = gradients of magnetic field strength in the Y directions

The above equation are the basis of computation of force by Ansys Maxwell Software and the value of force and torque is computed and is discussed below in next section.

## RESULTS AND DISCUSSION

The calculation of Magnetic field is done corresponding to 100 A by the software and the maximum value of magnetic field was 1.5078 x 10$^{-003}$ tesla. The magnetic force applied on the iron particle is 5.1072 x10$^{-007}$ N and that on abrasives is 3.0857 x10 $^{-011}$ N while the torque on the iron particle is 1.6256 x 10 $^{-010}$ Nm and that on the abrasive particle is 5.4168 x 10 $^{-013}$ Nm.

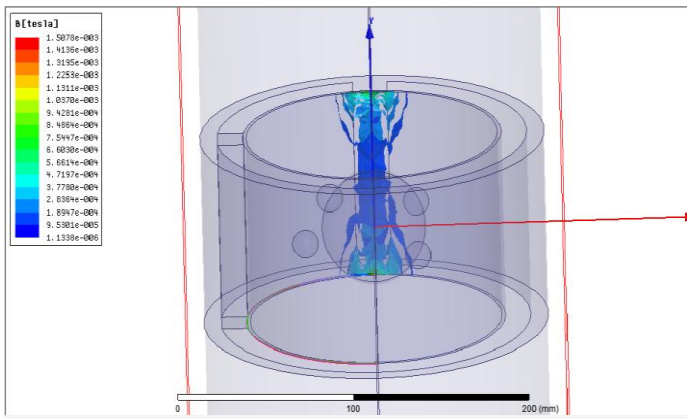



**Fig. 4** Magnetic field distribution on the work piece in Maxwell software

The above figure shows the distribution of the magnetic field on the work piece when the current is supplied at the 100 A on the work piece. results in the induction of magnetic field on the work piece. In order to determine the flow parameters, the geometry is imported to the ANSYS Fluent software where the viscous fluid along with the iron particle and the abrasive particle is reciprocated within the pressure difference of 20 to 40Mpa and the distribution of pressure is displayed on the work piece

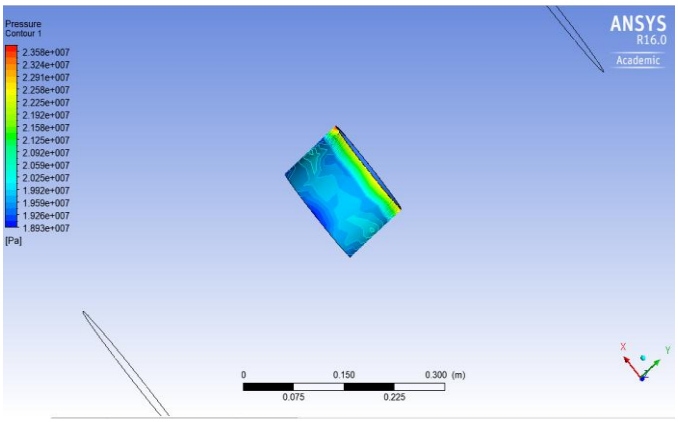



**Fig. 5** Pressure distribution on the work piece

The above figure shows a pressure of 23.83 MPa corresponding to the reciprocation of the fluid within the work piece between the specified pressure limit. The distribution is in accordance with the pressure variation achieved in abrasive flow machining bell Mouthing effect. [16].

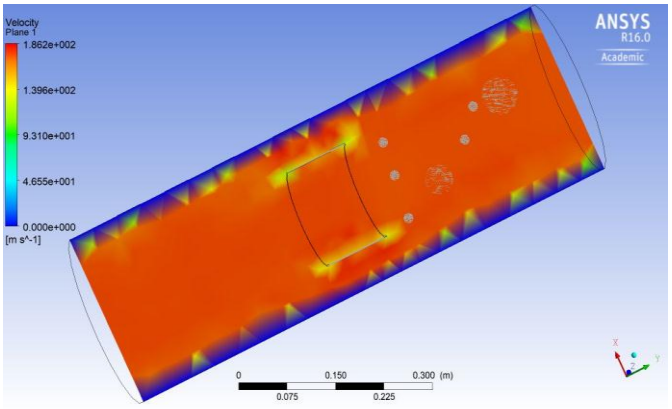



**Fig. 6** Velocity on the Plane

The above figure shows the velocity distribution in-side the working fluid. The plane was so selected that it covers the entire longitudinal section of the cylinder. The velocity is uniform through the length but at the entry of the work piece it shows variations. It is evident from the figure that on the work piece surface the velocity is maximum at the inlet side this is because of bell mouthing phenomenon [16].

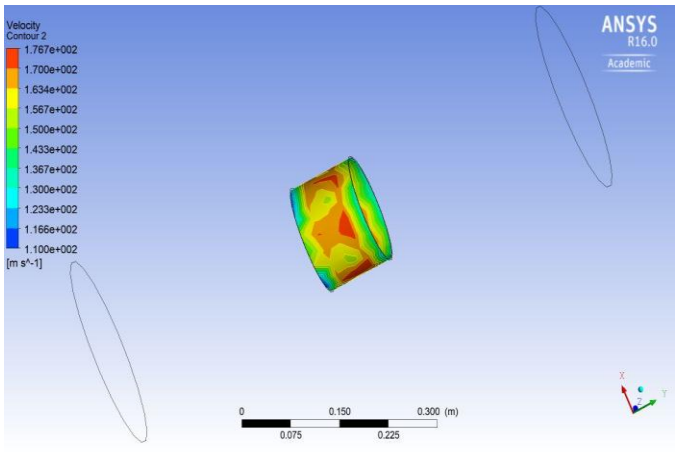



**Fig. 7** Velocity Distribution over work piece

It is evident from the above figure that the velocity is maximum in the areas where the abrasives are striking the work piece surface. The next diagram shows the streamline distribution on the work piece which shows machining of the work piece.

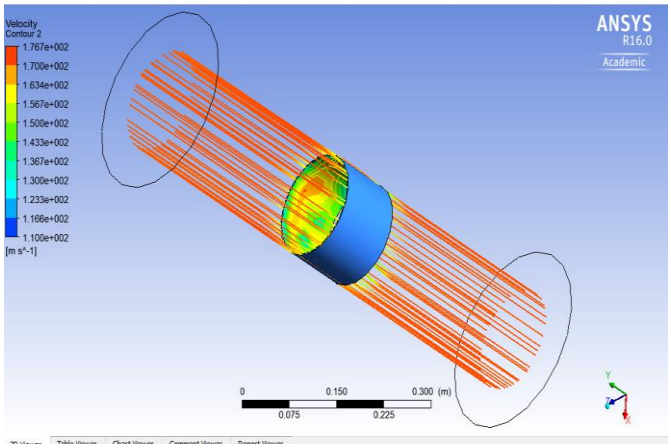
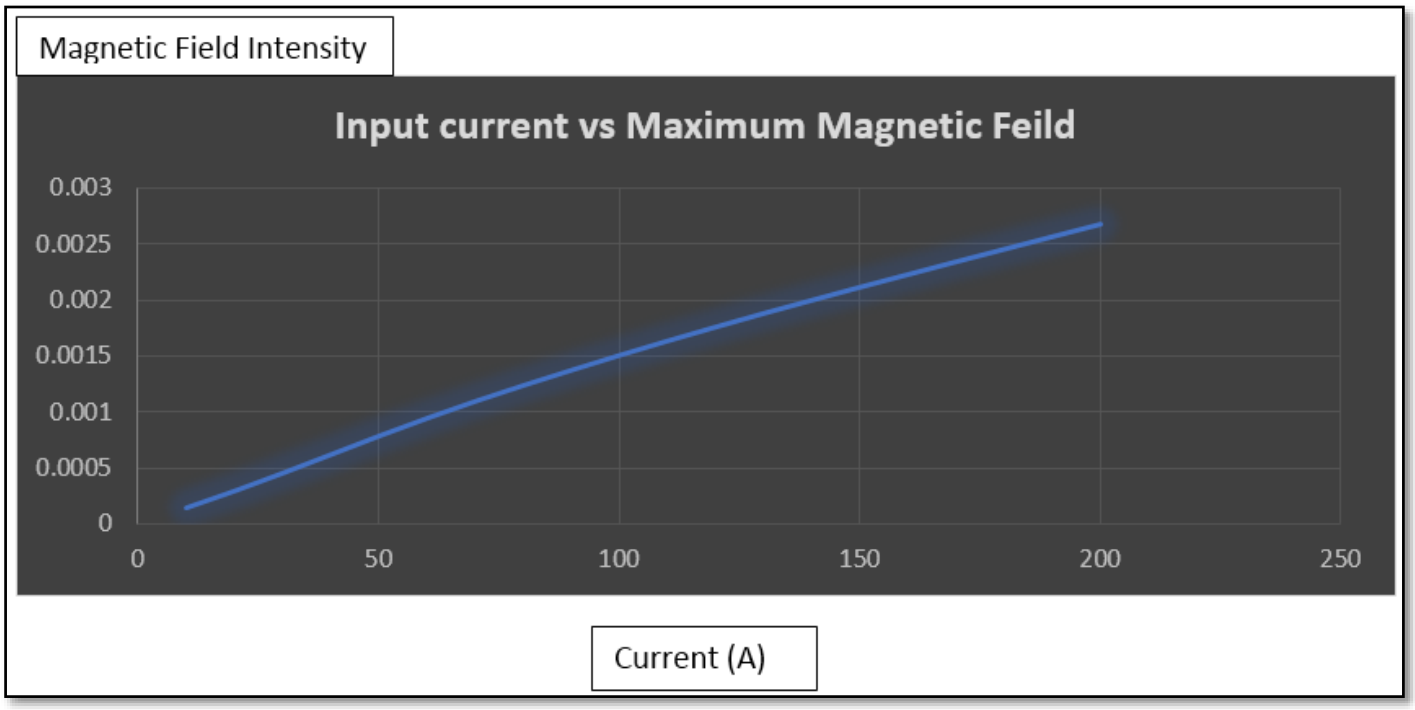



**Fig. 8** Streamline of the work piece

It can be concluded from the above figure that the streamline is uniform and smooth which shows the uniform machining of the work piece by the proposed viscoelastic fluids. It can be seen that there is distribution of velocity over the internal surface of the work piece which shows the material removal and machining of the internal surface. The graph plotted between the input current and the developed maximum field intensity is shown below with magnetic field intensity on Y axis and input current on the X axis. It can be seen the curve is almost a straight line which shows that the current is proportional to maximum magnetic field developed.




**PLOT 1** Current Vs Magnetic Field Intensity

## CONCLUSIONS

In the present work Modelling of magnetic flux density over the work piece and pressure distribution due to flow parameters have been done and following conclusion are drawn

- Viscoelastic Magnetic Finishing is suitable method for finishing of magnetic materials.
- The supplied current is an important parameter in the proposed set up.
- The pressure variation is an important parameter in the proposed set up.
- The proposed fluid is suitable for machining as the pressure variation on the work piece are in accordance with the pressure variation on the work piece as used in AFM.
- The maximum value of magnetic flux is 1.5078 x10 $^{-003}$ tesla corresponding to 100A current.
- The magnetic force applied on the iron particle is 5.1072 x $10^{-007}$ N and that on abrasives is 3.0857 x10 $^{-011}$ N while the torque on the iron particle is 1.6256 x 10 $^{-010}$ Nm and that on the abrasive particle is 5.4168 x 10 $^{-013}$ Nm.
- The graph plotted between the magnetic field intensity and the input current is approximately resembling a straight line.

## REFERENCES

[1] Shinmura T, Takazawa K, Hatano E, Matsunaga M. (1990) Study on magnetic abrasive finishing. Ann CIRP, 39(l): 325-328.

[2] Shinmura T., Takazawa K., Hatano E. and Aizawa T. (1985) Study on magnetic abrasive process-process principles and finishing possibility, Bull. Japan Soc. Prec. Eng. 19(1) 54–55.

[3] Kim JD., Choi MS. (1995) Simulation for the prediction of surface accuracy in magnetic abrasive machining. J Mater Process Tech 53:630–642.

[4] Fox M, Agrawal K, Shinmura T, Komanduri R (1994) Magnetic abrasive finishing of rollers. Ann CIRP 43(1): 181-184.Hou, Z.B., Komanduri, R Magnetic field assisted finishing of ceramics–part III On the thermal aspects of magnetic abrasive finishing (MAF) of ceramic rollers. J Tribology ASME 1998, 120, 660–667.

[5] Yadav V., Kumar G. Temperature distribution in the workpiece due to plane magnetic abrasive finishing using FEM. The International Journal of Advanced Manufacturing Technology, Volume 41, Numbers 11-12, 1051-1058, DOI: 10.1007/s00170-008-1557-7.

[6] Mulik RS., Pandey PM. (June 2010) Mechanism of surface finishing in ultrasonic assisted magnetic abrasive finishing process. Mat Manuf Proc, 25, 1418-1427.

[7] Srivastava, Ankit. Kumar., Komma, V. R: Finite Element Modelling Of Electrochemical Magnetic

Abrasive Finishing, Imperial Journal of Interdisciplinary Research (IJIR) Vol-2, Issue-12, 2016 ISSN: 2454-1362.

[8] Srivastava, Ankit. Kumar., Katiyar, Amit: Modelling and Simulation of Magnetic Abrasive Finishing for Thermal Analysis, Imperial Journal of Interdisciplinary Research (IJIR) Vol-2, Issue-12, 2016 ISSN: 2454-1362.

[9] Singh, Rajneesh. Kumar, Singh, D.K, Gangwar, Swati: Fuzzy logic optimization of process parameters affecting surface temperature in MAF, IOP Conf. Series: Materials Science and Engineering 404 (2018) 01 2017 doi:10.1088/1757-899X/404/1/012017

[10] Li, Xiuhong., Li, Wenhui., Yang, Shengqiang: Effect of coupling agent on interfacial bonding Properties of viscoelastic magnetic abrasive tools and finishing Performance, Advance Material research DOI 10.5013/IJSSST.a.17.28.17.

[11] Li, Wenhui., Li, Xiuhong., Yang, Shengqiang., Li, Weidong: A newly developed Media for Magnetic abrasive finishing process: Material removal behaviour and finishing performance, Journal of Materials Processing Tech. 260 (2018) 20-29.

[12] Kyung-In Jang, Jongwon Seok, Byung-Kwon Min: International Journal of Machine Tools & Manufacture, Vols. 50(2010), pp.869-881.

[13] A.C. Wang, S.H. Weng: Journal of Materials Processing Technology, Vols. 192-193 (2007), pp.486-490.

[14] Li, Wenhui., Li, Xiuhong: New Viscoelastic Magnetic Abrasive and its Surface Finishing Process. Advance Material research. ISSN:1662-8985. Vols. 314-316, pp 300-303. (2011).

[15] Mishra, Vivek., Goel, Harsha., Pandey, M. Pulak: Modelling of Magnetic Field and Magnetic Pressure in Magnetic Abrasive Finishing Process. World Academy of Science, Engineering and Technology 60 2011.

[16] Singh, D. K., Jain, V., Raghuram, V: Parametric study of magnetic abrasive finishing process, Journal of materials processing technology, Vol. 149, No. 1, pp. 22-29, 2004

[17] Walia, R.S., Shan, H.S., Kumar, P.K: Finite element analysis of media used in the centrifugal force assisted abrasive flow machining process. Department of Mechanical and Industrial Engineering, Indian Institute of Technology, Roorkee, India, DOI: 10.1243/09544054JEM325.