SCHOLARLY PUBLICATIONS

A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT

@DTU

(55th Edition)

JULY 2017

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 **Fifty fifth** 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 **July 2017**. 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 3 Degree of Freedom Quarter Car Active Suspension System Model Design Approach using PID, Fuzzy Logic and Fuzzy Tuned PID, ***6.Naman Garg*** and ***3.Bharat Bhushan***, Electrical, DTU

2. A big data driven sustainable manufacturing framework for condition-based maintenance prediction, ***7.1.Ajay Kumar***, Ravi Shankar and Lakshman S. Thakur, Operation & Information Management, DTU

3. A Context Based Algorithm for Sentiment Analysis, ***7.1.Srishti Sharma***, Shampa Chakraverty, Akhil Sharma and Jasleen Kaur, Computer Engineering, DTU

4. A thermodynamic analysis of ejector type vapour refrigeration system using eco-friendly refrigerants, ***3.R S Mishra***, ***6.Ankit Dwivedi*** and ***6.Shadab Ahmad***, Mechanical, DTU

5. Analysis of First Stage Gas Turbine Blade for Evaluating Damping Energy and Failure Mode, ***3.Sushila Rani***, ***3.Atul K. Agrawal*** and ***3.Vikas Rastogi***, Mechanical, Production, Industrial and Automobile, DTU

6. Column study for chloride removal from waste water by a low cost adsorbent (bio adsorbent), ***6.Lokesh Kumar*** and ***3.S K Singh***, Environment, DTU

7. Comparative study of parabolic trough collector through MWCNT/H2o nanofluid and water, ***3.R.S. Mishra*** and ***7.Harwinder Singh***, Mech, Industrial and Production, DTU

8. DEVELOPMENT OF PHONEME DOMINATED DATABASE FOR LIMITED DOMAIN T-T-S IN HINDI, ***7.1.Archana Balyan*** ECE, DTU

9. Device-Circuit Analysis of Ferroelectric FETs for Low-Power Logic, ***8.1.Shreya Gupta***, Mark Steiner, Ahmedullah Aziz, Vijaykrishnan Narayanan, Suman Datta and Sumeet Kumar Gupta, electrical and electronics , DTU

10. Effect of acidic aggressive environment on strength of mortars in early stage, Shobha Ram, Abhishek Singh and ***3.Alok Verma***, Civil, DTU

11. Effect of gas composition on morphological properties of graphene nanosheet, ***6.Neha Gupta*** and ***3.Suresh C. Sharma***, Applied Physics, DTU

12. Efficient Color Cast Correction Based on Fuzzy Logic, ***3.Om Prakash Verma***, Computer Science, DTU

13. Enhanced structural, dielectric, ferroelectric and piezoelectric properties of (1-x)Ba0.9Sr0.1TiO3 – (x)Ba0.7Ca0.3TiO3 ceramics derived using mechanochemical activation technique, ***7.Aditya Jain***, ***3.Amrish K. Panwar***, ***7.Rakesh Saroha*** and A. K. Jha, Applied Physics, DTU

14. Evaluating the Feedback Effects of Brand Extension on Parent Brand Equity: A Study on Indian FMCG Industry, ***6.Richa Joshi*** and ***3.Rajan Yadav***, DSM, DTU

15. Feasible star – delta and delta – star transformations for reliability networks, ***3.V.C. Prasad***, Electrical, DTU

16. Generating tunable THz radiation using rippled density plasma driven by density modulated relativistic electron beam, ***6.Pratibha Malik, 3.Suresh C. Sharma*** and ***3.Rinku Sharma***, Applied Physics, DTU

17. Immobilization studies of cellulase on three engineered polymer surfaces, Rajesh Ahirwara, ***3.Jai G. Sharma***, Pradip Nahara and ***6.Saroj Kumar***, Biotechnology, DTU

18. Linguistic Lyapunov reinforcement learning control for robotic manipulators, Abhishek Kumar and ***7.1.Rajneesh Sharma***, Control and Instrumentation, DTU

19. LiTaO3 based metamaterial perfect absorber for terahertz spectrum, ***6.Nishant Shankhwar***, ***3.Yogita Kalra*** and ***3.Ravindra Kumar Sinha***, Applied Physics,, DTU

20. Location based Twitter Opinion Mining using Common-Sense Information, Amita Jain and ***3.Minni Jain***, Computer Science, DTU

21. Modeling the emission of high power terahertz radiation using Langmuir wave as a wiggler, ***6.Jyotsna Panwar*** and ***3.Suresh C. Sharma***, Applied Physics,, DTU

22. Optimization of process parameters in Abrasive Flow Machining, ***6.Parvesh Ali***, Abhinav Sri, ***3.R.S Walia*** and Sehijpal Singh, Mechanical, DTU

23. Rebuttal to 'Fully-uncoupled independent control of frequency and condition of oscillation: A caution, Raj Senani, ***3.Data Ram Bhaskar***, Shanti Swaroop Gupta and Bhopendra Singh, Electronics and Communication, DTU

24. Review of some Case Studies of Concentrated Solar Thermal Power Technologies: Status of CSP in India, ***7.NirmalPratap Singh*** and ***3.Dr. J P Kesari***, Mechanical, DTU

25. Role of Magnetic Exchange Interaction due to Magnetic Anisotropy on Inverse Spin Hall Voltage at FeSi3%/Pt Thin Film Bilayer Interface, Jyoti Shah, Saood Ahmad, Rishu Chaujar, ***3.Nitin K. Puri***, P.S. Negi and R.K. Kotnala, Applied Physics, DTU

26. Simulation for Position Control of DC Motor using Fuzzy Logic, ***6.Pankaj Kumari Meena*** and ***3.Dr. Bharat Bhushan***, Electrical, DTU

27. Software Defined Internet of Things (IoT) in 5G Millimeter Wave (mmWave) Communication System, S. K. Agrawal and ***3.Dr. Kapil Sharma***, Computer, DTU

28. Stabilization of nonlinear inverted pendulum system using MOGA and APSO tuned nonlinear PID controller, ***3.Sudarshan. K.Valluru*** and ***3.Madhusudan Singh***, Electrical, DTU

29. Theoretical modeling to study the impact of different oxidizers (etchants) on the plasma-assisted catalytic carbon nanofiber growth, ***6.Ravi Gupta*** and ***3.Suresh C. Sharma***, Applied Physics, DTU

30. Thermodynamic (Energy-Exergy) Analysis Of Nine MW Coal Based Thermal Power Plant Using Entropy Generation Principle, Kaushalendra Kumar Dubey and ***3.R. S. Mishra***, Mechanical, DTU

31. Thermodynamic study of R134a in Vapour Compression Refrigeration System in Summer Climate, Kaushalendra Kumar Dubey and ***3.R.S Mishra***, Mechanical, DTU

32. Understanding the role of internal microstructure in capsule-based healing of polymeric composites, ***6.Manorama Tripathi***, ***3.Devendra Kumar*** and Prasun Kumar Roy, Applied Chemistry and Polymer Technology, DTU

33. Vulnerability to a Bioterrorism Attack and the Potential of Directed Evolution as a Countermeasure, ***6.Ritwik Raj***, Ananya Saxena and Ritcha Saxena, Biomedical, DTU

34. Wideband high gain conical dielectric resonator antenna: An experimental study of superstrate and reflector, Sounik Kiran Kumar Dash and ***3.1.Taimoor Khan***, Electronics and Communication, 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*

International Journal of Engineering Technology, Management and Applied Sciences

IJETMAS

www.ijetmas.com June 2017, Volume 5, Issue 6, ISSN 2349-4476

A 3 Degree of Freedom Quarter Car Active Suspension System Model Design Approach using PID, Fuzzy Logic and Fuzzy Tuned PID

Naman Garg
Department of Electrical Engineering
Delhi Technological University
Delhi,New Delhi, India

Bharat Bhushan
Department of Electrical Engineering
Delhi Technological University
Delhi,New Delhi, India

Abstract—The primary objective of any suspension system in automobiles is to isolate the road disturbances experienced by the tires frombeing transmitted to the passengers due to road profile. The passive suspension system can hardly improve the ride quality and handling performance at the same time. The aim of suspension system is to improve ride comfort, road handling and stability of vehicles.In this paper the study is extended to a quarter car (3-DOF) seat-suspension model instead of 2 DOF suspension systems. Three different active controllers are designed to compare the performance of three controllers with the passive suspension system. The three controllers designed are PID controller,Fuzzy logic controller and fuzzy tuned PID controller. In this work, MATLAB/SIMULINK software is used for simulation purpose and simulation result demonstrates that active suspension system shows better result in comparison to passive suspension system.

Keywords—suspension system,3-DOF,2 DOF,passive suspension system,PID,Fuzzy logic, fuzzy tuned PID,MATLAB/SIMULINK,ride quality.

I. INTRODUCTION

A suspension system is of three types: passive, semi-active and active suspension system. Passive suspension system is the simplest and consists of springs and dampers. In semi active suspension system consists of a variable damper that can adapt to actual demands. The active suspension system contains separate actuator to exert extra force on the suspension system.Researchers have already investigated active suspension system using models like 1/4 car model, 1/2 car model, full car model etc.Previously various works have been done on 2 degree of freedom quarter car model. Mat Hussin Ab. Talib and Intan Z. Mat Darus[1] investigated the performance of PID controller for ½ car model and PID parameters are tuned using heuristic method, Ziegler Nichols method and iterative method. In paper [2], Fuzzy logic has been studied and successfully implemented on a 2-DOF quarter car model.Muhamadetal.[3] proposed composite nonlinear feedback control for a linear 2-DOFquarter car model and was compared with LQR controller. In [4] the research work investigates the design of observer for 2 DOF model using sliding mode control methodology. In [5], LQR, optimised LQR and H infinity control was discussed and implemented in quarter car model and it was found that optimised LQR control scheme give the best result compared to other control methods. When the degree of freedom of the system increases, complexity of the system also increases along with the difficulty level to control the system. A 3-DOF system is designed using a fuzzy logic controller for the suspension system and analyzed to comparewith that of passive suspension system [6].In [7], two optimal controllers PID and LQR are designed for2 DOF model. The Fuzzy PID control scheme was proposed in previous research work in many systems. In fuzzy PID control scheme, Kp, Ki and Kd parameters are automatically tuned using fuzzy rules. Fuzzy PID control technique was successfully applied in 2 DOF quarter car model based active suspension system [8].

In this paper,3 DOF quarter car model is modeled and three different control methods namely

PID,fuzzy and fuzzy tuned PID are successfully designed and verified in simulation. The response of the control methods are compared with that of passive suspension system. The overall system is designed using MATLAB's Fuzzy Logic Toolbox and SIMULINK toolbox.

Section II gives description modeling of 3-DOF quarter car model for active and passive suspension system. Section III gives the formulation and design of the proposed controllers. Section IV gives description of simulation result and comparison of control schemes. Section V gives the conclusion.

II. MODELLING OF ACTIVE AND PASSIVE SUSPENSION

The 3 DOF quarter car model used is a simplified version of the full car model. It consists of passenger mass(m_p), sprung mass (m_s) and unsprung mass(m_{us}). Spring (k_p) and damper (c_p) represents the seat. The sprung mass is known as mass of the car body and it is supported on spring (k_s) and damper (c_s)of suspension system. Unsprung mass represents mass of car wheel. Tyre of the car is replaced by using spring (k_t). Fig. 1 represents the 3 DOF quarter car model for active suspension system.

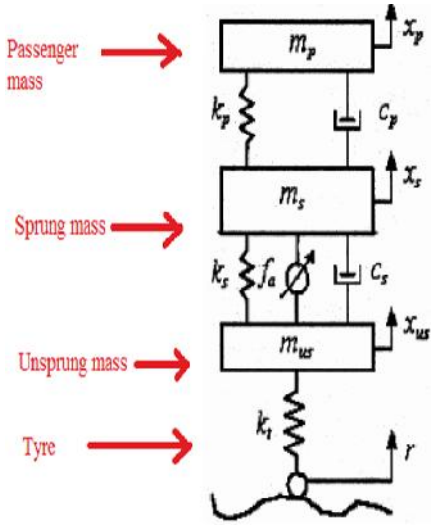


Fig.1. 3 DOF quarter car model of active suspension system.

A. *Equations*

Newton's second law of motion is applied to the seat, sprung mass and unsprung mass and the differential equations obtained are as follows:

$$m_p\ddot{x_p} + k_p(x_p - x_s) + c_p(\dot{x}_p - \dot{x}_s) = 0 \quad (1)$$

$$m_s\ddot{x_s} - k_p(x_p - x_s) - c_p(\dot{x}_p - \dot{x}_s) + k_s(x_s - x_u) + c_s(\dot{x}_s - \dot{x}_u) = f_a \quad (2)$$

$$m_u\ddot{x_u} - k_s(x_s - x_u) - c_s(\dot{x}_s - \dot{x}_u) + k_t(x_u - r) = -f_a \quad (3)$$

Here, $x_1 = x_p, x_2 = \dot{x}_p, x_3 = x_s, x_4 = \dot{x}_s, x_5 = x_u, x_6 = \dot{x}_u$, the equations can be written in terms of state variable $x_1, x_2, x_3, x_4, x_5, x_6$ as follows

$$\dot{x} = A + B + G \quad (4)$$

Where X=state input matrix, U=control input matrix

W=road input matrix.

$$\dot{x_1} = x_2 \quad (5)$$

$$\dot{x_2} = -1/m_p[k_p(x_1 - x_3) + c_p(x_2 - x_4)] \quad (6)$$

$$\dot{x_3} = x_4 \quad (7)$$

$$\dot{x_4} = {}^{-1}\!/_{m_s}\left[k_p(x_3 - x_1) + c_p(x_4 - x_2) + k_s(x_3 - x_5) + c_s(x_4 - x_6)\right] + f_a/m_s \quad (8)$$

$$\dot{x_5} = x_6 \quad (9)$$

$$\dot{x_6} = {}^{-1}\!/_{m_u}\left[k_s(x_5 - x_3) + c_s(x_6 - x_4) + k_t(x_5 - r)\right] - f_a/m_u \quad (10)$$

For passive suspension system f_a=0.Table I shows value of all the system parameters used in the model [6].

TABLE I. SYSTEM PARAMETER VALUES FOR QUARTER CAR MODEL

Parameter		
Parameter	***Value***	***Unit***
Passenger mass, m_p	100	Kg
Sprung mass, m_s	2050	Kg
Unsprung mass, m_{us}	100	Kg
Stiffness of the seat, k_p	100000	N/m
Seat Damping coefficien, c_p	6000	Ns/m
Stiffness of suspension system, k_s	400000	N/m
Damping coeffcient of suspension system, c_s	5000	Ns/m
Stiffness of the tyre, k_t	2000000	N/m

IJETMAS
www.ijetmas.com June 2017, Volume 5, Issue 6, ISSN 2349-4476

III. CONTROLLERS DESIGN

Auto tuned PID, Fuzzy logic controller and fuzzy tuned PID are proposed to control the vehicle dynamic model. A step input of 10 cm or 0.1 m is used as disturbance to test the system performance.

A. PID controller

PID controller is a commonly used control scheme in many industrial control systems. In this work, PID controller is auto tuned using MATLAB Simulink's auto tune box designed.

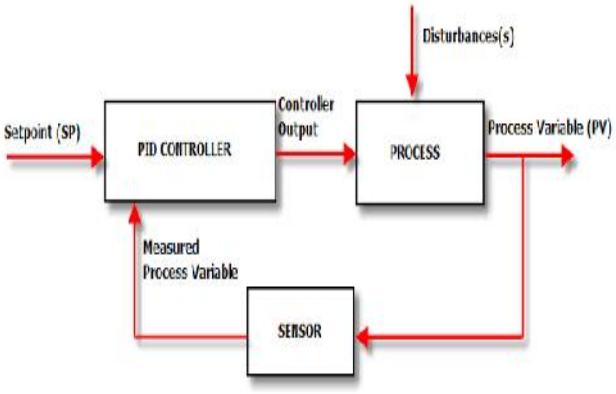


Fig.2.PID controller block digram for vehicle system.

The PID controller is represented by:

$$u(t) = Kp \times e(t) + Ki \int_0^t e(\tau)d\tau + Kd\frac{d}{d}$$ (11)

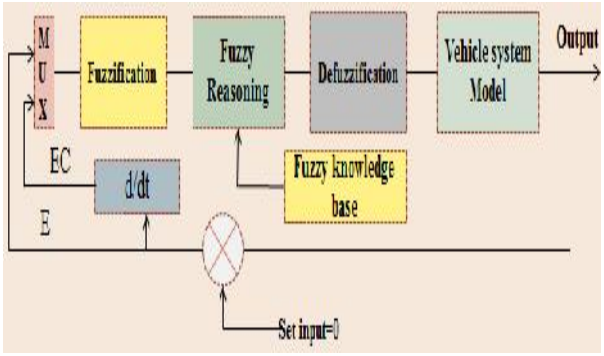
Where Kp is the proportional gain, Ki is the integral gain and Kd is derivative gain, e is the positional error.

B. FUZZY logic controller

Figure 3 represents the block diagram of fuzzy logic controller.

The steps in designing fuzzy logic controller are as follows:

J Trapezoidal membership functions are used and Linguistic variables assigned to these fuzzy sets are, NB, NM, NS, ZE, PS, PM and PB. Total 49 rules are made.



.Fig.3. Fuzzy logic controller block.

J .Mamdani type inference method and centroid defuzzification method are used. Output of fuzzy controller is actuator and inputs are the sprung mass displacement error(E) and the rate of change of velocity(EC) of the suspension system.

Fuzzy control rules are shown in table II.

TABLE II. FUZZY CONTROL RULES

		E						
		PB	**PM**	**PS**	**ZE**	**NS**	**NM**	**NB**
EC	**PB**	PB	PM	PM	PB	PM	PS	ZE
	PM	PM	PM	PM	PM	PS	ZE	NS
	PS	PM	PM	PM	PS	ZE	NS	NM
	ZE	PM	PM	PS	ZE	NS	NM	NM
	NS	PM	PS	ZE	NS	NM	NM	NM
	NM	PS	ZE	NS	NM	NM	NM	NM
	NB	ZE	NS	NM	NM	NM	NM	NM

C. FUZZY tuned PID controller

Fuzzy self-adjusting PID controller is the combination of PID controller with adjustable parameters and fuzzy controller. Figure 4 represents the schematic block diagram of fuzzy tuned PID controller for vehicle suspension system.

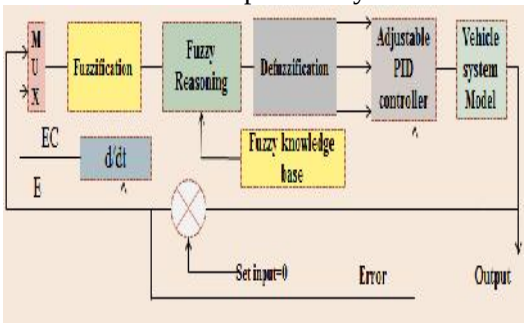


Fig.4. Fuzzy tuned pid controller block.

IJETMAS

www.ijetmas.com June 2017, Volume 5, Issue 6, ISSN 2349-4476

Relative displacement andrelative velocity are taken as error(E) and rate of change of error(EC) and using fuzzy logic rules given in table III Kp, Ki and Kd values are determined. Mamdani type inference method, centroid defuzzification method and Trapezoidal membership functions are used. Total of 27 rules are made and linguistic variables assigned are Low, Medium and High.Range of error, rate of error, Kp, Ki and Kd are estimated as:

E=[-0.13 0.07], EC=[-11 7],Kp=[0 1e+04], Ki=[0 1e+04],Kd=[1000 3e+03]

TABLE III. FUZZY CONTROL RULES FOR KP, KI AND KD

Kp		E		
		Low	Medium	High
EC	Low	P	P	ZE
	Medium	ZE	ZE	ZE
	High	N	N	ZE

Ki		E		
		Low	Medium	High
EC	Low	P	P	ZE
	Medium	ZE	ZE	ZE
	High	N	N	ZE

Kd		E		
		Low	Medium	High
EC	Low	N	N	ZE
	Medium	N	ZE	P
	High	ZE	P	P

IV. RESULT AND COMPARISON

A. Result

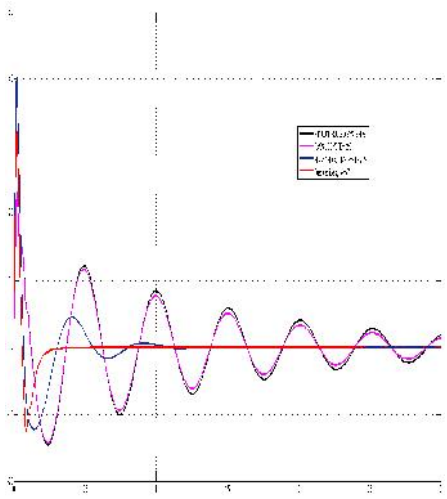


Fig.5. Passenger acceleration (m/sec^2) vs. time (sec).

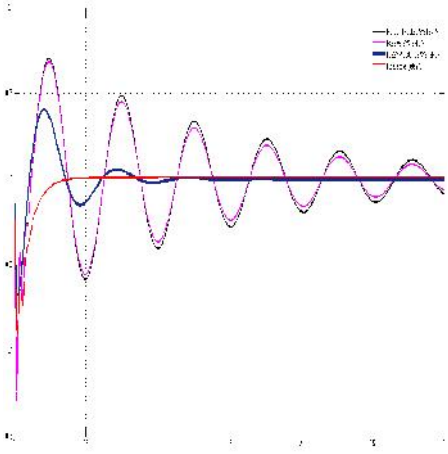


Fig.6. Suspension deflection (m) vs. time (sec).

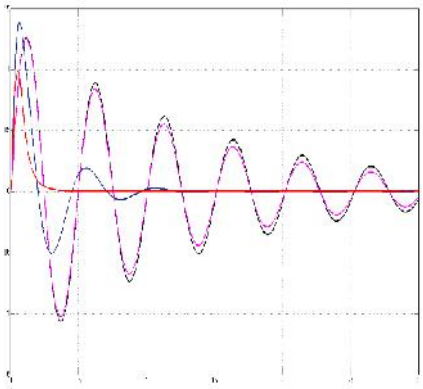


Fig7. Seat velocity (m/sec) vs. time (sec).

www.ijetmas.com June 2017, Volume 5, Issue 6, ISSN 2349-4476

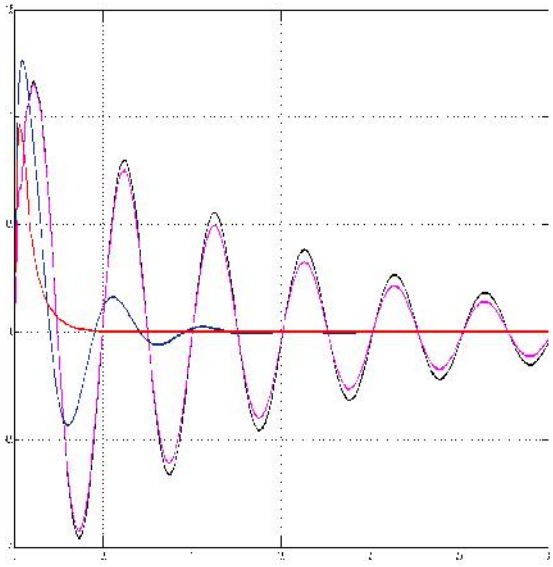


Fig.8. Sprung mass velocity (m/sec) vs. time(sec).

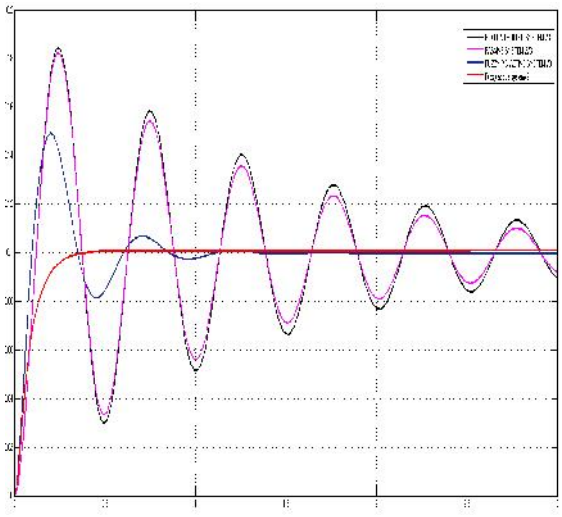


Fig.9. Sprung mass displacement (m) vs. time (sec).

In the figure black, blue, red and pink colors are used to denote Auto tuned PID, Fuzzy PID, Fuzzy and Passive Suspension system. Fig. 5,6,7,8,9 show the comparison between the PID, Fuzzy and Fuzzy tuned PID active suspension system and passive suspension system.

B. Comparison

Table IV, V, VI,VII and VIII show the comparison of the proposed controllers with passive suspension system.

TABLE IV. PERFORMANCE CHARACTERISTICS OF SUSPENSION DEFLECTION

	Suspension deflection			
	PID	***Fuzzy PID***	***FUZZY***	***PASSIVE***
Maximum peak value (m)	0.06	0.0405	0	0.0685
Settling time (sec.)	7	1.8	0.6	7.5
Steady state error (%)	0	0	0	0

TABLE V. PERFORMANCE CHARACTERISTICS OF SEAT ACCELERATION

	Seat acceleration			
	PID	***Fuzzy PID***	***FUZZY***	***PASSIVE***
Maximum peak value (m/sec^2)	23.609	41	32	23.7147
Settling time (sec.)	8	1.5	0.6	8.5
Steady state error (%)	0	0	0	0

TABLE VI. PERFORMANCE CHARACTERISTICS OF SEAT VELOCITY

	Seat velocity			
	PID	***Fuzzy PID***	***FUZZY***	***PASSIVE***
Maximum peak value (m/sec)	1.2	1.38	0.95	1.2525
Settling time (sec.)	8.5	1.5	0.8	9
Steady state error (%)	0	0	0	0

TABLE VII. PERFORMANCE CHARACTERISTICS OF SPRUNG MASS VELOCITY

	Sprung mass velocity			
	PID	***Fuzzy PID***	***FUZZY***	***PASSIVE***
Maximum peak value (m/sec)	0.9919	1.24	0.94	1.1545
Settling time (sec.)	2.1	1.5	0.9	9
Steady state error (%)	0	0	0	0

TABLE VIII. PERFORMANCE CHARACTERISTICS OF SPRUNG MASS DISPLACEMENT

	Sprung mass displacement			
	PID	***Fuzzy PID***	***FUZZY***	***PASSIVE***
Maximum peak value (m/sec)	0.18	0.13	0.1	0.1823
Settling time (sec.)	7.8	1.6	0.8	8
Steady state error (%)	0	0	0	0

V. CONCLUSIONS

Active suspension system has been successfully designed using the three controllers. The resulting closed-loop control systemcan be able to eliminate the effect of road disturbance to some extent. Auto tuned PID is designed using MATLAB Simulink's auto tune tool. Fuzzy tuned PID gives better result than Simulink's auto tuned PID model. In case of performance parameters like seat acceleration, suspension travel, sprung mass velocity, seat velocity, sprung mass displacement Fuzzy control scheme gives better result as compared to other control methods.

REFERENCES

[1] Mat Hussin Ab. Talib, Intan Z. Mat Darns, "Self-Tuning PID Controller for active suspension system with hydraulic actuator,"IEEE Symposium on Computers & Informatics(ISCI),pp.86-91,April 2013.

[2] Salah G. Foda, "Fuzzy control of a quarter-car suspension system,"12th International Conference on Microelectronics,2000

[3] Ismail etal"A linear model of quarter car active suspension system using composite nonlinear feedback control,"2012 IEEE Student Conference on Research and Development(SCOReD).

[4] Elnaz Akbari, Morteza Farsadi, Intan Z.Mat Darus,Ramin Ghelichi, " Observer Design for Active Suspension System Using Sliding Mode Control," 2010 IEEE Student Conference on Research and Development (SCOReD 2010), 13 - 14 Dec 2010, Putrajaya, Malaysia K. Elissa.

[5] A.H. Shirdel ,E. Gatavi, Z. Hashemiyan, "Comparison of H-∞ and optimized-LQR controller in active suspension system,"Second International Conference on Computational Intelligence, Modelling and Simulation ,2010.

[6] Tinnavelli Ramamohan Rao and Punjala Anusha, "Active Suspension System of a 3 DOF Quarter Car Using Fuzzy Logic Control for Ride Comfort," 2013 International Conference on Control, Automation, Robotics and Embedded(CARE).

[7] B.Pratheepa, "Modeling and simulation of automobile suspension system," Frontiers in Automobile and Mechanical Engineering (FAME),2010.

[8] Sangzhi Zhu, Haiping Du, and Nong Zhang,"Development and Implementation of Fuzzy, Fuzzy PID and LQRControllers for an Roll-plane Active Hydraulically InterconnectedSuspension"2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)July 6-11, 2014,Beijing,China.

Accepted Manuscript



Title: A big data driven sustainable manufacturing framework for condition-based maintenance prediction

Authors: Ajay Kumar, Ravi Shankar, Lakshman S. Thakur

PII: S1877-7503(16)30512-9
DOI: http://dx.doi.org/doi:10.1016/j.jocs.2017.06.006
Reference: JOCS 705

To appear in:

Received date: 23-12-2016
Revised date: 27-3-2017
Accepted date: 13-6-2017

Please cite this article as: Ajay Kumar, Ravi Shankar, Lakshman S.Thakur, A big data driven sustainable manufacturing framework for condition-based maintenance prediction, Journal of Computational Sciencehttp://dx.doi.org/10.1016/j.jocs.2017.06.006

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 big data driven sustainable manufacturing framework for condition-based maintenance prediction

Ajay Kumar[a*] ajay.tomar@dmsiitd.org, Ravi Shankar[b], Lakshman S. Thakur[c]

[a]Bharti School of Telecommunication Technology & Management, Indian Institute of Technology Delhi, New Delhi, India-110016
[b]Department of Management Studies, Indian Institute of Technology Delhi, New Delhi, India-110016
[c]Operation & Information Management Department, School of Business, University of Connecticut, Storrs, CT, USA

[*]Communicating author.

ACCEPTED MANUSCRIPT

Highlights

- This research develops a big data analytics framework to quantify the remaining life prediction uncertainty considering the prediction accuracy improvement, and an effective CBM optimization approach to optimize the maintenance schedule.
- The proposed big data analytics framework in our study uses a CBM optimization approach that utilizes a new linguistic interval-valued fuzzy reasoning method for predicting the information of remaining life prediction uncertainty.
- The experimental results are performed on a big dataset which is generated from a sophisticated simulator of a gas turbine propulsion plant and our results show that the method used in the proposed framework outperform the traditional ones in terms of classification accuracy and other statistical performance evaluation metrics.

ACCEPTED MANUSCRIPT

Abstract

Smart manufacturing refers to a future-state of manufacturing and it can lead to remarkable changes in all aspects of operations through minimizing energy and material usage while simultaneously maximizing sustainability enabling a futuristic more digitalized scenario of manufacturing. This research develops a big data analytics framework that optimizes the maintenance schedule through condition-based maintenance (CBM) optimization and also improves the prediction accuracy to quantify the remaining life prediction uncertainty. Through effective utilization of condition monitoring and prediction information, CBM would enhance equipment reliability leading to reduction in maintenance cost. The proposed framework uses a CBM optimization method that utilizes a new linguistic interval-valued fuzzy reasoning method for predicting the information. The proposed big data analytics framework in our study for estimating the uncertainty based on backward feature elimination and fuzzy unordered rule induction algorithm prediction errors, is an innovative contribution to the remaining life prediction field. Our paper elaborates on the basic underlying structure of CBM system that is defined by transaction matrix and the threshold value of failure probability. We developed this framework for analysing the CBM policy cost more accurately and to find the probabilistic threshold values of covariate that corresponds to the lowest price of predictive maintenance cost. The experimental results are performed on a big dataset which is generated from a sophisticated simulator of a gas turbine propulsion plant. A comparative analysis confirms that the method used in the proposed framework outpaces the classical methods in terms of classification accuracy and other statistical performance evaluation metrics.

Keywords- data driven sustainable enterprise, fuzzy unordered induction algo, big data analytics, condition-based maintenance, machine learning techniques, backward feature elimination.

ACCEPTED MANUSCRIPT

1. Introduction

Sustainability presents itself as a pressing concern for both, the present as well as the future generations and it has become a necessity due to the regulations which are imposed by the stakeholders. Performance evaluation is an important component of sustainability initiatives in manufacturing organizations. Our everyday lives are replete with references to energy efficiency, sustainable development, triple bottom line and many others. And in the said context, several companies have realized the challenges posed by the modern industrial and ecological frameworks, forcing them to expand their focus from just being competitive to sustainable (Chouikhi et al., 2014; Shaw et al., 2013; Liao et al., 2006; Bhattacharya et al., 2014). The contemporary civilised lifestyle counts manufacturing as one of its significant pillars that stands to play a crucial role in establishing sustainable practices. Presently, a vast majority of manufacturing models rely on conventional standards, and it is technology, combined with culture and economy, that is increasingly being expected to develop efficient tools and options for newer approaches towards a sustainable manufacturing concept. Simultaneously, high-tech manufacturing plants and sophisticated systems represent a considerable heritage that requires a comprehensive and sustainable management. A new sustainable world is envisioned on the foundation of new technology, new strategies and solutions, and new business models, and this vision holds particularly true for the manufacturing sector (Emovon et al., 2015; Validi et al., 2015). Predictive maintenance can be considered to illustrate the primary discipline to be used for this goal. Condition based maintenance minimizes economic loss caused by unexpected breakdowns by lifetime prognosis, and therefore addresses the sustainability concerns by optimizing maintenance planning to enhance the quality of customer service. The performance of the system, in terms of product quality, scalability, reliability, productivity, costs and capacity, is deeply impacted by different configurations. However, application of data science for such a complication remains a challenge, especially when the failures, even if rare, are catastrophic

ACCEPTED MANUSCRIPT

(Khatab,2015; Dey, 2001;Wang et al.,2015). CBM is a maintenance strategy which conducts the maintenance process using the information gathered via condition monitoring. This strategy is responsible to reduce the maintenance cost through maintenance actions when the machine failure process is approaching. The implementation of optimal CBM is effective when there is accurate prediction of the component. As of now, there are several condition-based prediction methods that is used to determine the equipment's health condition at some inspection point. These prediction techniques can be segregated into different model and data driven ways. Both have different function. On one hand, model-based method helps to determine the equipment's health condition by using different damage prediction models, which is, dependent on damage prediction mechanics (Vachtsevanos et al, 2006). The propagation process is usually complex and it is troublesome to accurately represent different damage prediction process. While building a physics-based framework or model, a large number of aspects need to be taken under careful consideration. For example, reciprocity, dynamics, etc. However, the prediction or prognosis accuracy of health condition can be incredibly enhanced if an authentic and legitimate physics-based framework or model is built successfully. The techniques that are used for predicting health condition at present emphasise on building physical frameworks/models for bearings and gears. A method for predicting health condition of gear system was formulated in Kacprzynski et al (2002) and Tian et al. (2014). This method or approach was based on the physical models of gear tooth crack/fracture initiation and crack propagation. Condition-based maintenance (CBM) has emerged as an option for various industries wherein unforeseen failures are reduced, equipment reliability is improved, and maintenance cost is shrunk by enhancing maintenance optimization. Information and data that correspond to the state of a machine are detected and collected by Condition Monitoring (CM). And the data accumulated from CM forms the basis of preventive maintenance, of which CBM is a type. CBM is, therefore, a procedure that reduces the maintenance costs of machines by utilizing a systematic programme of supervising key parameters that could caution against failure, and by

ACCEPTED MANUSCRIPT

implementing corrective maintenance once the parameters achieve the target values (Jonathan 2002). Other applications of CBM are: monitoring, diagnostics and prognostics. Diagnostics is used for identifying the reasons for equipment failure and we use the prognostics for calculating the time period of a potential failure. Conventionally, we classify the maintenance into condition-based, preventive and corrective, wherein CBM relies on condition monitoring and age data data in order to make maintenance decisions. Condition monitoring data thus employed may often be collected from vibration, oil and other analysis. By acting only when a failure is predicted, CBM intelligently avoids any unnecessary maintenance tasks.

1.1 Types of Maintenance

Maintenance actions can be broadly classified into the following four categories: corrective maintenance, preventive maintenance, planned maintenance and CBM. In corrective maintenance, the asset is operated to its breaking-down point, followed by maintenance activities with the aim of speedily restoring the system. This procedure passively waits for the failure and then proceeds to repair it. It is also referred to as run-to-failure maintenance (Goyal et al., 2016). In preventive maintenance, it is executed prior to breakdowns so that they can be avoided and any issues that might arise out of failures can be reduced. Adjustments, replacements, renewals, and inspections are organized in line with well-calculated planning and schedule. This is in stark opposition to CM where random failure patterns are observed due to unpredictability (Singh et al., 2015). In planned maintenance, the machine is periodically inspected and the suitable parts are replaced on the basis of a fixed timetable. For this reason, it is also referred to as time-based maintenance (Bengtsson, 2004). The CBM approach helps in ascertaining the conditions of in-service maintenance assets in order to calculate potential degradations, and to calculate effectively when predictive maintenance techniques will be performed to decrease the disruptions. It has been proven that CM and CBM technologies have significantly reduced the cost of maintenance and improved operational safety (Rao, 1996; Dey, 2004). Three major elements constitute the CBM system are sensors, signal converters and

ACCEPTED MANUSCRIPT

processors. Sensors are connected to the machinery to evaluate selected physical parameters of temperature or vibration. Signal converters change the sensor signals into an electrical, digital form and processors implement CBM logic on the signals while providing data and decision support information to the user. Owing to its popularity, recent years have been witness to the rise of numerous approaches which have several elements similar to those of CBM (Bhattacharya et al., 2015). Some examples to be noted are: predictive maintenance (PdM), reliability centred maintenance (RCM), plant asset management system (PAM) and total productive maintenance (TPM).

Predictive Maintenance (PdM) measures the condition of the equipment, determines whether it will fail in a specified future period, and then proceeds to take an action to avoid the ramifications of that failure (Dunn 2001). Reliability Centred Maintenance (RCM) is an elaborate seven-step structured process that regulates equipment maintenance strategies, and may be inclusive of condition-based, predictive, and planned maintenance (Kennedy 2001). This approach evolved in the aircraft industry around the 1950s. Plant Asset Management System (PAM) delivers well-timed information to the operating and maintenance personnel in order to synchronize most favourable decisions related to process operations and assess maintenance. This, as a result, conveniently expands the total production output of a plant at a much lower cost per unit of output with no personnel increase (Bever 2000). Total Productive Maintenance (TPM) is a company-wide equipment management program which is a part of lean manufacturing due to its relationship with just-in-time and total quality management. TPM takes into account the eight pillars of maintenance, human resource management, safety, new equipment management, training and process quality management, instead of just focusing on the equipment itself (Kennedy 2001, Dunn 2001). CBM and CM (Condition Monitoring) are significant for the functioning of a TPM system. Therefore, the reasons for executing the CBM approach in a manufacturing plant are increasing the safety, development of

ACCEPTED MANUSCRIPT

quality standards, timely failure rate prediction, increasing the reliability, continuous flow support, lower maintenance support and reduced unplanned down time.

Lu and Meeker (1993) devised an approach for predicting time-to-failure probability distribution from an organized degradation experimental set by applying random coefficient growth models and examining numerous degradation patterns. Wang (2000) determined the failure threshold by following a random coefficient model and diminished the total cost by supervising interval in condition-based maintenance. Many studies have been conducted to understand the failure diagnosis using machine learning techniques, for example, Pong et al (2000) developed a framework based on Auto-Associative Neural Networks (AANN) for predicting the engine faults. Kharoufeh and Cox (2005) developed an approach for combining degradation exploratory data with probabilistic failure models that assisted in estimating the lifetime distribution contingent on Markovian deterioration. Later, Gebraeel et al. (2005) formulated a Bayesian approach that made optimum use of condition monitoring data in closed form which overhauls the residual-life distribution. Grall et al. (2002) and Be´renguer et al. (2003) deliberated on preventive strategies for progressively deteriorating maintenance systems, and Grall et al. (2002) suggested a stable inspection CBPM approach. This involves reducing the overall cost of failure and maintenance expenses which determine the inspection intervals and preventive maintenance threshold. What makes CBM a favourable maintenance approach is its adoption of a prompt and proactive method instead of the traditional reactive interventions. In order to administer a significant change, predictive modelling is assigned a crucial position that allows automatic triggering of alarms minus any requirement of complex, usually unknown (and, to an extent, unknowable) models. With this objective in mind, a firm use of Machine Learning (ML) models is suggested and analysed; these models have exhibited appealing modelling and predictive capacities even in the extremes of complex problem arenas. Specifically, we have focused on a fuzzy unordered rule induction algorithm wherein FURIA models are developed solely from the data assembled from the field on the equipment's that require modelling.

ACCEPTED MANUSCRIPT

The predictive model is not contingent on any theoretical knowledge of the turbines, their configuration or of the general system. In general, there is a lack of studies that have attempted to develop condition-based maintenance model without major implementation effort in big data environment. This paper standardizes the performance and potential of fuzzy unordered rule induction algorithm (FURIA) model by upscaling the data to predict the performance decay of Gas Turbines (GTs) positioned on a naval vessel.

The objective of this work is to develop a big data analytics framework initiate a robust designing of a predictive CBM methodology for naval propulsion plants stocked with GTs. In this big data analytics work particularly, a naval vessel run by a GT is given an appearance of propulsive performance degradation—between two dry docks—so that a data set of experiments can be collected. This data is further utilized to design a predictive model that automatically diagnoses the system decay state, thus aiming for the most effective CBM in the considered domain. In this article, a fuzzy induction rule extraction growth model is employed for the purpose of model degradation. A degradation experiment is made productive by extracting the failure time distribution of the parts. The distribution is later streamlined with condition data from an individual part to predict its failure time, and the effect of the amount of condition data for updating the maintenance actions is investigated. Further in the conclusion section, we elaborate on some significant managerial insights which would prove to be very effective for supply chain managers.

The current study offers interesting insights in three ways.

1. First, to the best of our knowledge, this study is the first one which investigates the contributions of failure rate prediction of mechanical components for developing decision support system and a big data analytics framework of predictive based maintenance is developed and fuzzy rules induction–based approach is harnessed to handle more regular models, inclusive of unknown error variance.

ACCEPTED MANUSCRIPT

2. Second, our advanced analysis indicates that the proposed framework gives a superior result on condition-based maintenance prediction because of having better adjustment because a backward feature elimination approach is employed with FURIA to augment reliability predictions and maintenance decisions by incorporating prior information on the degradation behaviour.

3. Third, our proposed framework enhances the performance of FURIA algorithm by upscaling the data using advanced feature selection method to predict the performance decay in big data environment. We attempt to utilise a subset of components and prepare a model utilizing these feature selection properties which is used to empower the fast and accurate machine learning calculation and avoiding the overfitting.

2. Literature Review

Failure detection is defined as the process of isolating, identifying and detecting a component that has ceased to operate. So in failure detection, when we talk about fault detection: detection of abnormal condition and reported, Failure isolation is the process of determining of component which has failed and the failure identification is the process of estimating the nature of fault in whole automated manufacturing process. There are lots of machine learning methods in available literature that have been developed in previous years for gas turbine diagnosis. Multiple architectures, framework and algo of ANNs have been introduced in literature for gas turbine maintenance prediction, for example, Volponi et al. (2003) developed a method based on Kalman filters and Feed Forward Back-Propagation (FFBP) ANN for gas turbine performance diagnostics. Romessis & Mathioudakis (2003) proposed a probabilistic ANN model in operation with parts failure probability for sensor fault detection. An ANN method was developed by Ogaji and Singh (2003) and Self-Organizing Maps (SOM) based approach which is unsupervised ANN is used by Yun Peng (2012). In addition, Genetic Algorithms (GA), which is firstly used by Sampath and Singh (2004) for making an integrated fault diagnostics model. Li and Lee (2005) proposed another method to predict gears

ACCEPTED MANUSCRIPT

that have a fatigue tooth crack. This method was also based on physical models. Li and Lee (2005) used fracture mechanics model, a model that identifies the stiffness of gear meshing, and a gear dynamic model for proposing this method. Marble and Morton (2006) proposed a method that can be used for health condition prediction of propulsion system bearings. This method was based on the finite element model and bearing spall propagation model. Unline model-based approaches, data-driven approaches are used for health condition prediction for component or equipment based on the gathered condition monitoring data. This data may be oil analysis data, fuel consumption data, acoustic emissions data, environmental conditions data, vibration analysis data, etc. Liao et al. (2006) established the ideal CBM threshold value for a residual degrading part by resolving imperfect maintenance. Romessis & Mathioudakis (2006) and Eustace (2008) developed stochastic machine learning methods based on probabilistic Bayesian function, which is a hybrid combination of fuzzy logic and probabilistic function that is used to measure the falsity degree of all parameters. A large number of methods like ANN-based methods (Wu et al, 2012), multicomponent system condition-based maintenance methods (Tian and Liao, 2011) and PHM-based methods (Lugtigheid et al, 2008) have been proposed for optimising maintenance of CBM by minimising the overall predicted maintenance expenses.

To optimise condition-based maintenance (CBM), it is important to predict the uncertainty if health condition prediction is carefully used. Paik et al. (2010) engineered a real-time monitoring system to be used for a full-scale ship which deployed a wireless sensor network and data transmitted over power lines. Degradation modelling is viewed as a coherent method of predicting the reliability of highly dependable system and rarely susceptible to failures. Feng et al. (2010) constituted an approach to calculate replacement intervals so as to simultaneously upgrade the quality in microelectromechanical systems in the case of wear degradation. Loboda (2011) developed a model based on Radial Basis Function (RBF) networks for gas turbine fault identification and Li et al. (2012) conceived a condition monitoring and fault diagnostic system for marine diesel engines that

ACCEPTED MANUSCRIPT

made use of information fusion technology. Chouikhi et al. (2014) proposed a condition-based maintenance model for a single-unit production system of goods and services to determine optimal inspection dates by Nelder-Mead method. Emovon et al. (2015) developed a hybrid MCDM method for selecting the optimum maintenance strategies by using Delphi-AHP and Delphi-AHP-PROMETHEE in ship machinery systems. The replacement policy based on ANN in Wu et al (2012) also makes use of prediction error for estimating the uncertainty of prediction (Tian et al, 2010). The policy assumes that the prediction error is always the same through the complete process i.e. the prediction appropriateness is not enhanced during a component's prediction process. This is also assumed in other reviewed researches (Coraddu et al., 2017; Roy et al., 2016; Jiang et al., 2015; Tian and Liao, 2011; Lugtigheid et al., 2008; Castanier et al, 2005). However, Gebraeel et al. (2005) indicates that the accuracy of prediction often enhances with the component's age as it approaches or moves towards the failure time. Prediction outcome based on the exploration of trial data used in this research shows that accuracy of prediction gets better with time. Singh et al. (2015) proposed a sustainability evaluation method for manufacturing using integrated (FAHP) and fuzzy inference system (FIS) approach and balanced scorecard framework is used to categorize the indicators which are identified from literature. Wang et al. (2015) worked on a cloud-based predictive maintenance framework for intelligent manufacturing and a low-cost cloud computing node is developed with embedded mobile agent middleware and numerical libraries to enhance the system flexibility and adaptability. Khatab (2015) solved an imperfect preventive maintenance optimization problem and developed a hybrid hazard rate model for finding the threshold probabilistic optimal value of reliability together with condition-based preventive maintenance. Goyal et al. (2016) developed a framework to predict and optimize the manufacturing processes using soft computing in condition-based maintenance environment. In this research paper, we put forward CBM optimisation method wherein the uncertainty of health condition prediction is calculated on the basis of prediction errors. We consider that prediction accuracy improves over time in the whole prediction process. By way of

ACCEPTED MANUSCRIPT

shaping the relationship between tool's life percentage and average value of prediction error, and the relationship between tool's life percentage and the value of prediction error which is calculated form standard deviation, the existing life prediction uncertainty can be quantified considering the improvements in prediction accuracy.

3. Proposed Framework and Model Development

3.1 Framework of the Research

CMB consists of five parts: data collection and acquisition, data pre-processing, failure detection, failure isolation and failure identification. Figure 1 presents the big data analytics framework of train fuzzy unordered rule- based machine learning algorithm for reducing the maintenance cost and improving the equipment reliability in CBM program. The proposed framework develops an intelligent hybrid approach based on backward feature elimination and FURIA method for predicting the remaining life uncertainty. Backward feature elimination method generally used to compute the sum of square of error after eliminating each variable. Then, we try to identify those variables (with n-1 input features) whose removal produce smallest increase in the sum of square of error and finally we remove it. We repeat this process in this feature elimination phase until no other instances can be dropped and then the dataset is used to train the fuzzy rule-based classifier for remaining life prediction information on a gas turbine fault diagnosis case study dataset. First of all, we add some needed features and dimensions in the dataset using feature engineering and then we use the dataset for predicting the failure probability, valuable information across diverse metrics, which is used for providing quicker operational support service to operators.

Feature selection algorithm is the combination of search technique and evaluation measure which is used for proposing the feature subsets to find the subset with minimum error rate. Backward feature selection is used in data pre-processing and it is used to train ML faster and easier to interpret. Backward feature selection algorithm improves the model accuracy and reduces the

ACCEPTED MANUSCRIPT

overfitting if we choose the right subset of features and then train the model using these subsets. It is a covetous streamlining computation, which expects to find the best performing highlight subset. These are the steps in feature selection algorithm for finding the subset in a big dataset. Firstly, it adds irregularity to the given informational index by making rearranged duplicates of all elements (which are called shadow highlights). Then, it prepares an irregular fuzzy classifier on the amplified informational collection and applies a component significance measure to assess the significance of every element where higher means more imperative. At each cycle, it checks whether a genuine component has a higher significance than the best of its shadow highlights and continually evacuates highlights, which are regarded exceptionally irrelevant. Finally, the calculation stops either when all elements get affirmed or rejected or it achieves a predefined farthest point of irregular timberland runs. The proposed CBM framework is described in Fig. 1, and is divided into two phases. A method for calculating the FURIA based remaining life prediction uncertainty is proposed in this framework to address the key challenge in existing ML literature. Then we can easily calculate the maintenance cost for optimal CBM policy by using proposed framework. The FURIA prediction method used in the framework gives the predicted failure time information which is required in CBM optimization. In this article, we develop a framework for predicting the failure time using backward feature elimination and FURIA prediction errors which are obtained from training phase. In the FURIA training phase, our proposed model is trained under historical failure input dataset including measurements collection of condition monitoring at all the inspection data points (current and historical). In training process, the weights of the FURIA model are adjusted between the FURIA classifier output and the actual life percentage values to minimize the error. When we complete the FURIA classifier training, we can start the testing of the prediction performance of the trained FURIA model and calculate the FURIA prediction error which is defined as the difference between the FURIA predicted failure time and actual failure time on the test dataset. Finally we obtain the set of FURIA lifetime prediction error values. We assume that the prediction error which is calculated

ACCEPTED MANUSCRIPT

by FURIA, is normally distributed and the probabilistic value of prediction uncertainty is the main expertise of our proposed FURIA classifier based maintenance predictive model. Now we can calculate easily the statistical measures of the FURIA lifetime prediction error by using these FURIA prediction error values. Now we can use this proposed framework for decreasing the total cost of maintenance, material and spare parts inventories, unplanned downtime and downtime cost and increasing the failure rate prediction, quality standards, equipment's reliability and uptime and support of continuous flow.

3.2 Dataset description and model validation

Coraddu et al. (2014) worked on this data and developed a meta-classifier for CBM optimization and uploaded the Gas Turbines (GT) data on freely accessible UCI machine learning and artificial Intelligence repository. They collecetd the data from a sophisticated simulator of a GT, which is mounted on a Frigate characterized by a COmbined Diesel eLectric And Gas (CODLAG) propulsion plant type. For generating the data, they worked on multiple module forming the five simulators including the propeller simulator and GT simulator. The others simulators gear box simulator, hull simulator and the last controller simulator which have been developed over the year. After the releasing of these simulator's dataset it is possible to calculate the performance decay on all observations and other components like turbine decay and compressor decay and simply we can say it is an agreement with available data with possible vessel value. Compressor and turbine degradation coefficient are the parameters which have been described for propulsion system behaviour so that a combination of triple can be explained with possible degradation state. The compressor and turbine decay coefficient tested with the precision 0.001 for getting good scale of presentation. The compressor and turbine decay coefficient have been investigated in the domain for implementing the condition based maintenance policy. The stored dataset has sixteen features vector which represents the system performance decay and contains the physical asset for the gas turbine

ACCEPTED MANUSCRIPT

measures. Position of lever, GT revolution rate, Propeller Torque of starboard and port, GT compressor temperature, GT compressor decay coefficient, speed of ship, revolution rate of gas generator, shaft torque of GT, fuel flow, turbine decay coefficient, Inlet and outlet air pressure of GT compressor, turbine injection control, Turbine exit temperature, turbine exit pressure and GT exhaust pressure. The main goal of the proposed framework is to train the classifier of machine learning for predicting the remaining life information and to develop the CBM approach for optimizing the maintenance prediction schedule. The GT dataset have 11934 instances for binary classification task and labelled by different fault types like decay coefficients and temperature inputs. Each instance of this classification task of GT simulator dataset has one fault type and 16 independent feature variables. The fuzzy rule-based classifier model utilizes the failure histories based on inspection data collected in this period. If we utilize the historical failure information in proper way then we can use this information to build the accurate predictive model and we can achieve the better remaining life prediction. We start the process to train the FURIA model by using the historical failure dataset with optimal failure times. After being trained, we can use the proposed FURIA model for predicting the remaining life uncertainty subject to collected condition-based monitoring measurements.

ACCEPTED MANUSCRIPT

4. Approaches used in Big Data Analytics based CBM Prediction Framework

An intense interest has been recently attracted by Big Data Analytics for its attempt or bid to extract wisdom, knowledge and information from Big Data. With the coming up of ICT and sensor technology in the industry, reams of nonlinear, streaming and high-dimensional data are being gathered and organised for supporting decision making. The faults detection or diagnosis in these data sources is one of the major implementations in eMaintenance solutions, since it could enable maintenance decision making. Timely detection of any kind of fault in the system may reduce chances of accidental breakdowns, and ensure the safety as well as the reliability of industrial systems. Data complexities, such as high nonlinearity, fast-flowing streams, and high dimensionality hamper fault detection/diagnosis applications. From the perspective of data modelling, high dimensionality deteriorate the correctness of fault diagnosis algorithms. Data complexities like fast-flowing data streams require fault diagnosis algorithms to provide near real-time or real-time responses as new samples arrive. High nonlinearity of data requires methods of fault detection to have adequately expressive potential, and to prevent or avoid problem of overfitting. In the existing literature, the big data concept by the 3 Vs i.e. high velocity, volume and variety, along with one C that denotes complexity. Velocity refers to data speed in and out, volume includes both the number of dataset dimensions and instance size, Variety refers to the range of the types and sources of data, and the Complexity refers to the high nonlinearity, fast-flowing streams, and high dimensionality, poor data quality and various other dataset complexities. In general, a dataset can be termed as Big Data if it difficult to acquire, curate, analyse and visualise it making use of the various technologies that exist (Kumar et al. 2016). Data is of numerous forms in the maintenance area. They can be generic like maintenance objectives and strategies data, and specific like data on maintenance work orders. They can originate or emanate from various internet sources, printed user manuals or Information System. They can be unstructured or well-structured.

ACCEPTED MANUSCRIPT

4.1 FURIA & Fuzzy Rules Representation

Huhn & Hullermeier (2009) have developed a fuzzy classification method which is called FURIA (Fuzzy Unordered Rule Induction Algorithm) and it extends the machine learning classification capabilities of RIPPER algorithm (Cohen, 1995) by adding more comprehensible rule sets. When we complete the FURIA classifier training, we can start the testing of the prediction performance of the trained FURIA model and calculate the FURIA prediction error which is defined as the difference between the FURIA predicted value and actual value on the test dataset. We assume that the prediction error which is calculated by FURIA, is normally distributed and can calculate the FURIA prediction error values. In particular, FURIA uses a rule stretching method for classifying the data and uses the fuzzy and unordered rules instead of conventional non-fuzzy rules because fuzzy rules have lots of benefits compare to non-fuzzy rules. When we try to develop the model by using conventional rules then it gives "sharp" decision boundaries but with abrupt transitions between different classes so we prefer the fuzzy membership function. A fuzzy set is obtained through replacing set intervals by trapezoidal membership function. We can specify this kind of fuzzy interval by four types of parameters

$$X^F = (Z^{s,L}, Z^{c,L}, Z^{c,U}, Z^{s,U}) \tag{1}$$

$$X^f(y) = \begin{cases} 1 & z^{c,L} \le y \le z^{c,U} \\ \dfrac{y - z^{s,L}}{z^{c,L} - z^{s,L}} & z^{s,L} \le y \le z^{c,L} \\ \dfrac{z^{s,U} - y}{z^{s,U} - z^{c,U}} & z^{c,U} \le y \le z^{s,U} \end{cases} \tag{2}$$

Where $Z^{c,L}$ and $Z^{c,U}$ are the lower bound and upper bound of the membership function and $Z^{s,L}$ and $Z^{s,U}$ are the support value of upper bound and lower bound. We can define a fuzzy rule having n selectors, i= 1, 2.,.n covers z:

ACCEPTED MANUSCRIPT

$$f(x)=\prod_{i=1...n} I^{f}(x_i) \quad (3)$$

Now we try for obtaining the fuzzy rules from RIPPER algo, which is basically used for fuzzification and searching for best combination of fuzzy extension in the same structure replaced by fuzzy extension intervals. For a single antecedent in training dataset T_d :

$$T_d = \left\{ n=(n_1...n_n) \in T_d \; \square I^{f}(x_n) \geq 0 \text{ for all } n \neq i \right\} \subseteq T_d \quad (4)$$

Now we partition T_d function into positive and negative instances and try to measure the fuzzification quality using purity test:

$$P_u = \frac{P_x}{P_x + N_x} \quad (5)$$

Where

$$P_x = \sum_{x \subseteq T_{d+}} \mu_A(x) \quad (6)$$

$$P_x = \sum_{x \subseteq T_{d-}} \mu_A(x) \quad (7)$$

In each iterations, rules are fuzzified and try to find out the best fuzzification computed for each and every antecedents.

Fuzzification algo for single rule

1: X= set of antecedents of single rule c

2: while $X \neq \phi$ do

3: $x_{max} \leftarrow null \{ x_{max} = highest\ purity\ antecedent$

4: $Pu_{max} \leftarrow 0 \{ Pu_{max} = highest\ purity\ value \}$

5: for i← 1 to size (X) compute Fuzzification of X_i

6: $Pu_{max} \leftarrow$ purity of Fuzzification

7: if $Px_i > Pu_{max}$ then

8: $Pu_{max} \leftarrow Pu_{X\{i\}}$

9: $x_{max} \leftarrow Px_i$

10: end if

ACCEPTED MANUSCRIPT

11: end for

12: $X \leftarrow X \backslash x_{max}$

13: Update the Rule r with x_{max}

14: end while loop

All the ties having larger distance from the centre are broken and this is repeated until all antecedents with the largest purity of the whole dataset have been identified and fuzzified. For best fuzzification we test all values:

$$\left\{ \frac{n_i}{n=(n_1,...,n_n)} \in T_d, n_i \leq Z_i^{c,L} \right\} \quad (8)$$

$$\left\{ \frac{n_i}{n=(n_1,...,n_n)} \in T_d, n_i \leq Z_i^{c,U} \right\} \quad (9)$$

For calculating the output of classifier we take a new instance n and the class support defined by:

$$S_n(x) = \sum_{i=1...n} \mu_r n(x).CF(r_n^{(j)}) \quad (10)$$

Where $CF(r_n^{(j)})$ = Certainty Factor of the Fuzzification rule and defined by:

$$CF(r_n^{(j)}) = \frac{2\frac{T_d^{(j)}}{T_j} + \sum_{x \in T_d^{(j)}} \mu_r n(x)}{2 + + \sum_{x \in T_d} \mu_r n(x)} \quad (11)$$

Where $T_d^{(j)}$ = subsets of training instances. We estimate the lifetime distribution of components using maximum likelihood method on failure and suspension dataset. We have the distribution parameters $\propto$ which is Weibull scale parameter and shape parameter β. We can write likelihood function as follows (Wu et al., 2013):

$$C_r = \frac{P_n(t+T) - P_n(t)}{1 - P_n(t)} \quad (12)$$

Where t= age of components, T= length of inspection interval and P_n = normal distribution function of failure time prediction using FURIA.

ACCEPTED MANUSCRIPT

Now we calculate the expected total replacement cost as follows:

$$C(t_m)=\int_0^{\infty} f_m(t_m).C_f.I(t_{pr}(t_n)\geq t_m)dt_n \qquad (13)$$

And the expected total replacement time can be calculated as follows:

$$T(t_m)=\int_0^{\infty} f_m(t_n).t_f.I(t_{pr}(t_n)\geq t_m)dt_n \qquad (14)$$

So the total expected replacement cost with respect to failure probability threshold value C_{ta} as follows:

$$C_{ta}=\int_0^{\infty}\frac{\beta}{\alpha}\left(\frac{t_m}{\alpha}\right)^{\beta-1}\exp\left[-\left(\frac{t_m}{\alpha}\right)^{\beta}\right]*C_t(t_m)dt_m \qquad (15)$$

And the total expected replacement time with respect to failure probability threshold value T_{ta} as follows:

$$T_{ta}=\int_0^{\infty}\frac{\beta}{\alpha}\left(\frac{t_m}{\alpha}\right)^{\beta-1}\exp\left[-\left(\frac{t_m}{\alpha}\right)^{\beta}\right]*T_t(t_m)dt_m \qquad (16)$$

Now we can calculate the total replacement cost per unit of time in CBM framework considering threshold probability value as follows:

$$C_{exp}(P_r)=\frac{C_{ta}}{T_{ta}} \qquad (17)$$

In data exploration phase there are various things we will need to do for selecting the appropriate feature selection method. What is information in the dataset, how much and what is the nature of this dataset. Wrappers, filters and embedded are the three methods of feature selection. In preprocessing step, filter methods are usually made use of. No machine learning algorithms are applied for selecting the features. Selection of features is done based on their scores in different statistical examinations for their correlation or interdependence on outcome variables. ANOVA, Chi Squared Test and correlation coefficient scores are examples of filter methods. It is important to bear in mind that multicollinearity is not removed through filter methods. So, it is also important to deal with or handle multicollinearity of features ahead of training or directing models for our data. A subset of features is usually used for training a model in wrapping methods. We decide to remove or add

ACCEPTED MANUSCRIPT

features to the subset on the basis of the conclusions that we make from the previous model. The problem is fundamentally reduced to a search problem. The key differences that exist between wrapper methods and filter methods for selection of features are as follows. Wrapper methods measure the relevance of a subset of features by way of training a model on that subset and filter methods calculate the appropriateness or relevance of features by their correlation or connection with dependent variable. Wrapper methods are slow and very expensive and filter methods are much faster than wrapper methods because they do not involve in training of a model. Wrapper methods use the cross validation for evaluating a subset of features and provide the most appropriate or the best subset. Filter methods use statistical methods for evaluating the features but usually fail to discover the best subset of features. Using the subset provided by wrapper methods often make the model vulnerable to overfitting but this is not the case with filter methods. First of all we work on a Canadian Kraft Mill company (Stevens, 2006) dataset which is collected from Gould pump bearings to give an overview of our proposed big data fuzzy analytics maintenance framework. The main objective of this dataset is to minimize the frequency of pump failure. As we know that bearings are the critical components for failing of the pumps in this dataset. We extract the failure dates information, operating starts data, event data and out-of-service intervals event data from historical event database. When we start data pre-processing and exploration phase then we classify the data into two categories- inspection and event dataset. Event dataset further can be classified into three categories- beginning, suspension and failure data event. We collect the total 30 histories (mainly focus on failure and preventive replacements dataset) from 7 different pump locations. So based on this input dataset we fist calculate the actual inspection & fitted value measurements and then we compare the maintenance cost. We record the total 49(=7*5+7*1+7*1) vibration measurements records for working with inspection level data. We divide whole CBM optimization process into four steps. The first step is data preparation and exploration including significant analysis and parameter

ACCEPTED MANUSCRIPT

estimation and the other steps are: building the transition probability matrix, estimation of cost data and the last is condition-based optimization (Wu at al. 2013).

Step 1- Data exploration (significant analysis and parameter estimation)

Variable identification, transformation and the new variable creation are the steps which are used to prepare our data for building the predictive analytics and significant analysis. We perform the data exploration and significance analysis for all 49 measurements and we get the information that P1H_Par5 and P1V_Par5 are the covariates which have greatest effect on bearing's health condition. Based on this analysis we try to calculate the four parameters: scale parameter ($\propto$), shape parameter ($\boldsymbol{\beta}$), covariate weight for P1H_Par5 ($\boldsymbol{\gamma}$) and weight for P1V_Par5 ($\boldsymbol{\delta}$) as follows:

$$x_1(t) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1} e^{(\gamma z_{P1H}(t)+\delta z_{P1V}(t))} \tag{18}$$

$$= \frac{3.12}{2650}\left(\frac{t}{2650}\right)^{3.12-1} e^{(20.09 z_{P1H}(t)+54.26 z_{P1V}(t))} \tag{19}$$

And the others fitted measurements values are given as follows:

$$x_2(t) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1} e^{(\gamma z_{P1H}(t)+\delta z_{P1V}(t))} \tag{20}$$

$$= \frac{3.42}{2670}\left(\frac{t}{2670}\right)^{3.42-1} e^{(21.19 z_{P1H}(t)+53.46 z_{P1V}(t))} \tag{21}$$

Now we start the process to build the transition probability matrix for calculating the maintenance cost which indicates the probabilistic values of covariate at next inspection time. Based on actual inspection and fitted measurements and taking the value 28 days of inspection interval we build a table 1 for all measurements.

ACCEPTED MANUSCRIPT

Step 2- Maintenance cost estimation

In maintenance services, cost estimation or evaluation is an essential component. It directly affects the economic performance of businesses. Underestimation or overestimation of services can harm the performance of enterprises. While overestimation scars the image of the firm in the market, underestimation results in financial losses. In the manufacturing industry, cost estimation and evaluation has been the basis of design. The importance of cost estimation is very much evident when examining or investigating business policies and strategic issues in many companies as accurate or reliable cost estimations are significant for competitive bidding as well as pricing (H'midaa et al., 2006). After obtaining the transition matrix we start the process for estimating the failure and preventive replacement cost. Based on historical data and previous experience we estimate the cost of replacement to be (C=$3000) and the cost of failure is $10000 so the penalty cost is (K=$7000).

Step 3- CBM policy optimization

For performing the optimal Condition-based Maintenance policy, we begin with assessing the condition of a new component or equipment at fixed intervals. At every point of assessment, the probability of conditional failure during the next interval computed and thereafter compared with the optimal threshold failure probability. A failure replacement needs to be performed every time the failure occurs. Here, a total of five test histories have been used for demonstrating the proposed method for optimising CBM. The same procedure can be used for calculating the probability of failure at each point of inspection/assessment for all test histories and the decisions for replacement can be made. Now, we have calculated all the required values of transition matrix and other cost parameters. So using these parameters based on inspection measurements, we obtain these values:

$$\propto = 2654,\ \beta = 2.246,\ \gamma = 23.07,\ \delta = 61.24$$

And the CBM optimal policy is as:

ACCEPTED MANUSCRIPT

Optimal risk threshold level (D) = 17.32$/day

Optimal maintenance cost (C) = 15.74$/day and the average replacement interval = 802 days.

Now we will calculate the CBM policy based on the fitted measurements parameters:

$\propto$ =2624, β = 2.466, γ = 22.07, δ = 68.74

And the CBM optimal policy is as:

Optimal risk threshold level (D) = 19.32$/day

Optimal maintenance cost (C) = 14.27$/day and the average replacement interval = 892 days. Now we calculate the average maintenance cost using the Weibull-FR function and by comparing the CBM maintenance results before and after inspection measurements. Based on the inspection measurement and fitted measurement the final result is shown in table 2.

5. Result & Discussion

We propose a two phase prediction-based maintenance big data analytics framework for optimizing the maintenance schedule based on feature engineering and fuzzy unordered rule based induction algorithm. FURIA is an extension of RIPPER algorithm and produces the better decision boundaries between different classes. In the first phase of our proposed framework, we apply the feature engineering on the available gas turbine dataset. It creates the new variables and then we focus on those observations which are much higher than rest of samples and in last Fuzzy classifier is trained with best prediction accuracy of backward feature elimination method. In the second phase, we delete the outlier value and the target value will be replace by the predicted value of trained fuzzy classifier. To demonstrate the effectiveness of our proposed fuzzy classifier based framework, we use a big dataset, which is generated from GT simulator. After applying the proposed method on

ACCEPTED MANUSCRIPT

required dataset, our research framework can calculate the remaining life prediction uncertainty and a threshold value of failure probability. Table-3 presents the result of statistical measures and distribution ratio of classes using this two phase backward feature elimination and fuzzy unordered classifier method. Column 1 and 2 in Table 3 present sensitivity and accuracy measures of SVM, ANN, logistic regression, KStar algorithm, BayesNet, LogitBoost, Random Forest, AdaBoost, RIPPER algo and our proposed framework based on backward feature elimination and FURIA combination. Table 4 presents the results when the best hybrid method is selected and then it trained by backward feature elimination and FURIA and we show the results in terms of F-1 score, sensitivity, precision, accuracy, specificity, false positive & negative rate and negative predicted value. It is observed from the results that pre-processing of data using the feature selection and backward feature elimination method gives the better prediction of instances in terms of sensitivity and accuracy. In last phase, this modified data is used for training the FURIA and other machine learning algorithms on a case study of gas turbine predictive maintenance dataset for predicting the lifetime prediction errors and failure probability threshold value by using our proposed framework. We can see the result in ROC curve, lift chart and confusion matrix. Lift chart is simply use to calculate the ratio between obtained results for measuring the effectiveness of proposed model. Receiver operating characteristic (ROC) curve which is shown in figure 2 is a plot that provide the information about the classifier when we change the threshold value continuously. Confusion matrix is an error table that is used for measuring and visualizing of the performance of any machine learning algorithms (Kumar et al. 2016). In error confusion matrix, each column represents the predicted class measures and each row represents actual class measures. We calculate the predicted values in lift curve when classifier gives probability of each class in error matrix. In ROC we built the graph by adding the total number of cases on x- axis and TPN values (correctly classified instances) on y- axis. ROC curve generally uses the same variable with maximum percentage on y-axis and true negatives values on x- axis. Our hybrid approach starts with the feature selection

ACCEPTED MANUSCRIPT

approach in all n dimensions which is a backward feature elimination method. First of all we compute the sum of square of error after eliminating each variable and we repeat the process n times. Then, we try to identify those variables (with *n-1* input features) whose removal produce smallest increase in the sum of square of error and finally we remove it. We repeat this process in this feature elimination phase until no other instances can be dropped. In the second phase, when we apply the FURIA method on this dataset, our proposed framework works on growing set with a rule specialization by adding multiple antecedents. When RIPPER learns a rule by removing antecedents for a particular class, the examples denoted as positive instances in this particular class, whereas the other classes examples are denoted as negative. Now we generate a new rule on this growing data and adds selectors with an empty conjunction until the rule covers those negative instances who do not belong to the target class. The dependent variable in this method can take the value 1 with the probability q because of it dichotomous nature and value 0 with the probability 1-q so we use the binary prediction of faults in 'Yes' and 'No' category. We analyse the statistical measures and estimation results of FURIA classifier by using ROC, error matrix and lift curve in big data environment. We performed backward feature elimination and FURIA method on gas turbine dataset and obtained the result in Table 3, 4 and 5.The ROC provides the graphical representation between the true positives instances and the false positives instances into the relative adjustments. The ROC provides the graph between the faults that are correctly classified as faults, and percentage of non-faults that are wrongly classified as faults. In a standard ROC curve, each point corresponds to a prediction class result of a confusion matrix. ROC curve for SVM, ANN, BayesNet, LogitBoost, Random Forest and FURIA algorithms on validation data has been shown in Figure 2. In ROC curves, we consider the best predictive model for which the ROC curve passes close to 1 and it is used to measure the performance of classifier. If it close to 1 then we can say the proposed model has 100% sensitivity and specificity, it means classifier output don't have any false positive and false negative values. In FURIA and RIPPER models, generally a score is produced, which uses a

ACCEPTED MANUSCRIPT

probabilistic threshold value for classifying the binary class in faults or no-faults cases. If the classifier's predicted output value is equal or greater than the probabilistic threshold value, it will be considered in fault category, otherwise not. AUC is other performance indicator of any classifier which gives the information about the best predicted classes of model and represents the probabilistic discrimination between yes and no category and comes under the ROC curve. In general we can say that predictive models having AUC value equal or greater than 0.5 are considered better compare to other ML models.

6. Conclusions, Limitations & Future Direction

Presently, condition monitoring is capable of identifying when machine problems will arise and, if it possesses enough experience, of locating the precise cause. However, it remains a challenge to forecast the machine's remaining life uncertainty during condition-based maintenance schedule. Similarly, it is tough to discern whether to change or to maintain the machine. Existing documentation on prediction based remaining life estimation has focused on simply reliability-based models, and the need for an uncomplicated, well-organized prediction model which can be promptly used by industries remains unaddressed. This paper is an attempt to locate this deficiency and to introduce a suitable big data analytics condition-based predictive maintenance model. Condition monitored measurements have stable and failure zone and if the measurements are normal then we use a reliability-based model. An increase in the condition measurements indicates a potential problem, and the remaining machine life estimation is calculated on the basis of reliability and condition monitoring information. In order to critically analyse the model, a gas turbine case study was carried out which demonstrated encouraging initial results with every machine failure being predicted beforehand. The predictive accuracy of the model on a hold-out sample. After considering the FURIA effect and using our proposed approach, the performance of model also improved exceptionally in the form of sensitivity and accuracy in this study. We developed the big data

ACCEPTED MANUSCRIPT

analytics framework for analysing the maintenance data from an eMaintenance point of view and explored the potential opportunities and challenges to support CBM decision-making. Our result shows that traditional machine learning models like SVM, ANN, logistic regression and random forest don't have the desired accuracy for a given range compare to our proposed hybrid model based on MAPE, MSE and MAD. This made it amply clear that the proposed model relied upon the precision and quality of the measurements. The model is expected to be utilized in numerous condition monitored situations given that the failure lead time is adequately long and CM accurately displays the health of the machine. This article presented a hybrid approach which is a combination of backward feature elimination and FURIA for estimating the reliability of degrading components in big data environment on a condition monitoring dataset. An efficient utilization of backward feature elimination and degradation experiment combined with a FURIA predictive model have used to obtain the failure time distribution on condition monitoring data. Our proposed CBM framework at certain inspection points is summarized in two ways: (1) failure generally occurs during the previous and next inspection interval. (2) If it occurs during next inspection interval then we perform preventive replacement and if it occurs during previous inspection interval then we work on failure replacement Otherwise, the operation can be continued. A fuzzy unordered-based method makes room for more general modelling and involves prior knowledge as informative priors. This method is quite in contrast to the oft-repeated prevailing methods that resort to approximate assumptions in order to acquire distribution in closed form. That the maintenance decisions can be managed in a more economical manner by including CBM data or prior knowledge has been distinctly demonstrated by the simulation examples. Updating methodology allows the user to choose the condition dataset to be assembled so as to make reliability estimations confidently. Moreover, if there is ample information available about the rate at which the asset will age prior to its use, the aforementioned methodology supplies a synthesized framework to integrate with probability distributions in order to polish reliability estimations and to revamp maintenance decisions.

ACCEPTED MANUSCRIPT

Backward feature elimination model and FURIA have been used by the proposed method for the purpose of predictive maintenance and reliability modelling. High geometric mean and F-measure values cannot be obtained concurrently or simultaneously due to the overlapping of the observed values under faulty and normal conditions. A majority of research studies on predictive maintenance deduce that the maintained component of equipment is reinstated 'as good as new'. Only important equipment or components are practically monitored or examined, and maintenance of only faulty critical equipment/components is carried out. It is unreasonable to presume that the replaced maintenance equipment or failure component is restored or put back to 'good as new' state or condition. It is necessary to consider faulty maintenance in times ahead of this research work. As techniques of manufacturing develop or pioneer, manufacturing systems grow more complicated. Certain types of manufacturing lines or systems may be made up of several critical components which need to be continuously checked and maintained. In any multicomponent system, stochastic and economic dependences prevail. There exist dependences that cannot be quantified. Although dependence can be modeled and defined, it is possible that the modeled dependences will multiply and intensify the computation complexities and difficulties. Future study will be conducted considering multicomponent systems. A further augmentation of these assumptions to exponential models and auto-correlated error distributions can be considered for exploration in future CBM research.

ACCEPTED MANUSCRIPT

References

Berenguer, C., Grall, A., Dieulle, L.,& Roussignol, M.;1; (2003). 'Maintenance policy for a continuously monitored deteriorating system'. *Probability in the Engineering and Informational Sciences*, 17 (1):235–250.

Bever, K.;1; (2000). 'Enterprise Systems integration: opportunities and obstacles developing plant asset management systems'. *National Manufacturing Week*, Chicago.

Bengtsson, M.;1; (2004). 'Condition Based Maintenance System: an Investigation of Technical Constituents and Organizational Aspects. Malardalen University Licentiate Thesis.

Bhattacharya, A., Mohapatra, P., Kumar, V., Dey, P.K., Brady, M., Tiwari, M.K., & Nudurupati, S.S.;1; (2014). "Green supply chain performance measurement using fuzzy ANP-based balanced scorecard: a collaborative decision-making approach". *Production Planning and Control*, 25(8):698-714.

Bhattacharya, A., Dey, P.K., & Ho, W.;1; (2015). 'Green manufacturing supply chain design and operations decision support'. *International Journal of Production Research*, 53(21):6339-6343.

Castanier B, Grall A and Berenguer C;1; (2005). "A condition-based maintenance policy with non-periodic inspections for a two-unit series system". *Reliability Engineering and System Safety*, 87(1): 109–120.

Cohen, W.;1; (1995). "Fast effective rule induction", Proceedings of the *Twelfth International Conference on Machine Learning*, pp. 115–123.

Coraddu, A., Oneto, L., Baldi, F. & Anguita, D.;1; (2017). "Vessels fuel consumption forecast and trim optimisation: A data analytics perspective". *Ocean Engineering*, 130(1):351-370.

Chouikhi, H., Khatab, A. & Rezg, N.;1; (2014). 'A condition-based maintenance policy for a production system under excessive environmental degradation'. *Journal of Intelligent Manufacturing*, 25(4):727-737.

Coraddu, A., Oneto, L., Ghio, A., Savio, S., Anguita, D.,& Figari, M.;1; (2014). 'Machine learning approaches for improving condition-based maintenance of naval propulsion plants'. Proceedings of the Institution of Mechanical Engineers, Part M: *Journal of Engineering for the Maritime Environment,* 1(1):1–18.

Dey, P.K.;1; (2004). 'Decision support system for inspection and maintenance: a case study of oil pipelines'. *IEEE Transactions on Engineering Management*, 51(1):47-56.

ACCEPTED MANUSCRIPT

Dey, P.K.;1; (2001). 'A risk-based model for inspection and maintenance of cross-country petroleum pipeline'. *Journal of Quality in Maintenance Engineering*, 7(1):25-43.

Eustace R.W.;1; (2008). 'A Real-World Application of Fuzzy Logic and Influence Coefficients for Gas Turbine Performance Diagnostics', *ASME Turbo Expo*, 29(1):116-119.

Emovon, I., Norman, R.A. & Murphy, A.J.;1; (2015). 'Hybrid MCDM based methodology for selecting the optimum maintenance strategy for ship machinery systems'. *Journal of Intelligent Manufacturing*, 26(3):1-13.

Feng, Q., Peng, H., & Coit, D.W.;1; (2010). 'A degradation-based model for joint optimization of burn-in, quality inspection, and maintenance: A light display device application'. *International Journal of Advanced Manufacturing Technology*, 50(1):801–808.

Grall, A., Dieulle, L., Berenguer, C., & Roussignol, M.;1; (2002). 'Continuous-time predictive maintenance scheduling for a deteriorating system'. *IEEE Transactions on Reliability*, 51(2):141–150.

Gebraeel, N.Z., Lawley, M.A., Li, R., & Ryan, J.K.;1; (2005). 'Residual-life distributions from component degradation signals: A Bayesian approach'. *IIE Transactions*, 37(1):543–557.

Goyal, D., Pabla, B.S., Dhami, S.S., & Lachwani, K.;1; (2016). 'Optimization of condition-based maintenance using soft computing'. *Neural Computing and Applications*, 27(1):1-16.

H'midaa, F., Martinb, P. and Vernadata, F.-O.;1; (2006), "Cost estimation in mechanical production: the cost entity approach applied to integrated product engineering", *International Journal of Production Economics*, 103(1):17-35.

Huhn, J.C., & E. Hullermeier, E.,;1; (2009). "FURIA: an algorithm for unordered fuzzy rule induction", *Data Mining and Knowledge Discovery*, 19 (1): 293–319.

Jiang X, Duan F, Tian H & Wei X.;1; (2015). "Optimization of reliability centered predictive maintenance scheme for inertial navigation system". *Reliability Engineering & System Safety*, 140(1):208–17.

ACCEPTED MANUSCRIPT

Kacprzynski GJ, Roemer MJ, Modgil G, Palladino A and Maynard K;1; (2002). "Enhancement of physics-of-failure prognostic models with system level features". In: Proceedings of the *2002 IEEE Aerospace Conference*, Big Sky, MT.

Khatab, A.;1; (2015). 'Hybrid hazard rate model for imperfect preventive maintenance of systems subject to random deterioration'. *Journal of Intelligent Manufacturing*, 26(3):601-608.

Kharoufeh, J.P., & Cox, S.M.;1; (2005). 'Stochastic models for degradation based reliability'. *IIE Transactions*, 37(1):533–542.

Kennedy, R.;1; (2001). 'Examining the Processes of RCM and TPM'. www.plant-maintenance.com/articles/RCMvTPM.shtml.

Kumar, A., Shankar, R., Choudhary, A., & Thakur, L.S.;1; (2016). 'A Big Data MapReduce Framework for Fault Diagnosis in Cloud-Based Manufacturing'. International Journal of Production Research, 54(23):7060-7073.

Li, Z., Yan, X., Guo, Z., Zhang, Y., & Yuan, C.;1; (2012). 'Condition monitoring and fault diagnosis for marine diesel engines using information fusion techniques', *Electron & Electrical. Engineering*, 123 (7):109-112.

Li CJ and Lee H;1; (2005). "Gear fatigue crack prognosis using embedded model, gear dynamic model and fracture mechanics". *Mechanical Systems and Signal Processing*, 19(4): 836–846.

Loboda I., Feldshteyn Y., & Ponomaryov V.;1; (2011). 'Neural Networks for Gas Turbine Fault Identification:Multilayer Perceptron or Radial Basis Network, *ASME Turbo Expo*, 29(1):116-119. .

Lugtigheid D, Banjevic D and Jardine AKS;1; (2008). "System repairs: When to perform and what to do?" *Reliability Engineering and System Safety,* 93(4): 604–615.

Lu, C.J., & Meeker, W.Q.;1; (1993). 'Using degradation measures to estimate a time-to-failure distribution'. *Technometrics*, 35(2):161–174.

Liao, H., Elsayed, E.A., & Chan, L.Y.;1; (2006). 'Maintenance of continuously monitored degrading systems'. *European Journal of Operational Research*, 175(1):821–835.

Marble S and Morton BP;1; (2006). "Predicting the remaining life of propulsion system bearings". In: *Proceedings of the 2006 IEEE Aerospace Conference*, Big Sky, MT.

ACCEPTED MANUSCRIPT

Ogaji, S.O.T., & Singh, R.;1; (2003). 'Gas Path Fault Diagnoses Framework for a 3- Shaft Gas Turbine', Proceeding of the Institution of Mechanical Engineers.' *Journal of Power and Energy*, 217(3):149-157. .

Paik, B.G., Cho, S.R., Park, B.J., Lee, D.K.,& Bae, B.D.;1; (2010). 'Development of real-time monitoring system using wired and wireless networks in a full scale ship'. *International Journal of Naval Architecture & Ocean Engineering*, 2 (3):132-138.

Pong J. L., Ming C. Z., Tzu C. H., & Jin Z.;1; (2000). 'An evaluation of engine faults diagnostics using artificial neural networks', *ASME Turbo Expo*, 29(1):116-119. .

Rao, B.;1; (1996). 'Handbook of Condition Monitoring'. Elsevier Science- 1st edition.

Ross M. Jonathan,;1; (2002). 'Condition-Based Maintenance -A Tool for Improving Productivity in Shipyards'. *Journal of Ship Production*, 18(3):175-184.

Romessis C., & Mathioudakis K.;1; (2006). 'Bayesian Network Approach for Gas Path Fault Diagnosis', *ASME Journal of Engineering for Gas Turbines and Power*, 128(1):64-72. .

Romessis C., & Mathioudakis K.;1; (2003). 'Setting Up Of a Probabilistic Neural Network for Sensor Fault Detection Including Operation with Component Faults', ASME Journal of Engineering for Gas Turbines and Power, 125(3):634-641.

Roy, R. , Stark, R., Tracht, K., Takata, S. and Mori, M.,;1; (2016). "Continuous maintenance and the future – Foundations and technological challenges," *CIRP Annals - Manufacturing Technology*, 65(2): 667-688.

Sampath S., & Singh R.;1; (2004). 'An integrated fault diagnostics model using Genetic Algorithm and Neural networks', *ASME Turbo Expo*, 29(1):116-119.

Shaw, K., Shankar, R., Yadav, S.S., & Thakur, L.S.;1; (2013). 'Modelling a low-carbon garment supply chain'. *Production Planning and Control*, 24(9):851-565.

Singh, S. Olugu, E.U. Musa, & S.N. Mahat, A.B.;1; (2015). 'Fuzzy-based sustainability evaluation method for manufacturing SMEs using balanced scorecard framework'. *Journal of Intelligent Manufacturing*, 27(1):1-18.

ACCEPTED MANUSCRIPT

Stevens, B.;1; (2006). 'EXAKT Reduces Failures at Canadian Kraft Mill,' www.omdec.com.

Tian Z, Wong L and Safaei N;1; (2010). "A neural network approach for remaining useful life prediction utilizing both failure and suspension histories". *Mechanical Systems and Signal Processing*, 24(5): 1542–1555.

Tian Z and Liao HT;1; (2011). "Condition based maintenance optimization for multi-component systems using proportional hazards model". *Reliability Engineering and System Safety*, 96(5): 581–589.

Tian ZG, Wu BR, Chen MY.;1; (2014). "Condition-based maintenance optimization considering improving prediction accuracy". *Journal of the Operation Research Society*, 65(9):1412–22.

Validi, S., Bhattacharya, A. & Byrne, P.J.;1; (2015). 'A solution method for a two-layer sustainable supply chain distribution model'. Computers and Operation Research, 54(1):204-217.

Vachtsevanos G, Lewis FL, Roemer M, Hess A and Wu B;1; (2006). "Intelligent Fault Diagnosis and Prognosis for Engineering Systems". *Wiley: New York*.

Volponi, A.J., DePold, H., Ganguli, R., & Daguang, C.;1; (2003). 'The Use of Kalman Filter and Neural Network Methodologies in Gas Turbine Performance Diagnostics: A Comparative Study', *Journal of Engineering for Gas Turbines and Power,* 125(1):917-924. .

Wu BR, Tian Z and Chen MY;1; (2012). "Condition based maintenance optimization using neural network based health condition prediction". *Quality and Reliability Engineering International*, 29(8):1151-1163.

Wang, W.;1; (2000). 'A model to determine the optimal critical level and the monitoring intervals in condition-based maintenance'. *International Journal of Production Research*, 38(6):1425–1436.

Wang, J., Zhang, L., Duan, L. & Gao, R.;1; (2015). 'A new paradigm of cloud-based predictive maintenance for intelligent manufacturing'. *Journal of Intelligent Manufacturing*, 28(1):1-13.

Wu, B., Tian, Z. and Chen, M.;1; (2013). 'Condition-based Maintenance Optimization Using Neural Network-based Health Condition Prediction.' *Quality and Reliability Engineering International,* 29(1): 1151–1163.

Yun-Peng C., Shu-Ying L., Shuang Y., & Ning-Bo Z.;1; (2012). 'Fault Diagnosis of a Gas Turbine Gas Fuel System Using a Self-Organizing Network', *Advanced Science Letters*, 8(7):386-392. .

ACCEPTED MANUSCRIPT

Author-1 (Ajay Kumar): Ajay Kumar is a senior PhD scholar at Bharti School of Telecom Technology & Management, Indian Institute of Technology Delhi, New Delhi, India. He has joined the full time PhD program in January 2012 at IIT Delhi. He received his Bachelors of Technology (B.Tech) in year 2008 and Masters of Technology (M.Tech) in Electronics & Computer engineering from DCE (Delhi College of Engineering) in Year 2011. He has published various articles in reputed journals, including Telematics & Informatics, International Journal of Production Research etc. His research interests include big data analytics, Business Analytics, Data Mining and Operations Research. He is a member of Institute for Operations Research and Management Science (INFORMS), Decision Science Institute (DSI) and Association of Information Systems (AIS).

Author-2 (Ravi Shankar): Ravi Shankar is a Professor of Decision Science, Operation Management & Business Analytics at Indian Institute of Technology Delhi, India. His areas of interest are supply chain analytics, business analytics, operations research, big data analytics, fuzzy modelling, sustainable logistics, etc. He has published over 300 research papers in reputed journals, including Omega, European Journal of Operations Research, Expert System with Applications, Applied Soft Computing, International Journal of Production Research, International Journal of Production Economics, IEEE Systems Man and Cybernetics Part C, Computers and Industrial Engineering, Computers and Operations Research, etc.

Author-3 (Lakshman S. Thakur): Lakshman S. Thakur is a Professor of Operation & Information Management Department at School of Business, University of Connecticut, USA. Dr. Thakur has previously been a Visiting Professor of Operations Research at Yale School of Management 1985-1987. His primary research interests are in the development and applications of linear, nonlinear, and integer programming methods in Management Science and function approximations in optimization mathematics. He has published in Management Science, Mathematics of Operations Research, SIAM Journal on Applied Mathematics, SIAM Journal on Optimization and Control, Journal of Mathematical Analysis and Applications, Naval Research Logistics, and Computers and Operations Research. Dr. Thakur has served as a consultant to IBM Corporation on their Manpower Planning with Risk Assessment System and New Product Warranty System as well as a senior consultant and Director of Management Science in a consulting organization. He is an associate editor of Naval Research Logistics. His current research focuses on production scheduling, product design and facility location problems. His research on scheduling (with Dr. P.B. Luh) is supported by a National Science Foundation grant. He is a member of Operations Research Society of America, The Institute of Management Sciences, and Mathematical Programming Society.

ACCEPTED MANUSCRIPT

Figures Legend

Figure1. Proposed big data driven sustainable framework for CBM Prediction

Figure 2. ROC Curve for different Classifiers

ACCEPTED MANUSCRIPT

Tables

Table1. Transition probability matrix

P1H_Par5	**0 to 0.0592**	**0.0592 to 0.0156**	**0.0156 to 0.0521**	**0.0521 to 0.1675**	**Above 0.1675**
0 to 0.0592	0.5821	0.5231	0.2322	0.1122	0.1458
0.0592 to 0.0156	0.4622	0.4982	0.3422	0.1323	0.1227
0.0156 to 0.0521	0.4742	0.4242	0.4522	0.1482	0.2154
0.0521 to 0.1675	0.4832	0.3622	0.5622	0.2385	0.1874
Above 0.1675	0	0	1	1	1

ACCEPTED MANUSCRIPT

Table 2. CBM optimization results before and after fitting the data

Result/ Proposed Framework	Average CBM cost ($/day)	Average Replacement Interval (days)
Before	15.74	802
After	14.27	892
Change	9.34%	11.22%

ACCEPTED MANUSCRIPT

Table3. Performance Matrices Result

	Sensitivity	Accuracy	Specificity	F-1 Score	Negative Predictive Value	False Positive Rate	False Negative Rate	Precision
Support Vector Machine (SVM)	0.9989	0.9104	0.8533	0.8973	0.9992	0.1467	0.0011	0.8145
Multilayer Perceptron (ANN)	0.9268	0.9585	0.9922	0.9584	0.9274	0.0078	0.0732	0.9922
Logistic Regression	0.9424	0.9439	0.9452	0.9416	0.9467	0.0548	0.0576	0.9408
KStar Algorithm	0.9522	0.9627	0.9729	0.9616	0.9548	0.0271	0.0478	0.9713
Bayes Net	0.9326	0.9384	0.9439	0.9362	0.9370	0.0561	0.0674	0.9399
LogitBoost	0.7345	0.7817	0.8419	0.7903	0.7135	0.1581	0.2655	0.8554
Random Forest	0.9772	0.9749	0.9727	0.9738	0.9790	0.0273	0.0228	0.9704
AdaBoost	0.8745	0.6041	0.5695	0.3340	0.9726	0.4305	0.1255	0.2064
JRip (RIPPER Algorithm)	0.9703	0.9707	0.9710	0.9695	0.9726	0.0290	0.0297	0.9686
FURIA	0.9756	0.9761	0.9766	0.9752	0.9774	0.0234	0.0244	0.9747

Table 4. Confusion matrix and Error Report of FURIA

<table>
<tr><th colspan="4">Error Report of Fuzzy Unordered Induction Algo (FURIA)</th></tr>
<tr><td colspan="2">Kappa statistic 0.9522</td><td colspan="2">Confusion Matrix</td></tr>
<tr><td colspan="2">Mean absolute error 0.0244</td><td></td><td colspan="2">Predicted Class</td></tr>
<tr><td colspan="2">Root mean squared error 0.1372</td><td>Actual Class</td><td>True</td><td>False</td></tr>
<tr><td colspan="2">Relative absolute error 4.8757%</td><td>True</td><td>1119</td><td>29</td></tr>
<tr><td colspan="2"></td><td>False</td><td>28</td><td>1211</td></tr>
</table>

ACCEPTED MANUSCRIPT

Table 5. Detailed Accuracy Report

Detailed Accuracy by Class							
TP Rate	**FP Rate**	**Precision**	**Recall**	**F-Measure**	**MCC**	**ROC Area**	**PRC Area**
0.975	0.023	0.976	0.975	0.975	0.952	0.986	0.978
0.977	0.025	0.977	0.977	0.977	0.952	0.986	0.982
0.976	0.024	0.976	0.976	0.976	0.952	0.986	0.980**(Weighted Av.)**

ResearchGate

See discussions, stats, and author profiles for this publication at:
https://www.researchgate.net/publication/315668777

A Context Based Algorithm for Sentiment Analysis

Article *in* International Journal of Computational Vision and Robotics · January 2018

DOI: 10.1504/IJCVR.2017.10005890

CITATIONS
0

READS
38

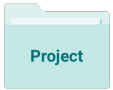
4 authors, including:



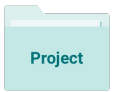
Shampa Chakraverty
Netaji Subhas Institute of Technology
73 PUBLICATIONS **150** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Emotion based trust networks, recommender systems and NLP View project



Enhancing the quality of E-Learning and E-Governance systems View project

All content following this page was uploaded by Shampa Chakraverty on 24 June 2017.

The user has requested enhancement of the downloaded file.

A Context Based Algorithm for Sentiment Analysis

Srishti Sharma*, **Shampa Chakraverty, Akhil Sharma**
Department of Computer Engineering,
Netaji Subhas Institute of Technology,
Dwarka, New Delhi- 110 078, India
E-mail: srish.060788@gmail.com
E-mail: apmahs.nsit@gmail.com
E-mail: akhil.sharma92@yahoo.com
*Corresponding author

Jasleen Kaur
Department of Electronics and Communication Engineering,
Netaji Subhas Institute of Technology,
Dwarka, New Delhi- 110 078, India
E-mail: jasleenkaurbatra@yahoo.co.in

***Abstract*— With netizens continuing to express a range of opinions and making assessments online, it has become a challenge to mine sentiments accurately from the ever-multiplying Big Data. We present a context-driven Sentiment Analysis scheme with the objective of refining the degree of subjectivity during Sentiment Analysis. The essence of our scheme is to capture in a stable manner, the mutual influence of the sentiments of neighboring words on the sentiment of each word in a document. A parametric Influence Function combines the native sentiment score of each word with the context-derived sentiment score obtained from surrounding words. We apply a Genetic Algorithm to fine tune the parameters of the Influence Function so as to obtain the best possible accuracy for a given corpus. Experimental results on hotel reviews extracted from Tripadvisor.com show an average accuracy of 73.2% which is 3.6% more than the results obtained from the Baseline Sentiment Analysis approach using native scores obtained from SentiWordNet. Though the improvement is small, it re-affirms our belief that**

contextual information provides valuable reinforcement of sentiment scores especially with regard to the borderline cases where words show near neutral sentiments. We also present a comparison with alternative Sentiment Analysis approaches that shows the strength of our proposed Context Based Sentiment Analysis approach.

***Keywords*-- Sentiment Analysis, Opinion Mining, Context Based Sentiment Analysis, SentiWordNet, Word Sense Disambiguation**

Biographical Notes: Srishti Sharma is a PhD research scholar in the Computer Engineering Department at Netaji Subhas Institute of Technology, New Delhi, India. She holds a B.E. degree in Computer Science and Engineering from Maharishi Dayanand University, Rohtak and M.Tech degree in Information Systems from Delhi Technological University, Delhi. Her research interests include Sentiment Analysis and Biometrics.

Shampa Chakraverty is Professor and Head of the Computer Engineering Division at Netaji Subhas Institute of Technology, New Delhi, India. She completed her Bachelor's degree in Electronics and Communication ('83) from Delhi College of Engineering, M.Tech in Integrated Electronics and Circuits from IIT-Delhi ('92) and PhD in Computer Engineering from Delhi University, India. Her research interests include Analysis of Big Social Data and Recommender Systems, E-Learning and E-Governance, Machine Learning and Document Classification, Digital Database Watermarking and Design and Exploration of Multiprocessor Architectures.

Akhil Sharma is a Computer Science Graduate from Netaji Subhas Institute of Technology, Delhi. He is presently working as a Software Development Engineer at Amazon and has previously worked at D.E. Shaw. He is passionate about programming, algorithms and technology.

Jasleen Kaur has done B.Tech from Department of Electronics and Communication Engineering, Netaji Subhas Institute of Technology, Delhi. Ideas that bridge the gap between humans and technology interest her. Passionate to code, she is presently working at Amazon as a Software Development Engineer.

I. INTRODUCTION

As social beings, we often seek opinions from each other on topics on which we have only partial knowledge and value the inputs received from experienced peers. The easy accessibility of the Internet opens up the option of availing the opinion of innumerable others who, though geographically apart,

share mutual interests. Today, we can visualize a scenario where a person starts browsing reviews written by others in order to shape his/her decision, whether the object of interest is a product or an investment option or choice of a movie or voting for a candidate in an election. Indeed, there is a huge explosion of 'sentiments' available from social media such as Twitter and Facebook, blog sites, message boards, review websites, and user forums. The unstructured opinionated data available on these websites is a rich source of information that can be tapped to mine real time feedback on any issue.

According to (Liu 2010), an opinion is a quintuple (o_j, f_{jk}, so_{ijkl}, h_i, t_l) where o_j is the target object, f_{jk} is a feature k of the object j, so_{ijkl} is the sentiment value of the opinion of the opinion holder h_i on feature f_{jk} of object o_j at time t_l. The sentiment factor so_{ijkl} may be positive, negative, neutral or it may express a more granular assessment. In terms of Liu's definition, Sentiment Analysis (SA) is the task of determining correctly the sentiment so_{ijkl} at time l as expressed by the author i on a feature k of the entity j which is discussed in any opinionated document d.

Marketing managers, PR firms, campaign managers, politicians, equity investors and online shoppers are direct beneficiaries of Sentiment Analysis (SA) technology and highlight its growing prominence. Significant research work has been carried out in the domain of SA since the past decade. However, the problem of tapping local contextual information in SA has largely gone unnoticed. The main focus of this paper is on developing a scheme for SA that takes into account the textual context presented by words within the same sentence as well as beyond the boundaries of a sentence and allows their subjectivity scores to exercise influence on the sentiment expressed by any given word. We capture the synergy between the words comprising a document and investigate its impact on the overall sentiment conveyed by its contents.

II. RELATED WORK

Some of the earliest works in the area of SA can be traced back to the efforts of (Turney 2002) in discovering sentiments in product reviews. Around the same time, (Pang et al., 2002) focused on extracting sentiments from movie reviews. In the beginning, SA was mostly limited to document analysis to determine their sentiment polarity to be either positive or negative. Currently, SA can be carried out at the document or review level (Grabner et al., 2012);

(Kasper and Vela 2011), (Pang et al., 2002), (Turney 2002); at the sentence level (Celikyilmaz et al., 2010), (Esmin et al., 2012), (Fiadhi et al., 2012), (Shimada et al., 2011), (Wu and Ren 2011) and at the feature level (Singh et al., 2013). Broadly speaking, opinion mining methods may be grouped as (a) lexicon based methods wherein SA is carried out by developing a lexicon or using an existing lexicon that separates the words into positive, negative and neutral (Esuli and Sebastiani 2005); (Esuli and Sebastiani 2007); (Hatzivassiloglou and McKeown 1997) (b) machine learning based methods wherein computers perform SA by training a machine learning based text classifier like Naïve Bayesian, SVM, k-NN, etc. with a suitable feature selection scheme and then using it to classify unknown data (Gamon 2004), and (c) a combination of the above two approaches, as in (Weichselbraun et al., 2010). In this paper, we adopt the lexicon based approach to perform SA at the document level.

Context Based SA Techniques are those that take into account the contextual and semantic knowledge encapsulated within documents. Some approaches where contextual information is utilized to improve the performance of SA systems have been reported. (Nasukawa and Yi 2003) attribute sentiment to specific subjects in the text. They determine the document sentiment by identifying subjective expressions, determining their polarity and the relationship of these expressions to their subject. They label "verbs" as the sentiment transmitters that transfer sentiment from one argument to the other. For example: in the sentence "this book is good", the verb "is" transmits the sentiment extracted from "good" to the subject "book" and associates them. Terms having part-of-speech tags different from "verb" are dealt with in an easier way - they directly transfer their sentiment to the related argument.

Many of the Context Based SA systems focus on a single application domain and are based on the assumption that words always have the same connotation in that domain. For example, the word "*unpredictable"* may invariably imply a negative sentiment when used in the context of a car's engine but is always positive when used in the context of a thriller movie's plot. However, this may not be true always. Consider a camera's description that says "*long* battery life" and "*long* time to focus". One can easily see that within the same domain, the word "long" expresses positive sentiment in the first case and negative sentiment in the other. Thus, it is reasonable to conclude that domain considerations alone cannot suffice for performing context based SA.

(Polanyi and Zaenen 2006) argue that a simple summation of word polarities may not be the ideal way of calculating the sentiment of a document. They identify contextual interactions in a document and categorize the concepts responsible for context switches into Sentence based contextual valence shifters and Discourse based contextual valence shifters. Furthermore, they discuss the case when a document can describe more than one entity/topic/fact and propose that in such cases, the calculation of point of view must be done separately with respect to each entity and must take into account higher order factors such as genre that influences the document structure.

(Weichselbraun et al., 2010) propose a context dependent approach to SA in large textual databases that addresses the problem of polarity shifts of terms depending on their usage context. Most sentiment detection approaches have no way to deal with language ambiguity and ignore it altogether, although the same word may have more than one different meaning and different polarities that change according to the context in which a statement is made. To address this shortcoming, they use Naïve Bayes Method to identify ambiguous terms in text. They develop a contextualized lexicon that stores the polarity of these ambiguous terms, together with a set of co-occurring context terms. They further show that considering the usage context of ambiguous terms improves the accuracy of high-throughput sentiment detection methods.

In (Context Assisted Sentiment Analysis n.d.), the authors suggest that context may affect sentiment at two levels; at the domain level of the content and at the sentence-structure level. Accordingly, they first improve the performance of sentiment-lexicon based methods through domain ontology support. Secondly, they adopt a machine learning approach to extract frequently appearing substrings or expressions and learn the sentiments of these expressions. Finally, they combine these approaches using strategies for voting and chaining.

Taking cue from the above works, we explore how context may affect sentiment across larger units of discourse like sentences, paragraphs and full texts or narratives rather than simply the words and phrases contained within a sentence. The proposed context-based approach builds up on the lines suggested by these researchers and factors in the possibility that indeed, words may have different meanings in different contexts. We also delve deeper to understand how the effective sentiment of words may change due to the presence of other

polar words in their vicinity. The proposed algorithm presents a mathematical model to represent this effect and evaluate the enhanced sentiment.

III. A SCHEME FOR CONTEXT BASED SENTIMENT ANALYSIS

The block diagram shown in Fig. 1 describes the general flow of our proposed scheme for context driven SA. We now describe each of its steps.

A. Data Collection

For conducting SA with an acceptable degree of accuracy, we need a corpus of sentiment carrying textual data. As a case study for our experiments, we compiled a dataset of 500 hotel reviews extracted from the hotel review website 'Tripadvisor.com'. Out of these, 250 reviews were labelled as positive and remaining 250 were labelled as negative. This forms our labelled corpus in Fig. 1. This case study serves to illustrate the usefulness of our approach. It can be easily extended to other domains such as movie review websites, blogs, newspaper articles etc.

B. Pre-Processing

The pre-processing steps are illustrated in Fig. 2. This module begins by converting the entire review text to lower case to obtain a uniform casing and tokenizing the resultant text so that each review document is split into distinct words or tokens. Subsequently, punctuations such as full-stops and commas are removed. Next, stop-words like 'he', 'she', 'the', etc. are removed because such words do not contribute to sentiments. Finally, hyphens are removed from words that are hyphenated. For example: 'well-appointed' is changed to well appointed. This is necessary because Word Sense Disambiguation (WSD) considers hyphenated words as undefined and does not assign any sense to them.

C. Word-Sense Disambiguation

Many English words have more than one senses or meanings depending on the context in which they are used. Word Sense Disambiguation is an important pre-requisite to resolve the true implication of a word in its presented context. For example, a simple word like "fly" can mean either "an insect" or "lifted and moving in air". Thus, "fly" has at least two senses and it is of utmost

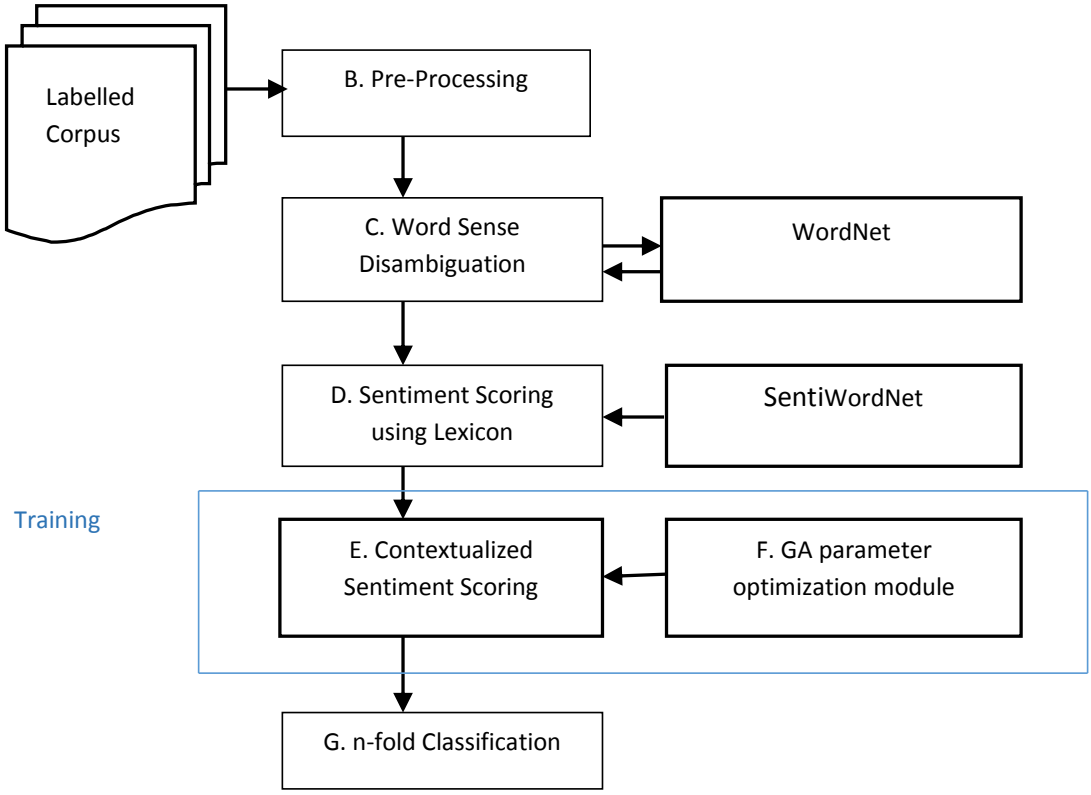


Fig. 1: Block Diagram of the Proposed Context Based SA Approach

importance to correctly identify which sense of a word has been used in a given context.

WSD checks for the meaning of each word in a document, given a window of neighboring words and using this, assigns a sense number to every word as listed in WordNet (Princeton University 2010). This has also been shown in Fig.1. Given a sentence, "Birds fly in the sky", the output after WSD for "fly" is *verb* and the score corresponding to this sense is used. For a sentence, "She was bit by a fly", the output after WSD for "fly" is *noun*. Nouns generally don't carry sentiments. The significance of WSD is that the sentiment score assigned for "fly" is different for these two senses. And since our approach works on sentiment scores of words, this step is critical to the working of the proposed algorithm. Once a word is assigned a sense number, other senses of that word are not considered when disambiguating subsequent words.

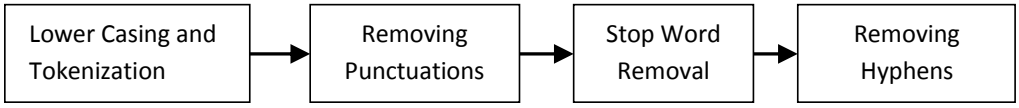


Fig. 2: The Pre-processing steps

D. Sentiment Scoring using Lexicon

After WSD, the next step is to assign raw scores to all the words in a review document. These are the native sentiment scores of all words independent of all other words. As represented by Fig.1, this step is carried out making use of the lexicon SentiWordNet (Esuli and Sebastiani 2006). SentiWordNet is a lexical resource for SA that assigns three scores to each set of synonyms (synset) of WordNet: positivity score, negativity score and objectivity score. All the words in this set of synonyms, called synset or synonym ring are considered to semantically equivalent and interchangeable in some context by WordNet. The sentiment scores assigned by SentiWordNet to any word are in the range *0* to *1* where the sum of all three scores for every word is unity.

The process *GetSentiWordNetScore*(.), described in Fig. 3, takes as input a word *w* and assigns to it a raw score, *rawsc*(*w*) and a label, denoted by *label*(*w*) which is a subjectivity indicator. Word labels are either objective i.e. factual or bearing no sentiment indicated by *obj*, subjective-positive indicated by *1* or subjective-negative indicated by *0*. In Line 1, the process extracts the positive and negative scores of word *w*, represented as *pos_score(w)* and *neg_score*(*w*) respectively, from SentiWordNet. If both the positive and negative scores of a word are zero, the word is labeled as objective *obj* and its raw score is set to zero (lines 2-5). If the positive score is greater than negative score for the word, the word is positive and its label is set to *1* and its positive score is taken as its raw score (lines 6-9). Otherwise, the word is negative, labelled *0* and its negative score is taken as the raw score (lines 10-13).

GetSentiWordNetScore(.) is run taking all the words in a document turn by turn for assignment of raw scores and subjectivity indicators to all words comprising the document. While further processing, none of the words having label=’obj’ are taken into consideration for SA.

E. Contextualized Sentiment Scoring

The proposed approach for context based SA makes use of the interrelationships between words in a review document in order to evaluate their semantic relevance to a particular sentiment. The intuition is that the relevance of a word is calculated by taking into account its intrinsic relevance as well as the relevance of terms that appear within the same context. That is, along with

the raw scores of all the words, the proposed algorithm also incorporates context scores of all the words, which are computed based on the neighboring words and their influences on one another.

```
GetSentiWordNetScore(.)

Input: Word w
Output: Native score of w: rawsc(w), Label of w: label(w)
Begin
    Extract pos_score(w), neg_score(w) from SentiWordNet
    If ((pos_score(w)==0) && (neg_score(w)==0)){
       rawsc(w)=0;
      label(w)=obj;
    }
    Else if (pos_score(w) >neg_score(w)) {
        rawsc(w)=pos_score(w);
        label(w)=1;
    }
    Else {
        rawsc(w)=neg_score(w);
        label(w)=0;
    }
End
```

Fig 3: Pseudocode for obtaining native scores and labels of words

The influence exercised by a word on another word is computed using a parametric function called the Influence Function, *IF*. This function computations make use of context. "Context" refers to the text surrounding a given word. The idea is that, in order to measure the relevance of a particular word to a given sentiment, the features of the context in which the word appears, are as important as the word itself. Thus, for any positive word, say, if the context words are also positive, the score of the positive word increases, representing a stronger probability towards positivity and if the context words are negative, the score of the positive word decreases. Similarly, for a negative word, the reverse is true. The context *ctxt*(*w*) of a word *w* is the set of words that surround *w* within the dimensions z_L(left margin) and z_R (right margin). Thus, if

" . . . $u-zl-1$,$u-zl$, . . . $u-1$,w, $u+1$. . . $u+zr$, $u+zr+1$ . . ." is a text passage of any document, then $ctxt(w) = \{u-zl\ ,\ ...,\ u-1,\ u+1,\ ...,\ u+zr\}$. In our experiments, we set z_L as the start and z_R as the end of the document.

Two factors used to calculate the change in the native scores of individual words are:

1) Distance of the neighboring word: Of course, the farther a neighboring word is, the lesser is its impact on the present word. We reasonably assume that the distance of neighboring word affects the present word in an inverse manner.

2) Score of the neighboring words: If the neighboring word is highly subjective, it affects the present word more strongly. That is to say, the absolute value of the sentiment score of neighboring words affects the present word in a direct relationship.

The effects of the absolute score of a neighboring word *u* and its distance from the target word *w* are computed by using two separate influence functions- IF_{score} and IF_{dist} respectively, as given below:

$$IF_{score} = \frac{1}{\left(1+ e^{-(\alpha_1 * rawsc(u)+\beta_1)}\right)} \quad (1)$$

$$IF_{dist} = [1 - \frac{1}{\left(1+e^{-(\alpha_2 * reldist(w,u)+\beta_2)}\right)}] \quad (2)$$

The factor *reldist*(*w*,*u*) denotes the distance between the present word *w* and the neighboring word *u* in terms of number of intervening words, divided by the total number of words in the document. The factor *rawsc*(*u*) is the native score of the neighboring word *u*. The factors *rawsc*(*u*) and *reldist*(*w*,*u*) have either zero or positive values.

Note that the above equations have a sigmoidal form $1/(1+ e^{-x})$ in the R.H.S. This logistic sigmoid is a monotone increasing function which approaches *0* when $x \to -\infty$ and approaches *1* when $x \to \infty$. These functions thus act as two soft switches to enable or disable the effects of the distance and score respectively controlled by the model's parameters α_1, α_2 and β_1, β_2. Parameters α_1 and α_2 control the slope of the distance and score curves respectively. Parameters β_1 and β_2 control the point where the switch assumes a median value ½. As shown in Fig.1, these parameters will be optimized by Genetic Algorithms (GA), which is described later.

The context score of a word, given by *surr*(w,*u*) is the resultant effect after multiplication of the two influence functions. This is represented

mathematically in equation 3. We have multiplied the effects of score and distance to obtain the final context score because we want that the context score of a word is high only when both the distance and score effect values are high i.e. when there are other words in the close proximity of the present word having high subjective scores. If the ex-or of labels of *u* and *w* is zero, that is, both the words have the same subjectivities, then context score will be added else subtracted from the native score of *w*.

$$surr(w,u) = \begin{matrix} IF_{score} * IF_{dist} \ ; if\ label(w) xor\ label(u) = 0 \\ -IF_{score} *\ IF_{dist} \ ; \qquad otherwise \end{matrix} \qquad (3)$$

Thus, the final revised score of a word will result from the combination of its native score and its context score $\sum_{u\in\ Q(w)} surr(w,u)$ and is given in equation 4. Here *rawsc*(*w*)is the native score of word *w*, *Q*(*w*) is the set of words that belong to the context of *w*, *surr*(*w*,*u*) is the context effect of the surrounding word *u* in the vicinity of *w* and $\lambda \in [0, 1]$ is the damping factor.

$$s(w) = [(1-\lambda) * rawsc(w)] + \lambda \sum_{u\epsilon Q(w)} surr(w,u) \qquad (4)$$

F. Optimization using Genetic Algorithms

Genetic algorithms (GA) generate solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover. As shown by Fig. 1, the optimum values of α_1, α_2 and β_1, β_2 used in Equations 1 and 2 for contextualized sentiment scoring are obtained by fine-tuning using Genetic Algorithms. The objective function of the GA is set to maximize the accuracy of the classification task, which is described in Equation 5.

G. Classification

As represented in Fig.1, the last step of the proposed context based SA approach is classifying a review file into positive or negative. To calculate the subjectivity for each file, process *GetFileSubjectivity*(.), explained in Fig. 4 is used. The positive score count and negative score count of the file denoted by *sc_pos*, *sc_neg* are initialized to zero (line 1). Then for every word encountered, its label is checked to determine whether the word is positive or negative. For a positive word, the positive score count *sc_pos* is incremented by the score of the word. For a negative word, the negative score count *sc_neg* is incremented by

the score of the word (lines 2-6). At the end, the normalized positive score *normsc_p* is calculated which represents the ratio of positive to total scores and similarly the *normsc_n* represents the ratio of negative to total scores (lines 7-8).

If normalized positive score is greater than the normalized negative score, the file is labelled positive, else it is labelled negative (lines 9-11).

IV. EXPERIMENTS AND RESULTS

All review files are pre-processed as shown in Fig. 2 using the Python Natural Language Toolkit. WSD is performed using the module WordNet::SenseRelate::AllWords in Perl. Next, a SentiWordNet Baseline is established which provides a reference point for the purpose of comparing the performance of the proposed context oriented classification algorithm. The baseline and the context based approach are as described:

A. Baseline approach

This approach makes use of the raw scores of the words in a document as obtained from SentiWordNet using process *GetSentiWordNetScore*(.) explained in Fig. 3. The baseline established ignores contextual considerations completely. The subjectivities of files for this approach, are computed replacing each occurrence of the revised score *s(w)* by raw score *rawsc(w)* in process *GetFileSubjectivity*(.). Then, the factors *normsc_p* and *normsc_n* indicate the baseline positive and negative subjectivities respectively.

B. Proposed Context Based approach

The Python library Pyevolve is used for fixing the parameter values using GA. The model's parameter values α_1, α_2 and β_1, β_2 in Equations 1 and 2 are fixed keeping a population size=10 and training for 300 generations. Then, context scores of all words are computed using Equations 1, 2 and 3 and then the final revised scores of words, taking into account the raw scores as well as context scores, are computed as described in Equation 4. The subjectivities of all review files are computed using process *GetFileSubjectivity*(.). The factors *normsc_p* and *normsc_n* indicate the positive and negative subjectivities of the review files.

```
GetFileSubjectivity(.)
Input: File F
Output: Subjectivity of file F: Subj(F)
Begin
Set sc_neg=0; sc_pos=0;
    For each word w in F {
    If (label(w)==1)
        sc_pos+= s(w);
    Else sc_neg +=s(w);
    }
    normsc_p=(sc_pos/sc_pos+sc_neg)*100;
    normsc_n=(sc_neg/sc_pos+sc_neg)*100;
    If (normsc_p>normsc_n)
        Subj(F)=pos;
    Else Subj(F)=neg;
End
```

Fig. 4: Pseudocode for Computing the Subjectivity of an Input File

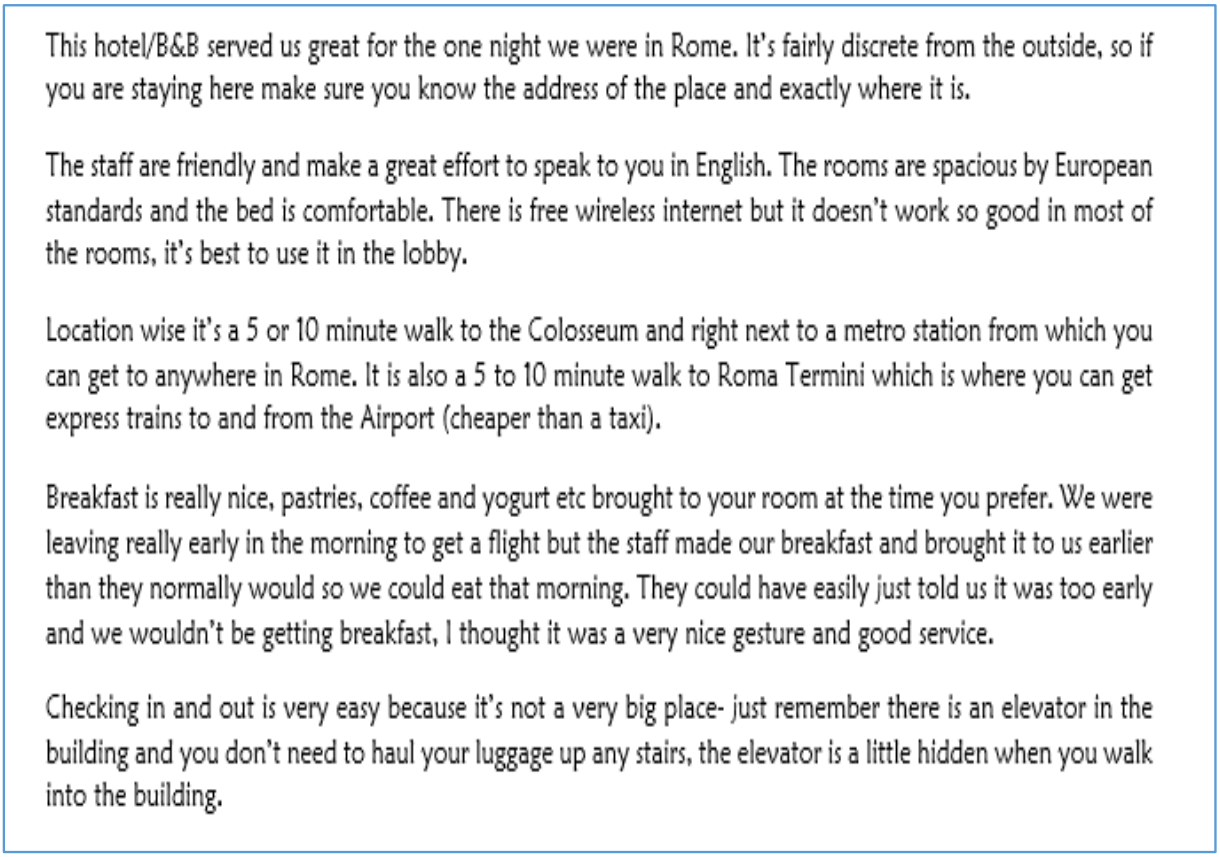
The performance of the proposed algorithm is tested on a total of 500 hotel reviews of varying lengths, extracted from the hotel review website 'Tripadvisor.com'. A sample review is shown in Fig 5. In order to determine the accuracy of our approach and facilitate comparison with Baseline, we apply 10-fold cross validation, without randomizing the training and test sets. Thus, in each fold the training and test sets are completely disjoint. For all folds we keep a total of 450 review files containing 225 positive and 225 negative files for tuning the parameters α_1, α_2 and β_1, β_2 using GA. The remaining 50 review files are used for testing the performance of the proposed context-based algorithm.

C. Evaluation Metrics

The following well-known metrics were used for evaluation of SA:
Accuracy: The accuracy of classification is as represented in Equation 5.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (5)$$

TP: True Positive, TN: True Negative, FP: False Positive and FN: False Negative. True positives indicate the number of positive files which are rightly



This hotel/B&B served us great for the one night we were in Rome. It's fairly discrete from the outside, so if you are staying here make sure you know the address of the place and exactly where it is.

The staff are friendly and make a great effort to speak to you in English. The rooms are spacious by European standards and the bed is comfortable. There is free wireless internet but it doesn't work so good in most of the rooms, it's best to use it in the lobby.

Location wise it's a 5 or 10 minute walk to the Colosseum and right next to a metro station from which you can get to anywhere in Rome. It is also a 5 to 10 minute walk to Roma Termini which is where you can get express trains to and from the Airport (cheaper than a taxi).

Breakfast is really nice, pastries, coffee and yogurt etc brought to your room at the time you prefer. We were leaving really early in the morning to get a flight but the staff made our breakfast and brought it to us earlier than they normally would so we could eat that morning. They could have easily just told us it was too early and we wouldn't be getting breakfast, I thought it was a very nice gesture and good service.

Checking in and out is very easy because it's not a very big place- just remember there is an elevator in the building and you don't need to haul your luggage up any stairs, the elevator is a little hidden when you walk into the building.

Fig. 5: A Sample hotel review extracted from Tripadvisor.com

classified as positive. Similarly True negatives indicate the number of negative files which are correctly classified as negative. False positives denote the number of negative files which are misclassified as positive. False negatives indicate the number of positive files that are misclassified as negative.

Precision: Precision, *Pr*, for a positive class is the number of items correctly labeled as belonging to the positive class, (True Positives) divided by the total number of elements labeled as belonging to the positive class (the sum of True Positives and False Positives). This is represented in Equation 6.

$$Pos\ Pr = \frac{TP}{TP+FP} \tag{6}$$

For a negative class, we consider True Negatives and False Negatives instead, as represented in Equation 7.

$$Neg\ Pr = \frac{TN}{TN+FN} \tag{7}$$

Recall: Recall, *Re,* is defined as the number of True Positives divided by the total number of elements that actually belong to the positive class (the sum of True Positives and False Negatives). This is represented in Equation 8.

$$Pos\ Re = \frac{TP}{TP+FN} \quad (8)$$

For a negative class, we consider True Negatives and False Positives instead, as represented in Equation 9.

$$Neg\ Re = \frac{TN}{TN+FP} \quad (9)$$

F1- Score: This is the harmonic mean of Precision and Recall. Positive and Negative F1 are calculated as shown in Equations 10 and 11 respectively.

$$Pos\ F1 = 2 * \frac{Pos\ Pr * Pos\ Re}{Pos\ Pr + Pos\ Re} \quad (10)$$

$$Neg\ F1 = 2 * \frac{Neg\ Pr * Neg\ Re}{Neg\ Pr + Neg\ Re} \quad (11)$$

We also report the macro-level F1 measure, which is the average of the Positive and Negative F1 scores.

The accuracies obtained by the Baseline approach and by the proposed Context based approach, are shown in Table I. The difference between the two is highest in Fold 3 where it is 8% and least in Folds 7 and 8 where the Context and Baseline accuracies are the same. In other folds, the difference between Context and Baseline accuracies varies from 2 to 6%. The Context accuracy is better or equal to the Baseline accuracy for each of the folds. The average accuracies over all the folds for Baseline and Context approaches are 69.6% and 73.2% respectively, thus giving an improvement of 3.6%. Thus, the proposed Context Based SA method shows a significant improvement as compared with the Baseline approach.

Table II compares the performance of the Baseline and the Context Based SA methods on both positive as well as negative reviews. It enlists the average values of Precision, Recall and F1-measure on all evaluation runs for both the Baseline and the Context Based approach. When we compare the Context Based approach with Baseline, we register an improvement of 3.46% in the Precision positive reviews, but at the cost of Recall, which decreases by 2%. Overall, we still record a gain of 1.81% in positive F1-score by using the Context Based approach. For negative reviews, we register an improvement of 9.20% in Recall

by the Context Based approach, with a slight loss of 1.45% in Precision. On the whole though there is a significant gain of 7.96% from Baseline F1-score to Context F1-score for negative reviews.

Table I: Classification accuracies of proposed Context Based SA scheme with cross-validation over different folds

	Accuracy in Percent(Baseline)	**Accuracy in Percent (Context)**
Fold 1	60	64
Fold 2	68	72
Fold 3	60	68
Fold 4	72	76
Fold 5	68	72
Fold 6	66	72
Fold 7	74	74
Fold 8	74	74
Fold 9	78	80
Fold 10	76	80
Average	**69.6**	**73.2**

Thus, for positive as well as negative reviews, the proposed Context Based SA approach works better than the Baseline approach. A closer examination revealed that several words with borderline sentimental scores gained strength when augmented with high contextual scores. Likewise, some words with high sentimental scores diminished in strength when surrounded by more objective neighboring words.

Table II: The average Precision, Recall and F1 values after 10-fold Cross-validation for Baseline and Context Based SA approaches

	Baseline Approach			**Proposed Context Based SA Approach**		
	Re %	**Pr %**	**F1 %**	**Re %**	**Pr %**	**F1 %**
Pos	96.00	63.16	76.08	94.00	66.62	**77.89**
Neg	43.20	91.32	57.81	52.40	89.87	**65.77**

V. COMPARISON WITH RELATED WORK

We compare the performance of our approach with the performance of two recent papers that propose alternative SA approaches. The first paper by (Hogenboom et al., 2011) does not make use of any contextual information from the corpus. The authors investigate a specific aspect of SA- negation and its impact on SA. They proposed and compared several approaches to account for negation in SA, differing in their methods of determining the influence of negation keywords. The authors also experimented with different values of sentiment inversion factor that controls to what extent the scores of negated words are inversed. They evaluated the performance of their approaches on a collection of 1,000 positive and 1,000 negative English movie reviews, which have been extracted from movie review web sites (Pang and Lee 2004).

Table III shows the best results reported in (Hogenboom et al., 2011) and those obtained by using our approach on the same corpus. On comparing the F1 scores of both approaches, we find that for positive reviews, the F1 score obtained by our Context Based Approach is 15.29% more than by that obtained by them. For negative reviews, again our F1 score is better by 6.39%. Overall, our context driven SA approach outperforms their negation-feature based SA approach by a significant margin of 10.84%. This clearly establishes our contention that including contextual information in SA improves its performance significantly.

Next, we compare the performance of our approach with that of another context-based SA approach proposed in (Weichselbraun et al., 2010). The authors tested their approach on hotel reviews extracted from Tripadvisor with 10-fold cross validation. The results reported in (Weichselbraun et al., 2010) and those obtained by running our context based approach, are reproduced in Table IV. On comparing the F1 scores, we can see that on an average, for positive reviews, their F1 score is higher by 1% than the F1 score obtained by our approach. On the other hand, for negative reviews, our approach achieves a 3% higher F1 score. On the whole, our approach registers a 1% increase in macro-F1 scores. The improvement is small, but on large datasets, even this value cannot be considered insignificant. The results re-affirm the superiority of our context based SA approach.

The principal advantage of our approach is the flexibility it offers in terms of two parametric functions that show the effects of score and distance. The

tunable parameters α and β are optimized for achieving the best possible results for different corpora. On the other hand, in (Weichselbraun et al., 2010) the authors have applied WSD but they have completely ignored the issue of corpus-specific fine tuning of the performance of SA.

Table III: Comparison of F1 scores of proposed context based and alternative SA approach for positive and negative review files

	Positive F1 %	**Negative F1 %**	**Macro F1 %**
SA approach in (Hogenboom et al., 2011)	54.70	51.90	53.30
Proposed Context Based SA	**69.99**	**58.29**	**64.14**

Table IV: Comparison of F1 scores over 10-folds between proposed and alternative Context Based SA approach for positive and negative review files

	Positive F1 %	**Negative F1 %**	**Macro F1 %**
Proposed Context Based SA	78	**64**	**71**
Context Based SA in (Weichselbraun et al., 2010)	**79**	61	70

VI. CONCLUSIONS AND FUTURE WORK

We developed a novel Sentiment Analysis algorithm utilizing the contextual information of the words comprising a document. Using this as a foundation we were able to measure the degree to which the words in a document influence each other and the impact of the dynamics of this relationship on the overall document sentiment. On the basis of the results obtained, we conclude that context plays a very significant role in determining a document's sentiment and should not be ignored. To extend the work, we plan to compute the contextual influence amongst words not only on the basis of their physical distances but also other features such as WordNet distances.

REFERENCES

Celikyilmaz, A, Hakkani-Tur, D and Feng, J (2010), 'Probabilistic Model-Based Sentiment Analysis of Twitter Messages', *in the 2010 IEEE International Conference on Spoken Language Technology, Workshop*, pp 79-84

Context Assisted Sentiment Analysis, Available from:<http://www.ceti.cse.ohio-state.edu/architectures-services/technical-reports/OSU-CISRC-6-09-TR28%20-%20Context-Assisted%20Sentiment%20Analysis.pdf>, [17 March 2014]

Esmin, Ahmed A.A. , Jr., Roberto L. de Oliveira ; Matwin, S (2012), 'Hierarchical Classification Approach to Emotion Recognition in Twitter', In *Proceedings of the 11th International Conference on Machine Learning and Applications, ICMLA*, Boca Raton, FL, USA, Volume 2, pp 381–385,IEEE.

Esuli, A and Sebastiani, F (2005), 'Determining the semantic orientation of terms through gloss classification', *In Proceedings of CIKM-05*, pp 617–624

Esuli, A and Sebastiani, F (2006), 'Sentiwordnet: A publicly available lexical resource for opinion mining*' In Proceedings of the 5th Conference on Language Resources and Evaluation (LREC '06)*, pages 417–422.

Esuli, A and Sebastiani, F (2007), 'Pageranking Wordnet Synsets: An application to opinion mining', *In Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics (ACL).*

Fiaidhi, J, Mohammed, O, Mohammed, S, Fong, S and Tai hoon Kim (2012) 'Opinion Mining over Twitterspace: Classifying Tweets Programmatically using the R Approach', in *Seventh International Conference on Digital Information Management*, pp 313-319, IEEE

Gamon, M (2004),'Sentiment classification on customer feedback data: Noisy data, large feature vectors, and role of linguistic analysis', in *Proceedings of the 20th International Conference on Computational Linguistics (COLING),* Geneva, Switzerland, Article no. 841

Gräbner, D, Zanker, M, Fliedl, G and Fuchs, M (2012),'Classification of Customer Reviews based on Sentiment Analysis', *19th Conference on Information and Communication Technologies in Tourism (ENTER), Springer,* Helsingborg, Sweden.

Hatzivassiloglou, V and McKeown, KR (1997), 'Predicting the semantic orientation of adjectives', in *Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics*, pp 174–181.

Hogenboom, A.; van Iterson, P.; Heerschop, B.; Frasincar, F.; Kaymak, U (2011), 'Determining negation scope and strength in sentiment analysis', in *Proceedings of the 2011 IEEE International Conference onSystems, Man, and Cybernetics (SMC)*, pp 2589-2594.

Kasper, W and Vela, M (2011), 'Sentiment Analysis for Hotel Reviews', in *Proceedings of the Computational Linguistics-Applications Conference,* Jachranka, *pp. 45–52*

Liu, B (2010), 'Sentiment Analysis and Subjectivity', A Chapter in *Handbook of Natural Language Processing*, 2nd Edition, 2010.

Nasukawa, T and Yi, J (2003), 'Sentiment Analysis: Capturing Favorability Using Natural Language Processing', in *Proceedings of the International Conference on Knowledge Capture,* New York, USA, pp 70-77

Pang, B, Lee, L and Vaithyanathan, S (2002), 'Thumbs up? Sentiment Classification using Machine Learning Techniques', *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*. pp. 79–86.

Pang, B. and Lee L (2004), "A Sentimental Education: Sentiment Analysis using Subjectivity Summarization based on Minimum Cuts," in 42nd Annual Meeting of the Association for Computational Linguistics ACL, 2004, pp. 271–280.

Polanyi, L and Zaenen, A (2006),' Contextual Valence Shifters', In *Shanahan J.G., Yan. Qu, and Wiebe J (Eds.), Computing Attitude and Affect in Text: Theory and Applications*, Chap.1 Springer, Heidelberg, pp. 1–10.

Princeton University "*About Wordnet*". WordNet. Princeton University, 2010. Available from: <http://wordnet.princeton.edu/>, [17 March 2014]

Shimada, K, Inoue, S, Maeda, H and Endo T (2011), 'Analyzing Tourism Information on Twitter for a Local City', *First ACIS International Symposium on Software and Network Engineering, SSNE'11,* pp-61-66.

Singh, V.K, Piryani, R, Uddin, A and Waila, P (2013), 'Sentiment Analysis of Movie Reviews', *in Proceedings of International Multi Conference on Automation, Computing, Control, Communication, and Compressed Sensing*, Kerala, India, pp 712-717

Turney, P (2002), 'Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews', *Proceedings of the Association for Computational Linguistics*. pp. 417–424

Weichselbraun, A, Gindl, S and Scharl, A (2010), 'A Context-Dependent Supervised Learning Approach to Sentiment Detection in Large Textual Databases', in *Journal of Information and Data Management, Vol. 1, No. 3, pp 329-342.*

Wu, Y and Ren, F (2011), 'Learning Sentimental Influence in Twitter', *International Conference on Future Computer Sciences and Application*, pp. 119-122

View publication stats

International Journal of Research in Engineering and Innovation (IJREI)
journal home page: http://www.ijrei.com
ISSN (Online): 2456-6934

A thermodynamic analysis of ejector type vapour refrigeration system using eco-friendly refrigerants

R S Mishra, Ankit Dwivedi, Shadab Ahmad

Department. of Mechanical Engineering, Delhi Technological University, Delhi
E. Mail. rsmishra@ dtu.ac.in, ankitdwivedi2008@gmail.com

Abstract

A thermodynamic analysis on Ejector Refrigeration Cycle (ERC), have been performed using eco-friendly refrigerants (i.e.R-404A, R-410A, R-407C, R-423A, R-500, R-502 and R-507C) A numerical computation is carried out for finding COP, Second Law Efficiency (η_{II}) , Refrigeration Effect (RE) and Heat Input (Q_{in}) of the ERC along with the variation in boiler temperature (T_b) , condenser temperature (T_c) and evaporator temperature (T_e) for ecofriendly refrigerants for the ranges of the temperatures where the evaporator temperature is varying from 253K to 273K, with varying condenser temperature from 313K to 328K and with varying boiling temperature from 320 K to 345K. The maximum first Law Efficiency (η_I) is found in the range of 1.5-1.8 with fluid R-404A and condenser temperature 328K. At 253K evaporator temperature R-404A is working with maximum Second Law Efficiency (η_{II}) 37%. For the entire range of boiling temperature R-410A is having maximum Second Law Efficiency (η_{II}) except at 345K where R-404A is highest. The refrigeration effect is the maximum and is in the range of 200kW to 220kW for R-410A for all the ranges of temperatures. The highest COP is associated with R-404A with temperature variations.
© 2017 ijrei.com. All rights reserved

Keywords: Ejector VCR, First law performance, Ecofriendly Refrigerants

1. Introduction

Ejector refrigeration system (ERS) runs on low grade energy and has been studied since the mid-1950s. Compared to other refrigeration systems, ERS has a few special advantages i.e. the simple construction, its reliability and it is economic. However, the COP for the conventional ERS is significantly low. This restricts its wide applications. In recent times, researchers have tried to develop wider application for ejector refrigeration systems in refrigeration and air conditioning by direct utilization of low-grade energy, such as solar and waste heat. In ERS an ejector is used instead of an expansion valve and the ejector performs two functions namely expansion of the liquid refrigerant and pumping of the refrigerant. The ejector has advantages compared to the expansion valve such as use of energy, which is dissipated as swirl in the expansion valve, reduction work of the compressor, as the ejector also works as a pump. The ejector saves fuel or electricity for operation of the compressor. Among the advantages are flexibility, no moving parts so free of vibration and light weight

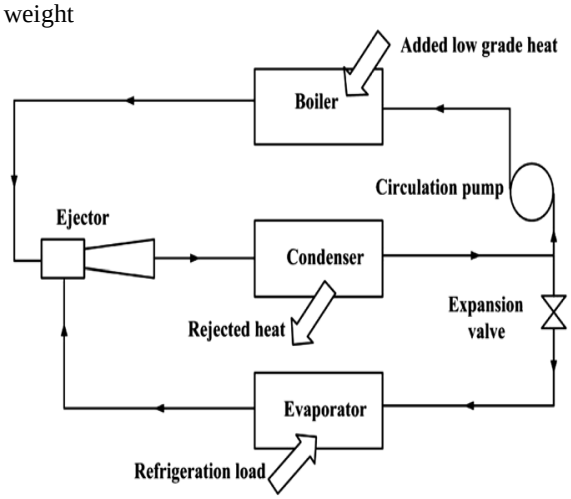


Figure 1: A schematic of the ERS.

Corresponding Author: R.S. Mishra
Email Id: rsmishra@dtu.ac.in

The inter-relation between the different parts can be expressed by the fig. 1. Boiler used low grade energy to produce vapours to be expanded inside the ejector. Without a compressor can this system perform saving high grade energy and cost.

1.1 Literature Review

Tischendorf et. al. [1] focused on the differences in energy dissipation in each component compared to the the whole ejector refrigeration cycle. With help of this analysis, improvement of energetic efficiency by using an ejector was set in relation to the potential improvement in efficiency of other components such as heat exchangers. The refrigerants R134a and R744 (CO2) were compared in regard to the entropy production of the heat pump system.
Kshirsagar et. al [2] proposed a combined vapour compression-ejector refrigeration system which used the waste heat of condenser of simple vapour compression system and this heat was utilized to drive the binary ejector refrigeration system. Cooling effect produced by this binary system considered as input to the cooling effect of basic vapour compression system. They also investigated the characteristics and the efficient design of the ejector to improve ejector refrigeration systems. The computational fluid dynamics (CFD) code, FLUENT, was employed to predict the flow phenomena and performance of CPM and CMA steam ejectors.
M. M. Rashidi et. al. [3] studied a combined power and refrigeration cycle that combined the Rankine cycle and the ejector refrigeration cycle and a solar energy heat source was used. This combined cycle produced both power output and refrigeration output simultaneously. The effects of the evaporator temperature, turbine inlet and outlet pressures on the thermal and the exergetic efficiencies were investigated. Simulation results showed that thermal and exergy efficiencies increase with increasing evaporator temperature. In addition, it was found that the increase in the turbine inlet and outlet pressure leads to the increase in the first law efficiency and reduction in the second law efficiency.
B. J. Huang et. al. [4] developed a high-performance solar ejector cooling system using R141b as the working fluid. Experimentally a COP of 0.5 for a single-stage ejector was obtained for a cooling system at a generating temperature of 90°C, condensing temperature of 28°C, and an evaporating temperature 8°C. For solar cooling application, an optimum overall COP obtained was around 0.22 at a generating temperature of 95°C, evaporating temperature of 8°C and solar radiation at 700 Wm .
R.Khajuria et. al. [5] investigated the performance analysis of ejector refrigeration system with R404A. The ejector refrigeration system used exhaust emission of automobile as thermal energy for providing heat to generator. The result showed that system using R404A as refrigerant can be used in the ejector refrigeration system for area ratio 7.84. Also Cooling capacity of the system increases with increase in evaporator temperature and generator temperature.
J.Yu et. al. [6] proposed a new ejector refrigeration system (NERS) with an additional liquid–vapor jet pump. The jet pump decreased the backpressure of the ejector, and the entrainment ratio and the coefficient of performance (COP) was increased for NERS. Two refrigerants namely R134a and R152a were considered. Calculations showed that COP of NERS could be improved more effectively and that happened at the cost of more pump work. The exergy of the new system was higher.
Kashyap et. al. [7] developed a dimensional mathematical model to analyse the performance ejector refrigeration cycle with working fluid R410a and compared with performance of R134a. The result showed that performance of R134a was better than R410a for area ratio 5.64 and 7.84.
Jain A. et. al [8] used solar cooling with the help of an ejector refrigeration system and did mainly two analysis in this system. By doing the analysis, the result were shown for the refrigerants under the steady state conditions. Z Huifan et. al. [9] designed an experimental solar ejector refrigeration setup using R-134a as working fluid. The average daily COP was found to be 0.18, and the average EER was up to 3.5.
There is a huge scope of studying different eco-friendly refrigerants that can be used as the main working fluid in the ERS. Different combinations of temperatures can be observed in the cycle and for all the refrigerants. Variations of the performance parameters against the temperature can be obtained by the software simulations. The second law efficiency is a measure of exergy destroyed during the cycle. Different working fluids can be compared on that basis too. An optimum combination of parameters should be there for the best performance. After such analysis one can think of modifying the system and enhance its capacity to utilize the exergy available more efficiently. Irreversibility can be reduced by using intra-cycle heat transfer equipment so that the heat lost will be reduced by this method.

1.2 Ejector Refrigeration System

The system can be inferred to fig.1 which has namely a condenser, boiler, evaporator, ejector, flash chamber (between evaporator and ejector), pumps and throttle valves.

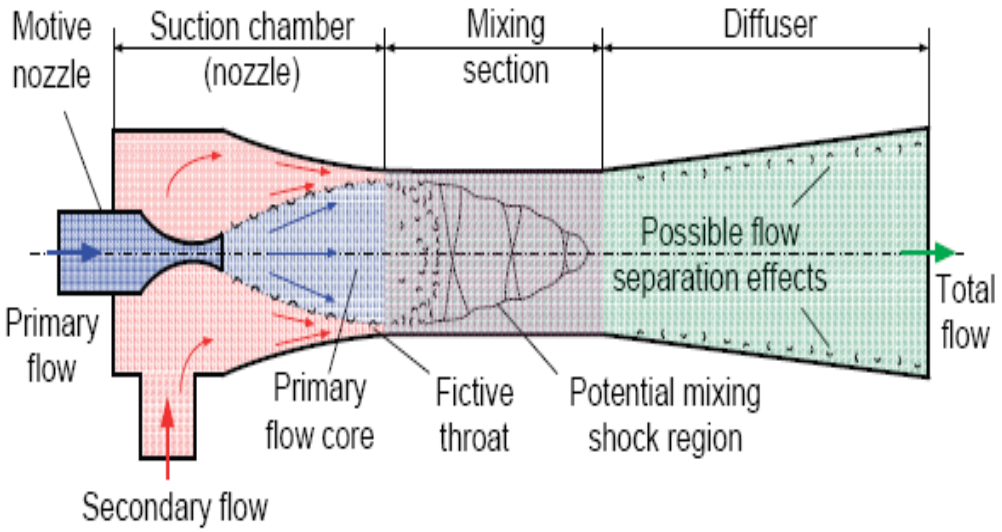


Figure 2: Schematic of a typical two-phase ejector design [10]

A typical ejector consists of a motive nozzle, a suction chamber, a mixing section, and a diffuser as shown in the fig 2. The ejector converts the internal energy and pressure of the motive fluid stream into kinetic energy. The motive nozzle is of a converging-diverging design allowing the jet exiting the nozzle to become supersonic. Depending on the state of the primary fluid, the flow at the exit of the motive nozzle might be two-phase. Flashing of the primary flow inside the nozzle might get delayed due to thermodynamic and hydrodynamic non-equilibrium effects.

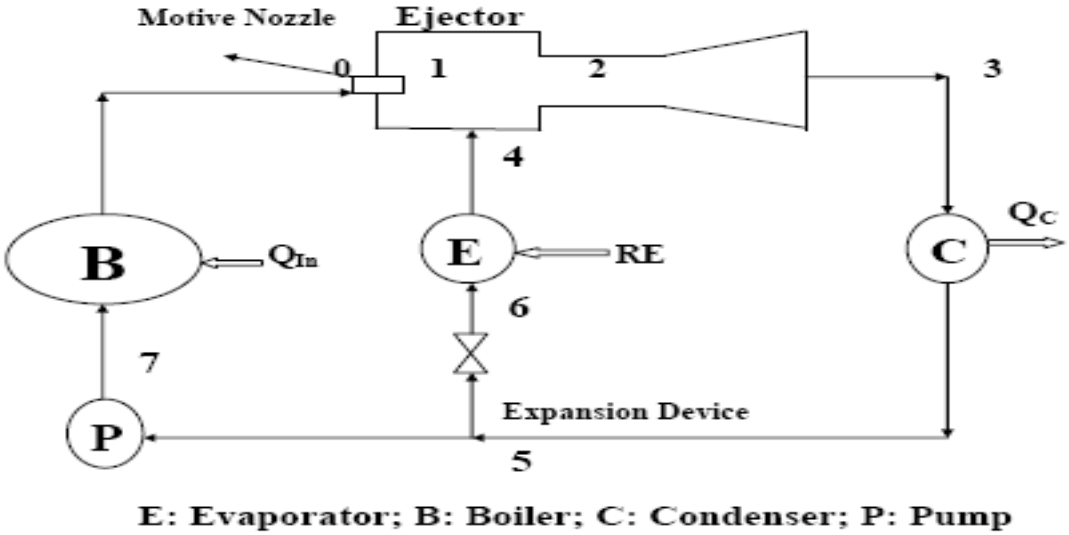


Figure 3: the schematic of the ERS

High pressure refrigerant vapour is supplied to the nozzle from the boiler and it is expanded in the ejector. Here, the vapour originated from the evaporator is entrained with the high velocity jet and it is further compressed in the thermo-compressor. Now the kinetic energy of the mixture is converted into static pressure and mass is discharged to the condenser. The condensate refrigerant is returned to the boiler to be re-circulated. Here RE is the refrigeration effect, Q_{In} is the heat input to the boiler and Q_C is the heat rejected in condenser.

2. Results & Discussions

As mentioned earlier the variation of performance parameters such as COP, Second Law Efficiency (η_{II}), Refrigeration Effect (RE) and Heat Input (Q_{in}) are studied with the operating temperatures in the boiler, the condenser and the evaporator through a computer simulation. The plots showing the results are as follows. Fig 4 shows that for the same range of boiling temperature R-423A and R-500 show a decline in the COP with the increase in T_b. The COP obtained for these refrigerants is also less than 1 for most of the range, whereas R-407C is showing a flat result for COP near 1. R502 and R507A show a more or less flat variation from 1.2 to 1.4. For the most of the temperature range R410A is giving highest COP from 1.4 to 1.8.

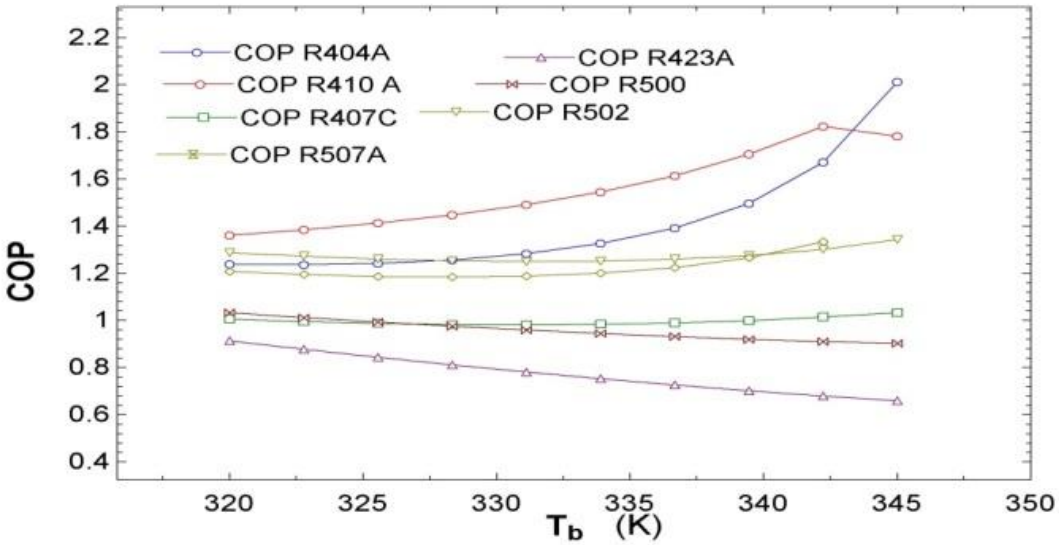


Figure 4: Variation of first law efficiency (COP) with boiler temperature (T_b)

The refrigeration effect in fig 5 is showing similar variations as the COP. except the RE for R410A is far more than the rest of the refrigerants in the analysis for the entire range.

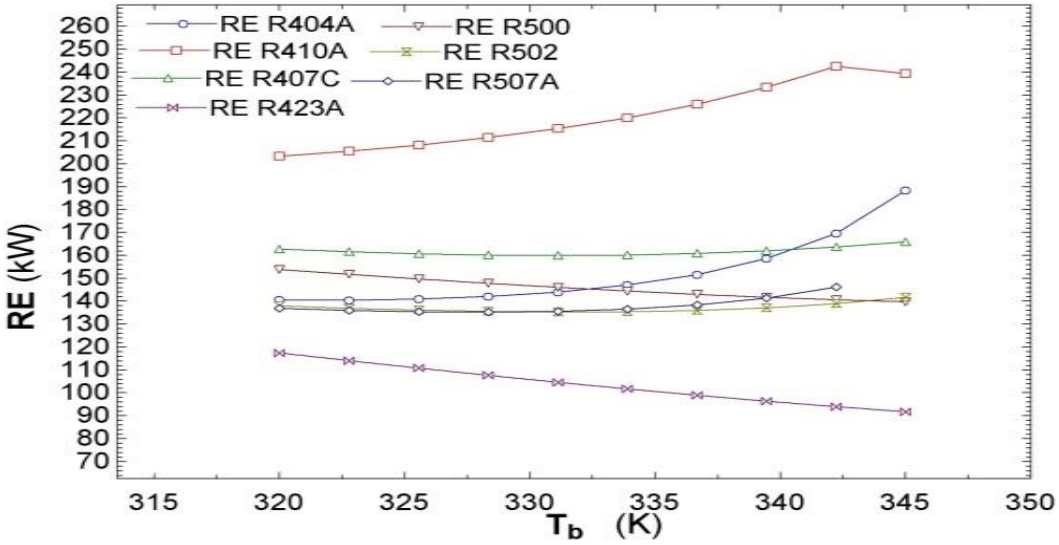


Figure 5: Variation of Refrigeration effect (RE) with boiling temperature (T_b)

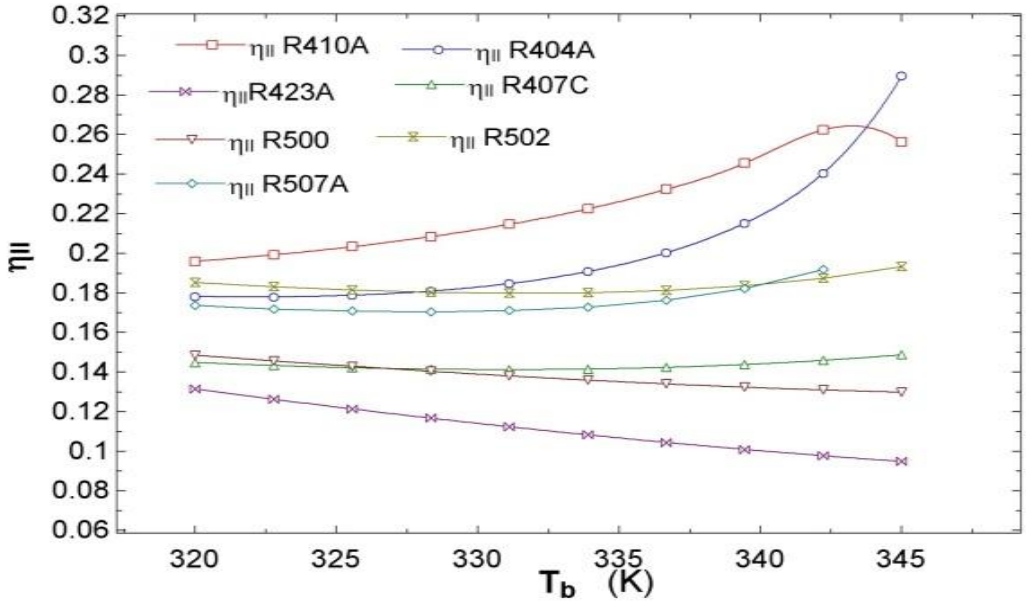


Figure 6: Variation of Second (Exergetic efficiency) with boiler temperature (T_b)

It can be interpreted from the fig 6 that based on the exergy loss R410A is performing the best where the second law efficiency varies from 20% to 26%. Rest of the refrigerants show similar variation as they show with COP.

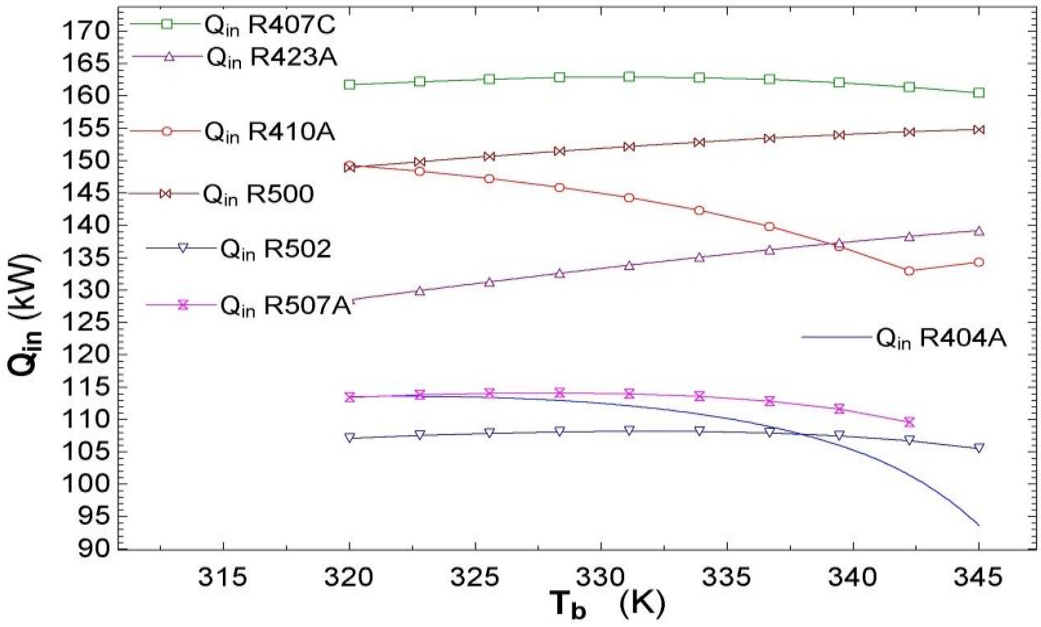


Figure 7: Variation of heat supplied to the boiler with boiler temperature (T_b)

The boiling temperature variation shows the least heat input required. For R502 the average heat input required is of a range of 100kW. Both R507A and R404A shows the similar starting heat input at 320K while for R404A it reduces quickly near 345K. From the fig 8 we can study the variation of COP with T_C. Most of the refrigerants are showing a flat variation for the entire operating range. R423A provides the least COP given as .65 to .75 and R404A has the maximum in the range of 1.8 to 2.5. R410A has the next best performance with 1.45 to 1.75 throughout the condenser operation

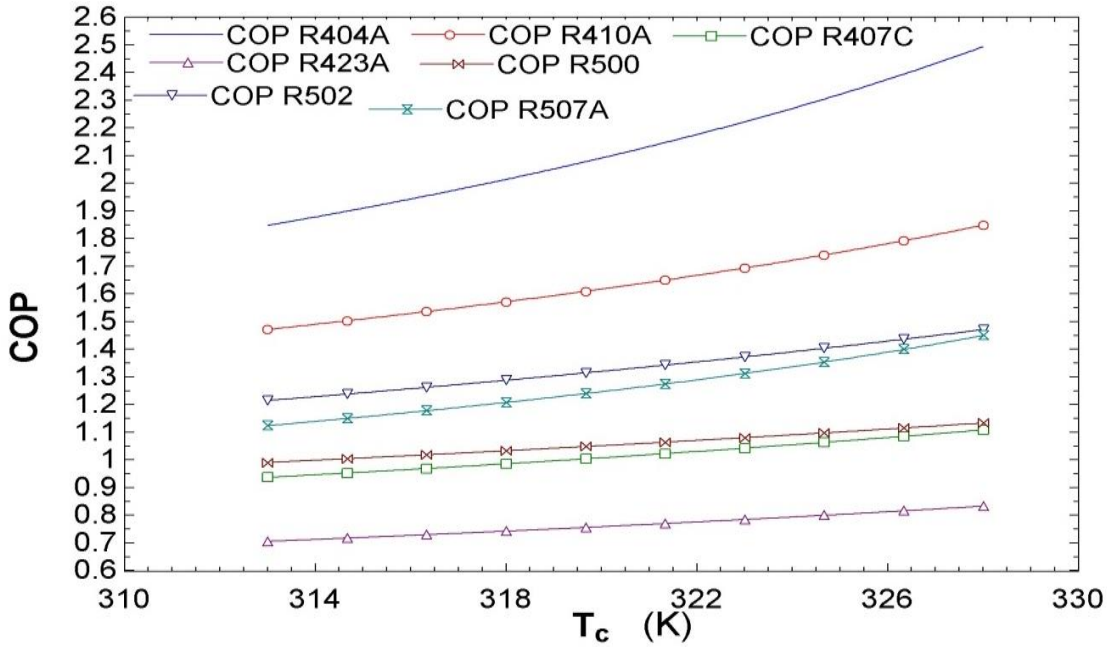


Figure 8: Variation of first Law Efficiency (COP) with Condenser operating temperatures. (T_c)

Fig 9 shows the requirement of heat input for the boiler for the chosen refrigerants. The heat input is low grade energy and instead of the high grade work it is used. So we get numerous options to arrange sources of for heat. Waste heat from several power generating cycles can be used as the source. The higher is the heat input the lower will be the COP of the cycle, so shows the fig 9 that the least heat input required is for R404A, for which with increase in the operating temperature of the condenser, the heat required is as low as 75kW. Although the trend is decreasing for all the working fluids.

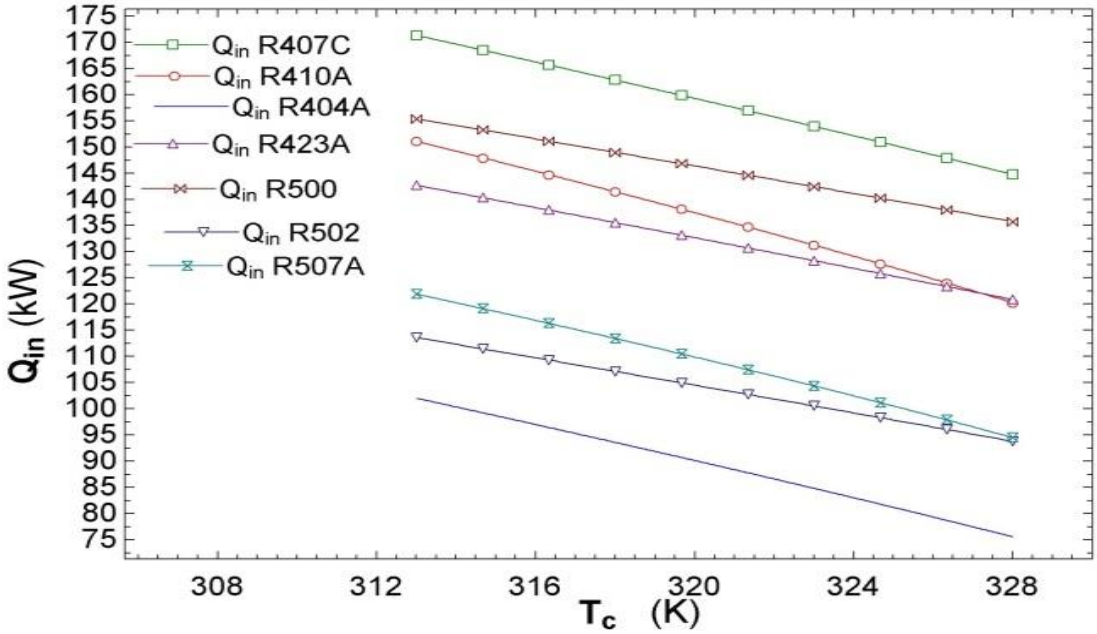


Figure 9: Variation of heat input required with Condenser temperatures.(T_c)

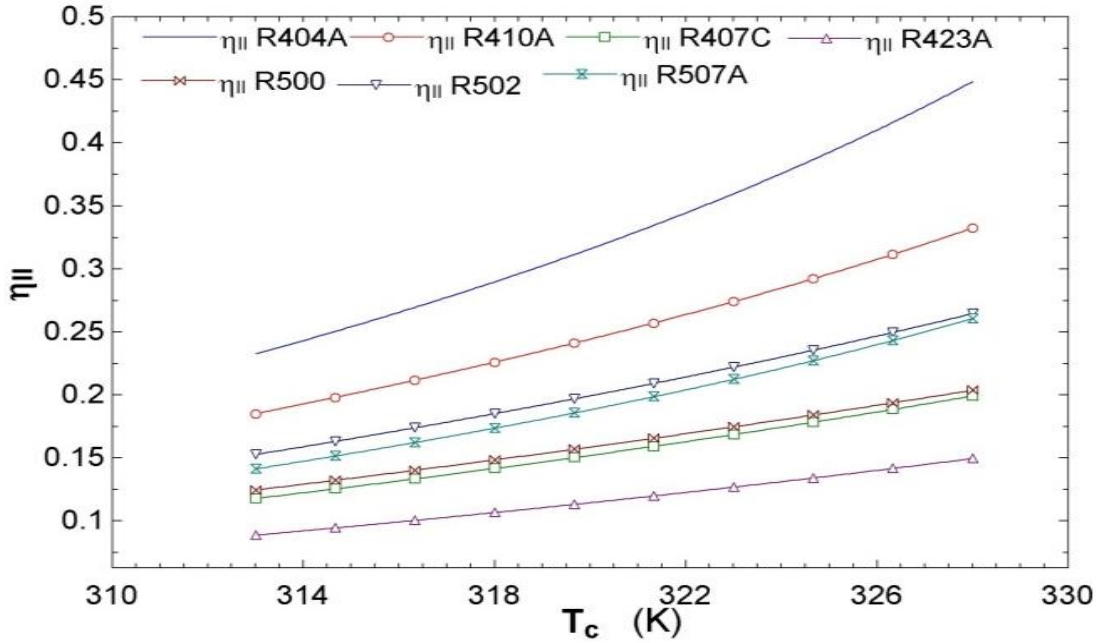


Figure 10: Variation of Exergetic efficiency with Condenser temperatures (T_c)

Through fig 10 we can observe that R404A gives the best exergetic performance with the efficiency ranging from 25% to as high as 45%. Next best option is R410A with the maximum efficiency 30%. It can be seen that R502 and R507A are having results very close and similarly R500 and R407C are very close so these are replaceable fluids for the cycle. The evaporator temperature shows the required temperature to be developed. Here a 20K variation of operating temperature has been taken. The minimum temperature that has been chosen is -20°C and the COP variations for the refrigerants are given in the fig 11. At 253K it can be seen that R404A has a considerably high COP of 1.443 and at 273K it has 1.843. Over the entire range of temperature R410 has a flat range of COP as 1.3 to 1.4

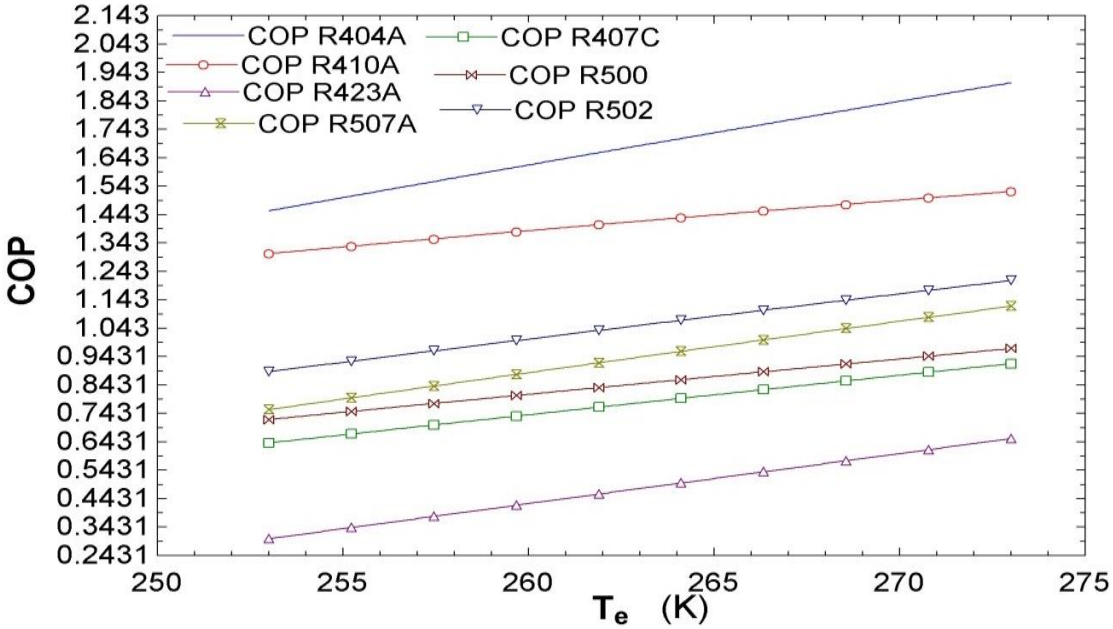


Figure 11: Variation of first Law Efficiency (COP) with evaporator temperatures (T_e)

The refrigeration effect is the highest for R410A which is 180kW at 253K to 210kW at 273K. Despite a better COP, R404A doesn't provide a greater RE. R500 and R407C are almost overlapping similarly R502 and R507A.

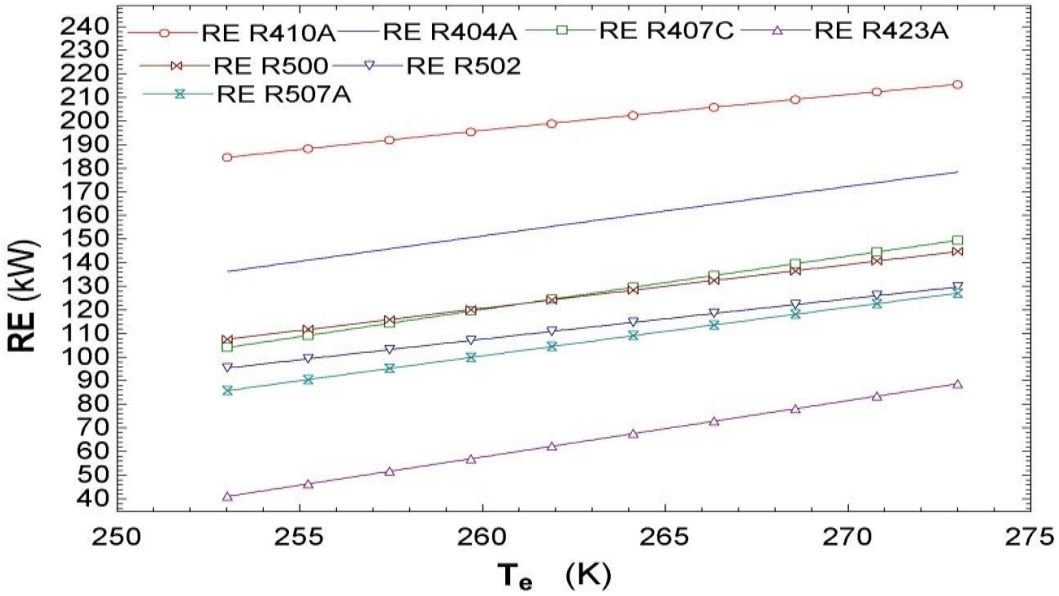


Figure 12: Variation refrigeration effect with evaporator temperatures (T_e)

In fig 13 a variation for second law efficiency is shown. It can be seen that as the temperature increases in the evaporator exergy losses increase and the second law efficiency decreases. At 253K the efficiency is the best and it is of range of 38% for R404A and for R410A the efficiency is of a range of 34%.

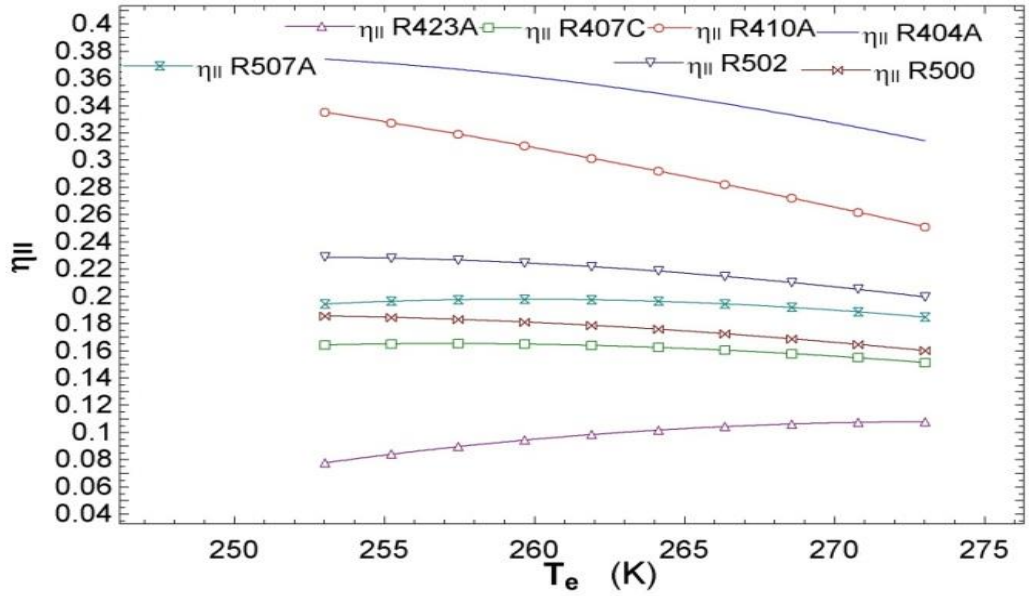


Figure 13: Variation of Exergetic efficiency with evaporator temperatures.

3. Conclusions

The Numerical computation is carried out for variations of the performance parameters with the operating temperatures in the system. The following conclusions have been drawn.

- For the variations in the boiling temperature R410A has the best COP ranging from 1.4 to 1.8, also R410A has the best RE.
- The R410 has a exergetic efficiency ranging in 20% to 26% when the boiling temperature is ranging from 320K to 345K. R502 has the least heat input required for the operation of the cycle.
- R404A has the maximum in the range of 1.8 to 2.5 when condenser temperature is observed to be from 313K to 328K. With the view of operation of condenser its R404A that requires least heat input for operation, also 45% of second law efficiency is obtained for the same working fluid.
- R502 and R507A are having results very close and similarly R500 and R407C are very close so these are replaceable fluids for the cycle.
- At the least evaporator temperature 253K, R404A is having COP of 1.443 and RE is generated by R410A. Second law efficiency of 38% for R404A and for R410A 34% is observed.

References

[1] Christian Tischendorf , Denise Janotte , Ricardo Fiorenzano , Wilhelm Tegethoff ," Investigation of Energy Dissipation in an Ejector RefrigerationCycle" Modelica Conference, Como, Italy, Sep. 20-22, 2009.

[2] Suhas D Kshirsagar, M M Deshmukh [2013]," Thermal Design & Performance Of Combined Vapour Compression-Ejector Refrigeration System Using R600a "International Journal of Engineering Research and Applications Vol. 3, Issue 2, March -April, pp.1368-1380.

[3] M. M. Rashidi, O. A. Bég a, A. Aghagoli," Utilization of waste heat in combined power and ejector refrigeration for a solar energy source" Int. J. of Appl. Math and Mech. 8(17): 1-16, 2012.

[4] B. J. Huang, J. M. Chang, V. A. Peterenko, K. B. Zhuk ," A solar ejector cooling system using refrigerant R141b" Solar Energy Vol. 64, Nos 4–6, pp. 223–226, 1998.

[5] Rohit Khajuria , Jagdev Singh , "Performance analysis of ejector refrigeration system with environment friendly refrigerant driven by exhaust emission of automobile" Advances in Applied Science Research, 2013, 4(5):232-237.

[6] Jianlin Yu , Hua Chen, Yunfeng Ren, Yanzhong Li ," A new ejector refrigeration system with an additional jet pump" Applied Thermal Engineering 26 (2006) 312–319.

[7] Sandeep Kashyap, R.C. Gupta, "Theoretical study of ejector refrigeration system with working fluid R410a" International Journal of Engineering Science and Technology.

[8] Aditya Jain, S.K.Agrawal, P.Pachorkar," Steady-State Analysis of the Solar-Driven Ejector Refrigeration System Using Water, Methanol, Ammonia As A Refrigerant" International Journal of Emerging Technology and Advanced Engineering ,Volume 2, Issue 9, September 2012.

[9] Zheng Huifan,Fan Xiaowei,Zhang Lihe ," Study on the Hourly Performance of A Solar Driven Ejector Refrigerant System" The 2nd International Conference on Computer Application and System Modeling (2012).

[10] Stefan Elbel, Predrag Hrnjak," Ejector Refrigeration: An overview of historical and present developments with an emphasis on air conditioning applications" International Refrigeration and Air Conditioning Conference at Purdue, July 14-17, 2008.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 6, Issue 3, May 2017

Column study for chloride removal from waste water by a low cost adsorbent (bio adsorbent)

Lokesh Kumar[1] and Prof. S K Singh[2]
[1]Research scholar, Department of Environment engineering, Delhi Technological University.
[2]Professor & Dean, Department of Environment Engineering, Delhi Technological University (formerly Delhi college of Engineering), Delhi, India

Abstract: - In this article, adsorption process has been found effective for removal of chloride. Water pollution and its control has become an important issue to deal with. The other filtration method of water treatment is expensive and not feasible. The requirement of low cost natural bio filter is being studied for the removal of water pollutants. Specifically for chloride removal the use of Bio adsorption is found to be very effective. The Parthenium has been extensively used for removal of chloride either by using batch mode adsorption or by column study. Here in this article, the removal of chloride has been studied by using column method.

Key words: **- Chloride, Bio-adsorption, Parthenium, Column Study.**

I. INTRODUCTION

The removal of chloride from water by Parthenium using fixed-bed column adsorption techniques was investigated. Fixed-bed column experiments were carried out for different influent chloride concentrations, bed depths, and various flow rates. The breakthrough time and exhaustion time decreased with increasing flow rate, decreasing bed depth and increasing influent chloride concentration. The bed depth service time model and the Thomas model were applied to the experimental results. Both model predictions were in good agreement with the experimental data for all the process parameters studied, indicating that the models were suitable up to some extent for parthenium fix-bed column design. Chloride is a salt compound resulting from the combination of chloride salts includes Sodium Chloride (NaCl), and Magnesium Chloride ($MgCl_2$). The chloride ion is negatively charge anion (Cl^-). It is formed when the elements chlorine gains an electrons or when a compound such as hydrogen chloride is dissolved in water or other polar solvent. Chloride salts such as NaCl are very soluble in water. It is an essential electrolyte located in all body fluids responsible for maintaining acid/base balance, transmitting nerve impulses and regulating fluid in and out of cells. The word chloride can also form part of the name of chemical compounds in which one or more chlorine atoms are covalently bonded. For example, methyl chloride, more commonly called chloromethane, (CH_3Cl) is an organic compound with a covalent C-Cl bond. It is not a source of chloride ion. Chloride in concentration above 600mg/l tends to give water a salty taste. The concentration of chloride content above 200mg/l is considered objectionable. Presence of high quality of chloride content in water resources indicates pollution due to human and industrial wastes and also from the earthen rocks in the subsurface. High chloride concentrations in freshwater can harm aquatic organisms by interfering with osmoregulation, the biological process by which they maintain the proper concentration of salt and other solutes in their bodily fluids. Difficulty with osmo regulation can hinder survival, growth, and reproduction.

II. LITERATURE REVIEW

The Bio adsorbents cover small plants, shrubs and trees etc. Low cost materials prepared from Bio- wastes such as grass, plants, shrubs, fruits and vegetable peels, nuts, shells, pulps, stones, barks, roots, leaves, fruit wastes and sawdust of certain timber trees have gained immense popularity because of their good metal removal efficiency and ease of operation in low cost. Bio- waste was studied in removal of Hg by chemically modified cotton (Roberts and Rowland, 1973), removal of lead ions using soymida fabrifuga bark (Banker and Dara 1985). Cr removal by rice husk carbon was studied by Srinivasan et al. (1988), Manju and Aniradhan (1990) used coconut fibre pith based pseudo activated carbon for Cr (VI) removal. Removal of Ni from aqueous solutions was studied using an agricultural waste peanut hulls (Periasamy and Namasivayam, 1995). The use of

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 6, Issue 3, May 2017

Bio waste based adsorbents became common in late 1990s and lot of materials have been studied so far, for example, sunflower plant dry powder for removal of heavy metals (Bhalke et al., 1999), eucalyptus globulus bark for removal of Ni and Cu (Dolultani et al., 1999), orange peel for Ni removal (Ajmal et al.,2000), coirpith activated carbon in Cd, Ni and Cu removal (Kadirvelu, 2001), hazelnut shell activated carbon in Ni removal (Demirbas et al., 2002), rice husk in Cr removal (Tang et al., 2003), Jack fruit in Cd removal (Inbaraj and Sulochana, 2004), black gram husk in Pb removal (Saeed et al., 2005), Coconut copra meal in Cd (Ofomaja and Ho ys, 2006), ficus religiosa leaves in Pb and Cr removal (Qaiser etal., 2007), pretreate arca shell biomass in removal of heavy metals and radionuclie (Dahiya et al.,2008), Rice husk ash in Pb removal (Naiya et al., 2009), Bamboo activated carbon in Pb removal (Lalhruaitluanga et al., 2010), onion skin in Pb removal (Saka et al., 2011), Pinus tree bark in Pb removal (Affonso et al., 2012),Aliathus excels tree bark in Pb removal (Waghmare and Chaudhary, 2013), Neem bark and potato peels in removal of Malachite green (Sharma et al., 2014), Chondruscrispus powder for removal of chromium ions (Elavarasan, 2015). Ashtikar and Parkhi (2014) studied removal of Cu (II) from aqueous solution using mango seed powder as low cost adsorbent. The results showed that the adsorption was rapid during first 45 minutes & equilibrium was reached in 90 minutes. With increase in the adsorbent doses & temperature percent removal of copper increases. Sethu et al. (2010) studied adsorption of Cu(II) on mango leaf bio-sorbent. Results showed that the maximum adsorption capacity was 206.85 mg/g at 100 mg/L Cu(II) ion concentration and 0.4 g/L of adsorbent dosage. Adsorption of Cu(II) on mango leaf biosorbent was found to fit the Freundlich isotherm. Mango leaf was utilized as low-cost adsorbent for the removal of potassium ion from aqueous solution (Chaudhary et al., 2015). A maximum adsorption capacity of 42.5mg/g was obtained at pH 6.5 and Langmuir model fitted the equilibrium data better, giving correlation coefficient of 0.9999.

Removal of Pb(II) ions from aqueous solution by adsorption using bael leaves (Aegle marmelos) was studied by Chakravarty et al. (2009).The maximum Pb removal capacity was reported as 104 mg/ g at 50 mg/ L initial Pb(II) concentration at pH 5.1. The sorption process was described by pseudo second order kinetics model and Langmuir isotherm equations. Bael tree leaf was studied for removal of Pb ions (Kumar and Gayathri, 2009). The maximum adsorption capacity obtained was 90.07 mg/g at pH 5.0 at 30^0 C in for the initial metal ion concentration of 25-100 mg/l and an adsorbent dose of 5-20 g/l. In my earlier research paper it was studied that during the batch study of the chloride by using parthenium a significant adsorption was observed at pH -7. Veried dose shows different adsorption. (Lokesh Kumar, S. K. Sigh ,2015, Removal of chloride from Ground Water by Bio-adsorption, Int. J of advance research.

III. MATERIAL AND METHODS

A. CHLORIDE SAMPLE & GAJARGHAS (Parthenium) AS BIOADSORBENT

Natural Wastewater Sample with Chloride Concentration 1000 ppm was taken and the preliminary experiments were done with this sample. A plant known as grass / Gajarghas / Parthenium sps. / Congress (Belongs to the family asteraceae) was used as a bio-adsorbent. Before using, the Plant, it was dried & powdered (containing leaves and stem). The word 'Parthenium' has been derived from the Latin word 'Parthenice', meaning reputed medicinal merits. It has been speculated by scientists that the unique properties of this weed can be exploited for different purposes like activated carbon making, biomass generation, pesticide use etc. Parthenium is considered to be one of the ten worst weeds in the world. Parthenium is an herbaceous annual or ephemeral member of the family Asteraceae. It can reach heights of up to 2 m in good soil, and attain flowering in less than 4– 6 weeks of germination. Seeds produced per plant can go up to 25000. Dispersion of the seeds occurs through various vectors like water, muddy surfaces, small animals, vehicles, machinery etc. Partheniumhas a variety of vernacular names, like Congress grass, White top, Star Weed, Carrot Weed, Gajar Ghaas (Hindi), Ramphool, feverfew etc. The pollen and seeds of these plants are known to be a major cause of asthma, and bronchitis. The weed shows the presence of toxins known as Sesquiterpene lactones, parthenin, phenolic acids such as caffeic acid, vanillic acid,anisic acid, panisic acid, and parahydroxy benzoic acid which are known to be lethal to humans as well as animals.

B. METHOD OF ACTIVATION OF BIO-ADSORBENT

Sample collection was done on the basis of the health of the plant through visual observations. Young plants having a fresh green shoot and sizable stem thickness were selected. On collection, the specimens were washed with tap water and then distilled water. Further, for biomass collection, the specimens were subjected to drying in an oven at 100C for three days. Drying was followed by careful crushing of the specimen,

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 6, Issue 3, May 2017

and then sieving the mixture through a 500 micron sieve. The obtained biomass was used for the study. The biomass was added to the test solutions in the ratio of 0.10 g: 100 ml test solution. The pH of the sample was maintained at 7, and the chloride concentration at 100% of the initial wastewater solution. The mixture was then incubated in a rotary shaker incubator at 150 rpm and 300 C for 60 minutes. On the termination of contact time, the solution was filtered using an ordinary filter paper. The sample was titrated for chloride content using the Argentometric titration method specified in the APHA handbook for water and wastewater analysis. The tests were repeated by carrying out variations in pH, by using different chloride concentration levels, and with variable contact times.

IV. MODEL ON COLUMN PERFORMANCE

Optimization of parameters in the continuous adsorption process by experimental methods is an expensive and time consuming. It would be much cheaper and faster to use mathematical modeling to predict the duration of the column before regeneration become necessary (Warchol and Petrus 2006). Moreover, it is hard to generalized operational design parameters, therefore, theoretical models taking in to account of chemical and physical conditions become important (chen et al. 2003). Some of the well known mathematical models applied for column mode operation are Bed Depth Service Time (BDST) model, Thomas Model, Theoretical breakthrough curve etc.

A. BED DEPTH SERVICE TIME (BDST) MODEL

The Bohart-Adams equation can be represented as (Bohart and Adams 1920):

$$Ln\left(\frac{Co}{Cb}-1\right)=Ln(e^{kadsN0(z/u)}-1)-kadsC0ts \quad (1)$$

Further, Hutchins (1973) modified the Bohart Adams equation and presented a linear relationship between the bed depth and service time which required only three fixed bed tests to collect the necessary data.

$$t_s=\frac{N_0 Z}{C_0 u}-\frac{1}{k_{ads}C_0}Ln\left(\frac{C_0}{C_b}|1\right) \quad (2)$$

Where,

t_s is the service time at breakthrough point (h),
N_0 the dynamic bed capacity (mg L^{-1}) ,
Z the packed- bed column depth (cm),
u the linear flow rate (cm/h) defined as the ration of volumetric flow rate Q(cm^3/h) to the cross section area of the bed A (cm2) ,
C0 and Cb are respectively the influent and the breakthrough adsorbate concentration (mg/L) and
Kads the adsorption rate constant (L/mg.h).

Plotting service time (tss) versus bed depth (Z) will generate a straight line equation having slope of (N_0/C_0.u) and the intercept of (-)$\frac{1}{k_{ads}C_0}Ln\left(\frac{C_0}{C_b}-1\right)$. Critical bed depth ($Z_0$) represents the theoretical minimum depth of column that would be able to prevent the adsorbate concentration from exceeding C_b. It is obtained when breakthrough is immediate and it can be calculated by substituting $t_s = 0$ in equation as below

$$Z_0=\frac{u}{K_{ads}N}\left(\frac{C_0}{C_b}-1\right) \quad (3)$$

According to BDST model equation, the data collected from one flow rate experiment can predict the system with different flow rate. When an experiment conducted at flow rate Q_1, yields an equation of the form

$$t_s = a_1 z + b_1 \quad (4)$$

The predicted equation for new flow rate Q_2 is given by :

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 6, Issue 3, May 2017

$t_s = a_2 z + b_1$ (5)

and $a_2 = a_1 \left(\frac{Q_1}{Q_2}\right)$ (6)

where a_1 and a_2 are the slop at flow rate Q_1 and Q_2 respectively. However the intercept b_1 remained same since it is independent of flow rate in line arized BDST equation. BDST equation can also be used to design system for treating other influent solute concentration using the data of a previous laboratory experiment of one influent solute concentration. When an experiment conducted at initial concentration C_1, yield an equation of the form

$t = r_1 X + S_1$ (7)

The predicted equation for new flow rate Q_2 , is given by:

$t = r_2 X + S_2$ and (8)

The new slop and intercept values can be determined as :

$r_2 = r_1 \left(\frac{C_1}{C_2}\right)$ (9)

$$s_2 = s_1 \frac{C_1}{C_2}\left(\frac{Ln\left[{}^{C_2}/_{C_F} - 1\right]}{Ln\left[{}^{C_1}/_{C_b} - 1\right]}\right) \quad (10)$$

Where,

r_1 and r_2 is slopes at influent concentration
C_1 and C_2 respectively, S_1 and S_2 are intercept at influent concentration C_1 and C_2 respectively ,
C_F is effluent concentration at influent concentration C_2 and
C_b is effluent concentration at influent concentration C_1.

V. COLUMN ADSORPTION EXPERIMENTS

Continuous flow adsorption experiments were conducted in PVC columns of 4.0 cm inside diameter. At the top of the column, the influent fluoride solution (1000 mg /L) was pumped through the packed column (22, 44 and 66 cm), at flow rates of 5, 10 and 15 mL/min, using a peristaltic pump. Samples were collected from the exit of the column at regular time intervals and analyzed for residual chloride concentration at pH 7. The parthenium first activated and then it was used as bio adsorbent .The saturation capacity for the Parthenium in these column studies was calculated from the following equation where qe is the chloride adsorbed (mg/g), C_0 is the influent chloride concentration (mg/L), C is the effluent chloride concentration (mg/L), VE is the volume of solution required to reach the exhaustion point (L), and m is the mass of adsorbent (g).

A. EFFECT OF FLOW RATE

The adsorption columns were operated with different flow rates (5, 10, and 15 mL/min) until no further chloride removal was observed. The breakthrough curve for a column was determined by plotting the ratio of the Ce/C_0 (Ce and C_0 are the chloride concentration of effluent and influent, respectively) against time. The column performed well at the lowest flow rate (5 mL/min). Earlier breakthrough and exhaustion times were achieved, when the flow rate was increased from 5 to 10 mL/min. The column breakthrough time (Ce/C_0=0.05) was reduced from 65 to 35 min, with an increase in flow rate from 5 to 10 mL/min. This was due to a decrease in the residence time, which restricted the contact of chloride solution to the Parthenium. At higher flow rates the chloride ions did not have enough time to diffuse into the pores of the Parthenium and they exited the column before equilibrium occurred.

Successful design of a column chloride adsorption process requires a description of the dynamic behavior of chloride in a fixed bed. Various simple mathematical models have been developed to describe and possibly

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 6, Issue 3, May 2017

predict the dynamic behavior of the bed in column performance. One model used for continuous flow conditions is the Thomas model, which can be written as:

$$C_e/C_0 = 1/1 + exp\ ((k_{th}/Q)(q_0M - C_0V_{eff})) \quad (11)$$

Eq. can be expressed in linear form as:

$$ln[C_0/C_e - 1] = k_{th}q_0M/Q - k_{th}C_0t \quad (12)$$

Where

V_{eff} is the volume of effluent (L),
k_{th} is the Thomas model constant (L/mg h),
$q_0$is the adsorption capacity (mg/g),
Q is the volumetric flow rate through column (L/h),
M is the mass of adsorbent in the column (g),
C_0 is the initial chloride concentration (mg/L) and
Ce is the effluent chloride concentration (mg/L) at any time t (h).

The Thomas model constants k_{th} and q_0 were determined from a plot of ln [C_0/Ce-1] versus t at a given flow rate. The model parameters are given in Table. The Thomas model gave a good fit of the experimental data, at all the flow rates examined, with correlation coefficients greater than 0.97, which would indicate the external and internal diffusions were not the rate limiting step. The rate constant (k_{th}) increased with increasing flow rate which indicates that the mass transport resistance decreases. The reason is that the driving force for adsorption is the chloride concentration difference between Parthenon and solution.

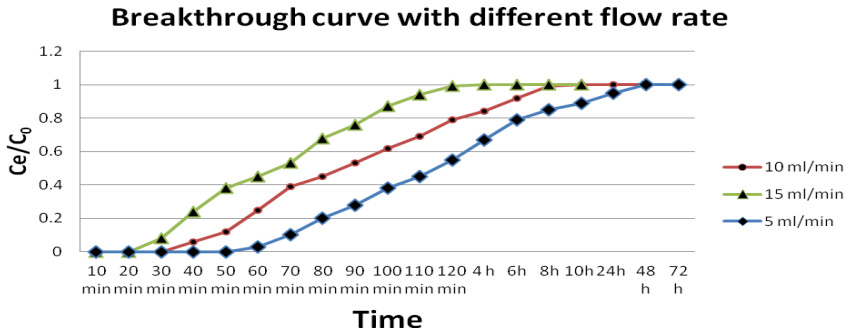


Fig 1 Breakthrough curves expressed as Ce/C_0 versus time at different flow rates (initial chloride Concentration 1000 mg/L, initial pH 7, bed depth 66 cm and temperature 30 ± 1°C)

B. EFFECT OF BED HEIGHT

The accumulation of chloride in a fixed-bed column is dependent on the quantity of adsorbent inside the column. In order to study the effect of bed height on chloride retention dry activated parthenium of three different bed heights, viz. 22, 44, and 66 cm, were used. A chloride solution of fixed concentration (1000 mg/L) was passed through the fixed-bed column at a constant flow rate of 5 mL/min. The breakthrough time varied with bed height. Steeper breakthrough curves were achieved with a decrease in bed depth. The breakthrough time decreased with a decreasing bed depth from 44 to 22 cm, as binding sites were restricted at low bed depths. At low bed depth, the chloride ions do not have enough time to diffuse into the surface of the Parthenium and a reduction in breakthrough time occurs. Conversely, with an increase in bed depth, the residence time of chloride solution inside the column was increased, allowing the chloride ions to diffuse deeper into the Parthenium. The breakthrough service time (BDST) model is based on physically measuring the capacity of the bed at various percentage breakthrough values. The BDST model constants can be helpful to scale up the process for other flow rates and concentrations without further experimentation. It is used to predict the column performance for any bed length, if data for some depths are known. It states that the bed depth, Z and service time, t of a column

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 6, Issue 3, May 2017

bears a linear relationship. The rate of adsorption is controlled by the surface reaction between adsorbate and the unused capacity of the adsorbent. The BDST equation can be expressed as follows

$$t_s = \frac{N_0 Z}{C_0 u} - \frac{1}{k_{ads} C_0} Ln\left(\frac{C_0}{C_b} - 1\right) \quad (13)$$

Where C_b is the breakthrough chloride concentration (mg/L),
N_0 is the adsorption capacity of bed (mg/L),
Z is depth of column bed (cm),
u is the linear flow velocity of chloride solution through the bed (mL/cm^{2}h),
Ka is the rate constant (L/mg h).

The column service time was selected as the time when the normalized concentration, Ce/C_0 reached 0.05. A plot of service time versus bed depth, at a flow rate of 5 mL/min was linear. The high correlation coefficient value (R^2= 0.988) indicated the validity of the BDST model for the present system. The values of N and Ka were evaluated from the slope (N/C_0u) and intercept ($(1/KaC_0)\ln[(C_0/C_b)-1]$) of the BDST plot. The values of BDST model parameters are presented in Table . The value of Ka characterizes the rate of transfer from the fluid phase to the solid phase. If Ka is large, even a short bed will avoid breakthrough, but as Ka decreases a progressively deeper bed is required to avoid breakthrough

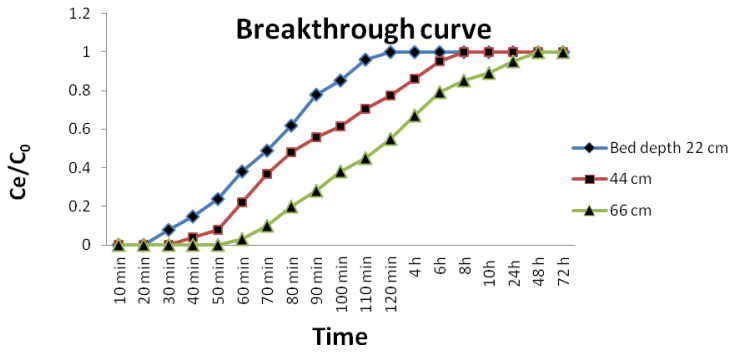


Fig 2 Breakthrough curves expressed as Ce/C_0 versus time at different bed depth (initial chloride Concentration 1000 mg/L, initial pH 7, flow rate 5 mL/min and temperature 30 ± 1°C)

C. THOMAS MODEL AND BDST MODEL PARAMETERS FOR THE ADSORPTION OF CHLORIDE ON PARTHENIUM

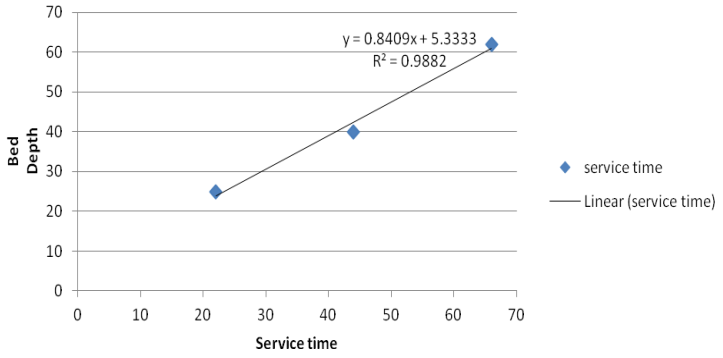


Fig 3 Parameters for BDST Models

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 6, Issue 3, May 2017

Table 1 BDST MODEL coefficient

N (mg/L)	Ka (L/mg h)	R^2
20063	5.52X 10-4	0.998

D. PARAMETERS FOR THOMAS MODEL

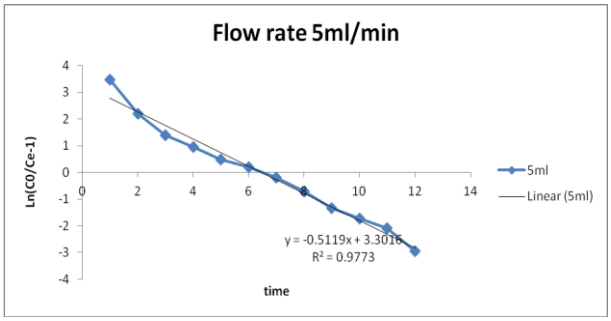


Fig 4 Parameters for Thomas model at flow rate 5ml/min

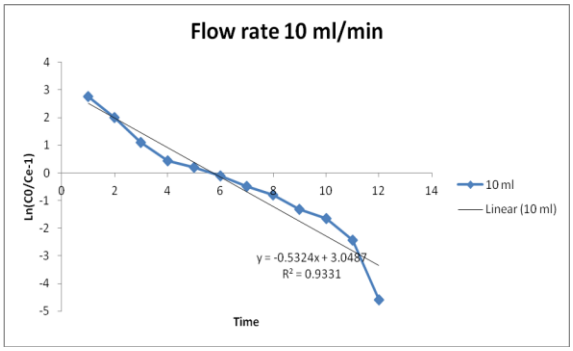


Fig 5 Parameters for Thomas model at flow rate 10 ml/min

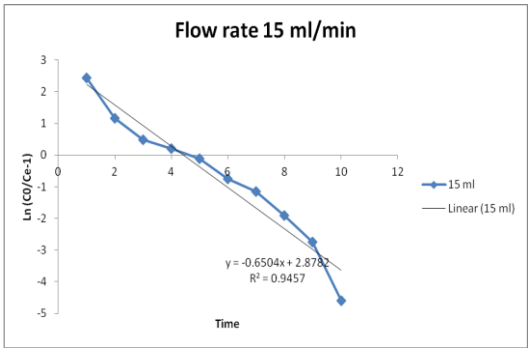


Fig 6 Parameters for Thomas model at flow rate 15 ml/min

Table 2 THOMAS MODEL COEFFICIENT

Flow Rate (ml/min)	**q0 (mg/g)**	**Kth (L/mg h)**	**R^2**
5	**3.87**	**0.000511**	**0.977**

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 6, Issue 3, May 2017

10	6.85	0.000532	0.933
15	7.95	0.00065	0.945

VI. RESULTS AND DISCUSSION

BDST model is best suited to the observation and adsorbent performance compare to Thomas model. The high correlation coefficient value (R^2= 0.998) indicated the validity of the BDST model for the present system. Thomas model also shows good results. The correlation coefficient value (R^2= 0.977 with flow rate of 5 ml/min.). In the Thomas model as the flow rate increases from 5 to 15 ml/min the efficiency of the adsorption reduces. The column breakthrough time (Ce/C_0=0.05) was reduced from 65 to 35 min, with an increase in flow rate from 5 to 10 mL/min. This was due to a decrease in the residence time, which restricted the contact of chloride solution to the Parthenium. At higher flow rates the chloride ions did not have enough time to diffuse into the pores of the Parthenium and they exited the column before equilibrium occurred. The concentrations of chloride in the treated waste water are expected to be well below the permissible limits for discharge of effluents as per Indian standards. Regeneration of adsorbents is expected to allow reuse of the adsorbents. A good recovery of metal ions is expected on desorption. Studied waste material is expected to be of significant use in water treatment.

VII. CONCLUSION

Parthenium can be utilized as a good bio-adsorbent for the removal of Chloride salts. It can also be used for the removal of heavy metals and other effects on the waste water treatments. The different column models have been effective for study of maximum efficiency of parthenium towards removals of different salts.

VIII. FUTURE SCOPE

The column study can be conducted with some different bio adsorbents for removal of pollutants in water and waste water according to the affinity of the bio adsorbent. Desorption can also be studied for recovery of heavy metals.

REFERENCES

[1] Azam, F. and Ifzal, M. (2006): Microbial populations immobilizing NH4 + - N and NO3 - -N differ in the sensitivity to sodium chloride salinity in soil, Soil Biology and Biochemistry, 38(8):2491 – 2494.

[2] Chu, W. (1999): Lead Metal Removal by Recycled Alum Sludge, Wat. Res., 33, 13, 3019. 3) Cleserl S. L., Greenberg E. A., Eaton D.A. (1999): Standard Methods for Examination of Water and Wastewater, American Public Health Association, 20th Edn.

[3] Department of National Health and Welfare (Canada) Guidelines for Canadian drinking water quality, supporting documentation, (1978), Ottawa.

[4] Elsener, B. and Angst, U. (2007): Mechanism of electrochemical chloride removal, Corrosion Science, 49(12):4504 – 4522.

[5] Fadeeva, V. K. (1971): Effect of drinking water with different chloride contents on experimental animals, Gigiena i sanitarija (in Russian), 36(6):1115 (Dialogue abstract No. 051634).

[6] Fork, D.C., (1986): The control by state transitions of the distribution of excitation energy in photosynthesis, Annual Review of Plant Physiology and Plant Molecular Biology, 37: 335 – 361.

[7] Giyeon, Y., Buchanan, I.D. (2006): Project Report on Colour and Chloride Removal from Pulp Mill Effluent Using Ion – Exchange, Sustainable Forest and Management Network Publication, University of Alberta, Edmonton, Alberta, T6G 2E9.

[8] Gryndler, M., Rohlenova, J., Kopecky, J., Matucha, M., Chloride concentration affects soil microbial community.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 6, Issue 3, May 2017

[9] Hajrasuliha, S. (1979): Accumulation and toxicity of chloride in bean plants, Plant and Soil, 55(1). Universal Journal of Environmental Research and Technology 422 Apte et al.

[10] Liang, Lv., Jing, He., Min, Wei., Evans, D.G., and Duan, X. (2006): Uptake of chloride ion from aqueous solution by calcined layered double hydroxides: Equilibrium and kinetic studies. Water Research, 40(4): 735 – 743.

[11] Lockhart, EE, Tucker, CL, Merritt, MC. (1955): The effect of water impurities on the flavor of brewed coffee, Food Research, 20:598.

[12] lsener, B. and Angst, U. (2007): Mechanism of electrochemical chloride removal, Corrosion Science, 49(12):4504 – 4522.

[13] Lu, Y.K., Chen, Y.R., Yang, C.M., and Ifuku, K. (1995): Influence of Fe- and Mg- deficiency on the thylakoid membranes of a chlorophylldeficient ch5 mutant of Arabidopsis thaliana. Botanical Bulletin of Academia Sinica, 36.

[14] Muller, M. (2000): Effects of Chloride transport inhibition and chloride substitution on neuron function and on hypoxic spreading – depression – like depolarization in rate hippocampal slices. Neuroscience; 97(1):33 – 45.

[15] Osinchak, J.E., Hultin, H.O, Zajicek, O.T., Kelleher, S. D., Huang, Chen-Huei. (1992): Effect of NaCl on catalysis of lipid oxidation by the soluble fraction of fish muscle. Free Radical Biology and Medicine, 12(1): 35 -41.

[16] Patil, Y.B., and Paknikar, K.M. (1999): Removal and recovery of metal cyanide using a combination of biosorption and biodegradation processes. Biotechnology Letters, 31: 913 – 919.

[17] Wesson, L.G. (1969): Physiology of the human.

[18] Abdus- Salam N. and Adekola F.A.,2005, The influence of pH and adsorbent concentration on adsorption of lead and zinc on a natural Goethite, African J. of Sci. and Tech., 6 (2), 55-66.

[19] Affonso C. G. J, Leonardo S, Cleber A. L, Herbert N, Daniel S and Edleusa P. S.,2012; Applicability of the Pinus Bark (Pinus elliottii) for the Adsorption of Toxic Heavy Metals From Aqueous solutions, Acta Scientiarum, 34 (1), 79- 87.

[20] Ajmal M., Rao R.A., Ahmad R., Ahmad J., 2000, Adsorption studies on Citrus reticulata (fruit peel of orange): removal and recovery of Ni(II) from electroplating wastewater, J of Haz Mater., 1, 79(1-2), 117-31.

[21] Akinbiyi A, 2000, Removal of lead from aqueous solutions by adsorption using peat moss; A thesis, Univ. of Regina, Saskat Chewan.

[22] Ashtikar S.V., Parkhi A.D., 2014, Adsorption of Copper from Aqueous Solution Using Mango Seed Powder,Int. Journal of Engineering Research and Applications, 4 (4), pp.75-77.

[23] Babel S and Kurniawan T. A., 2003, Low-cost adsorbents for heavy metals uptake from contaminated water: a review, Journal of Hazardous Materials, 97 (1-3), 219-243.

[24] Bankar D and Dara S., 1985, Effectiveness of soymida fabrifuga bark for scavenging lead ions, Proc. National seminar on “Pollution control and environment management”, organized by NEERI, Nagpur, 01, 121-125.

[25] Cafer S, Omer S, Ahin M. M., 2012, Applications on agricultural and forest waste adsorbents for the removal of lead (II) from contaminated waters, int. J. Environ. Sci. Technol., 9: 379-394.

[26] Chakravarty S., Mohanty A., Sudha T.N., Upadhyay A.K., Konar J., Sircar J., 2009, Removal of Pb(II) ions from aqueous solution by adsorption using bael leaves (Aegle marmelos), J or Haz Mater, 173 (1-3), 502-9.

[27] Chaudhary B., Desai V. and Shah M., 2015,Mango leaf utilized as low-cost adsorbent for the removal of potassium ion from aqueous solution International Journal of Research in Pharmacy And Chemistry, 5(1), 196-207.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 6, Issue 3, May 2017

[28] Dabrowski A, 2001, Adsorption—from theory to practice, Adv. Colloid Interface Sci.93, 135–224.

[29] Elavarasan A., Nandhakumar V. and Ravi B., 2015, Equilibrium and kinetic studies of adsorption system of chromium ions from aqueous solution using chondruscrispus activated carbon, Journal of Chemical and Pharmaceutical Research, 7(4):992-999.

[30] Ferhan C. and Ozgur A., 2011, Activated Carbon for Water and Wastewater treatment: Integration of Adsorption and Biological Treatment, First Edition, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

[31] Goel P.K., 2006, Water pollution, Causes, Effects and Control, New age international publishers, Revised second edition.

[32] Kumar P.S., Gayathri R, 2009, Adsorption of pb^{2+} ions from aqueous solutions onto bael tree leaf powder: isotherms, kinetics and thermodynamics study, Journal of Engineering Science and Technology, 4 (4), 381 – 399

[33] Lalhruaitluanga H., Jayaram K., Prasad M.N.V., Kumar K.K., 2010, Pb II adsorption from aqueous solutions by raw and activated charcoals of Melocanna baccifera Roxburgh (Bamboo)- A comparative study, J Hazard Mater, 175: 311-318.

[34] Lee C.K., Low K.S., 1989, Removal of copper from solution using moss, J. Envi. Tech. Letters, 10: 395-404.

[35] Lopez-Cortes A., Ochoa J.L., 1999, in: A. Dabrowski (Ed.), Adsorption and its Applications in Industry and Environmental Protection, vol. 120B, Elsevier, Amsterdam, p. 903.

[36] Muhammad N., 1997, Removal of heavy metals in slow sand filters, Proceeding of the 23th WEDC Int. Conf. on water supply and sanitation for all, Durban, SA. 167-170.

[37] Muhammad N., 1998, Adsorption of heavy metals in slow sand filters, Proceeding of the 24th WEDC Int. Conf. on sanitation and water for all. Islamabad, Pakistan, 258-261

[38] Naiya T.K., Bhattacharya A.K., Mandal S, Das S.K., 2009, The sorption of Pb(II) ions on rice husk ash., J Hazard Mater 163, 1254–1264.

[39] Periasamy K, Namasivayam C, 1995, Removal of Ni II from aqueous solution and ni plating industry waste water using an agricultural waste: peanut hulls. Waste manag, 15 (1): 63-8.

[40] Qaiser S, Saleemi A.R., Ahmad M.M., 2007, Heavy metal uptake by agro based waste materials, Electronic j of Biotech, 10(3).

[41] Roberts E.J. and Rowland S.P., 1973, Removal of mercury from aqueous solutions by nitrogen containing chemically modified cotton. Environ. Sci.Technol., 7(6), 552-555.

[42] Saeed A. Iqbal M. and Akhtar M.W., 2005, Removal and recovery of lead(II) from single and multimetal (Cd, Cu, Ni, Zn) solutions by crop milling waste (black gram husk), J. Hazard. Mater., 117(1), 65–73.

[43] Saka C. S., Ahin O, 2011, Removal of methylene blue from aqueous solutions by using cold plasma and formaldehyde treated onionskins, Color Technol, 127, 246–255.

[44] Schmidt K, 1977, Behavior of special pollutants in slow sand filters used in artificial recharge of ground water , XVII Congress, Baden Baden.

[45] Sethu V.S.,2010, Goey K.S., Iffah F.R., Khoo C.M., and Andresen J. M., 2010,adsorption characteristics of cu (ii) ions in Aqueous solutions using mangifera indica(mango) leaf biosorbents, Journal of Environmental Research And Development, 5 (2).

[46] Sharma N, Tiwari D; Singh S. K., 2014, The Efficiency Appraisal for Removal of Malachite Green by Potato peel and Neem Bark: Isotherm and Kinetic Studies, Int. J of Chem. and Envi. Engineering, 5(2), 83-88.

[47] Tang P.L., Lee C.K., Low K.S., Zanal Z., 2003, Sorption of Cr VI and Cu II in aqueous solution by ethylenediamine modified rice hull, Environ Tech, 24: 1243- 51.

[48] Treybal, R. E., "Mass Transfer Operation", 3rd ed., McGraw Hill, 1980.

[49] Vigneswaran H., Moon H., in: A. Dabrowski (Ed.), Adsorption and its Applications in Industry and Environmental Protection, vol. 120B, Elsevier, Amsterdam, 1999, p. 533.

[50] Waghmare V.H. and Chaudhari U.E., 2013, Adsorption of Pb(II) from aqueous solution on Ailanthus Excelsa Tree Bark , Current World Environment, vol. 8(3), 489- 492.

[51] Lokesh Kumar, S. K. Sigh ,2015, Removal of chloride from Ground Water by Bio-adsorption, Int. J of advance research(2015) vol 3, Issu 5, 140-154.

[52] Neetu Sharma, D.P. Tiwari, S.K. Singh (2014), "The Efficiency Appraisal for Removal of Malachite Green by Potato Peel and Neem Bark: Isotherm and Kinetic Studies", International Journal of Chemical and Environmental Engineering, Vol.5, No. 2, April 2014, PP 83-88

[53] D.P.Tiwari, S.K.Singh, Neetu Sharma (2014), "Sorption of Methylene Blue on Treated Agricultural Adsorbents: Equilibrium and Kinetic Studies", Journal of Applied water Science, 4 (1), 2014.

[54] Abhas Jain, S.K. Singh (2014), "Defluoridation of Water Using Alum Impregnated Brick Powder And Its Comparison with Brick Powder", International Journal of Engineering Science and Innovative Technology (IJESIT), Vol.3, Issue2, July

[55] Neetu Sharma, D.P.Tiwari, S.K.Singh, (2014), " efficiency of chemically treated potato peel and neem bark for Sorption of direct Red Dye-81 from aquous solution", Rasayan , 7 (4), 399-407

[56] Neetu Sharma, DP Tiwari and S.K. Singh (2012), "Decolorisation of synthetic dyes by Agricultural Waste a Review"., international Journal at Scientific & Engineering Research, 3 (2), 1-10.

[57] SK Singh, Lokesh Kumar (2014), "Characterization of rural drinking water sources in Bhiwani District, Haryana: A Case Study". International Journal of Interdisciplinary Research and Innovations Vol. 2, Issue 4, pp: (27-37), Month: October - December 2014, Available at: **www.researchpublish.com**.

[58] *S.K.Singh(2017),Present Scenario & Challenges of Solid Waste Management in India,* Proceedings of International conference on Emerging areas of Environmental Science and Engineering EAESE-2017,February 16-18,2017.

[59] A Kaushal, SK Singh (2017), Removal of heavy metals by nanoadsorbents: A review, Journal of Environment & Biotechnology Research 6 (1), 96-104

[60] A.A. Gour, S. K. Singh, A. Mandal (2016). "A glance at the world: Composition, trends and impacts of municipal solid waste in India", Waste Management – **Elsevier**. 55 (2016), I-III . doi.org/10.1016/S0956-053X(IG) 303999-3.

[61] Achla Kaushal, S. K. Singh (2016). "Critical Analysis of adsorption data statistically", Appl. Water Science -**Springer**, September 2016, DOI 10.1007/s13201-016-0466-4.

[62] Achla Kaushal, S.K. Singh(2016),"Removal of Zn (II) from Aqueous Solution Using Agro-based Adsorbents", Imperial Journal of interdisciplinary research , 2(6).

International Journal of Research in Engineering and Innovation (IJREI)
journal home page: http://www.ijrei.com
ISSN (Online): 2456-6934

Comparative study of parabolic trough collector through MWCNT/H_2o nanofluid and water

R.S. Mishra, Harwinder Singh

Department of Mechanical, Industrial and Production Engineering, Delhi Technological University, India-110042
E.Mail:harrymehrok14@gmail.com, rsmishra@dtu.ac.in

Abstract

In this present work MWCNT nanofluid and water were used as working fluid to compare the thermal performance of solar parabolic trough collector. Both the fluids were flowing through receiver at different volume flow rates 160L/h and 100L/h. Experimental tests was performed only during sunny weather and temperature at outlet of receiver was measured through thermometer after every half an hour of total testing time period. MWCNT nanofluid with weight fraction 0.01% and 0.02% and water were used to find efficiency of system and it has been seen that MWCNT nanofluid 0.02wt% with 160L/h showed better results for overall thermal efficiency among other and also an application of surfactant Triton X-100 with in MWCNT nanofluid was used to enhance the quantity of heat absorption capability of base fluid. © 2017 ijrei.com. All rights reserved

Keywords: MWCNTs, Solar Parabolic Trough Collector, Surfactant Triton X-100, Thermal performance, overall thermal efficiency

1. Introduction

Solar parabolic trough collector is a prominent way to generate the electric power for both commercial and industrial purpose among other concentrating solar power technologies. Conventional fluids are having a limited capacity to extract and carrying heat in the receiver of parabolic trough collector, therefore concept of using nano fluids instead of conventional fluids become more popular in solar thermal collector. In this present study MWCNT nanofluid at different concentration and water as conventional fluid were used to examine and compare the performance of solar parabolic trough collector. Tooraj Yousefi et al [1] was conducted an experimental study through MWCNT nanofluid with 0.2wt% and 0.4wt % at different mass flow rates 0.0167kg/sec to 0.05kg/sec and also studied the effect of surfactant Triton X-100 on the stability of nanofluid. Xiang-Qi Wang et al [2] studied the heat transfer and fluid flow characteristics of nano fluids in forced and free convection. He also studied that nano fluids has a capability to enhance the heat transfer characteristics and concluded that nano fluids are superior to conventional fluids in term of thermal physical properties. It has been seen from review study that suspended carbon nanotubes possess higher thermal conductivity and aspect ratio as compare to other existing nano particles. Nor Azwadi Che Sidik [3] et al concentrated his review study upon synthesis process for nano fluids. Both metallic and nonmetallic particles can be used to prepare nano fluids, Apart from this it has been noticed in this review work that metallic particles acquire higher thermal conductivity value as comparison to nano fluids containing oxide particles and nano fluids possess also less specific heat and higher viscosity as compare to conventional fluids. John Philip et al [4] studied in his review work that enhancement in thermal conductivity with incremental change in particle volume fraction and effect of temperature on thermal conductivity of nanofluid. Tooraj yousefi et al [5] conducted an experimental study with the use of Al_2O_3-H_2O nanofluid as working fluid in flat plate collector. Volume concentration of nano particles were taken as 0.2% and 0.4% with the application of surfactant Triton X-100 and nanofluid was flowing through collector tubes with different mass flow rates i.e. 1, 2 and 3Lit/min and it was concluded in this experimental work that 28.3% enhancement in efficiency of collector along with 15.63% efficiency enhancement with the

Corresponding author: R.S.Mishra
Email Address: rsmishra@dce.ac.in

application of surfactant, which is due to enhancement in heat transfer. Zhongyang Luo et al [6] was proposed a simulation model of nano fluid solar collector based upon direct absorption solar collection process by solving the radiative heat transfer for particulate media along with conductive and convective heat transfer in collector for evaluation of efficiency and to achieve the purpose of improvement in efficiency various nanoparticles like TiO_2, Al_2O_3, Ag, Cu, SiO_2, graphite nanoparticles, and carbon nanotubes were used in process of manufacturing nano fluid. Further this study concludes that thermal conductivity decrease with increase in temperature and improvement in outlet temperature around 30k-100k along with enhancement in efficiency 2-25% through nanofluid than base oil. Omid Mahian [7] et al conducted a theoretical to evaluate the performance of mini channel-based flat plate collector by using four different types of nanoparticles like Cu, Al_2O_3, TiO_2 and SiO_2 in water as base fluid. Mass flow rate in this work was assumed 0.1% and 0.5kg/s along with 4% volume fraction of nanoparticles. This study was concluded that heat transfer coefficient generally higher in all types of Nano fluids except SiO_2-H_2O nanofluid and higher efficiency with maximum entropy generation was measured through SiO_2-water based nanofluid, while lower efficiency and lowest entropy generation was measured through Cu/water based nanofluid. Ali Ijam [8] et al examine the stability of nanofluid with the help of thermal conductivity and UV-measurement technique and further thermos physical properties and electrical conductivity of graphite oxide nano sheets-deionized water and ethylene glycol were investigated at different temperatures and it was concluded in this research study that enhancement in thermal conductivity in a linear way with increasing concentrations and temperature and maximum enhancement in thermal conductivity was measured from 6.67% to 10.47% with volume fraction 0.10% and range of temperature varies between 25 to 45^0C. Further an enhancement in electrical conductivity was noticed linearly with increasing the temperature of nanofluid and nonlinear enhancement in electrical conductivity was measured with increasing the concentration of graphite oxide nano sheets. Tooraj Yousefi et al [9] also conducted an experimental study related to the investigation of pH value MWCNT-H_2O nanofluid on the performance of the flat plate collector and results outcomes from the experimentation concluded that enhancement in efficiency of the collector due to high difference between *pH* of nanofluid and pH of isoelectric point.

2. Experimental Procedure

2.1 Working fluids

MWCNT nanofluid and water were used as working fluid in solar parabolic trough collector. MWCNTs with 97% purity mixed with base fluid (Distilled water) used to manufacture nanofluid. After reviewing all past research work and going through lot of test run, sonication time was decided up to 45 minutes with proper stirring for certain 8-10 minutes. Application of Triton X-100 was used for the proper dispersion of nanoparticles in base fluid during experimental work. Water as working conventional fluid was also used for evaluation and compares the performance of collector.



(a)



(b)

Figure 1(a) & (b) showed sonicated MWCNT nanofluid at different 0.01wt% and 0.02wt%

2.2 Mathematical equations

Important mathematical equations were used to evaluate the performance and properties of nano fluids are discussed as under:

Specific heat or heat capacity of nanofluid

$$c_{nf} = \{(1 - \varphi_p)\, \rho_f\, c_f + \varphi_p\, \rho_{np}\, c_{np}\} / \rho_{nf} \quad (1)$$

Density of nanofluid

$$\rho_{nf} = (1 - \varphi_p)\,\rho_f + \varphi_p\,\rho_{np} \quad (2)$$

Instantaneous efficiency

$$\eta_i = Q_u / (G_T\, R_b\, W\, L) \quad (3)$$

Thermal efficiency

$$\eta_{th} = \frac{mCnf(Tout - Tin)}{A_{aper} G_T t} \quad (4)$$

Overall thermal efficiency

$$\eta_{ot} = \frac{mCnf(Tmax - Tmini)}{A_{aper} G_{avg} T} \quad (5)$$

2.3 Experimental set up

Parabolic trough collector has a receiver tube at focal point, which was covered by glass tube to maintain vacuum around absorber tube and it was coated black to absorb the maximum amount of solar radiations. Mirror strips were provided on parabolic shape of collector to reflect and concentrate the sun light on absorber tube. Inlet and outlet temperature measured through thermometer and storage tank of certain capacity around 10 Liters was also provided to store working fluid. Further proper aluminium foil and glass wool insulation was used to cover piping system and storage tank to reduce the heat losses. Pump was used for continuous forced flow of liquid through the receiver tube. Experiment was performed during the day time 9:30am to 3:00pm according to Indian standard time and during the conditions of continuous solar flux. Global solar flux and wind speed throughout the day measured with solar power meter and anemometer. Figure 2 showed experimental set up of solar parabolic collector, which is situated in Thapar University (Patiala).



Figure: 2: Experimental set up of parabolic through collector at Thapar University, Patiala

Table 1: Specification of parabolic trough collector (Thapar University)

Length of collector	1.2m
Breadth of collector	0.915m
Aperture area	1.0188m^2
Rim angle	90^o
Focal Length	0.30m
Receiver internal diameter	0.027m
Receiver outer diameter	0.028m
Glass cover internal diameter	0.064m
Glass cover outer diameter	0.066m

2.4 Experimental findings

This section contains inlet and outlet temperatures along with total solar intensity data related to performance of solar parabolic collector. Transient analysis was done on solar collector and it was assumed that piping and storage tank are fully insulated; it means that heat losses from these system components are negligible.

Table 2: MWCNT nanofluid with 0.01wt% at 160L/h on (24/3/2015)

Time Interval	T_{in} (℃)	T_{out} (℃)	Solar Intensity (W/m^2)	Wind Speed (m/s)
9:30-10	24.9	27.9	941.454	0.9
10-10:30	27.9	33.1	1008.504	1
10:30-11	33.1	39.5	1123.949	0.2
11-11:30	39.5	45.5	1085.224	0.9
11:30-12	45.5	50.8	974.809	1
12-12:30	50.8	54.8	960.654	0.95
12:30-1	54.8	58	956.929	0.8
1-1:30	58	59.6	942.029	0.9
1:30-2	59.6	60.5	893.604	1.2
2-2:30	60.5	60.9	887.644	2.2
2:30-3	60.9	61.2	874.239	2.3

Table 3: MWCNT nanofluid with 0.01wt% at 100L/h (25/3/2015)

Time interval	T_{in} (^{0}C)	T_{out} (^{0}C)	Solar Intensity (W/m^2)	Wind Speed (m/s)
9:30-10	26.6	29.9	690.964	0.4
10-10:30	29.9	35.6	725.234	0.6
10:30-11	35.6	41.9	759.504	0.8
11-11:30	41.9	47	834.749	0.7
11:30-12	47	49.9	860.079	0.55
12-12:30	49.9	52.9	854.864	0.7
12:30-1	52.9	55.4	881.684	1
1-1:30	55.4	57.5	883.919	1.15
1:30-2	57.5	58.6	868.274	1.45
2-2:30	58.6	59	865.294	3.55
2:30-3	59	59.2	712.569	1.3

Table 4: MWCNT nanofluid with 0.02wt% at 160L/h (27/3/2015)

Time interval	T_{in} (℃)	T_{out} (℃)	Solar Intensity (W/m^2)	Wind Speed (m/s)
9:30-10	26.2	29.2	755.779	0.2
10-10:30	29.2	33.3	881.109	0.8
10:30-11	33.3	38.7	971.884	0.8
11-11:30	38.7	44.9	1089.134	1.1
11:30-12	44.9	50.5	990.309	0.6
12-12:30	50.5	55.7	976.739	0.75
12:30-1	55.7	59.5	980.619	0.8
1-1:30	59.5	62.6	950.769	0.85
1:30-2	62.6	64	936.814	2.5
2-2:30	64	64.8	922.659	2.1
2:30-3	64.8	65.1	901.799	1.4

Table 5: MWCNT nanofluid with 0.02wt% at 100L/h (28/3/2015)

Time interval	T_{in} (℃)	T_{out} (℃)	Solar Intensity (W/m^2)	Wind Speed (m/s)
9:30-10	26.4	29.4	777.384	0.95
10-10:30	29.4	33.9	859.334	2.05
10:30-11	33.9	40.9	929.364	2.35
11-11:30	40.9	47.4	959.909	3.2
11:30-12	47.4	52.3	969.594	3.65
12-12:30	52.3	55.4	974.809	0.3
12:30-1	55.4	58.9	953.204	3.25
1-1:30	58.9	61	1003.119	2.1
1:30-2	61	62	1013.549	1.7
2-2:30	62	62.5	975.554	2
2:30-3	62.5	62.7	958.419	2.65

Table 6: Water at 160L/h (12/3/2015)

Time interval	T_{in} (℃)	T_{out} (℃)	Solar Intensity (W/m^2)	Wind Speed (m/s)
9:30-10	17.6	19.8	765.464	0.5
10-10:30	19.8	23.4	788.559	0.25
10:30-11	23.4	26.9	833.259	1.15
11-11:30	26.9	31.5	864.549	0.5
11:30-12	31.5	34.8	893.604	0.65
12-12:30	34.8	37.8	924.149	1.1
12:30-1	37.8	40.9	960.654	0.4
1-1:30	40.9	43.3	940.539	0.65
1:30-2	43.3	45.1	934.579	1.1
2-2:30	45.1	46	948.734	0.7
2:30-3	46	46.3	946.499	1.9

Table 7: Water at 100L/h (13/3/2015)

Time interval	T_{in} (℃)	T_{out} (℃)	Solar Intensity (W/m^2)	Wind Speed (m/s)
9:30-10	15.3	18	873.489	1.65
10-10:30	18	21.6	891.369	0.8
10:30-11	21.6	25.8	900.309	0.7
11-11:30	25.8	29.3	907.759	1.45
11:30-12	29.3	32.6	946.499	1.55
12-12:30	32.6	35.4	950.969	2.5
12:30-1	35.4	38	959.164	1.4
1-1:30	38	40.9	922.659	0.55
1:30-2	40.9	42.2	944.264	2.65
2-2:30	42.2	43	940.539	1.05
2:30-3	43	43.5	999.394	1.4

3. Results and Discussions

This section includes the comparative study between the results achieved from MWCNT nanofluid (0.01wt% & 0.02wt %) and water at dissimilar volume flow rate 160L/h and 100L/h and results of comparative study are discussed and shown graphically in this area.

(a) It has been noticed during experimental work that maximum instantaneous efficiency through MWCNT nano fluid at 0.01% with 160L/h comes out 96.49% in the time interval 10:30-11:00am and also maximum instantaneous efficiency through water 90.17% was measured at 160L/h during the time interval 11:00-11:30am shown graphically in figure 3. After that value of instantaneous efficiency starts to decrease continuously and it can be due to increasing instability and agglomeration of MWCNTs in base fluid and also can be due to decreasing solar intensity after around 12:00am and increasing wind speed. Thermophysical properties of nanofluid and water also affect the performance characteristics like thermal conductivity of MWCNT nanofluid is higher as comparison to water, while other properties for MWCNT nanofluid like specific heat and density posses' lower value than water.

(b) Figure 4 showed graphical comparison between instantaneous efficiency from MWCNT nanofluid (0.01wt %) at 100L/h and water. Maximum instantaneous efficiency comes out 87.85% through MWCNT nanofluid during the time interval 10:30-11:00am and for water maximum instantaneous efficiency was 49.40% was measured in same time interval. Suspension of nano particles in to base fluid is the better solution to increase the heat transfer properties or to improve the thermos physical properties of the base fluid.

(c) Increasing concentration of nano tubes in base fluid has an ability to increase the thermos physical properties of nano fluids, which is further very helpful to enhance the performance of system. In figure 5 the maximum instantaneous efficiency 96.46% was measured by MWCNT nano fluid (0.02%) at 160L/h showed in the time interval 11:00-11:30am and water showed 90.17% in the same time interval, which is less in amount as comparison to nano fluid. It was also seen that increasing concentration of carbon nanotubes also increases the viscosity of nanofluid. Increasing viscosity of nanofluid also has an ability to increase pressure loss in the system, which directly affects to the efficiency of the collector system. Graphical figure 6 showed Maximum instantaneous efficiency from solar parabolic trough collector was 79.77% measured during the time interval 10:30-11:00am through MWCNT. Nano fluid (0.02%) at 100L/h and also water showed maximum instantaneous efficiency was 49.40% measured at same volume flow rate and also in same time interval.

(d) Maximum thermal efficiency through MWCNT nano fluid (0.01wt %) at 160L/h was 7.79% during the time interval 10:30-11:00am, which is larger than maximum thermal efficiency was achieved from water i.e. 7.28% in the time interval 11:00-11:30am as shown figure 7. Further Maximum thermal efficiency possessed by nano fluid (0.01wt %) at 100L/h was 11.36% during the time interval 10:30-11:00am, which is greater than maximum thermal efficiency of water 6.39% was measured at 100L/h in the same time interval is shown graphically in figure 8.

(e) MWCNT (0.02%) nanofluid at 160L/h possess maximum value of thermal efficiency was 7.79% during the time interval 11:00-11:30am as comparison to maximum thermal efficiency achieved from water i.e. 7.28% was measured in the same time interval as shown in figure 9 graphically, while MWCNT nano fluid (0.02%) at 100L/h showed Maximum thermal efficiency was 10.31% at 100L/h during time interval 10:30-11:00am and water showed maximum thermal efficiency 6.39% was measured in the same time interval as shown in graphical figure 10.

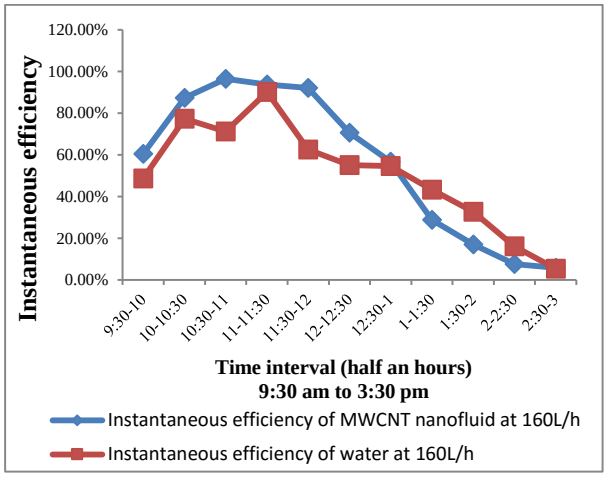


Figure. 3 variations in instantaneous efficiency with respect to time

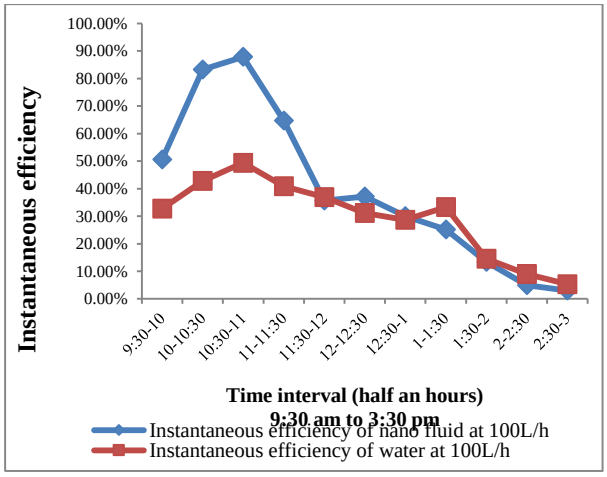


Figure. 4 variations in instantaneous efficiency with respect to time

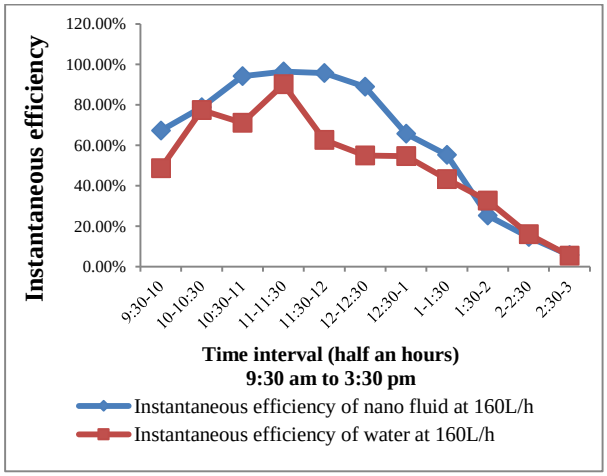


Figure: 5 variations in instantaneous efficiency with respect to time

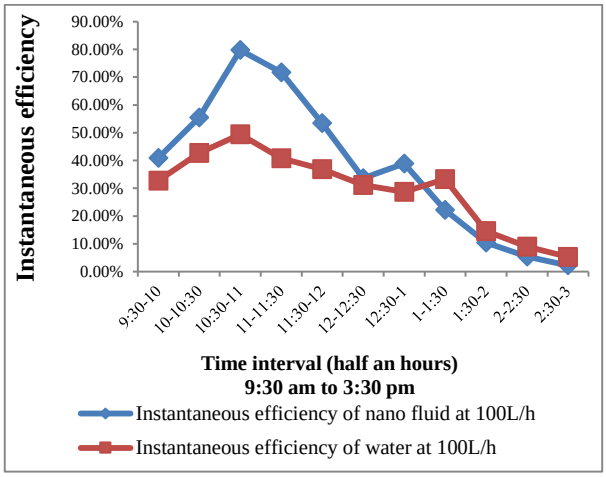


Figure: 6 variations in instantaneous efficiency with respect to time

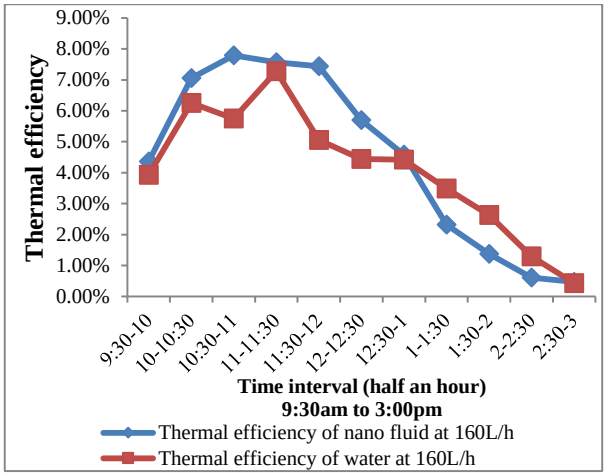


Figure. 7 variations in thermal efficiency with respect to time

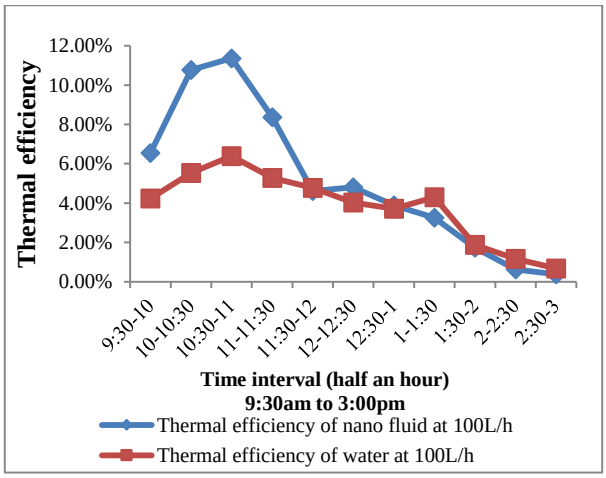


Figure. 8 variations in thermal efficiency with respect to time

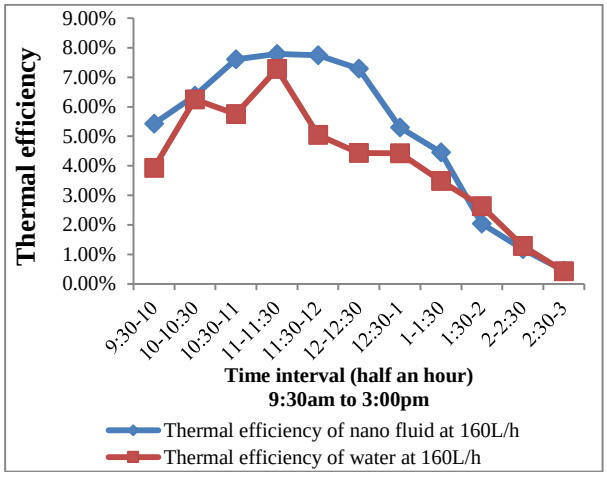


Figure. 9 variations in thermal efficiency with respect to time

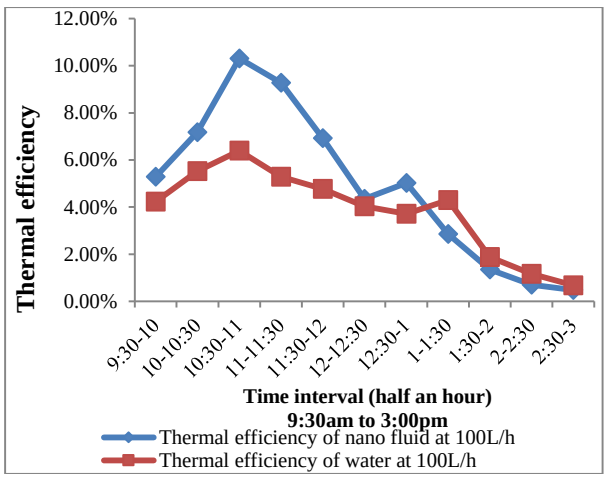


Figure. 10 variations in thermal efficiency with respect to time

4. Conclusion

In this research work comparative study has done between performance of solar collector due to MWCNT- Distilled H_2O nano fluid and simple water. Total amount of fluid was taken 6 litres and MWCNTs was used in 0.01wt% and 0.02wt% with water as reference fluid. Instantaneous efficiency and thermal efficiency showed better results than water up to certain time limit. Thermal conductivity is the only thermo physical property of nano fluid, which is responsible to enhance the efficiency of solar collector. MWCNT nano fluid 0.02wt% at 160L/h volume flow rate showed highest overall thermal efficiency, which can be due to higher heat transfer rate in nano fluid at high concentration of nanoparticles and high mass flow rate.

Nomenclature

$\dot{m}$ = mass flow rate (Kg/s)
C = Specific heat of water $\left[\frac{J}{kg-k}\right]$
C_{nf} = Specific heat of MWCNT nano fluid $\left[\frac{J}{kg-k}\right]$
G_T = Total solar intensity (W/m^2)
t =Time period (half an hour)
T_{out}= Outlet temperature (^0C)
T_{in}= Inlet temperature (k)
W = Width of collector (m)
ρ = Density of water $\left[\frac{kg}{m^3}\right]$
η_i = Instantaneous efficiency
η_{th} = Thermal efficiency
η_{ovt} = Overall thermal efficiency
T = Total experimental time period
Tmax= Maximum temperature (K)
Tmini = Minimum temperature (K)

Greek symbols

φ_p = Weight fraction of MWCNTs
ρ_{nf} = Density of nanofluid $\left[\frac{kg}{m^3}\right]$
ρ_{np} = Density of MWCNTs $\left[\frac{kg}{m^3}\right]$

Subscript

th = Thermal
ovt = Overall thermal
np = Nano particles
in = Inlet

Acknowledgement

One of Author (Er. Harwinder Singh) is thankful to Dr. Kundan Lal, Deptt. of Mechanical Engineering, Thapar University Patiala for guidance

References

[1] Tooraj Yousefi, Farzad Veisy, Ehsan Shojaeizadeh, Sirus Zinadini (2012), "An experimental investigation on the effect of MWCNT-H_2O nanofluid on the efficiency of flat-plate solar collectors," Experimental Thermal and Fluid Science, vol. 39: pp. 207–212.
[2] Xiang-Qi Wang, Arun S. Mujumdar (2007), Heat transfer characteristic of nano fluids: a review," International Journal of Thermal Science, vol. 46: pp. 1-19.
[3] Nor Azwadi Che Sidik, H.A. Mohammed Omer A. Alawi, S. Samion (2014), A review paper on preparation methods and challenges of nanoluids," International Communications in Heat and Mass Transfer, vol. 54: pp. 115–125.
[4] John Philip, P.D. Shima (2012), "Thermal properties of nanofluids," Advances in Colloid and Interface Science, vol. 183–184: pp. 30–45.
[5] Tooraj Yousefi, Farzad Veysi, Ehsan Shojaeizadeh, Sirus Zinadini (2012), "An experimental investigation on the effect of Al_2O_3-H_2O nanofluid on the efficiency of flat-plate solar collectors," Renewable Energy, vol. 39: pp. 293-298.
[6] Zhongyang Luo, Cheng Wang, Wei Wei, Gang Xiao, Mingjiang Ni (2014), "Performance improvement of a nanofluid solar collector based on direct absorption collection (DAC) concepts," International Journal of Heat and Mass Transfer, vol. 75: pp. 262-271.
[7] Omid Mahian, Ali Kianifar, Ahmet Z. Sahin, Somchai Wongwises (2014), "Performance analysis of a minichannel-based solar collector using different nanofluids," Energy Conversion and Management, vol. 88: pp. 129-138.
[8] Ali Ijam, R. Saidur, P. Ganesan, Moradi Golsheikh (2015), "Stability, thermo-physical properties, and electrical conductivity of graphene oxide-deionized water/ethylene glycol based nanofluid," International Journal of Heat and Mass Transfer, vol. 87: pp. 92–103.
[9] T. Yousefi, E. Shojaeizadeh, F. Veysi, S. Zinadini (2012), "An experimental investigation on the effect of pH variation of MWCNT–H_2O nanofluid on the efficiency of a flat-plate solar collector," Solar Energy, Vol. 86: pp. 771-779.
[10] S.P. Sukhatme, J.K. Nayak (2012), "Solar Energy principles of thermal collection and storage," Mc Graw Hill, ISBN 0-07-026064-8.

DEVELOPMENT OF PHONEME DOMINATED DATABASE FOR LIMITED DOMAIN T-T-S IN HINDI

Archana Balyan

Department of Electronics and Communication Engineering, Maharaja Surajmal Institute of Technology, Affiliated to GGSIPU, New Delhi, India

ABSTRACT

Maximum digital information is available to fewer people who can read or understand a particular language. The corpus is the basis for developing speech synthesis and recognition systems. In India, almost all speech research and development affiliations are developing their own speech corpora for Hindi language, which is the first language for more than 200 million people. The primary goal of this paper is to review the speech corpus created by various institutes and organizations so that the scientists and language technologists can recognize the crucial role of corpus development in the field of building ASR and TTS systems. This aim is to bring together all the information related to the recording, volume and quality of speech data in speech corpus to facilitate the work of researchers in the field of speech recognition and synthesis. This paper describes development of medium size database for Metro rail passenger information systems using HMM based technique in our organization for above application. Phoneme is chosen as basic speech unit of the database. The result shows that a medium size database consisting of 630 utterances with 12,614 words, 11572 tokens of phonemes covering 38 phonemes are generated in our database and it cover maximum possible phonetic context.

KEYWORDS

Speech database, Phonemes, Metro Rail Passenger Information System, Hidden Markov Model, Text- to-Speech

1. INTRODUCTION

The objective of speech data collection is to primarily build speech recognition system (ASR), Text – to- speech synthesis systems (TTS) and speech translation system for various Indian languages. As there is ever growing demand for customized and domain specific voices for use in corpus based synthesis systems, it is essential that standard methods should be established for creating these databases as it directly affects the quality and accuracy of the system.

In a country like India, with over 22 officially recognized languages, the language barriers can impose hurdles in the development of related fundamental Information communication technologies (ICT). In such multilingual environment, if speech technology is integrated with cross- language speech to speech translation systems, it could be of great help to remove this blockade so that services and information can be provided across languages more easily. This shall be of great cultural and economic value for a diverse country like India. This basic infrastructure will speed up the creation of large size spoken language corpora and subsequently lead to the development multilingual technologies such as multi-lingual speech translation, multi-lingual speech transcription, and multi-lingual information retrieval. In the field of development of ASR, TTS and speech translation systems, advances have been largely promoted by the Ministry of the Information and Communication Technologies (MCIT), Govt. of India (GoI) via selected set of Indian academic and research institutions in consortium mode. The resources

including speech and text corpora collected in these efforts abide by the copyright restrictions of the sponsor [1].

Hindi is the National language of India and is spoken by majority of population (about 41%) residing mostly in northern, central, western and eastern regions of the country [2]. Hindi is written in the Devanagari script which contains 13 vowels, 33 consonants and 3 consonant clusters with special symbols [3]. In section 2, the paper focuses on the Hindi resources being developed, which can be used for research in computational linguistics. In section 3, the methodology adopted for creating our own speech database (phoneme) is presented. In section 4, the challenges faced during database creations are discussed. In section 5, conclusions are presented.

2. DEVELOPMENT OF SPEECH CORPORA: A SURVEY

In modern linguistics, a corpus is referred to as a large collection of structured text in written or spoken form to be used mainly for linguistic research [4]. If imperative linguistic information such as tags, labels etc is added to the text corpus, it is called annotated corpus. Development of TTS requires formalized knowledge for speech generation on all linguistic levels (morphological structure, the spelling variations and word sense analysis), it can be the best resource to acquire various language phenomenon. A speech database for automatic speech recognition system for travel domain has been developed at C-DAC, Noida. The training data consists of recording by 30 female speakers in the age group of 17-60 years. The recording environment was noise-free and echo cancelled. The duration of recorded sentences is approximately 26 hrs and consists of 8,567 sentences comprising of 74,807 words [6].

At IIT, Kharagpur, Hindi corpora have been developed from Broadcasted news bulletin . The total duration of speech corpus is 3.5 Hrs. The recording was done in the studio in a Noise free environment for 19 speakers (6 Male and 13 Female) [5].

The IIIT Hyderabad has collected speech databases for usage in speech synthesis systems for multiple Indian languages. The Wikipedia text corpus for Hindi language consists of 44100 sentences, 1361878 words, 942079 syllables and 1466610 phones. The optimal text selection comprises of 1000 sentences, 8273 words, 19771 syllables and 30723 phones. The duration of the recorded speech corpus is 1.12 hours with average duration of sentence being 4.35 sec. The speech data was recorded in noise free studio environment using a laptop and a headset microphone connected to a zoom handy recorder [6].

At TIFR Mumbai and CDAC Noida , a general purpose database has been created for a set of 10 phonetically rich sentences. The signals were recorded directly in the digital format using two microphones in a noise free environment. Each sentence consists of two parts; the first part consists of two sentences which cover the maximum phonemes occurring in Hindi and the second part consists of eight sentences to cover the maximum possible phonetic context [8]. ILCI (Indian Languages Corpora Initiative) started by Technology Development for Indian Languages (TDIL) for building parallel corpora for major Indian languages including English. The Central Institute of Indian Languages (CIIL) also has been its resources towards resource building; the focus being on development of raw corpora for the language community rather than on the annotated corpora.

2.1. DEVELOPMENT OF CORPORA SPECIFIC TO MOBILE COMMUNICATION ENVIRONMENT

KIIT college of Engineering has collected a text corpus of 2 million words of raw messages in 12 different domains. A speech corpus of duration of approx. 4 hours has been created by digitally recording 630 phonetically rich sentences by 100 speakers, where each speaker spoke 630

sentences. The recording was done through 3 channels simultaneously: a mobile phone, a headset and a desktop mounted microphone. The annotated data sets are available for various applications such as development of language models, language features, and language translation etc in field of mobile communication. This project was undertaken in collaboration by Nokia Research Centre China [6].At KIIT Bhubaneswar, a total of 600 phonetically rich sentences, consisting of 42,801 unique words was created for Hindi language after collecting text messages in Hindi. The sentences were recorded by 100 speakers. The digital recording has been performed through 3 channels simultaneously. In the developed database, the ratio of female voice recording to male voice recording is 6:4 [9].

At the Linguistic Data Consortium for Indian Languages LDCIL [1], a database was collected in two states ; Uttar Pradesh and Bihar . In collecting the database, 650 different native speakers in were recorded distributed across three age groups (16-20, 21-50 , 51 and over). The recording of news text corpus was conducted in noisy environment at home, office and public places through recorder with an inbuilt microphone by each speaker. The speech signal files from all speakers were transcribed and labeled at the sentence level [10].

3. CREATING SPEECH DATABASE FOR METRO RAIL PASSENGER INFORMATION SYSTEM (MRPIS)

Delhi Metro Rail Corporation (DMRC) was established by the Government of India (GoI) and the Government of Delhi in March 1995 to build a new metro system in Delhi, capital city of India. For communication within metro, Intercoms are provided for emergency communication between the passengers and the driver in each coach, and on-train announcements are in Hindi and English. Under circumstances such as such as technical breakdown, accidents, change of route or any unforeseen incidents etc there is need for special announcements other than regular ones. It requires certain specific approach to handle the communication between the Metro Rail operators and passengers to address this kind of emergency situation. The use of TTS shall improve the quality of service to the passengers on board. Development of natural sounding corpus-based TTS system for Delhi Metro requires database with large number of units having adequate variations in prosody and spectral characteristics created exclusively for metro rail passenger information system.

The major steps taken to build medium size database for use in development of TTS for metro rail passenger information system (MRPIS) are as given below.

1. Selection of text sentences as appropriate to selected domain
2. Recording of selected text corpus
3. Automatic segmentation of speech corpus at phonetic level

3.1 TEXT CORPUS: SELECTION OF SENTENCES

The domain specific text corpora was so collected that these sentences not only covers the most frequent words that are announced while travelling in metro rail but also to have an optimal text corpus balanced in terms of phonetic coverage and have sufficient coverage of prosodic contexts of each type of sound to capture the acoustics of Hindi speech. Phonetic rich sentences are needed for robust estimation of the parameters in building statistical models of context Independent HMM (CIHMM) for phonemes. A sentence set is considered to be phonetically rich if it contains most, if not all, phonemes of the Hindi language.

[1] http://www.ldcil.org/resourcesSpeechCorpHindi.aspx

In order to create a medium size phonetically rich speech database, 630 sentences were collected. There are 38 distinct phonemes in the corpus. It is seen that each sentence contain minimum of 11 phonemes and 2% of the sentences contain all the 38 phonemes. The percentages of sentence sets containing at least 30 and 34 distinct phonemes are 24% and 10% respectively.

3.2 RECORDING OF SENTENCES AND DIGITIZATION

The quality of digital sound is determined by discrete parameters such as the sample rate, bit capacity and number of channels [11]. One can be conclusive that our digitization standards should be able to faithfully represent acoustic signals. The LDC-IL under CIIL has designed the standards for acquisition of the speech data according to application it is intended for and the recording devices that were used for recording the speech samples [12].

The corpus consists of 630 phonetically- balanced Hindi sentences spoken by single male speaker. The digitized recording was done at sampling rate of 16 kHz , stored in 16 bit PCM-encoded waveform format in mono mode. The duration of recording is about 1.5 hours. The speech utterances are manually transcribed into text using INSROT[2] , the Indian Script Roman Transliteration scheme for Hindi. The Table 1 below shows the statistical analysis of the recorded speech corpus.

Table 1. Statistics of speech Corpus

Sl. no	Process	Count
1	Total number of utterances recorded	630
2	Total number of words	12,614
3	Total number of unique phonemes	38
4	Total number of phonemes recorded	11572
5	Total number of nouns recorded	78

3.3 AUTOMATIC SEGMENTATION OF SENTENCES INTO PHONEMES

The basic and essential task in building speech database is the speech data segmentation and labelling. In this work, phoneme is chosen as the basic speech unit for segmentation. The accuracy of the segmentation and labelling of phoneme directly affects the quality of TTS. Manual segmentation is labour intensive, tedious, time consuming, error prone, inconsistent and requires much effort. In order to reduce manual efforts and speed up segmentation process, automatic phonetic segmentation is preferred [13].

3.3.1 PRE-PROCESSING AND FEATURE EXTRACTION PROCESS

The recorded speech waveform is digitized at a sampling rate of 16 KHz and is pre-emphasized (α=0.99). The speech signal is uniformly segmented at interval of 25 msec (overlapping factor = 0.5). Thus, each frame consists of 400 samples. Hamming window is then applied to each frame and then FFT is performed on each frame. For each frame, 12 mel-frequency cepstral coefficients (MFCC)[14] are extracted including 12 delta Δ and 12 Δ - Δ coefficients with cepstral mean normalization (CMN) . Hence, a 36 dimension feature vector is extracted for each frame that is used as one of the inputs for building Phone models.

[2] http://www.tdil.mit.gov.in

3.3.2 HMM BASED PHONETIC SEGMENTATION

This is the process of building a model for each phoneme in the spoken speech. The work presented in this paper is based on Hidden Markov Models [15] wherein forced alignment using Viterbi algorithm [16] is applied to find the most probable boundaries for the known sequence of phonemes.

3.3.3 BUILDING HMM MODELS FOR PHONEMES

The phone models of each phoneme involves creation of HMM file, finding and storing global mean and variance for each phoneme occurring in database. The block diagram for training with HMM is given in Figure 1 below.

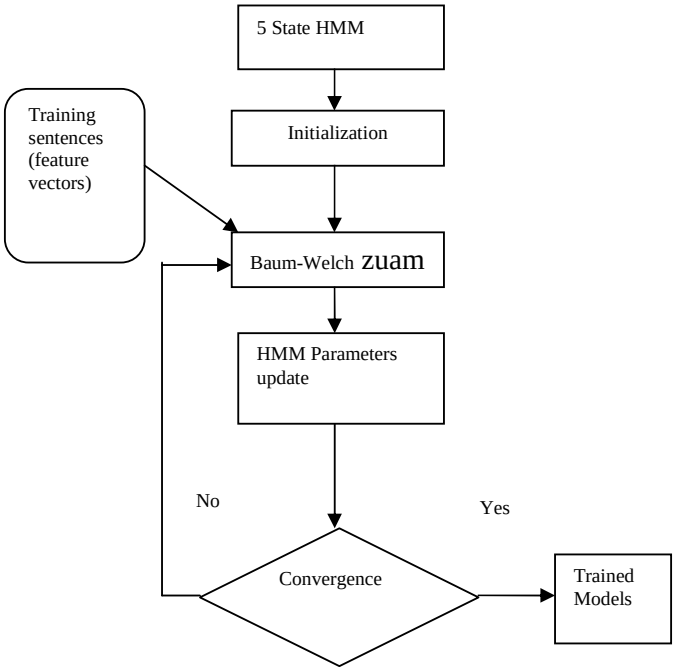


Figure 1: Block diagram for HMM Training

HMM dealing with continuous speech signals are characterized by a Gaussian pdf. In our experiment, Context Independent Phoneme HMM (CI Model) consists of M state (M=5) , for each phoneme unit and M state (M=3)for the non-speech region, i.e., the silence since silence is stationary and has no temporal structure to exploit. The phone models created so far are used to realign the training data (MFCC frames) and create new transcriptions. For alignment of frames with phones, the Viterbi re-estimation algorithm has been implemented , that starts with uniform segmentation (flat mode) and rapidly converges to the best model estimates in just few iterations and obtain optimized speaker-dependent HMM for each phoneme. Table 2 gives the parameters used for building HMM models for each phoneme.

Table 2. Parameters for Building Phone Model

Parameters	**Number**
Frame size	25msec
Feature vector dimension	36
No. of states to model each phoneme	5 for phonemes and 3 for silence
Topology	Left – to-right with no skip states
Number of Gaussian mixtures	05
No. of phone models	183
Convergence criterion	error less than 0.04
Training data	150-200 sentence
Test data	80-100 sentence
Platform	visual studio on Windows

3.4 DESIGN OF SPEECH DATABASE

The phonetic level segmentation for each recorded speech utterance is performed using HMM. In the database generated, each phoneme is tagged with its left and right context, information about start time and end time, duration of the phoneme and the location of the wave file to which it belongs. The speech database is prepared and stored in the format given below as shown in Figure 2.

{Prefix phoneme; phoneme; suffix phoneme}{Start time}{End time}{Duration of phoneme}{Wave file}

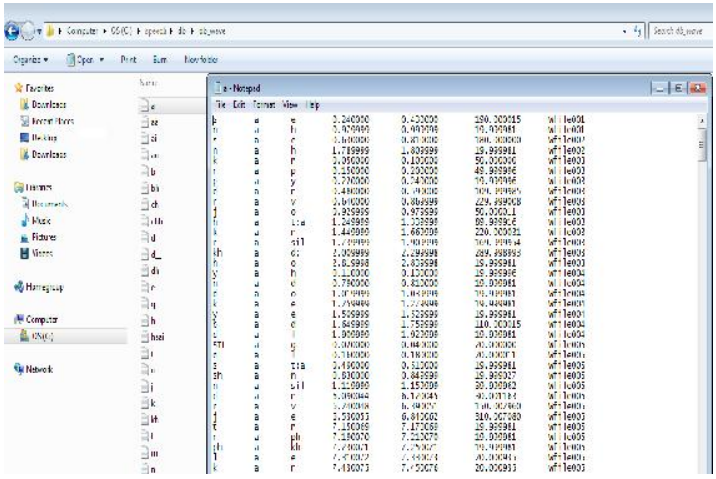


Figure 2: Screen shot of database generated for phoneme "a"

3.5 VOLUME OF DATABASE PREPARED

An analysis has been performed to calculate the volume of database prepared by taking a count of tokens generated for each phoneme occurring in the sentences selected to prepare the speech corpus. The distribution of tokens in the database is given in Table 3 given below.

Table 3. Number of tokens generated in our database

Phones	Tokens	% age	Phones	Tokens	%age
/ə/	752	6.498	t	362	3.1282
/ɑ/	1100	9.505	t^h	82	0.7086
/ɪ/	505	4.363	d	291	2.5146
/ii/	1285	11.104	d^h	97	0.83823
/U/	67	0.5789	n’	568	4.908
/u/	228	1.9702	p	301	2.6011
/ɛ/	112	0.9678	p^h	49	0.4234
/e/	123	1.0672	b	177	1.5295
/O/	331	2.806	b^h	78	0.6740
/ɔ/	56	0.4839	m	381	3.2924
/k/	756	6.5330	j	129	1.114
/k^h/	44	0.3801	r	880	7.604
/g/	217	1.8752	l	295	2.594
/g^h/	67	0.57898	v	344	2.972
/ng/	95	0.8209	s’a	208	1.797
/c/	95	0.8209	sa	87	0.7518
/c^h/	58	0.50121	s	446	3.8541
/j/	221	1.9097	h	348	3.0072
/j^h/	52	0.4493			
/n’a/	85	0.7345			

From design of MRPIS database, there are altogether 11572 phoneme units from 630 sentences. It can be seen that /ɑ/ and /ii/ are the two phonemes with highest occurrence in database. They contribute to almost 11% and 9.5 % of total phoneme unit extracted. The least occurring phonemes are /p^h/ and /$k^{h/}$ contributing to 0.4234% and 0.38% of the phonemes occurring in database.

4. CHALLENGES IN DATABASE PREPARATION

Important issues involved in database preparation for development of various speech technologies are (i) creating a generic acoustic database that covers language variations (ii) designing of the recording prompts and recording of speech databases to be used by corpus – based speech synthesizers and (iii) Lack of availability of phonetically balanced material. Also, the recording of the inventory requires large number of recording sessions with long durations under strict supervision to ensure good voice quality during recording sessions. . The recording process is expensive and puts a restriction on the capability to create databases in large number of voices for each restricted domain application.

5. CONCLUSIONS

Development of TTS systems for DMRC will be extremely useful for metro rail passenger information systems (MRPIS). Here, a technique based on HMM has been proposed for automatic segmentation and identification of phoneme like speech units. The phoneme database is created by using the state of art technique based on HMM. The design of the database is such that each speech unit (phoneme) is available in various phonetic contexts. A medium size database has been prepared that consists of 630 utterances with 12,614 words, 11572 tokens of phonemes covering 38 phonemes occurring in database. This database is believed to augment database for application in development of TTS for metro rail. In addition to above, an in- depth

survey of efforts made in database developments for Hindi language has been performed. It discusses some core linguistic resources of Hindi language, available through various resources developed for usage in text-to-speech synthesis and speech recognition technology.

It is suggested that the recordings from various resources can be grouped into application domains that can be combined to generate inventories which can be integrated with speech synthesizers to develop TTS. Unfortunately many of the existing corpora or resources lack features that are highly desirable for their use in the scientific context. These shortcomings include problems with availability (in some cases the use of very specific interfaces is required), high costs or strict licenses that allow reuse and data collection. This paper highlights the issue of distribution constraints which needs to be solved. As some of the problems such as copyright cannot be eliminated, it would be advantageous to have more resources available electronically that can be used with fewer limitations. This shall make involvement of extensive number of institutions and industries possible in research and development of TTS in Hindi language.

REFERENCES

[1] Prahallad, Kishore, et al. "The IIIT-H Indic Speech Databases." Interspeech, 2012
[2] https://en.wikipedia.org/wiki/List_of_languages_by_number_of_native_ speakers_in_India
[3] Kachru,Y. Hindi. John Benjamins Publishing company,Philadelphia, 2016.
[4] Dash, N S and B B Choudhary, “Why do we need to develop corpora for Indian languages”, International conferences on SCALLA , Bangalore, 2001,Vol. 11 in Elsevier.
[5] Sunita Arora, Babita Saxena, Karunesh Arora, S S Agarwal, “ Hindi ASR for Travel Domain,” In Proceedings of OCOCOSDA 2010, Kathmandu, Nepal.
[6] Agrawal, S. S. "Recent developments in speech corpora in Indian languages: Country Report of India." O-COCOSDA, Kathmandu (2010).
[7] Anandaswarup V, Karthika M et al., “Rapid Development of Speech to Speech Systems for Tourism and Emergency Services in Indian Languages”, In Proceedings of International Conference on Services in Emerging Markets, Hyderabad, India.
[8] Samudravijay K., P.V.S Rao and S.S. Agrawal, “Hindi Speech Data”, In proceedings of Sixth International Conference on Spoken Language Processing (ICSLP 2000),Beijing, China
[9] Shyam Agrawal, Shweta Sinha, Pooja Singh, Jesper Olsen, “ Development of Text and Speech Database for Hindi and Indian English specific to Mobile Communication Environment. In Proceeding of International Conference on The Language Resources and Evaluation Conference, LREC, Istanbul, Turkey, 2012.
[10] K.Srenivasa Rao,“ Application prosody model for Developing speech system”, International Journal of Speech Technology, 2011,Vol 11 in Elsevier
[11] Primer on Audio PC various issues associated with PC based audio technology, Online available on ttp://www.totalrecorder.com/primerpc.htm, 2016.
[12] Standards for Speech Data Capturing and Annotation, available online http://www.ldcil.org/download/samplingstandards.pdf, 2016.
[13] Balyan, Archana, S. S. Agrawal, and Amita Dev. "Automatic phonetic segmentation of Hindi speech using hidden Markov model" AI & society , Vol 27, no. 4 : 543-549,2000.
[14] Molau,S, Pitz, M, et al.,2001: Computing Mel-Frequency Cepstral Coefficients on the Power Spectrum, proceedings, ICASSP, IEEE.
[15] Rabiner, L.R., 1989: A Tutorial on hidden markov models and selected applications in speech recognition, Proceedings. IEEE, vol. 77, no. 2, 257-286.
[16] Forney, J.D, 1978: The Viterbi Algorithm, proceedings, IEEE, vol.no.3, 268-278.

AUTHOR



Archana Balyan has obtained her BE (ECE) degree from Bangalore University and received the ME (ECE) degree from Delhi College of Engineering, University of Delhi. She has obtained her PhD degree from GGSIP University in the area of speech synthesis. She has more than 20 years of teaching experience and is an Associate Professor in Maharaja Surajmal Institute of Technology (a premier institute affiliated to GGSIP University, Delhi). She has published several papers in reputed International Journals and in the proceedings of leading conferences. Her research interests include speech synthesis, computational linguistics and Analog electronics.

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

Device-Circuit Analysis of Ferroelectric FETs for Low-Power Logic

Shreya Gupta, Mark Steiner, Ahmedullah Aziz, *Student Member, IEEE*, Vijaykrishnan Narayanan, *Fellow, IEEE*, Suman Datta, *Fellow, IEEE*, and Sumeet Kumar Gupta, *Member, IEEE*

***Abstract*— Ferroelectric FETs (FEFETs) are emerging devices with an immense potential to replace conventional MOSFETs by virtue of their steep switching characteristics. The ferroelectric (FE) material in the gate stack of the FEFET exhibits negative capacitance resulting in voltage step-up action which entails sub-60 mV/decade subthreshold swing at room temperature. The thickness of the FE layer (T_{FE}) is an important design parameter, governing the device-circuit operation. This paper extensively analyzes the impact of T_{FE} on the device-circuit characteristics in conjunction with the interactions between FE and gate/drain capacitances. While it is well known that increasing T_{FE} yields higher gain albeit with the possibilities of introducing hysteresis, our analysis points to other unconventional effects emerging in circuits as T_{FE} is increased. Depending on the attributes of the underlying transistor, increasing T_{FE} beyond a certain value may lead to loss in saturation and/or negative differential resistance in the output characteristics. While the former effect results in the loss in gain of a logic gate, the latter may yield hysteretic voltage transfer characteristics. We also discuss the effect of T_{FE} on the circuit energy–delay. Our analysis shows that for high T_{FE}, the delay of the circuit may increase with an increase in supply voltage. However, for voltages <0.25 V, FEFINFETs show an immense promise yielding 25% lower energy at iso-delay.**

***Index Terms*— Ferroelectric FETs (FEFETs), hysteresis, negative capacitance, negative differential resistance (NDR).**

I. Introduction

AN EVER-INCREASING growth of mobile computing devices, wireless systems, and biomedical implantable devices has led to a surge in the demand for low power circuits [1]. Supply voltage (V_{DD}) reduction and technology scaling have been the critical driving factors to improve the energy efficiency. However, V_{DD} scaling is limited to near-threshold (V_{TH}) regions since reducing V_{DD} much below V_{TH} leads to a significant loss in speed and energy efficiency. This attribute is directly associated with the fundamental limit of 60 mV/decade on the subthreshold swing (SS) at room temperature for the MOSFETs. Surpassing this limit could enable more aggressive V_{DD} scaling leading to ultralow power dissipation. Therefore, devices that exhibit steep switching characteristics (i.e., SS $<$ 60 mV/decade at room temperature) are being actively explored.

Several steep slope devices have emerged in the recent past owing to their promise to deliver higher ON-state current at low voltages compared to the standard MOSFETs [3]–[17]. Amongst such devices, ferroelectric FETs (FEFETs) [7]–[17] are one of the leading exploratory devices. FEFETs utilize the negative capacitance of the ferroelectric (FE) material integrated in the gate stack to achieve steep switching [7]–[10] and higher ratio of ON and OFF currents (I_{ON}/I_{OFF}). However, because of the presence of FE in the gate stack, the gain in the ON current is accompanied by an increase in the gate capacitance. Thus, the overall benefits of FEFETs at the circuit level are not clear and therefore, need to be evaluated. Furthermore, the interactions of the highly nonlinear FE capacitance with other capacitances present in the devices and circuits lead to unconventional device characteristics, which, in turn, yield atypical circuit behavior. Such attributes need to be properly analyzed in order to harness the useful properties and avert the deleterious features.

In this paper, we perform a detailed analysis of FEFETs with a focus on the impact of the thickness of FE (T_{FE}) on the device-circuit performance. We present the relevant concepts from the perspective of digital logic. Analysis for memories and analog circuits requires a separate treatment; nevertheless, some of the trends discussed in this paper may be useful generally as well. T_{FE} is an important design parameter, directly governing the interactions of the FE with other device components and therefore, needs to be properly optimized. The analysis in this paper not only provides useful insights into circuit-driven optimization of T_{FE}, but also highlights some unconventional device-circuit features that arise due to the unique FE properties. This work makes the following contributions.

1) We discuss the effects of FE coupling with drain on the device-circuit characteristics, which becomes especially important at high T_{FE}. We discuss how such effect may

Manuscript received March 23, 2017; revised May 21, 2017; accepted June 9, 2017. This work was supported in part by the Semiconductor Research Corporation-Global Research Collaboration and in part by DARPA Young Faculty Award Program. The review of this paper was arranged by Editor W. Tsai. *(Corresponding author: Shreya Gupta.)*

S. Gupta, M. Steiner, A. Aziz, V. Narayanan, and S. K. Gupta are with the School of Electrical Engineering and Computer Science, The Pennsylvania State University, University Park, PA 16801 USA (e-mail: sfg5201@psu.edu; markrsteiner@gmail.com; afa5191@psu.edu; vijay@cse.psu.edu; skg157@engr.psu.edu).

S. Datta is with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556 USA (e-mail: sdatta@nd.edu).

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/TED.2017.2717929

0018-9383 © 2017 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.

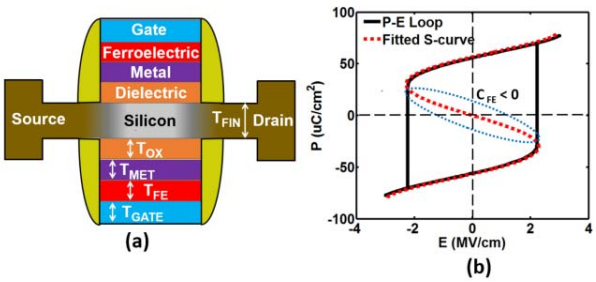


Fig. 1. (a) Schematic of the cross section of FE transistor. (b) P–E loop of FE capacitor showing the occurrence of negative capacitance in the FE material. In general, the intermediate metal layer is optional.

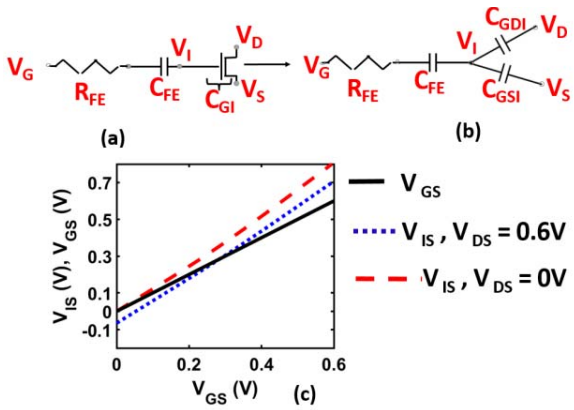


Fig. 2. (a) Equivalent circuit representation of FEFET. (b) Equivalent capacitance representation of FEFET with C_{GI} separated into C_{GSI} and C_{GDI}. (c) Voltage amplification in FEFET at $V_{DS} = 0$ and 0.6 V.

yield hysteretic voltage transfer characteristics (VTCs) of logic gates, even with non-hysteretic device operation.

2) We discuss the limits of increasing T_{FE} with regard to obtaining proper saturation in the output characteristics, which, in turn, determine the gain of the logic gates.
3) We report that at higher T_{FE}, the delay of FEFET-based circuits may increase with an increase in V_{DD} and describe the conditions for the same considering the effects of kinetic coefficient of the FE (ρ) and the wire capacitance (C_W).
4) We analyze the effect of T_{FE} and FE-drain coupling on the energy–delay characteristics of a ring oscillator (RO).

II. Basics of Ferroelectric FETs

FEFETs comprise of an FE layer in the gate stack, as shown in Fig. 1(a). The distinctive properties of the FEFETs are obtained by virtue of the interactions of the negative capacitance of the FE layer with the positive capacitance of the underlying transistor. As suggested in the first work on FEFETs [7], the negative capacitance arises because of a unique relation between the electric field (E) and polarization (P) of the FE [Fig. 1(b)]. This relation can be modeled with time-dependent Landau–Khalatnikov (LK) equation [11] given as follows:

$$E = \alpha P + \beta P^3 + \gamma P^5 + \rho dP/dt. \tag{1}$$

Here, α, β, and γ are static coefficients and ρ is the kinetic coefficient related to the time constant associated with the change of FE polarization. FE materials exhibit $\alpha < 0$, which results in hysteretic P–E characteristics shown in Fig. 1(b) [7], [8]. There is a region in this plot where the slope is negative which implies a negative FE capacitance (C_{FE}) [9], [10]. Steep switching FEFETs (to be employed as logic transistors), are designed to operate in this region [12], [13].

The capacitive interactions in FEFETs can be understood from a simple equivalent circuit representation of the device, as shown in Fig. 2(a) and (b). Here, V_G, V_D, and V_S are the applied voltages at the gate, drain, and source of the FEFET, respectively; and V_I is the voltage at the interface of FE and the dielectric of the underlying transistor (subsequently referred to as the gate of the underlying transistor). C_{GI} represents the positive gate capacitance associated with the underlying transistor, which can be further decomposed into gate to source capacitance (C_{GSI}) and gate to drain capacitance (C_{GDI}), as illustrated in Fig. 2(b). R_{FE} ($=\rho T_{FE}/A_{FE}$) models the term "$\rho dP/dt$" in (1), as discussed in [14]. Here, A_{FE} is the area of the FE layer cross section and it is same as the cross-sectional area of the underlying device. From this circuit, the following expression for the steady-state value of V_{IS} ($V_I - V_S$) is obtained when the FE operates in the negative capacitance region

$$dV_{IS} = \frac{dV_{GS}}{1 - \frac{C_{GI}}{|C_{FE}|}} - \frac{dV_{DS}}{\frac{|C_{FE}|-C_{GI}}{C_{GDI}}} = \eta dV_{GS} - \eta_D dV_{DS} \tag{2}$$

$$V_{IS} = \int_0^{V_{IS}} dV_{IS} = \int_0^{V_{GS}} \eta dV_{GS} - \int_0^{V_{DS}} \eta_D dV_{DS}. \tag{2a}$$

Here

$$\eta = \frac{1}{1 - \frac{C_{GI}}{|C_{FE}|}} = \frac{1}{1 - \frac{C_{GI}/A_{FE}}{|C_{FE}|/A_{FE}}} = \text{gain due to FE} > 1 \tag{2b}$$

$$\eta_D = \frac{1}{\frac{|C_{FE}|-C_{GI}}{C_{GDI}}} = \frac{C_{GDI}}{|C_{FE}| - C_{GI}} = \text{drain coupling factor.} \tag{2c}$$

It can be easily seen that

$$\frac{\eta}{\eta_D} = \frac{|C_{FE}|}{C_{GDI}}. \tag{3}$$

From the above equations, the following points can be deduced.

1) The change in V_{IS} (dV_{IS}) is dependent on the change in V_{GS}(dV_{GS}) and the change in V_{DS} (dV_{DS}). The value of V_{IS} is obtained by the integration of (2) and is depicted in (2a).
2) V_{IS} that determines the gain in I_{ON}/I_{OFF} is dependent on both V_{GS} and V_{DS} [Fig. 2(c)]. The dependence on the latter stems from the coupling between the gate and drain of the underlying transistor, represented by C_{GDI}. Thus, at $V_{GS} = 0$ and $V_{DS} > 0$, $V_{IS} < 0$ is obtained due the second term in (2a). We will discuss the implication of this effect in detail later.

3) As $|C_{FE}|$ reduces (e.g., with an increase in T_{FE}), the amplification (η) provided by the negative capacitance increases [3], which yields lower SS and higher I_{ON}/I_{OFF}. Concurrently, η_D also increases, which yields unconventional characteristics, as discussed later.
4) The total gate capacitance (C_{GFE}) of the FEFET is equal to $|C_{FE}|C_{GI}/(|C_{FE}| - C_{GI})$ or $(C_{GI} * \eta)$. Hence, FEFETs exhibit higher gate capacitance compared to the standard transistors. It is also easy to deduce that as $|C_{FE}|$ reduces, C_{GFE} increases.

It is also noteworthy that as T_{FE} is increased beyond a certain value, the FEFET device characteristics may show hysteresis [14] and even nonvolatility [15]. Since our work focuses on logic applications, we limit our discussions to the region of non-hysteretic device operation only.

With the understanding of the operation of FEFETs, we discuss the effect of T_{FE} on the device-circuit characteristics in the rest of the paper. Before describing the analysis, we present the simulation details in the next section.

III. Modeling and Simulation Framework

The analysis in this paper employs our in-house SPICE model for FEFETs [14] that is based on the time-dependent LK equation (1) solved self-consistently with the transistor equations. As a part of this self-consistent solution, we obtain the device currents, voltage at the internal gate (V_I), and FE polarization as a function of the applied gate/drain/source voltages. It is noteworthy that the FEFET model just has four assigned electrodes (gate, source, drain, and body) as in standard transistor models. Voltage at the internal gate (V_I) is obtained as a part of the device-circuit simulations. However, V_I can be monitored during device-circuit simulations, by virtue of the SPICE-based implementation of the model which enables the analysis of V_{IS} as a function of V_{GS} and V_{DS}, as discussed subsequently. We employ the high performance 10 nm FINFET predictive technology model (PTM) [18] for the standard/underlying transistor. It is worthwhile to clarify that the effects of diffusion barriers and fin tuck have not been taken into account in this work. However, this assumption does not impact the trends that we describe in the following sections. If we consider the additional capacitances due to fin tuck and other structural features of the FINFETs, the FE thickness at which the subsequently discussed unconventional effects occur will change, but the trends will remain the same. It may be noted that the objective of this paper is not to characterize FEFETs at the 10 nm node but to highlight generic unconventional aspects of the behavior of FEFETs in digital logic, for which we take the 10 nm PTMs as examples. To analyze the output characteristics of FEFETs more closely, we also employ 22 nm MOSFET model [18]. The 10 nm and 22 nm PTMs are referred as FINFETs/FEFINFETs and MOSFETs/FEFETs, respectively, in the rest of the paper. The list of parameters and their corresponding nominal values used for our analysis are given in Table I. The values of α, β, and γ are extracted from our experiments on hafnium zirconium oxide. The value of ρ is estimated as mentioned in [9]. We also analyze the trends for different values of ρ later. Note that we have assumed α, β, γ, and ρ to be independent of T_{FE} (as in other works [7]–17]) and the trends with respect to T_{FE} presented later need to be understood in the context of this assumption. The analysis of unconventional characteristics of FEFETs still remains valid (albeit those effects will occur at a different T_{FE} value if the dependence of the coefficients on T_{FE} is taken into account). It may also be mentioned that the calibrated value of γ (for which the best fit with experiments was achieved) is negative, which leads to non-physical FE characteristics for large $|P|$. However, as long as, we restrict the usage of model in a certain polarization range (in this case between -80 μC/cm^2 and 80 μC/cm^2), the model predicts the expected characteristics of FE (as shown in Fig. 1) and FEFETs. For the analysis in this paper, the polarization remains within -20 μC/cm^2 to 20 μC/cm^2 (i.e., in negative capacitance region) and therefore, the results do not have any non-physical trends even with a negative value of γ. We have verified that the trends we present in this paper remain the same for positive values of γ as well. The DC and time-dependent characteristics are obtained based on transient simulations [14]. Fig. 3 depicts an example of our simulation methodology. We begin our simulations with all the voltages at zero. We then ramp up the supply voltage from 0 to V_{DD} and once all the voltages in a circuit stabilize, we apply the input signals. This approach enables us to capture the effect of η and η_D on V_{IS} (2a) in a self-consistent fashion. Employing this

TABLE I
Simulation Parameters for FE

α	-1.17e9 m/F
β	4.9e9 m^5/F/C^2
γ	-3.9e9 m^9/F/C^4
ρ	0.05 Ω m

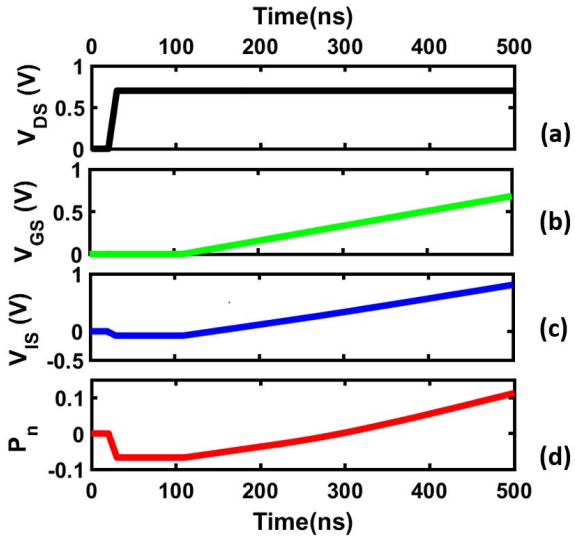


Fig. 3. Illustration of the simulation methodology for V_{GS} sweep of FEFINFETs at $V_{DS} = 0.6$ V. (a) Applied V_{DS} being ramped up from 0 to 0.6 V. (b) Applied input gate to source voltage V_{GS} being swept from 0 to 0.6 V. (c) Obtained voltage at the internal gate node of the FEFINFET. (d) Normalized polarization of the FE capacitor. Here, P_n is the normalized polarization defined as P/P_c, where, P_c is the polarization at which $dE/dP = 0$ in (1) at steady state.

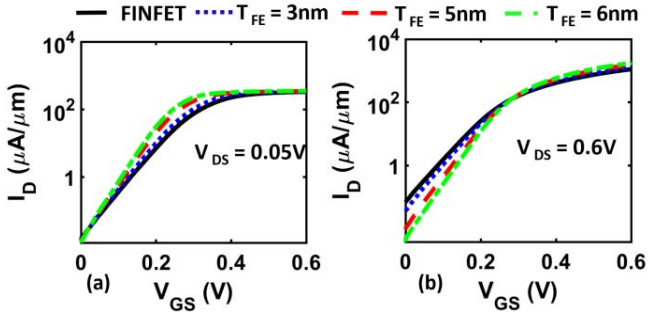


Fig. 4. I_D–V_{GS} characteristics of FEFINFETs for different values of T_{FE} showing steeper switching characteristics at (a) V_{DS} = 0.05 V and (b) V_{DS} = 0.6 V. FEFINFETs show reduction in OFF current for high V_{DS} values.

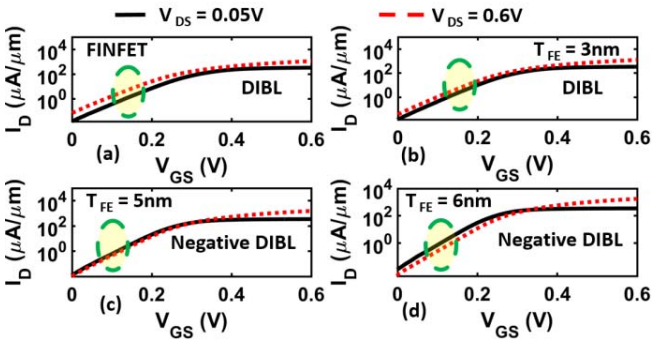


Fig. 5. DIBL in (a) FINFETs and (b) FEFINFETs at T_{FE} = 3 nm and Negative DIBL in FEFINFETs at (c)T_{FE} = 5 nm and (d) T_{FE} = 6 nm.

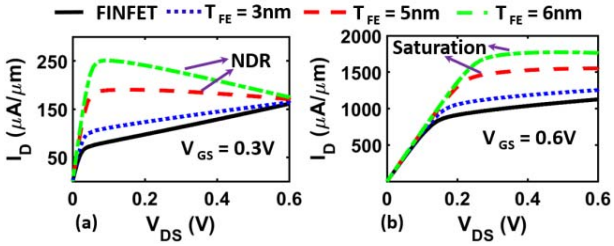


Fig. 6. I_D–V_{DS} characteristics of FEFINFETs for different values of T_{FE} for (a) V_{GS} = 0.3 V showing NDR and (b) V_{GS} = 0.6 V showing saturation in the ON current at T_{FE} = 5 nm and 6 nm.

simulation framework, we perform the device-circuit analysis of FEFETs in the following sections. We complement the self-consistent simulation-based quantitative analysis with the explanations employing the equivalent circuit in Fig. 2 and (2a) and (3) to provide qualitative insights

IV. Device Analysis

To explain the behavior of FEFET-based logic circuits, we discuss the relevant device characteristics and their dependence on T_{FE} in this section. Let us start with the transfer characteristics, shown in Fig. 4. It is evident that as T_{FE} increases, switching characteristics become steeper due to the increase in the voltage amplification provided by the FE layer (as discussed before and shown in previous works [3]–[8]). Another important observation is that FEFINFETs show reduction in I_{OFF} compared to FINFETs at high V_{DS}. This behavior is the result of the dependence of V_{IS} on V_{DS} [via C_{GDI} and η_D—see (2a) and (2c)]. As shown in Figs. 2(c) and 3, an increase in the drain voltage leads to a decrease in the polarization, which yields negative V_{IS} even though V_{GS} is 0. Furthermore, as T_{FE} increases, η_D increases (see Section II) yielding larger reduction in I_{OFF}. At low V_{DS} (=0.05 V), the OFF current is similar to that in FINFETs as the second term in (2a) becomes negligible. As a consequence of the FE-drain coupling, current gain due to steep switching is observed for a partial range of V_{GS}, as shown in Fig. 5(b). Moreover, increasing V_{DS} has two opposing effects on the subthreshold current due to: 1) larger electric fields which tend to increase the current and 2) negative impact on V_{IS}, which tends to reduce the current. These opposite effects lead to reduction in the drain-induced barrier lowering (DIBL), as can be seen in Fig. 5(a) and (b). In fact, at large T_{FE}, OFF current may reduce with increase in change in V_{DS}, due to the dominance of the second effect. This yields negative DIBL [Fig. 5(c) and (d)].

This dependence of current on V_{DS} is also manifested in the output characteristics (Figs. 6 and 7). However, the overall trends also depend on V_{GS}, T_{FE}, and the characteristics of the underlying transistor. To discuss this in detail, we consider different cases as outlined in the following.

1) At low values of T_{FE}, $|C_{FE}|$ is large compared to C_{GDI} and hence, $\eta >> \eta_D$ (3). Therefore, the positive effect of $\int_0^{V_{GS}} \eta dV_{GS}$ on V_{IS} dominates over the negative effect of $\int_0^{V_{DS}} \eta_D dV_{DS}$ (2a). Moreover, the voltage step-up is still sufficiently small such that saturation voltage (V_{DSAT}) < V_{DD}. Consequently, the behavior of I_D at lower T_{FE} values is similar to the conventional transistor, albeit with a larger I_{ON} (due to voltage step-up action) and lower output conductance (due to two opposing effects of V_{DS} on current, as explained before) (Figs. 6 and 7).
2) At higher T_{FE}, $|C_{FE}|$ decreases [14] and may become comparable to C_{GDI}. Consequently, η and η_D also become comparable (3). Therefore, the relative values of V_{GS} and V_{DS} control V_{IS}.
   2.1 For low V_{GS}, $\int_0^{V_{DS}} \eta_D dV_{DS}$ dominates over $\int_0^{V_{GS}} \eta dV_{GS}$ and therefore, increasing V_{DS} leads to reduction in V_{IS}. As a result, the current decreases with positive change in V_{DS} and the output characteristics exhibit negative differential resistance (NDR—Figs. 6 and 7) [17].
   2.2 In contrast, for high V_{GS}, $\int_0^{V_{GS}} \eta dV_{GS}$ becomes the dominant factor and offsets the effect of $\int_0^{V_{DS}} \eta_D dV_{DS}$. For this reason, NDR is not observed at high V_{GS}. (Figs. 6 and 7).

Also, V_{IS} at higher T_{FE} becomes large enough so as to significantly impact V_{DSAT}. The overall impact on the device characteristics is determined by the behavior of the underlying transistor with respect to the dependence of V_{DSAT} on the gate voltage (V_{IS}). In general, V_{DSAT} can be written as $(V_{IS}\text{-}V_{TH})^a$, where V_{TH} is the threshold voltage of the underlying transistor

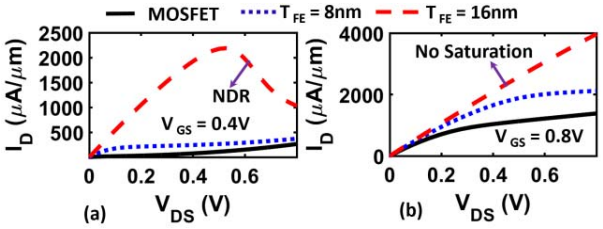


Fig. 7. I_D–V_{DS} characteristics of FEFETs for different values of T_{FE} for (a) $V_{GS} = 0.4$ V showing NDR and (b) $V_{GS} = 0.8$ V showing no saturation in the ON current at $T_{FE} = 16$ nm.

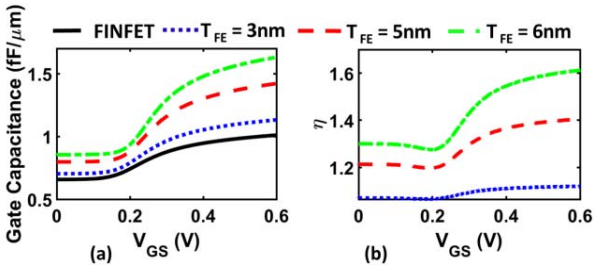


Fig. 8. (a) C_G–V_{GS} characteristics of FEFINFETs. (b) Variation of η with respect to V_{GS} in FEFINFETs for different values of T_{FE} at $V_{DS} = 0.6$ V.

and a is the exponent dependent on the mechanism of the current saturation in transistors [20].

It is well known that $1 < a < 2$, with the two extreme values corresponding to the extreme short channel FETs and long channel FETs, respectively. If V_{DSAT} is a weak function of V_{IS} (i.e., a is small, as in case of FINFETs considered in this work), increase in V_{IS} due to the negative capacitance of FE mildly affects V_{DSAT}, yielding $V_{DSAT} < V_{DD}$. Thus, the output characteristics exhibit saturation (Fig. 6). On the other hand, for larger a, (as in case of MOSFETs), we observe that under the influence of large V_{IS}, V_{DSAT} exceeds V_{DD}, thus, I_D fails to saturate (Fig. 7).

It may be noted that larger FE thickness is required for FEFETs (with a larger gate length) compared to FEFINFETs to achieve similar voltage step-up. This is because the gate capacitance per unit area (C_{GI}/A_{FE}) of the underlying FINFETs is more than that of MOSFETs (as per the transistor scaling trends). Hence, as per (2b), for the same T_{FE} or C_{FE}/A_{FE} value, the gain η for FEFETs will be lower than that of FEFINFETs. In other words, to get a similar gain, larger T_{FE} is required for FEFETs.

Besides the drain current, the delay and energy characteristics of a circuit also depend on the gate capacitance of the device. As discussed in Section II, the total gate capacitance increases by factor η due to the negative FE capacitance [Fig. 8(a)]. This factor η shows a non-monotonic trend with respect to V_{GS} [Fig. 8(b)]. As seen from (2b), η increases with increasing C_{GI}, but decreases with increasing $|C_{FE}|$. Now, the behavior of $|C_{FE}|$ with respect to V_{GS} can be understood by considering the following equation [derived from (1) for $C_{FE} < 0$]:

$$|C_{FE}| = \left|\frac{dQ}{dV}\right| = \frac{A_{FE}}{|\alpha| T_{FE} - (3\beta T_{FE} P^2 + 5\gamma T_{FE} P^4)}. \quad (4)$$

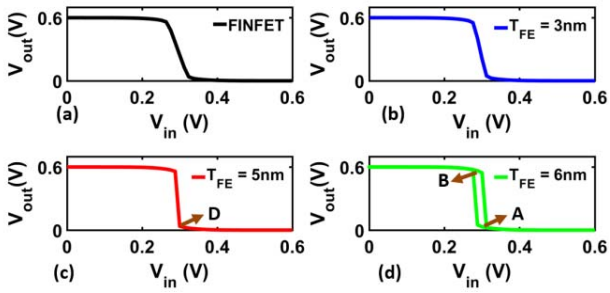


Fig. 9. Voltage transfer characteristics of inverter for (a) FINFET and FEFINFETs at (b) $T_{FE} = 3$ nm, (c) $T_{FE} = 5$ nm, and (d) $T_{FE} = 6$ nm showing hysteresis.

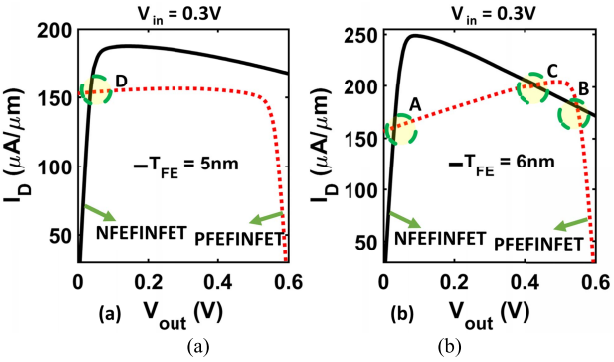


Fig. 10. Load line analysis of FEFINFET inverter for (a) $T_{FE} = 5$ nm with one operating point D and (b)$T_{FE} = 6$ nm with two operating points A and B.

As V_{GS} increases, charge and FE polarization (P) increase. From (4) and the parameters in Table I, one can deduce that increase in P leads to an increase in $|C_{FE}|$. The dependence of C_{GI} on V_{GS} can be observed from the solid line (corresponding to FINFETs) in Fig. 8(a). At low V_{GS}, increase in C_{GI} with V_{GS} is negligible. However, at high values of V_{GS}, the increase in C_{GI} is significant. Considering all the aforementioned factors, the following can be concluded: 1) at low V_{GS}, the ratio $C_{GI}/|C_{FE}|$ decreases with increasing V_{GS} (due to increase in $|C_{FE}|$ and negligible change in C_{GI}) and 2) at high V_{GS}, the ratio $C_{GI}/|C_{FE}|$ increases with increasing V_{GS} (because of a large increase in C_{GI}). From (3), one can observe that η increases with increasing $C_{GI}/|C_{FE}|$, which explains the non-monotonic trend observed in Fig. 8. An important point to be noted is that the trends of η with respect to V_{GS} are dependent on the signs of β and γ. The above discussion was for $\beta > 0$ and $\gamma < 0$ (see Table I). If an FE material has $\beta > 0$ and $\gamma > 0$, similar trends are observed as $|C_{FE}|$ increases with V_{GS} and η shows a non-monotonic trend, as discussed above. However, if $\beta < 0$ and $\gamma > 0$, then, depending on the values of β and γ, the term in the parenthesis of the denominator of the right-hand side of (4) becomes negative for small $|P|$, leading to decreasing $|C_{FE}|$ with increasing V_{GS}. As a result, η increases monotonically with increasing V_{GS}.

With the understanding of the device characteristics and their dependence on T_{FE} and gate/drain voltages, we now present their implications at the circuit level.

V. Circuit Analysis

In order to understand the behavior of FEFET-based circuits, we perform detailed DC and transient analysis. For DC analysis, our focus is on understanding the impact of the unconventional effects observed in FEFINFETs/FEFETs (such as NDR and non-saturation) on the VTCs of FEFINFET/FEFET-based inverter. The trends are also applicable for the stability of static random access memories. The objective of the transient analysis is to explain the energy–delay characteristics of FEFINFET-based circuits, for which we analyze a seven-stage RO.

A. DC Analysis

Fig. 9 compares the VTC of an FINFET-based inverter with the FEFINFET-based inverter at different T_{FE}. Due to the voltage amplification provided by the FE layer, the switching characteristics of the FEFINFET inverter become sharper i.e., the gain in the transition region increases. Moreover, as T_{FE} increases, the gain increases which enhances the noise margins. A new aspect that we observe in Fig. 9(d) is that at $T_{FE} = 6$ nm, VTC is hysteretic, even though the corresponding devices exhibit non-hysteretic behavior. The hysteretic circuit response is attributed to NDR in the output characteristics (Section IV). To substantiate this claim, we perform the load line analysis of the FEFINFET inverter for input voltage (V_{in}) $= V_{DD}/2 = 0.3$ V at $T_{FE} = 6$ nm [Fig. 10(b)] and $T_{FE} = 5$ nm [Fig. 10(a)]. From these plots, we observe the following.

1) At $T_{FE} = 6$ nm the load line plot shows three points of intersection due to NDR. When V_{in} is reduced from V_{DD} to 0, (i.e., the output voltage (V_{out}) increases from 0 to V_{DD}), the operating point is A. However, when V_{in} increases from 0 to V_{DD} (or V_{out} decreases from V_{DD} to 0), the operating point is B. [Fig. 10(b)]. This yields hysteretic VTC, as shown in Fig. 9(d). This hysteresis leads to higher noise margins and hence better noise immunity, similar to a Schmitt trigger [19].
2) However, performing a similar analysis at $T_{FE} = 5$ nm, we obtain only one point of intersection [D in Fig. 10(a)] despite the occurrence of NDR in the output characteristics [Fig. 7(a)]. Therefore, no hysteresis is observed in the VTC.

Thus, NDR may lead to hysteresis but it is not a sufficient condition for the same.

Let us now briefly discuss the effect of non-saturation in the output characteristics of FEFETs (observed at $T_{FE} > 16$ nm) on the inverter VTC. From Fig. 11, we observe that non-saturation leads to the loss of gain in the transition region and therefore, FEFET-based inverter fails to work as a binary logic gate. Thus, optimization of T_{FE} must be performed to ensure that the transistor shows saturation in the output characteristics at $V_{GS} \sim V_{DD}/2$ and $V_{DS} \sim V_{DD}/2$ to obtain regenerative action in the logic gates.

Before we conclude this section, let us summarize the differences between FEFETs and FEFINFETs. As discussed before in Section IV, we observe that the benefits obtained from FEFETs are at higher T_{FE} values as compared to FEFINFETs due to the differences in the capacitance per unit area of the underlying FETs. Furthermore, no saturation is observed in the output characteristics of FEFETs for large T_{FE} and high V_{GS} values whereas FEFINFETs exhibit saturation at the entire range of T_{FE} considered due to lower sensitivity of V_{DSAT} on V_{IS} in FEFINFETs. However, at low V_{GS} and high T_{FE}, NDR is observed in both the devices. As a result of these device characteristics, an inverter based on FEFINFET may show hysteresis whereas an FEFET inverter may fail to show the gain in the VTCs at large T_{FE}.

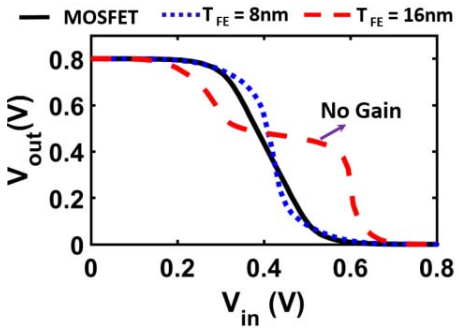


Fig. 11. VTCs of FEFET inverter for different values of T_{FE}. No gain in the VTC is observed at T_{FE}=16 nm.

B. Transient Analysis

In this section, we analyze a seven-stage inverter-based ROs to discuss the transient characteristics of FEFINFETs. Fig. 12(a)–(c) compare the delay and energy for FEFINFET- and FINFET-based ROs. It can be observed that FEFINFETs show reduction in delay, leakage energy, and total energy for $V_{DD} < 0.25$ V. However, for super-threshold voltages, all the three metrics are degraded with respect to FINFETs. The energy–delay plots [Fig. 12(d)] also show a crossover point. In addition, we report an atypical behavior of FEFINFET-based circuits in this paper. As can be observed in Fig. 12(a), the plot for $T_{FE} = 5$ nm shows an increase in delay with increasing V_{DD} in the super-threshold region.

To explain all these trends, including the atypical behavior, we consider an equivalent circuit of two contiguous FEFINFET-based inverters in an RO, as illustrated in Fig. 13. (Note the results shown in Figs. 12–16 have been obtained using proper simulations. The simple equivalent circuit is considered only to explain the trends.) The gate stack of the FEFINFETs are modeled as resistance R_{FE} in series with the capacitance $|C_{FE}|C_{GI}/(|C_{FE}|\text{-}C_{GI}) = \eta C_{GI}$ (see Section II). $R_{FE}(= \rho \cdot T_{FE}/A_{FE})$ models the $\rho dP/dt$ term in (1), as explained in Section II. In Fig. 13, we consider $C_{GI} = C_{GPI}$ for p-FEFINFETs and $C_{GI} = C_{GNI}$ for n-FEFINFETs. The drive strength of the transistors is modeled as R_{PI}/η_R and R_{NI}/η_R for p- and n-FEFINFETs, respectively. Here, R_{PI} and R_{NI} are the resistances of the underlying transistors (i.e., without considering the FE); and η_R is the factor by which the current increases in FEFINFETs in general due to the negative FE capacitance. Note, η_R depends on 1) η and η_D which determine V_{IS} [as per (2a)] and 2) the region of operation of the FET. Specifically, for $V_{IS} < V_{TH}$, $\eta_R > \eta$ since the current

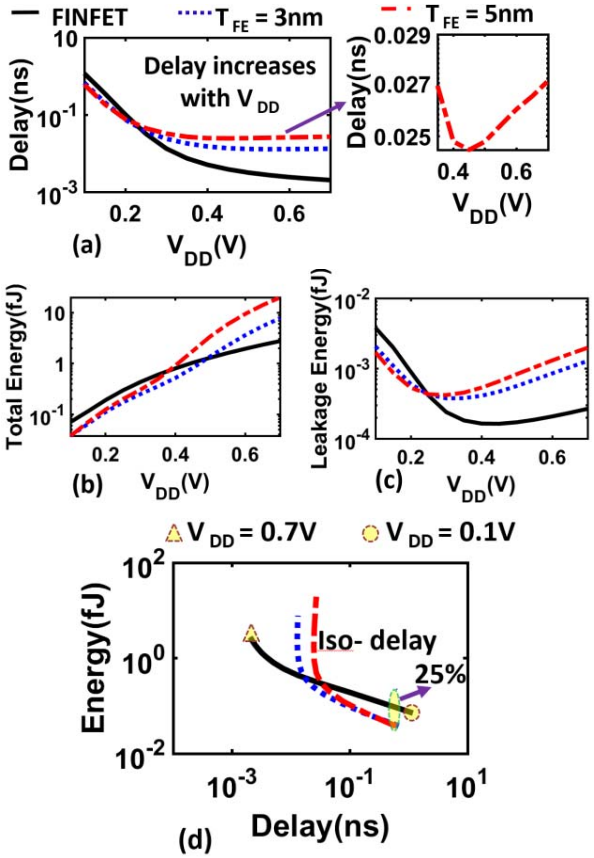


Fig. 12. (a) Delay–V_{DD} plots illustrating the increase in delay w.r.t. to V_{DD} for higher values of V_{DD}. (b) Total energy–V_{DD} plots. (c) Leakage energy–V_{DD} plots. (d) Energy–delay plots showing reduction in energy at iso-delay of an inverter stage in the seven-stage RO circuits for different values of T_{FE} at $C_W = 0$ fF, $\rho = 0.05\ \Omega \cdot$ m and FO $= 4$.

is exponentially dependent on voltage step-up provided by the negative capacitance. On the other hand, for $V_{IS} > V_{TH}$, η_R is comparable to η since the current is only subquadratically dependent on (V_{IS}-V_{TH}). Moreover, the effect of η_D further reduces η_R, which has a more pronounced effect at high voltages

We also consider the effects of wire capacitance C_W and fan out FO. In this paper, $C_W = 0$, signifies negligible wire capacitance. Our simulations not only treat the effect of wire capacitance as an additional load due to interconnects but also account for the self-consistent interaction of FE capacitance, C_{GSI}, C_{GDI}, and C_W. Both these effects impact the transient evolution of P and V_I during the circuit operation.

Now, employing Elmore's delay model on the circuit in Fig. 13, the time constants associated with the low-to-high and high-to-low transitions (τ_{LH} and τ_{HL}) are obtained as

$$\tau_{LH} = \text{FO}\cdot\frac{\eta}{\eta_R}R_{PI}(C_{GPI}+C_{GNI}) + \eta R_{FE}(C_{GPI}+C_{GNI}) + \frac{1}{\eta_R}R_{PI}C_W \quad (5a)$$

$$\tau_{HL} = \text{FO}\cdot\frac{\eta}{\eta_R}R_{NI}(C_{GPI}+C_{GNI}) + \eta R_{FE}(C_{GPI}+C_{GNI}) + \frac{1}{\eta_R}R_{NI}C_W. \quad (5b)$$

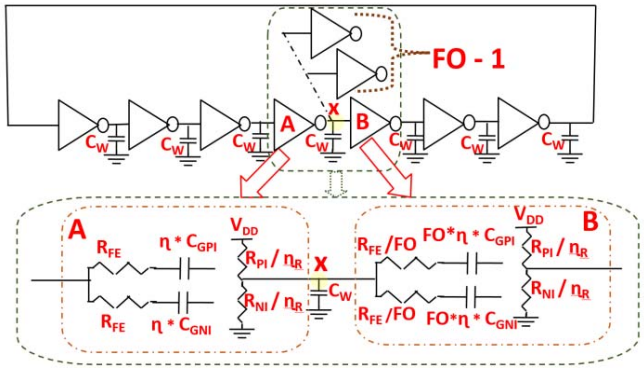


Fig. 13. Equivalent RC circuit of two adjacent FEFINFET-based inverters in a RO.

Also, since in case of FINFETs, $R_{FE} = 0$, $\eta = \eta_R = 1$ and therefore, (5) changes to

$$\tau_{LH} = \text{FO}\cdot R_{PI}(C_{GPI}+C_{GNI}) + R_{PI}C_W \quad (6a)$$

$$\tau_{HL} = \text{FO}\cdot R_{NI}(C_{GPI}+C_{GNI}) + R_{NI}C_W. \quad (6b)$$

From the equations above, we make the observations outlined below. For the time being, the wire capacitance is neglected in order to simplify the discussion. We will discuss its impact subsequently.

1) At high V_{DD}, $R_{NI}/\eta_R < R_{FE}$ and $R_{PI}/\eta_R < R_{FE}$ i.e., the inherent speed of the driving transistor is greater than the time constants associated with polarization transients. Thus, the second term in (5) dominates. As a result, the benefit of FEFINFETs in terms of higher ON current is offset by the extra resistance introduced by the FE layer, leading to lower circuit speed compared to FINFETs. As T_{FE} increases, $R_{FE}(=\rho T_{FE}/A_{FE})$ increases, yielding larger delay.
2) As V_{DD} is scaled to sub-/near-threshold values, R_{NI}/η_R and R_{PI}/η_R begin to dominate R_{FE}, as a result of which the first term in (5) starts to play a key role. Now, in the subthreshold region, $\eta_R/\eta > 1$, as discuss before. This yields superior performance in FEFINFET-based circuits. Our results show similar η_R and hence, similar delay for $T_{FE} = 3$ nm and 5 nm. This is because increasing T_{FE} not only boosts $\int_0^{V_{GS}} \eta dV_{GS}$ term in (2a) but also increases $\int_0^{V_{DS}} \eta_D dV_{DS}$, yielding similar V_{IS}. This leads to a mild dependence of ON current on T_{FE}.
3) To understand the atypical behavior of increasing delay with increasing V_{DD} in the super-threshold region, we need to examine the V_{DD} dependence of different terms (especially the dominant second term) in (5). Since R_{FE} dominates R_{NI} and R_{PI}, the reduction in delay that one expects typically with increasing V_{DD} (due to reduction in the inherent delay of the transistor) is nullified. On the contrary, η increases as V_{GS} (or V_{DD}) increases at higher values, as discussed before [see Fig. 8(b)]. This results in an increase in delay with V_{DD}. Note, this effect is observed when R_{FE} is large (and therefore, more dominant compared to the transistor's inherent resistance). This explains why the atypical behavior is observed at larger T_{FE}.

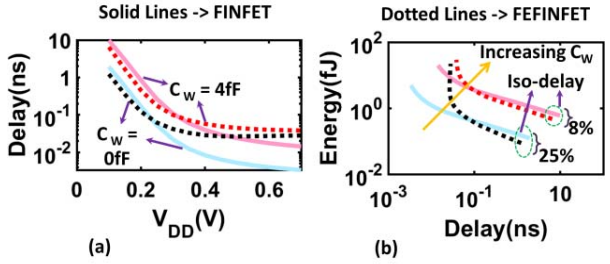


Fig. 14. (a) Delay–V_{DD} plots and (b) Energy–delay plots showing reduction in energy at iso-delay of an inverter stage in the seven-stage RO circuits for $C_W = 0$ and 4 fF at $\rho = 0.05\Omega m$, $T_{FE} = 5$ nm and FO = 4.

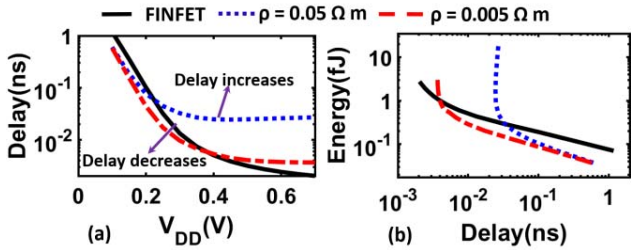


Fig. 15. (a) Delay–V_{DD} plots showing increase in delay w.r.t. to V_{DD} for higher value of ρ. (b) Energy–delay plots of an inverter stage in the seven-stage RO circuits for different values of ρ at $C_W = 0$ fF, $T_{FE} = 5$ nm and FO = 4.

Next, we discuss the comparison of energy consumption of FEFINFETs and FINFETs. Leakage energy (V_{DD}*I_{OFF}*Delay) primarily follows the trends of the delay and therefore shows an increase in FEFINFETs compared to FINFETs at high V_{DD}, but a decrease at low V_{DD}. The dynamic energy is higher in FEFINFETs due to increase in the gate capacitance. The total energy is higher in FEFINFETs at high V_{DD} due to the dominance of the dynamic energy. As V_{DD} is lowered, leakage energy begins to increase its contribution, thereby reducing the total energy in FEFINFETs as compared to FINFETs. As a result of all these factors, FEFINFETs show ~25% lower energy at iso-delay at $V_{DD} < 0.25$ V [Fig. 12(d)].

Now, we consider the effect of C_W. It is easy to deduce that the inclusion of the third term in (5) mitigates the loss in performance of FEFINFETs compared to FINFETs at high V_{DD}. This is because FEFINFETs have a better strength (lower resistance) to drive C_W. However, at the same time, the dynamic energy increases with increasing C_W, enhancing its contribution to the total energy. This reduction in the ratio of leakage to dynamic energy dwindles the energy benefits of FEFINFETs at low V_{DD} as C_W increases. As an example, [Fig. 14], at $C_W = 4$ fF, FEFINFETs exhibit 8% reduction in energy at iso-delay in the subthreshold region, as opposed to 25% energy reduction for $C_W = 0$. Another effect of increasing C_W is that the atypical behavior of FEFINFETs at high V_{DD} is not observed for sufficiently large C_W. This is because the third term in (5) starts to become comparable to the second term. This reinstates the decrease in delay due to reduction in R_{NI} and R_{PI} with increasing V_{DD}.

To analyze these trends further, we compare the energy–delay trends at different ρ (Fig. 15). As expected, lower ρ enhances the benefits of FEFINFETs to a larger range of V_{DD} due to the reduction in R_{FE}. Therefore, material optimization and device design for non-hysteretic steep switching operation of FEFINFETs should target the minimization of polarization lag in ferroelectrics with respect to the applied electric field. Moreover, FEFINFETs no longer exhibit anomalous behavior with respect to V_{DD} scaling at lower ρ. This is due to the fact that the second term in (5) does not remain as dominant anymore.

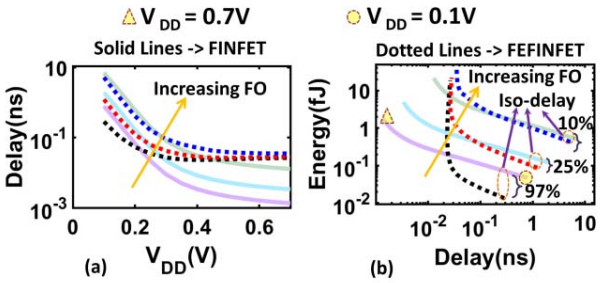


Fig. 16. (a) Delay-V_{DD} plots and (b) Energy–delay plots of an inverter stage in the seven-stage RO circuits for FO = 1, 4, and 16 at $C_W = 0$ fF, $T_{FE} = 5$ nm and $\rho = 0.05\ \Omega$ m.

Additionally, increasing the FO decreases the benefits of FEFINFETs as shown in Fig. 16. This is because as FO increases, the contribution of the dynamic energy component in the total energy increases. Since FEFINFETs exhibit larger dynamic energy compared to the standard FETs (due to higher gate capacitance), the energy benefits of FEFINFET (observed at lower V_{DD}) drops. At FO = 16, FEFINFETs exhibit 10% reduction in energy at iso-delay in the subthreshold region, as opposed to 25% energy reduction for FO = 4 and 97% energy reduction at FO = 1.

VI. Conclusion

In this paper, we extensively explored the effects of FE thickness and FE-drain coupling on the device-circuit characteristics of FEFETs in general. Our analysis showed the possibility of NDR in the output characteristics and the occurrence of negative DIBL, especially at high T_{FE}. This behavior was attributed to the coupling of the drain capacitance with the FE capacitance. Negative NDR may yield hysteretic VTCs of the logic gates yielding higher noise margins. We also discussed that if the underlying transistor shows a strong dependence of the saturation drain voltage on the gate voltage, the corresponding FEFET may fail to exhibit saturation in the output characteristics, leading to the loss in gain of the VTCs. Our analysis also shows that at high V_{DD} and high T_{FE}, FE transistors have degraded performance and may even show an increase in delay with increasing V_{DD}. However, FEFINFETs show a significant promise for low V_{DD} operation, yielding 8%–25% higher energy efficiency at iso-delay as compared to standard transistors.

References

[1] J. Rabaey, *Low Power Design Essentials*. Springer, 2009.
[2] N. S. Kim *et al.*, "Leakage current: Moore's law meets static power," *Computer*, vol. 36, no. 12, pp. 68–75, Dec. 2003, doi: 10.1109/MC.2003.1250885.

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

[3] Q. Zhang, W. Zhao, and A. Seabaugh, "Low-subthreshold-swing tunnel transistors," *IEEE Electron Device Lett.*, vol. 27, no. 4, pp. 297–300, Apr. 2006.
[4] W. Y. Choi, B.-G. Park, J. D. Lee, and T.-J. K. Liu, "Tunneling field-effect transistors (TFETs) with subthreshold swing (SS) less than 60 mV/dec," *IEEE Electron Device Lett.*, vol. 28, no. 8, pp. 743–745, Aug. 2007.
[5] H. Liu, S. Datta, and V. Narayanan, "Steep switching tunnel FET: A promise to extend the energy efficient roadmap for post-CMOS digital and analog/RF applications," in *Proc. Int. Symp. Low Power Electron. Des. (ISLPED)*, Beijing, China, Sep. 2013, pp. 145–150.
[6] S. Strangio *et al.*, "Impact of TFET unidirectionality and ambipolarity on the performance of 6T SRAM cells," *IEEE J. Electron Devices Soc.*, vol. 3, no. 3, pp. 223–232, May 2015, doi: 10.1109/JEDS.2015.2392793.
[7] S. Salahuddin and S. Datta, "Use of negative capacitance to provide voltage amplification for low power nanoscale devices," *Nano Lett.*, vol. 8, no. 2, pp. 405–410, 2007.
[8] A. I. Khan, C. W. Yeung, C. Hu, and S. Salahuddin, "Ferroelectric negative capacitance MOSFET: Capacitance tuning & antiferroelectric operation," in *IEDM Tech. Dig.*, Dec. 2011, pp. 11.3.1–11.3.4.
[9] A. I. Khan *et al.*, "Negative capacitance in a ferroelectric capacitor," *Nature Mater.*, vol. 14, no. 2, pp. 182–186, 2015.
[10] A. I. Khan *et al.*, "Experimental evidence of ferroelectric negative capacitance in nanoscale heterostructures," *Appl. Phys. Lett.*, vol. 99, no. 11, p. 113501, 2011.
[11] T. K. Song, "Landau-Khalatnikov simulations for ferroelectric switching in ferroelectric random access memory application," *J.-Korean Phys. Soc.*, vol. 46, no. 1, pp. 5–9, 2005.
[12] S. George *et al.*, "Device circuit co design of FEFET based logic for low voltage processors," in *Proc. IEEE Comput. Soc. Annu. Symp. VLSI (ISVLSI)*, Jul. 2016, pp. 649–654.
[13] S. George *et al.*, "NCFET based logic for energy harvesting systems," SRC TECHCON, 2015, pp. 1–4.
[14] A. Aziz, S. Ghosh, S. Datta, and S. K. Gupta, "Physics-based circuit-compatible SPICE model for ferroelectric transistors," *IEEE Electron Device Lett.*, vol. 37, no. 6, pp. 805–808, Jun. 2016, doi: 10.1109/LED.2016.2558149.
[15] S. George *et al.*, "Nonvolatile memory design based on ferroelectric FETs," in *Proc. Des. Auto. Conf. (DAC)*, Jun. 2016, pp. 1–6.
[16] G. Pahwa, T. Dutta, A. Agarwal, and Y. S. Chauhan, "Designing energy efficient and hysteresis free negative capacitance FinFET with negative DIBL and 3.5X ION using compact modeling approach," in *Proc. 42nd Eur. Solid-State Circuits Conf. (ESSCIRC)*, Lausanne, Switzerland, Sep. 2016, pp. 49–54.
[17] G. Pahwa *et al.*, "Analysis and compact modeling of negative capacitance transistor with high ON-current and negative output differential resistance—Part II: Model validation," *IEEE Trans. Electron Devices*, vol. 63, no. 12, pp. 4986–4992, Dec. 2016.
[18] (2016). *Arizona State University Predictive Technology Models*. [Online]. Available: http://ptm.asu.edu/
[19] L. A. P. Melek, A. L. da Silva, M. C. Schneider, and C. Galup-Montoro, "Analysis and design of the classical CMOS schmitt trigger in subthreshold operation," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 64, no. 4, pp. 869–878, Apr. 2017, doi: 10.1109/TCSI.2016.2631726.
[20] T. Sakurai and A. R. Newton, "Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas," *IEEE J. Solid-State Circuits*, vol. 25, no. 2, pp. 584–594, Apr. 1990, doi: 10.1109/4.52187.



Shreya Gupta received the B.Tech. degree in electrical and electronics engineering from Delhi Technological University, New Delhi, India, in 2015. She is currently pursuing the M.S. degree in electrical engineering with The Pennsylvania State University, University Park, PA, USA.

She joined GlobalFoundries, Burlington, VT, USA, as an Intern in 2016 and Intel Corporation, Folsom, CA, USA, as an Intern in 2017. Her current research interests include low-power VLSI circuit design and emerging device technologies.



Mark Steiner received the B.S. degree in electrical engineering from The Pennsylvania State University, University Park, PA, USA, in 2016.

In 2015, he joined the Integrated Circuits and Devices Laboratory under Dr. S. K. Gupta. Since 2016, he has been an Applications Engineer of power semiconductor modules with Mitsubishi Electric, Youngwood, PA, USA. His current research interests include semiconductor devices, digital logic systems, power electronic modules, and power electronic devices.



Ahmedullah Aziz (S'10) received the B.Sc. degree from the Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, in 2013, and the M.S. degree from The Pennsylvania State University, University Park, PA, USA, in 2016, where he is currently pursuing the Ph.D. degree with the Department of Electrical Engineering.

He is currently a Graduate Researcher with the Integrated Circuits and Devices Laboratory, The Pennsylvania State University.



Vijaykrishnan Narayanan (F'11) received the B.S. degree in computer science and engineering from the University of Madras, Chennai, India, in 1993, and the Ph.D. degree in computer science and engineering from the University of South Florida, Tampa, FL, USA, in 1998.

He is currently a Professor of Computer Science and Engineering and Electrical Engineering with The Pennsylvania State University, University Park, PA, USA.



Suman Datta (F'13) was with the Advanced Transistor Group, Intel Corporation from 1999 till 2007. He was a Professor at The Penn State University in Electrical Engineering, University Park, PA, USA, from 2007 to 2011. He is currently the Chang Family Chair Professor of Engineering Innovation with the University of Notre Dame, Notre Dame, IA, USA.



Sumeet Kumar Gupta (M'12) received the B.Tech. degree in electrical engineering from the IIT Delhi, New Delhi, India, in 2006, and the M.S. and Ph.D. degrees in electrical and computer engineering from Purdue University, West Lafayette, IN, USA, in 2008 and 2012, respectively.

He was a Senior Engineer at Qualcomm Inc., San Diego, CA, USA. He is currently a Monkowski Assistant Professor of Electrical Engineering with The Pennsylvania State University, University Park, PA, USA.

Available online at www.ijapie.org

INTERNATIONAL JOURNAL OF ADVANCED PRODUCTION AND INDUSTRIAL ENGINEERING

IJAPIE-2017-02-218, Vol 2 (2), 54-57

IJAPIE
Connecting Science & Technology with Management.
A Journal for all Products & Processes.

Effect of acidic aggressive environment on strength of mortars in early stage

Shobha Ram[1], Abhishek Singh[2], Alok Verma[3]
[1, 2](Department of Civil Engineering, Gautam Buddha University, Uttar Pradesh, India)
Email: abhishek180694@gmail.com
[3](Department of Civil Engineering, Faculty, Delhi Technological University, Delhi, India)
Email: alokverma_dce@hotmail.com

Abstract : This paper presents the result of experimental investigation on strength of mortars in early stage in acidic aggressive environment. Cement mortar is a building compound made by blending sand and cement with a predefined measure of water. Superplasticizer is also added in cement mortars to study its effect on compressive strength. Mortar specimens were prepared in different proportions with and without superplasticizer and their effect on compressive strength is noted in acidic aggressive environment. Superplasticizer is a admixture which reduced the water-cement ratio or increase the workability at same water content..

Keywords: Superplasticizer, admixture, compressive strength, aggressive environment

1. INTRODUCTION

Sulphuric acid attack on mortar can occur in many ways i.e. can be subjected to attacks by mineral acids which includes sulphuric, nitric, phosphoric and hydrochloric acid. Sulphuric acid is also important component in acid rain which can affect the durability of the mortar. Sewage system is another source of severe sulphuric acid attack. Since sulfur compounds are formed as a result of the sulfuric acid-cement paste reaction, the increase in sulfur content of concrete specimens could be used as a measure of the chemical manifestation of deterioration [1]. The spectrum of aggressive acidic media is wide. Acidic attack usually originates from industrial processes, but it can even be due to urban activity. Even natural exposure conditions may cause acid attacks. Free acids in natural waters are rare. Exceptions are carbonic waters and sulfurous and sulfuric acids in peat waters. Soils may contain huminous acids [2]. Effect of aggressive environment on compressive strength have been investigated [3, 4, 5, 6]. However, Impact on durability due to acidic aggressive environment on mortars have been studied [7, 8, 9, 10].

However, Superplasticizer is also added to check its effect on compressive strength. Superplasticizers are frequently used in order to improve the workability of mortar and concrete mixes for demanding applications. The addition of superplasticizers aims at two objectives: first, the addition of superplasticizers allows controlling the flow properties, which are of major importance for the design of self-compacting concretes, and second, superplasticizers allow the reduction of the water to cement ratio while maintaining workability in order to reach high strength and durability. The cement–water system is highly sensitive to the addition of superplasticizers. Small amounts of superplasticizers enhance the workability properties efficiently, but are often associated with strong, undesired retardation phenomena of the setting of the cement paste[11].

Impact of superplasticizer on compressive strength have been investigated [12, 13, 14, 15].

Adsorption, fluidity and hydration property have been studied and it has been reported that use of superplasticizers improves these properties [16, 17, 18, 19, 20]. Durability, pore size distribution and rheological properties have been investigated and beneficial effect of superplasticizer have been reported [21, 22, 23, 24, 25].

This paper presents the results of an experimental investigation carried out to access the effect of acidic aggressive environment on strength of mortars in early stage. Moreover, SP is also added to investigate compressive strength of mortars in early stage in aggressive environment.

2. EXPERIMENTAL PROGRAMME

It was intended to observe the effect of acidic aggressive environment on cement mortars and performance of two cement mortar mixes with varying amount of superplasticizers. The experimental programme included the consideration of effect of sulphuric acid environments on cement mortars in early age. The ratio in cement and sand in mortar were 1:4 and 1:6 by weight.

In present investigations, different superplasticizer dosages (1%, 2% of weight of cement) are provided in mortar mix and mixes are prepared with different water cement ratio. Abbreviations of some typical mortar mix designations are explained below in Table 1.

Mortar cube samples of 70.6 mm size were prepared as per standard procedure and immersed in two tanks of sulphuric acid of 1 N and 2 N concentrations respectively. The compressive strength of samples were observed after 7 and 14 days of curing. Variations of compressive strength with and without superplasticizer were noted with 7 and 14 days of curing.

Table 1 – Explanation of Some Mortar Mix Designations

Typical mix designation	Explanation
CM4W	Cement Mortar 1:4 Water cured
CM4N1	Cement Mortar 1:4 Normal 1 acid environment cured
CM4N2	Cement Mortar 1:4 Normal 2 acid environment cured
CM4S0%	Cement Mortar 1:4 with SP Dosage 0%
CM6W	Cement Mortar 1:6 Water cured
CM6N1	Cement Mortar 1:6 Normal 1 acid environment cured
CM6N2	Cement Mortar 1:6 Normal 2 acid environment cured
CM6S0%	Cement Mortar 1:6 with SP Dosage 0%

RESULTS AND DISCUSSION

Compressive strength of mortar mix designations 1:4 for both 7 days and 14 days with and without superplasticizer in all curing medium is shown in Table 2 and Table 3 respectively. However compressive strength of mortar mix designations 1:6 for both 7 days and 14 days with and without superplasticizer in all curing medium is shown in Table 4 and Table 5 respectively.

Table 2 – Compressive strength for 7 days

Mortar Proportion [1:4]			
SP Dosage	CM4W (N/mm^2)	CM4N1 (N/mm^2)	CM4N2 (N/mm^2)
CM4S0%	3.41	2.90	2.42
CM4S1%	4.11	3.37	2.82
CM4S2%	3.59	2.80	2.40

Table 3 – Compressive strength for 14 days

Mortar Proportion [1:4]			
SP Dosage	CM6W (N/mm^2)	CM6N1 (N/mm^2)	CM6N2 (N/mm^2)
CM6S0%	4.19	3.57	2.98
CM6S1%	5.05	4.13	3.49
CM6S2%	4.41	3.45	2.94

Table 4 – Compressive strength for 7 days

Mortar Proportion [1:6]			
SP Dosage	CM4W (N/mm^2)	CM4N1 (N/mm^2)	CM4N2 (N/mm^2)
CM4S0%	2.25	1.92	1.56
CM4S1%	2.76	2.18	1.80
CM4S2%	2.38	1.78	1.44

Table 5 – Compressive strength for 14days

Mortar Proportion [1:6]			
SP Dosage	CM6W (N/mm^2)	CM6N1 (N/mm^2)	CM6N2 (N/mm^2)
CM6S0%	2.78	2.36	1.90
CM6S1%	3.41	2.68	2.22
CM6S2%	2.92	2.18	1.76

Generally, the effect of sulphuric acid to degrade the mortar specimens with time and the degradation depends on many factors such as concentration of sulphuric acid, time of exposure to aggressive environment etc. It has been investigated that the superplasticizer dosage will improve the compressive strength, there is still an ideal farthest point for the utilization of admixture. At the point when the doses go past this farthest point, increment in dose will just diminish the compressive quality. This marvel happen since over dose of superplasticizer will bring about draining and isolation, which will influence the cohessiveness and consistency of the solid. Accordingly, compressive quality will diminish if the utilized measurements is past the ideal measurement.

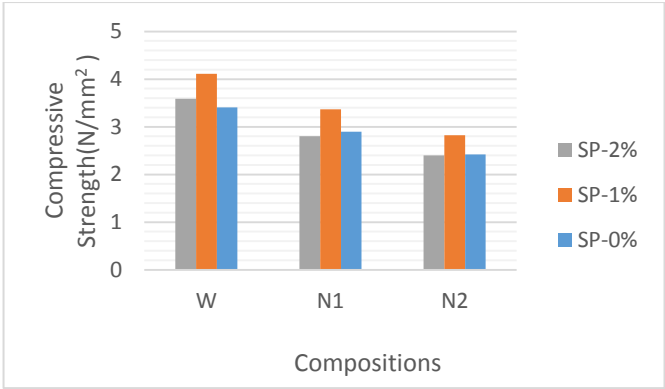


Figure 1 – Compressive strength for mortar 1:4 for 7 days

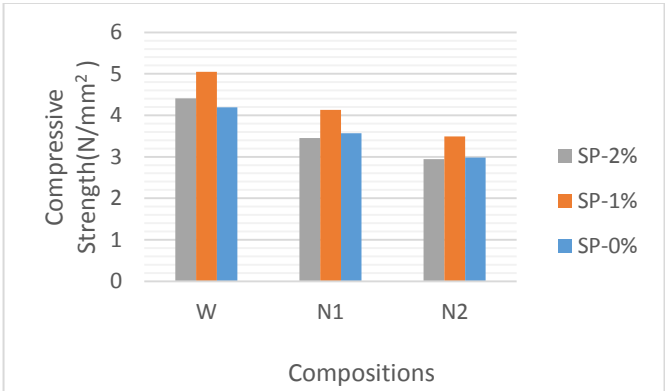


Figure 2 – Compressive strength for mortar 1:4 for 14days

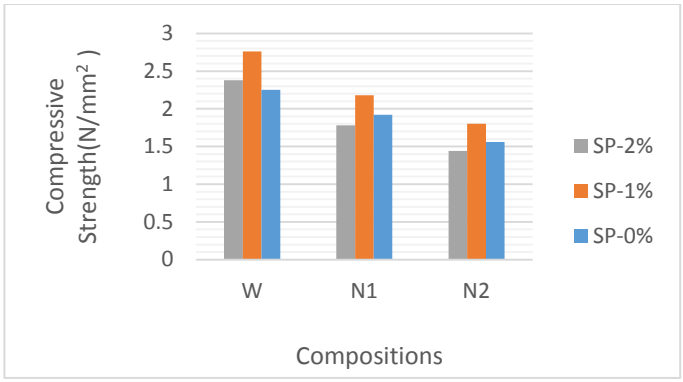


Figure 3 – Compressive strength for mortar 1:6 for 7days

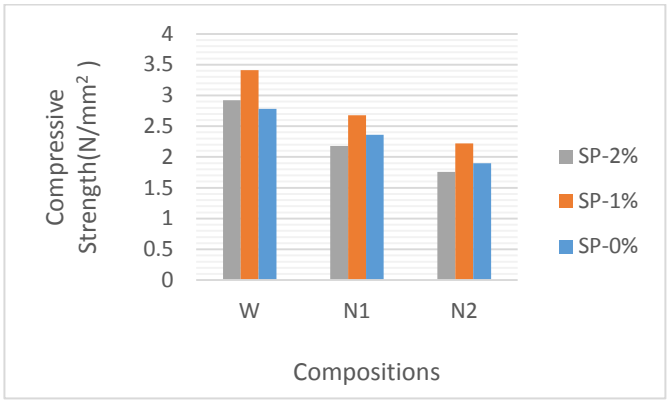


Figure 4 – Compressive strength for mortar 1:6 for 14days

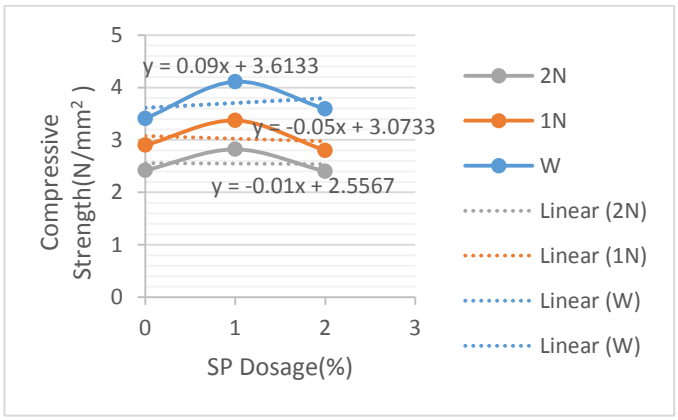


Figure 5 – Compressive strength for mortar 1:4 for 7 days

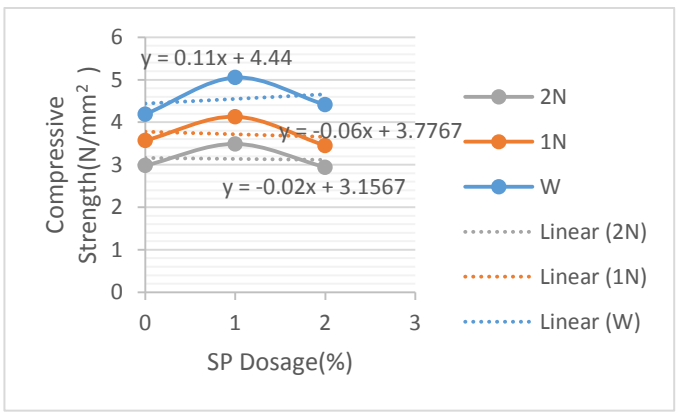


Figure 6– Compressive strength for mortar 1:4 for 14days

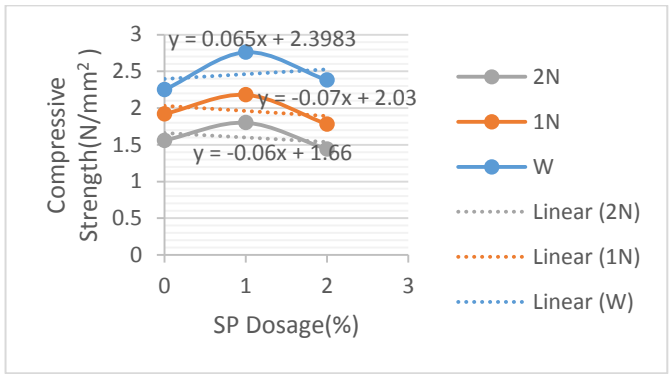


Figure 7 – Compressive strength for mortar 1:6 for 7days

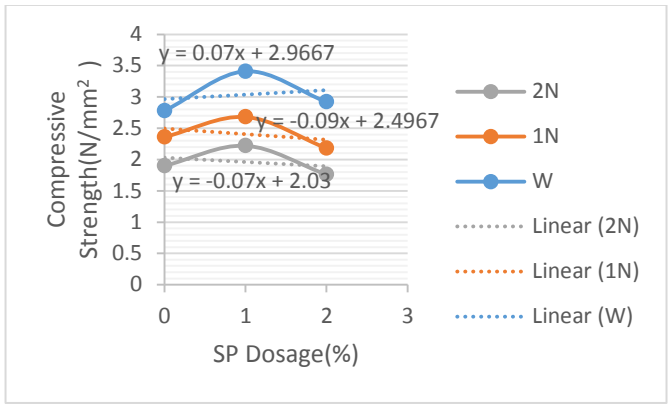


Figure 8– Compressive strength for mortar 1:6 for 14days

The figures are developed for all the above described results.

From the figure 1 to 4 it can be seen that there is approximately 17% decrease in compressive strength when curing medium changes from water to 1N and then finally to 2N concentrated sulphuric acid solution in both mortar mix designation 1:4 and 1:6 without superplasticizer. However, when 1% SP is provided, there is decrease of approximately 15% compressive strength. Moreover, there is decrease of approximately 18% when SP dosage exceeds to 2%.

From the figure 5 to 8 it can be seen that there is approximately 22% increase in compressive strength when curing days raised to 14 in both mortar mix designation 1:4 and 1:6. However, there is decrease of approximately 14% compressive strength when the curing medium changes from water to 1N and then finally to 2N concentrated sulphuric acid solution. It can also be concluded that there is increase of approximately 5% when dosage increases directly to 2% but when it compared with SP dosage 1%, it decreased. Moreover, in 2N concentrated solution compressive strength decreased at the rate of nearly 6% compared to 0% SP Dosage.

From the above figures it may be concluded that there is decrease of 15% compressive strength in increasing acidic aggressive environment. Moreover, due to increment of SP dosage there is decrease of 16% in compressive strength in both mortar mix designations 1:4 and 1:6.

CONCLUSIONS

An investigation has been performed to study the effects of acidic aggressive environment on strength of mortars in early stage in mortar proportion 1:4 and 1:6. The workability of mortar can be expanded by expansion of superplasticizer. Nonetheless, high doses of superplasticizer have a tendency to weaken the cohesiveness of concrete. Compressive strength decreases when curing medium changes. However, compressive strength increases when dosage exceed to 1% but there is decrease in compressive strength when dosage exceed to 2%. It indicates some optimum amount of SP which may be adopted for cement mortar mixes for best performances.

REFERENCES

1. Ying-Ming Wang, Kwai-Sang Chi: A neutral DEA model for cross-efficiency evaluation and its extension, in Expert Systems with Applications, An International Journal (2010),ref 3666–3675
2. Salah R. Agha, Ibrahim Kuhail, Nader Abdelnabi, Mahmoud Salem, Ahmed Ghanim: Assessment of academic departments efficiency using data envelopment analysis. Journal of Industrial Engineering and Management (2011), p301-325.
3. GU¨LSER KO¨KSAL & BURAK NALC¸ACI," The relative efficiency of departments at a Turkish engineering college: A data envelopment analysis ", International Journal of Operations & Production Management (2006) 51: 173–189
4. Rita Markovits-Somogyi," ranking efficient and inefficient decision making Units in data envelopment analysis ", Industrial Management & Data Systems, (2011)pp. 446 – 456
5. Abbott, M., & Doucouliagos, C. The Efficiency of Australian Universities: A Data Envelopment Analysis. Economics of Education Review, 22(1),(2003), 89-97..
6. Lopes, A., Lanzer, E. Data Envelopment Analysis-DEA And Fuzzy Sets to Assess The Performance of Academic Departments: A Study at Federal University Of Santa Catrina-UFSC. Pesquisa Operacional, 22(2), (2002), 217-230.
7. Moreno, A. & Tadepalli, R.. Assessing Academic Department Efficiency a Public University. Managerial and Decision Economics, 23(7), 385-397.8. Janusz Bucki Yvon Pesqueux, (2000),"Flexible workshop: about the concept of flexibility", International Journal of Agile Management Systems, Vol. 2 Iss 1, (2002) pp. 62 – 70
9. Tom Davis (1993) ,"Effective Supply Chain Management",Sloan Management Review, Summer 1993 edition.
10. Nand Kumar, Ranganath.M.S, Chandrashekhar, "Sustainability Analysis of Banking Sector in India using DEA", in National Conference on Inclusive & Sustainable Growth - IMT Nagpur
11. www.emeraldinsight.com
12. Banxia Frontier Analyst 4 Demo version.

1. Ying-Ming Wang, Kwai-Sang Chi: A neutral DEA model for cross-efficiency evaluation and its extension, in Expert Systems with Applications, An International Journal (2010),ref 3666–3675
2. Salah R. Agha, Ibrahim Kuhail, Nader Abdelnabi, Mahmoud Salem, Ahmed Ghanim: Assessment of academic departments efficiency using data envelopment analysis. Journal of Industrial Engineering and Management (2011), p301-325.
3. GU¨LSER KO¨KSAL & BURAK NALC¸ACI," The relative efficiency of departments at a Turkish engineering college: A data envelopment analysis ", International Journal of Operations & Production Management (2006) 51: 173–189
4. Rita Markovits-Somogyi," ranking efficient and inefficient decision making Units in data envelopment analysis ", Industrial Management & Data Systems, (2011)pp. 446 – 456
5. Abbott, M., & Doucouliagos, C. The Efficiency of Australian Universities: A Data Envelopment Analysis. Economics of Education Review, 22(1),(2003), 89-97..
6. Lopes, A., Lanzer, E. Data Envelopment Analysis-DEA And Fuzzy Sets to Assess The Performance of Academic Departments: A Study at Federal University Of Santa Catrina-UFSC. Pesquisa Operacional, 22(2), (2002), 217-230.
7. Moreno, A. & Tadepalli, R.. Assessing Academic Department Efficiency a Public University. Managerial and Decision Economics, 23(7), 385-397.8. Janusz Bucki Yvon Pesqueux, (2000),"Flexible workshop: about the concept of flexibility", International Journal of Agile Management Systems, Vol. 2 Iss 1, (2002) pp. 62 – 70
9. Tom Davis (1993) ,"Effective Supply Chain Management",Sloan Management Review, Summer 1993 edition.
10. Nand Kumar, Ranganath.M.S, Chandrashekhar, "Sustainability Analysis of Banking Sector in India using DEA", in National Conference on Inclusive & Sustainable Growth - IMT Nagpur
11. www.emeraldinsight.com
12. Banxia Frontier Analyst 4 Demo version.

Effect of gas composition on morphological properties of graphene nanosheet

Neha Gupta and Suresh C. Sharma

Citation: Physics of Plasmas **24**, 073510 (2017); doi: 10.1063/1.4993203
View online: http://dx.doi.org/10.1063/1.4993203
View Table of Contents: http://aip.scitation.org/toc/php/24/7
Published by the American Institute of Physics

Articles you may be interested in

Theoretical modeling to study the impact of different oxidizers (etchants) on the plasma-assisted catalytic carbon nanofiber growth
Physics of Plasmas **24**, 073504 (2017); 10.1063/1.4990556



Effect of gas composition on morphological properties of graphene nanosheet

Neha Gupta and Suresh C. Sharma
Department of Applied Physics, Delhi Technological University, Shahbad Daulatpur, Bawana Road, Delhi 110042, India

(Received 16 May 2017; accepted 27 June 2017; published online 13 July 2017)

A multiscale theoretical model to study the effect of different gas mixtures on the nucleation and growth kinetics of a graphene nanosheet in the reactive low-temperature plasma environment has been developed. The model includes the plasma sheath formalization, kinetics of all the plasma species, charging of the graphene sheet, plasma-surface interaction, clusters and graphene islands nucleation, and vertical growth of a graphene nanosheet. The three different gas mixtures, i.e., C_2H_2, CH_4, and CF_4 with hydrogen and argon, are considered in the present investigation to examine the variations in the number densities of carbon and hydrogen species generated on the catalyst surface and their consecutive effects on the dimensions (i.e., height and thickness) and number density profiles of the graphene nanosheet. It is found that the thickness and height of the graphene sheet are maximum for C_2H_2 gas mixtures and least for CH_4 and CF_4, respectively. On the basis of the results obtained, the field emission characteristics of the graphene sheet have been analyzed, and it is estimated that C_2H_2 contained gas mixture enhances the field emission characteristics of the graphene sheet followed by CH_4 and CF_4. The presented results are in good agreement with the existing experimental observations. *Published by AIP Publishing.*
[http://dx.doi.org/10.1063/1.4993203]

I. INTRODUCTION

Since the earliest reports on the isolation of graphene by the micromechanical cleavage of bulk graphite in 2004,[1] various pathways for graphene synthesis have been adopted by researchers, such as chemical exfoliation of graphite,[2] high temperature annealing of SiC surface,[3] reduction of graphene oxide sheets,[4] arc discharge method,[5] chemical vapor deposition (CVD),[6] and plasma enhanced chemical vapor deposition (PECVD).[7,8] Low temperature PECVD is the most intriguing technique for the synthesis of graphene used for field emission devices. It has been demonstrated that the vertical orientation, high surface area, and sharply exposed ultra thin and long edges of the graphene are the distinctive features of the PECVD route, which renders graphene to be a promising candidate for field emission applications.[7–10] Therefore, control over the growth and structural properties of graphene is requisite for the effective realization of graphene-based field emitters.

PECVD is considered to be the most feasible technique for large scale production of vertical graphene with enhanced controllability over the dimensions and structural properties at relatively low temperatures due to the presence of diverse combinations of energetic ions, electrons, radicals, neutrals, and other active species within the plasma region.[11] Subsequently, extensive studies have been carried out to control the morphology, growth rate, dimensions, and structure of graphene by effectively tuning the plasma parameters,[12] input power,[13] operating pressure,[14] type, and composition of feedstock gases.[15–17] With a few exceptions of solid or natural precursors (e.g., honey, milk, cheese, butter, and sugar),[18] a variety of carbon containing gases frequently diluted with Ar and/or H_2 is used for graphene sheet growth using PECVD method. Hydrogen plays an essential role of surface passivation, i.e., termination of surface dangling bonds and also acts as an effective carbon etching agent to remove the amorphous carbon layer formed during vertical graphene nucleation and growth, thus promoting the crystalline graphitic structure and sharp edges of the growing vertical graphene sheet.[19,20] The addition of Ar to the carbon source gas and hydrogen mixture increases the plasma stability by enhancing the electron energy,[21] and in addition, these energetic Ar ions also facilitates the activation of the initially hydrogen terminated surface bonds.[20] The non-uniformities of the fluxes of hydrogen and argon to the substrate can significantly affect the nanofabrication process. In this regard, Ostrikov *et al.*[20] have shown that by choosing the appropriate process parameters such as working pressure and composition of the reactive gas feedstock, one can prepare a uniform surface for the nanostructures growth from the non-uniform plasmas.

Hydrocarbons (C_2H_4, C_2H_2, or CH_4), fluorocarbons (C_2F_6, CF_4, CHF_3), and carbon dioxide/monoxide are the most common choices of carbon containing precursors for graphene sheet growth via PECVD method. The morphology and growth rate of the graphene sheet is closely related to the type of carbon feedstock gas used for graphene formation. For instance, it was reported that by using C_2H_2 carbon feedstock, vertical graphene sheets synthesized by RF-PECVD get thickened and exhibit higher growth rate than CH_4 grown vertical graphene sheets.[10,15] This observation was accredited to the fact that higher carbon bearing species are accessible in C_2H_2 plasmas due to the low dissociation energies in contrast to CH_4 plasmas.[15] Moreover, the

presence of carbon dimers (C_2) in plasma plays a significant role in the vertical graphene sheet growth.[22] However, in the plasma aided catalytic growth process of vertical graphene, carbon species are more effectively generated due to dissociation of precursor gases at the catalyst active surface while their dissociation in plasma is not much in principle.[23] The reactive C_2 radicals are readily formed in C_2H_2 containing feedstock than in CH_4 due to direct dissociation of $C \equiv C$ bond in C_2H_2.[24] Shiji *et al.*[16] have compared the morphology, growth rates, thickness, and spacing between the carbon nanowalls (CNWs) grown by RF-PECVD using $C_2F_6 + H_2$, $CH_4 + H_2$, $CF_4 + H_2$, $CHF_3 + H_2$, and $C_4F_8 + H_2$ gas mixtures. It was observed that the CNWs were not grown for plasmas employing $C_4F_8 + H_2$ precursor; however, the growth rates of CNWs for $C_2F_6 + H_2$ is highest and lowest for $CF_4 + H_2$ gas mixtures. Moreover, CNWs grown using $CH_4 + H_2$ gas mixtures were reported to be very thin and wavy in contrast to other fluorocarbon/hydrogen gas mixtures grown CNWs.[16] Furthermore, the trace amount of oxygen addition to C_2F_6+H_2 gas mixture increases the graphitization and control the CNWs nucleation by reducing the amorphous carbon layer under the CNWs.[17]

These unique structural variations achieved by using different carbon precursors in PECVD process make graphene a very attractive candidate for field emission devices. Zhu *et al.*[10] reported that the carbon nanosheets (CNSs) grown by using C_2H_2 feedstock are much higher and more vertically oriented than CH_4 synthesized carbon nanosheets; as a result of which C_2H_2 CNSs exhibit low threshold field, high current density, and high field enhancement factor than CH_4 CNSs. Likewise, Hiramatsu and Hori[25] have demonstrated that CNWs synthesized via $C_2F_6 + H_2$ system exhibit lowest onset electric field for emission of electrons among $CH_4 + H_2$, $CF_4 + H_2$, $CHF_3 + H_2$, and $C_4F_8 + H_2$ gas mixtures.

Our previous studies suggest that the plasma-based catalytic graphene growth and its field emission characteristics are affected by the plasma parameters, plasma power, and by nitrogen and boron doping.[12,26] Motivated by the above works[10,15–17,25] of controlling the growth and field emission properties of the graphene sheet by using different carbon feedstocks, we try to develop a theoretical model to understand the effect of different gas mixtures on the number density profiles of the carbon and hydrogen species generated per unit area on the catalyst surface and further their ramifications on the growth kinetics and number densities of the graphene sheet. Moreover, the present model describes the key processes involved in the catalyst-aided growth of vertical graphene sheet in its entirety, from formation of carbon clusters, nucleation and growth of graphene islands and growth of vertical graphene sheet in the presence of different gas mixtures which have not been presented earlier as far as by authors' knowledge. In addition, the present model considers the kinetics of various plasma species, i.e., electrons, ions, and neutrals, charging rate of the graphene sheet, number density profiles of carbon clusters and graphene islands, growth rate equation of graphene islands, and vertical graphene sheet.

The rest of the paper is organized as follows: Sec. II describes the growth processes of the vertical graphene sheet in detail including the kinetic equations of all the plasma species and charging rate of the graphene sheet. In Sec. III, the results of the contributions of different gas mixtures during the vertical graphene sheet growth are presented and discussed. The repercussions of different gas mixtures on the field enhancement factor of the graphene sheet are also addressed. Finally, the paper is concluded in Sec. IV.

II. MODEL

This section describes the nucleation and catalyst-assisted growth model of the graphene sheet for a typically PECVD case. As considered in our previous works,[12,26] we assume the growth of graphene sheet on the copper catalyst over the silicon substrate surface in an ionized gas mixture of carbon source gas, etchant gas, and carrier gas. However, in the present paper, we investigate the growth of the graphene sheet in three different gas mixtures, i.e., $C_2H_2 + H_2 + Ar$, $CH_4 + H_2 + Ar$, and $CF_4 + H_2 + Ar$. Our previous works[12,26] rely only on selected neutrals and ions of the reactive plasma; thus, in order to augment the significance of our previous works, we consider the growth of graphene sheet in the multi-component plasma consisting of electrons, various ions, and neutrals of $C_2H_2/CH_4/CF_4$+H_2+Ar gas mixtures. Table I shows the plasma composition of these gas mixtures that have been accounted on the basis of literature[27–33] specifying the plasma chemistry of these gas mixtures.

The catalyst-assisted growth scenario of graphene nanosheet in plasma environment has been presented earlier.[12] For clarity, the schema of the formation of vertical graphene in the presence of plasma is shown in Fig. 1. Figure 1(a) shows the initial stage in which the planar growth of 2D graphitic layers takes place due to coalescence of graphene islands formed on account of diffusion and collision of carbon clusters formed on the catalyst surface.[34] However, during the coalescence of graphene islands, internal stress may develop at the graphene island boundaries due to temperature gradients, ion bombardment, and lattice mismatch between the catalyst-substrate surface and the graphitic material.[11] This internal stress is the tensile stress (T)

TABLE I. Plasma species of different gas mixtures considered in the model.

	Neutrals		Ions	
Gas mixture	Type A	Type B	Type A	Type B
CH_4+H_2+Ar	CH_3, CH_4, C_2H_2, C_2H_4	H, H_2	CH_3^+, CH_4^+, C_2H_2+, C_2H_4+	H^+, H_2^+
C_2H_2+H_2+Ar	C_2H_2, C_4H_2, C_4H_3, C_6H_4, C_2H	H, H_2	$C_2H_2^+$, $C_4H_2^+$, $C_4H_3^+$, $C_6H_4^+$, C_2H^+	H^+, H_2^+
CF_4+H_2+Ar	CF, CF_2, CF_3	H, H_2	CF^+, CF_2^+, CF_3^+	H^+, H_2^+

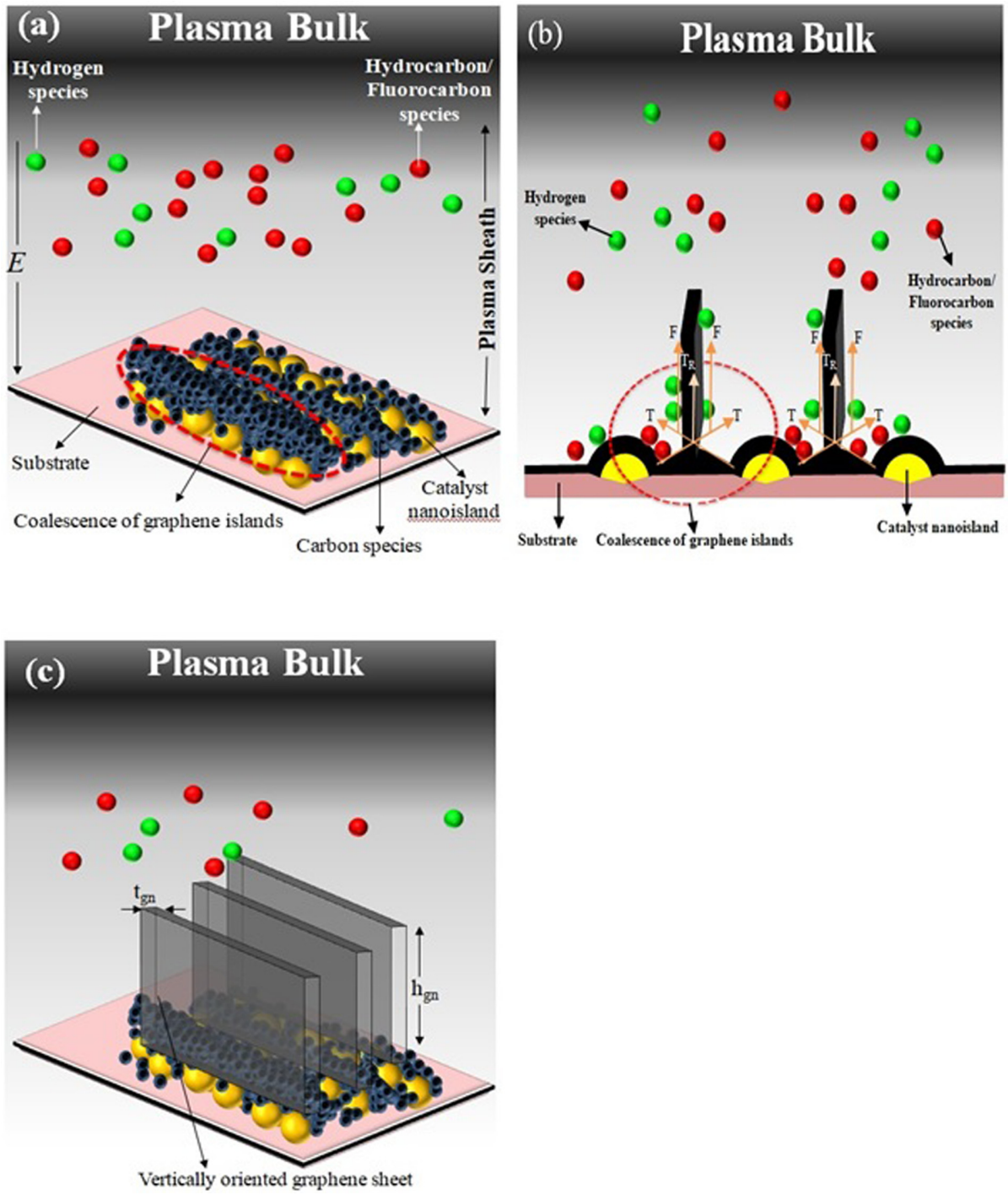


FIG. 1. Schema showing the (a) coalescence of graphene islands and (b) upward growth of the graphene sheet, and (c) scheme of the vertical graphene sheet considered in the model.

tangential to the planar graphitic layers, which tends to bend the free edge of the planar graphitic layers in the upward direction as shown in Fig. 1(b).[35] Further, under the influence of resultant tensile stress (T_R) and electric force (F) due to the plasma sheath at the graphene island boundaries [see Fig. 1(b)], and due to continuous delivery of hydrocarbon species the graphene sheet continues to grow in the vertical direction.[35] Finally, the graphene sheet growth terminates due to the competition between the hydrocarbon deposition and etching effects by hydrogen.[11] The scheme of the graphene sheet considered in the model is shown in Fig. 1(c).

The growth process is strongly affected by the energies and fluxes of hydrocarbon ions and etching gases near the catalyst-substrate surface. These fluxes can be effectively controlled by the plasma parameters and the electric fields induced by the plasma sheath formed near the boundaries of plasma-exposed catalyst-substrate surface, which is maintained at a negative potential U_s. It is assumed that the sheath

electric field is directed along the z-axis, which accelerates all the ions toward the catalyst-substrate surface. In order to investigate the sheath structure formed near the catalyst-substrate surface and the dynamics of the plasma species, the following sheath equations[36,37] are used.

$$\frac{d}{dz}(n_r u_{rz}) = \upsilon_{iz} n_e, \tag{1}$$

$$m_r n_r u_{rz} \frac{du_{rz}}{dz} = e n_r \frac{d\phi}{dz} - m_r n_r \upsilon_{rn} u_{rz}, \tag{2}$$

$$\frac{d^2\phi}{dz^2} = -4\pi e \sum_r \kappa_r n_r, \tag{3}$$

where Eqs. (1)–(3) are the continuity equation, ion-momentum equation, and Poisson's equation, respectively; r indicates either the electron (e), or the positively charged ions considered in the present model; n_r, m_r, and u_{rz} are their number density, mass, and fluid velocity, respectively. υ_{iz} and υ_{rn} are the ionization frequency and collision frequency of electrons and ions with electrons, respectively. ϕ is the sheath potential, κ_r is the rth ion to electron number density ratio, $\sum_r \kappa_r = 1$, and $0 < \kappa_r > 1$.

A. Kinetic equation of number of carbon species generated on the catalyst surface

Initially, we assume that the applied plasma power segregates the catalyst (Cu) film of thickness 4 nm into N_a ($=6\times10^8\,\mathrm{cm}^{-2}$) catalyst nanoislands of size r_c ($=50$ nm) each.[38] Carbon and hydrogen species are then generated from the particular composition of plasmas (i.e., $C_2H_2/CH_4/CF_4+H_2+Ar$) on these catalyst nanoislands via several processes, which have been incorporated in Eqs. (4) and (5), respectively.

Following Gupta *et al.*[39] and Marvi *et al.*[40]

$$\begin{aligned}\frac{dN_C}{d\tau} &= \sum_A J_A(1-\theta_t) + 2\sum_A n_{sA}\upsilon \exp\left(\frac{-\delta E_i}{k_B T_S}\right) \\ &+ 2\sum_{iA}\left(\sum_A \frac{n_{sA} y_{di}}{\nu_0}\right) J_{iA} + \sum_{iA} J_{iA} + J_C \\ &+ \sum_{iA}\sum_{iB} \frac{J_{iA}\sigma_{ads} J_{iB}}{\upsilon} - \sum_A n_{sA}\sigma_{ads} J_B \\ &- \sum_A n_{sA}\upsilon \exp\left(\frac{-E_{ads}}{k_B T_S}\right) - N_c \upsilon \exp\left(\frac{-E_{ev}}{k_B T_S}\right),\end{aligned} \tag{4}$$

$$\begin{aligned}\frac{dN_H}{dt} &= \sum_B J_B(1-\theta_t) + \sum_B n_{sB}\nu \exp\left(\frac{-\delta E_i}{k_B T_S}\right) \\ &+ \sum_{iA}\left(\sum_A \frac{n_{sA} y_{di}}{\upsilon_0}\right) J_{iA} + \sum_{iB} J_{iB} \\ &- \sum_B n_{sB}\upsilon \exp\left(\frac{-\delta E_{aH}}{k_B T_S}\right) - \sum_B n_{sB}\sigma_{ads} J_B \\ &- \sum_B n_{sB}\sigma_{ads} J_{iB},\end{aligned} \tag{5}$$

where N_c and N_H are the number densities of carbon and hydrogen species generated per unit area on the catalyst surface, respectively. $J_j = \left(\frac{n_j v_{thj}}{4}\right)$ is the flux of impinging neutral species j from the plasma (where j stands for species of type A or B); n_j and v_{thj} are the volume number density and thermal velocity of jth type of species in the plasma,[36] $n_{sj}(=\theta_j\nu_0)$ is the surface concentration of neutral species; θ_j is the surface coverage by species of type j, $\nu_0(\approx10^{15}\,\mathrm{cm}^{-2})$ is the number of adsorption sites per unit area,[36] $y_{di}(\approx 2.49\times10^{-2} + 3.29\times10^{-2}\times E_i$; where E_i is the ion energy in eV) is the stitching probability,[41] $\upsilon(\approx10^{13}\,\mathrm{Hz})$ is the thermal vibration frequency,[36] $J_{ij} = (n_{ij}(\frac{k_B T_e}{m_{ij}})^{\frac{1}{2}})$ is the flux of jth type of ion from the plasma (where, j refers to A or B type of ion),[36] $\sigma_{ads}(\approx10^{-16}\,\mathrm{cm}^2)$ is the cross-section for the reactions of atomic hydrogen,[36] T_s is the substrate or catalyst temperature (catalyst and substrate temperature are assumed to be same in the absence of heating effect),[40] E_{ads} is the hydrocarbon/fluorocarbon adsorption energy, E_{ev} is the carbon evaporation energy, δE_{aH} is the adsorption energy of etching species, i.e., hydrogen, and δE_i is the thermal energy barrier.

Equation (4) describes the rate of generation of carbon species per unit area on the catalyst surface via adsorption of hydrocarbon/fluorocarbon neutral species on the catalyst surface (first term), thermal and ion-induced dissociation of adsorbed hydrocarbon/fluorocarbon neutral species (second and third terms), decomposition of hydrocarbon/fluorocarbon ions (fourth term), direct carbon flux towards the catalyst surface (fifth term), and due to interaction of hydrocarbon/fluorocarbon ions and hydrogen ions, i.e., due to ion-ion interaction (sixth term).[40,42] The last three terms in Eq. (4) quantify the loss of carbon from the catalyst surface due to the interaction of adsorbed hydrocarbon species with the neutral hydrogen from the plasma, desorption, and evaporation of carbon atoms, respectively.[36,40]

Equation (5) describes the rate of generation of hydrogen species on the catalyst nanoparticle surface. The first four terms account for hydrogen adsorption, hydrogen atom generation due to thermal dissociation of hydrocarbons, ion-induced dissociation of hydrocarbons, and decomposition of ions of type B, respectively. The last three terms account for the hydrogen loss due to desorption (fifth term), and interaction of adsorbed hydrogen atoms with incoming hydrogen atoms and ions from the plasma (sixth and seventh terms).[40]

B. Kinetic equation of number densities of carbon clusters and graphene nuclei/islands

The carbon species created on the catalyst surface diffuse and attach each other to form carbon clusters, which in turn diffuse and collide with each other to form graphene nuclei. These graphene nuclei diffuse and coalesce (or stitch) together to form large graphene islands (with radius r_{Is}).[34] This leads to the lateral extension or growth of the planar graphene sheet, and we assume the length (l_{gn}) of the graphene sheet to be 0.81 μm (Ref. 43)

$$\begin{aligned}\frac{dN_{Cl}}{d\tau} &= N_a N_C D \exp\left(\frac{-E_{sdC}}{k_B T_s}\right) - N_{Cl} D' \exp\left(\frac{-E_{sdCl}}{k_B T_s}\right) \\ &\times \frac{1}{\pi r_c^2} - N_{Cl} N_{Is} D' \exp\left(\frac{-(E_{sdCl} + E_{incCl})}{k_B T_s}\right),\end{aligned} \tag{6}$$

$$\frac{dN_{Is}}{d\tau} = N_{Cl}D' \exp\left(\frac{-E_{sdCl}}{k_B T_s}\right) \times \frac{1}{\pi r_c^2} - \left(N_{Cl}N_{Is}D' \exp\left(\frac{-(E_{sdCl}+E_{incCl})}{k_B T_s}\right) + N_c N_{Is} D \exp\left(\frac{-(E_{sdC}+E_{incC})}{k_B T_s}\right) + N_{Is}D'' \exp\left(-\frac{E_{sdIs}}{k_B T_s}\right) \times \frac{1}{\pi r_{Is}^2}\right), \quad (7)$$

where N_{Cl} and N_{Is} are the surface concentration of carbon clusters and graphene islands, respectively. $D\exp\left(\frac{-E_{sdC}}{k_B T_s}\right)$, $D'\exp\left(\frac{-E_{sdCl}}{k_B T_s}\right)$, and $D''\exp\left(-\frac{E_{sdIs}}{k_B T_s}\right)$ are the surface diffusion coefficients of carbon atoms, carbon clusters, and graphene islands, respectively, with corresponding energy barriers, $E_{sdC} \approx 0.1\,\text{eV}$, $E_{sdCl} \approx 0.82\,\text{eV}$, and $E_{sdIs} \approx 2.6\,\text{eV}$, respectively.[34]

Equations (6) and (7) describe the growth rate of surface concentrations of carbon clusters and graphene islands, respectively. The first term on the right side of Eq. (6) specifies the generation of carbon clusters on account of surface diffusion of carbon atoms generated on the catalyst surface, while the second and last terms account for the cluster loss due to diffusion and incorporation of clusters into the graphene nuclei (see supporting information of Mehdipour and Ostrikov[34]). Furthermore, the first, second, third, and last terms of Eq. (7) indicate the formation of graphene nuclei due to surface diffusion of clusters, growth of graphene islands due to incorporations of clusters (with energy barrier $E_{incCl} \approx 0.8\,\text{eV}$) and carbon atoms (with energy barrier $E_{incC} \approx 0.85\,\text{eV}$), and coalescence (or ripening) of graphene islands, respectively (see supporting information of Mehdipour and Ostrikov[34]).

C. Growth rate equation of the graphene nuclei/islands

$$\frac{d(\pi r_{Is}^2)}{d\tau} = \left(\left(N_{Cl}D' \exp\left(\frac{-E_{sdCl}}{k_B T_s}\right) \times \frac{1}{\pi r_c^2} + N_{Cl}N_{Is}D' \exp\left(\frac{-(E_{sdCl}+E_{incCl})}{k_B T_s}\right) + N_c N_{Is} D \times \exp\left(\frac{-(E_{sdC}+E_{incC})}{k_B T_s}\right)\right) \times 2\pi r_c + N_{Is}D'' \times \exp\left(-\frac{E_{sdIs}}{k_B T_s}\right) \times \frac{1}{2\pi r_{Is}}\right) \times \left(\frac{M_c \times j_{iAgIs}}{\rho_c \upsilon}\right) + \gamma_A \pi r_{Is}^2 j_{AgIs}, \quad (8)$$

where $M_c (= 63.54\,\text{amu})$ and $\rho_c (= 8.96 g/\text{cm}^3)$ are the mass and density of the catalyst, and r_{Is} is the radius of the graphene island. $j_{ijgIs} = \pi r_{Is}^2 \left(\frac{8k_B T_{ij}}{\pi m_{ij}}\right)^{\frac{1}{2}} n_{ij}(x)\left[1 - Z\gamma_{ji}\right]\left(\exp\left[-\frac{eU_s}{k_B T_s}\right]\right)$ is the ion collection current at the surface of the graphene island[12] (subscript j refers to either A or B type of positively charged ion as explained earlier); Z is the charge over the graphene sheet [can be depicted from Eq. (11)], T_{ij} and m_{ij} are the mass and temperature of ion j, respectively, $\gamma_{ji} = \frac{e^2}{r_{Is} k_B T_s}$, $n_{ij}(x) = n_{ij0}\left(1 - \frac{2e\phi(x)}{m_{ij} v_{i0}^2}\right)^{-1/2}$ is the ion density at any point within the plasma sheath; v_{i0} is the ion velocity, $\phi(x) = \phi_0 \exp\left(-\frac{|x|}{\lambda_d}\right)$ is the electrostatic potential; ϕ_0 is the negative potential at the surface and λ_d is the Debye length, and $j_{jgIs} = \left(\pi r_{Is}^2 \left(\frac{8k_B T_n}{\pi m_j}\right)^{\frac{1}{2}} n_j\right)$ is the neutral collection current at the surface of the graphene island;[12] T_n, m_j, and n_j are the temperature, mass, and number density of neutral atom j, respectively, and γ_A is the sticking coefficient of neutral of type A.

Equation (8) outlines the increase in area of the graphene island due to surface diffusion of clusters (first term), incorporation of clusters and carbon into the graphene island (second and third terms), diffusion or coalescence of graphene island to form large graphene island (fourth term), and due to sticking of neutral atoms of type A on the growing graphene island (last term). The factor j_{iAgIs} accounts for the collection currents of ions of type A on the surface of growing graphene island.

D. Growth rate equation of the graphene sheet

$$l_{gn}\frac{d(h_{gn} \times t_{gn})}{d\tau} = \left(\frac{N_c}{\nu} \times D \exp\left(-\left\{\frac{E_{sdC}+E_{incC}}{k_B T_s}\right\}\right)\frac{2\pi r_{Is}}{\pi r_{Is}^2} \times (j_{iAgn}) + (\gamma_A j_{Agn})\right) \times \frac{M_{gn}}{\rho_{gn}}(1 - \xi_s \xi_c), \quad (9)$$

$$\left[h_{gn}(\tau) + l_{gn}\right]\frac{d(t_{gn})}{d\tau} = (\gamma_{iB} j_{iBgn} + \gamma_B j_{Bgn}) N_H h_{gn}(\tau)\frac{M_{gn}}{\rho_{gn}}, \quad (10)$$

where $l_{gn}, t_{gn}, and\ h_{gn}$ are the length, thickness, and height of the graphene sheet, respectively, ρ_{gn} and $(M_{gn} \approx 12g)$ are the density and mass of the growing graphene sheet, respectively, $j_{ijgn} = n_{ij}(x)(l_{gn}t_{gn} + h_{gn}t_{gn} + l_{gn}h_{gn})\left(\frac{k_B T_{ij}}{2\pi^2 m_{ij}}\right)^{\frac{1}{2}}\left\{\frac{2}{\sqrt{\pi}}\left(\frac{eV_s}{k_B T_i}\right)^{\frac{1}{2}} + \exp\left[\frac{eV_s}{k_B T_{ij}}\right] erfc\left[\left(\frac{eV_s}{k_B T_{ij}}\right)^{\frac{1}{2}}\right]\right\} \times \exp\left[-\frac{eU_s}{k_B T_s}\right]$, $j_{jgn} = n_j(l_{gn}t_{gn} + h_{gn}t_{gn} + l_{gn}h_{gn})\left(\frac{k_B T_n}{2\pi^2 m_j}\right)^{\frac{1}{2}}$ are the ion and neutral collection currents at the graphene sheet surface, respectively,[12] and V_s is the surface potential of the graphene sheet [refer Eq. (A2) of Sharma and Gupta[12]].

Once the planar graphitic layers are formed due to coalescence of graphene islands, internal stress or tensile stress is developed at the graphene island boundaries. These internal stresses may cause defects within the initial planar graphitic layer, which trigger the growth of graphene in the upward direction.[11] Equation (9) describes the increase in the area of the vertical growing graphene sheet. Under the influence of inbuilt electric field associated with the plasma, the graphene sheet grows vertically due to continuous supply of carbon species, i.e., due to ion (j_{iAgn}) and neutral collection current (j_{Agn}) at the graphene island boundaries and the surfaces of the graphene sheet. Moreover, the diffusion and attachment of carbon atoms at the peripherals of the growing

graphitic platelets additionally contribute to the vertical graphene sheet growth [first term of Eq. (9)]. The factors $\xi_s(=J_{iA}y_{sp}(1-\theta_t)\times\frac{4\pi r_c^2}{\upsilon})$ and $\xi_c(=4\pi r_c^2 N_C)$ in Eq. (9) are the carbon sputtering[41] and amorphous carbon terms, where y_{sp} is the sputtering yield. Finally, the growth of vertical graphene terminates due to etching of the growing graphene sheets by the hydrogen. Equation (10) describes the reduction in thickness of the graphene sheet due to etching of carbon atoms by hydrogen. The first and second terms on the right side of Eq. (10) represent the ion and neutral collection currents of species of type B, i.e., hydrogen on the surface of the graphene sheet.

E. Charging of the graphene sheet

$$\frac{d(Z)}{d\tau}=\sum_A j_{iAgn}+\sum_B j_{iBgn}-\gamma_e j_{egn}, \tag{11}$$

where Z is the charge over the graphene sheet, γ_e is the sticking coefficient of the constituent electrons, and $j_{egn}=\left(\frac{k_BT_e}{2\pi^2 m_e}\right)^{\frac{1}{2}}n_e(x)(l_{gn}t_{gn}+h_{gn}t_{gn}+l_{gn}h_{gn})\exp\left[\frac{eV_s}{k_BT_e}+\frac{eU_s}{k_BT_s}\right]$ is the electron collection current at the graphene sheet surface.

The first two terms on the right side of Eq. (11) describe the charge developed on the graphene sheet surface due to the accretion of positively charged ions of types A and B, respectively. The last term specifies the attenuation in charge due to the accretion of electrons at the surface of the graphene sheet.

F. Kinetic balance equation of all the plasma species

$$\frac{d(n_e)}{d\tau}=\sum_A\beta_A n_A+\sum_B\beta_B n_B-\sum_A\alpha_A n_e n_{iA}-\sum_B\alpha_B n_e n_{iB}-\gamma_e n_{gn}j_{egn}, \tag{12}$$

$$\frac{d(n_{iA})}{d\tau}=\beta_A n_A-\alpha_A n_e n_{iA}-n_{gn}j_{iAgn}+\sum_{lAB}k_l n_A n_{iB}, \tag{13}$$

$$\frac{d(n_{iB})}{d\tau}=\beta_B n_B-\alpha_B n_e n_{iB}-n_{gn}j_{iBgn}-\sum_{lAB}k_l n_A n_{iB}, \tag{14}$$

$$\frac{d(n_A)}{d\tau}=\alpha_A n_e n_{iA}-\beta_A n_A+n_{gn}(1-\gamma_{iA})j_{iAgn}-n_{gn}\gamma_A j_{Agn}-J_{adA}+J_{desA}-\sum_{lAB}k_l n_A n_{iB}, \tag{15}$$

$$\frac{d(n_B)}{d\tau}=\alpha_B n_e n_{iB}-\beta_B n_B+n_{gn}(1-\gamma_{iB})j_{iBgn}-n_{gn}\gamma_B j_{Bgn}-J_{adB}+J_{desB}+J_{th}-\sum_{lAB}k_l n_A n_{iB}, \tag{16}$$

where β_j is the coefficient of ionization of the constituent neutral atoms due to external field (subscript j refers to either A or B type of neutral atoms), $\alpha_j(T_e)=\alpha_{j0}\left(\frac{300}{T_e}\right)^k \mathrm{cm}^3/\mathrm{s}$ is the coefficient of recombination of electrons and positively charged ions; $k(=-1.2)$ is a constant[44] and j refers to A and B types of ions, $J_{adj}=\frac{P_j}{(2\pi m_j k_B T_n)^{\frac{1}{2}}}\times\frac{n_j}{J_j}$ is the adsorption flux onto the catalyst-substrate surface; P_j is the partial pressure of adsorbing species,[45] $J_{desj}=n_j\upsilon\exp\left(-\frac{\varepsilon_{ads}}{k_BT_n}\right)$ is the desorption flux from the catalyst-substrate surface;[45] ε_{ads} is the adsorption energy, $J_{th}=n_B\nu\exp\left(\frac{-\delta\varepsilon_{th}}{k_BT_s}\right)$ is the flux of type B neutral species (namely, hydrogen) on account of thermal dehydrogenation;[45] $\delta\varepsilon_{th}$ is the activation energy of thermal dehydrogenation.

Equation (12) describes the growth rate of electron density in plasma. The first and second terms on the right side in Eq. (12) account for the gain in electron density per unit time because of ionization of neutral atoms of types A and B, respectively. The third term and fourth terms symbolize the reduction of electron density per unit time due to recombination of electrons and ions of types A and B, respectively. The last term is the electron collection current at the graphene sheet surface.

Equations (13) and (14) are the particle balance equations of positively charged ions of types A and B, respectively. The first terms on the right hand side in Eqs. (13) and (14) describe the growth rate of ion density due to ionization of neutral atoms. The second and third terms represent the decay rate of positively charged ion density due to electron-ion recombination and ion collection current at the graphene sheet surface, respectively. The last term accounts for the gain/loss of ion density due to ion-neutral reactions[28] and their reaction rate coefficients, i.e., k_l have been taken from the experimental works of Mao and Bogaerts.[30]

Equations (15) and (16) indicate the balance of neutrals of types A and B in the plasma, respectively. The first term on the right side in Eqs. (15) and (16) depict the rate of augmentation of neutral atom density on account of electron–ion recombination. The second and third terms represent the rate of decrease in neutral atom density due to ionization and increase in the neutral atom density due to neutralization of ions collected at the graphene sheet surface, respectively. The fourth term is the accretion of neutral atoms on the graphene sheet surface. The next two terms denote the adsorption and desorption of neutrals to/from the catalyst-substrate surface, respectively. The term J_{th} in Eq. (16) represents the increase in hydrogen number density in plasma because of thermal dehydrogenation.[45] The last term is the decay in neutral atom density on account of neutral-ion reactions.[28]

III. RESULTS AND DISCUSSION

In the present work, a theoretical model has been developed to provide better insight into the catalyst aided and plasma assisted growth process of the graphene sheet in the presence of different carbon source gases. The plasma-sheath related effects along with the kinetics of the various plasma species, i.e., electrons, neutrals, and positively charged ions and various elementary processes such as generation/loss of carbon species on the catalyst surface, surface diffusion of species, formation of carbon clusters, diffusion of carbon clusters to form graphene islands/nuclei, coalescence of graphene islands to form graphene layer parallel to the catalyst-substrate surface and further the nucleation of vertical graphene sheet due to stress release at the interface

of the graphene island boundaries during coalescence have been incorporated in the model.

In this section, the numerical solutions of the model equations developed in Sec. II are used to establish the influence of different gas mixtures on the catalyst aided PECVD growth process and field emission characteristics of the graphene sheet. In particular, we investigate the consequence of different carbon feedstock gases, i.e., C_2H_2, CH_4, and CF_4, on the graphene sheet growth parameters (e.g., graphene sheet dimensions, i.e., height and thickness, growth rate of carbon species and etchant species on the catalyst surface, and graphene sheet nucleation density) with time and further its repercussions on the field enhancement factor of the graphene sheet.

Using MATHEMATICA software, we have simultaneously solved the first order differential equations (1)–(16) numerically with the appropriate default set of parameters, viz., electron number density $n_{e0}(=1\times10^9\,\mathrm{cm}^{-3})$, electron temperature $T_{e0}(=2.4\,\mathrm{eV})$, ion temperature $T_{i0}(=2400\,\mathrm{K})$, neutral temperature $T_{n0}(=2000\,\mathrm{K})$, coefficient of recombination of electrons and ions $\alpha_{A0}\approx\alpha_{B0}(=1.12\times10^{-7}\,\mathrm{cm}^3/\mathrm{s})$, $\kappa=-1.2$, and substrate temperature $T_s(=550\,^\circ\mathrm{C})$. Using Mao and Bogaerts,[30] for C_2H_2 contained gas mixture, ion number density $n_{iA0}(=4\times10^9, 3.8\times10^9, 3.6\times10^9, 1\times10^9$ and 3×10^7 for $C_2H_2^+, C_4H_2^+, C_4H_3^+, C_6H_4^+$ and C_2H^+, respectively), $n_{iB0}(=2\times10^8$ and 9×10^8 for H^+ and H_2^+, respectively), neutral atom number density $n_{A0}(=5\times10^{13}, 1\times10^{12}, 1\times10^{11}, 1.5\times10^{12}$, and 1.5×10^{10} for C_2H_2, C_4H_2, C_4H_3, C_6H_4, and C_2H, respectively), $n_{B0}(=4\times10^{13}$ and 1×10^{15} for H and H_2, respectively). For CH_4 contained gas mixture,[30] ion number density $n_{iA0}(=6\times10^8, 6.1\times10^8, 6\times10^9$, and 7×10^8 for $CH_3^+, CH_4^+, C_2H_2^+$, and $C_2H_4^+$, respectively) n_{iB0} $(=8\times10^7$ and 9×10^7 for H^+ and H_2^+, respectively), neutral atom density n_{A0} $(=5\times10^{12}, 1\times10^{14}, 7\times10^{13}$, and 8×10^{12} for CH_3, CH_4, C_2H_2, and C_2H_4, respectively), $n_{B0}(=9\times10^{13}$ and 1×10^{15} for H and H_2, respectively). The mass of ion A $m_{iA}(=(12\times x)+(1\times y)$ a.m.u for $C_xH_y^+$ and $12+(19\times x)$ a.m.u for CF_x^+), mass of ion B $m_{iB}(=(1\times x)$ a.m.u for H_x^+), mass of neutral A mA $(=(12\times x)+(1\times y)$ a.m.u for C_xH_y and $12\times(19\times x)$ a.m.u for CF_x), and mass of neutral B $(=(1\times x)$ a.m.u for H_x). For CF_4 contained gas mixture, $n_{iA0}(=2.1\times10^8, 3.1\times10^9$, and 2.6×10^9 for CF_3^+, CF_2^+, and CF^+, respectively), $n_{iB0}(=2.4\times10^8$ and 8.1×10^8 for H^+ and H_2^+, respectively), $n_{A0}(=3.5\times10^{12}, 3\times10^{12}$, and 0.7×10^{12} for CF_3, CF_2, and CF, respectively), $n_{B0}(=4.1\times10^{13}$ and 1×10^{15} for H and H_2, respectively). Other parameters used in the calculations are sticking coefficient of electron γ_e (=1), sticking coefficient of ion γ_{ij} (=1), thermal energy barrier on the catalyst surface δE_i (=1.3 eV for $C_{2n}H_x$, 2.1 eV for CH_x, and 3.3 eV for CF_x species),[46,47] carbon evaporation energy E_{ev} (=1.8 eV), adsorption energy E_{ads} (= 2.9 eV of $C_{2n}H_x$, 1.8 eV for CH_x, and 1.2 eV for CF_x species),[46] energy due to dehydrogenation δE_{sdh} (=1.7 eV), and adsorption energy of hydrogen species δE_{aH} (=1.8 eV).

Figure 2 shows the time evolution of the height of the graphene sheet for different carbon feedstocks, i.e., C_2H_2, CH_4, and CF_4 in the hydrocarbon/fluorocarbon + hydrogen + argon gas mixtures. From Fig. 2, it can be seen that the graphene sheet height is maximum for C_2H_2 followed by CH_4 and CF_4. Considering the experimental data on gas composition of CH_4 contained gas mixture for a typically PECVD case,[27–29] one may observe that the species (neutrals and ions) mainly present in the plasma are H_x (x = 1, 2), CH_x (x = 3, 4), and C_2H_x (x = 2, 4) types which have also been accounted in the present model calculations. The CH_x species can hardly dissociate on the catalyst due to high dissociation energies, whereas C_2H_x species are characterized by low dissociation energies.[10,30] Moreover, the adsorption energies for C_2H_x type species is higher than CH_x type species,[45] as a result of which the desorption term $\sum_k n_k \upsilon \exp\left(\frac{-E_{ads}}{k_B T_s}\right)$ in Eq. (4) becomes dominant for CH_x type species. Thus, C_2H_x type species present in the $CH_4/H_2/Ar$ plasmas are the main contributors of carbon species generation on the catalyst surface.[30] The similar effect can be seen from Fig. 3, which indicates that the decay rate of C_2H_x species in plasma is higher than the CH_x species.

However, according to the experimental observations,[27–32] C_2H_2 contained plasma is predominated by long chain hydrocarbon species (both ions and neutrals) of type $C_{2n}H_x$ (where, n = 1, 2 and x = 1–4). As can be seen, these species are carbon rich and hydrogen poor as compared to the species (both ions and neutrals) of CH_4+H_2+Ar plasma, as a result, they contribute more adequately to the carbon generation on the catalyst surface in contrast to CH_4 contained gas mixture. Additionally, these long chain species enhance the physical sputtering of the amorphous carbon layer formed on the catalyst surface and thus expected to prevent the poisoning of catalyst surface and maintain the catalytic activity.[30] Furthermore, there is a significant lack of carbon generation on the catalyst surface for CF_4 contained gas mixture in contrast to CH_4 and C_2H_2 [cf. Fig. 4(a)] due to their differences in dissociation energies.[48] In addition, in CF_4 growth, the catalytic activity is hampered since most of the H atoms that are responsible for removing the amorphous layer of carbon are significantly lost in the F abstraction from the CF_x species migrating on the growing surfaces and edges of the graphene sheet,[25] as a result of which the height of the graphene sheet is significantly lower than CH_4 and

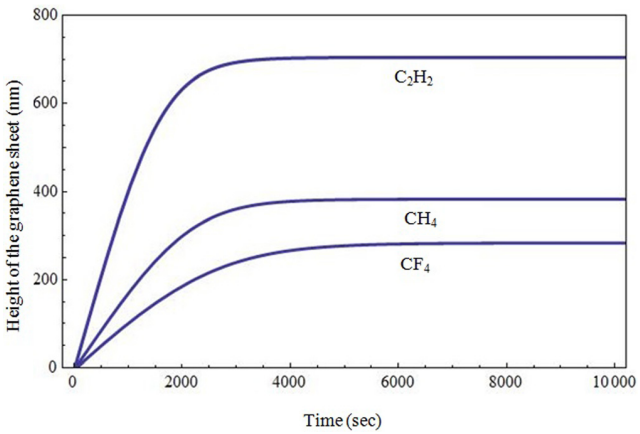


FIG. 2. The time evolution of the height (in nm) of the graphene sheet for different carbon precursor gases C_2H_2, CH_4, and CF_4. The other parameters are given in the text.

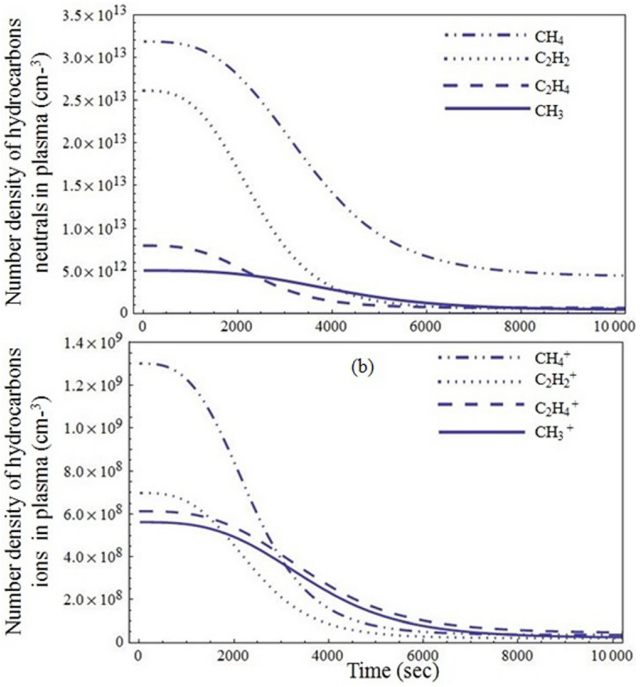


FIG. 3. The time evolution of densities of (a) hydrocarbon neutrals and (b) ions in $CH_4/H_2/Ar$ gas mixture. The other parameters are given in the text.

C_2H_2 contained gas mixtures. Since the graphene sheet grows in height due to diffusion and attachment of carbon atoms generated at the catalyst surface, one can infer that the graphene sheet grows higher for C_2H_2 contained gas mixture followed by CH_4 and CF_4 due to plentiful supply of carbon on the catalyst surface in the case of C_2H_2 contained gas mixture followed by CH_4 and CF_4.

The results of Fig. 2 are in compliance with the experimental observations of Zhu *et al.*[10] [cf. Figs. 1(b) and 1(d) of Zhu *et al.*[10]], Cai *et al.*,[15] and Shiji *et al.*[16] (cf. Fig. 4 of Shiji *et al.*[16]).

Figure 5 shows the time evolution of the thickness of the graphene sheet for various carbon gas precursors (i.e., C_2H_2, CH_4, and CF_4). It can be seen from the Fig. 5 that the thickness of the graphene sheet is highest for C_2H_2 contained gas mixture followed by CF_4 and CH_4. The reason for the growth of thicker graphene sheets in the case of C_2H_2 contained gas mixture is the plentiful generation of carbon [indicated in Fig. 4(a)] and the lesser amount of hydrogen incorporation in the C_2H_2 growth [cf. Fig. 4(b)]. From Fig. 4(b), it can be seen that the number density of hydrogen species on the catalyst surface (i.e., N_H) is utmost for CH_4 contained gas mixture followed by C_2H_2 and CF_4. This is evidence of the fact that the plasma species of CH_4 contained gas mixture are hydrogen rich and carbon poor in contrast to the plasma species of C_2H_2 and CF_4 contained gas mixtures; therefore, CH_4 contained gas mixture lead to thinner graphene sheets in contrast to C_2H_2 and CF_4. The thinner graphene sheets for CH_4 gas mixture in contrast to C_2H_2 has also been reported experimentally[10,15] [cf. Figs. 1 and 2 of Zhu *et al.*[10] and Figs. 2(b), 2(c), and 4(d) of Cai *et al.*[15]]. In CF_4+H_2+Ar plasma, N_H is least [cf. Fig. 4(b)] because of its gas composition and due to subtractive reaction of hydrogen species with F from the adsorbed CF_x species. However, the thickness of the graphene sheets is intermediate of the thickness of CH_4 and C_2H_2 contained gas mixtures due to the fact that C_2H_2 contained gas mixture is extremely carbon rich, and CH_4 contained gas mixture is exceptionally hydrogen rich in comparison with CF_4 contained gas mixture. Shiji *et al.*[16] have also reported in their experimental work that CF_4 synthesized graphene sheets are much thicker than the graphene sheets grown by using CH_4 gas mixture (cf. Fig. 3 of Shiji *et al.*[16]).

Now, we investigate the effect of different gas mixtures on the number density of vertical graphene sheets grown on catalyst-substrate surface via PECVD process. As described earlier in Sec. II, the vertical graphene sheet grows due to continuous delivery of carbon atoms, diffusion, and attachment of carbon atoms at the graphene nuclei/islands boundaries; therefore, in order to depict the impact of different gas mixtures on the number density of the graphene sheets, we study their consecutive effects on the graphene nuclei number density (i.e., N_{is}). At the early stage of growth, graphene nuclei are formed due to diffusion of carbon clusters, and with further increase in growth time, these graphene nuclei diffuse and stitch together to form large graphene islands, and thus, the number density of graphene nuclei decreases

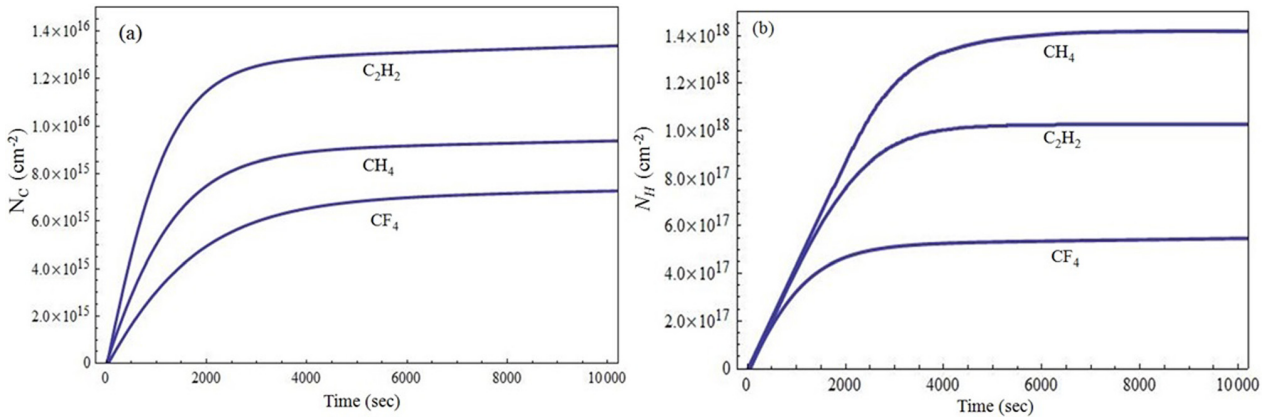


FIG. 4. The time evolution of number densities of (a) carbon species (i.e., N_C) and (b) hydrogen species (i.e., N_H) generated per unit area on the catalyst surface for different carbon precursor gases C_2H_2, CH_4, and CF_4. The other parameters are given in the text.

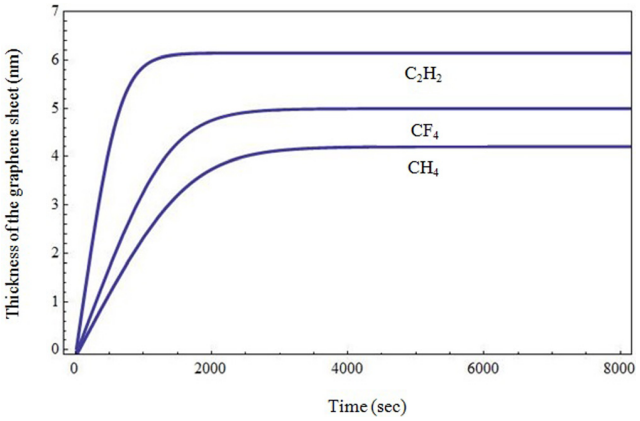


FIG. 5. The time evolution of the thickness (in nm) of the graphene sheet for different carbon precursor gases C_2H_2, CH_4, and CF_4. The other parameters are given in the text.

with growth time as can be seen from Fig. 6. Figure 6 illustrates the variation of graphene nuclei density as a function of growth duration for different gas mixtures (i.e., C_2H_2+H_2+Ar, CH_4+H_2+Ar, and CF_4+H_2+Ar). It can be seen from Fig. 6 that C_2H_2 contained gas mixture leads to more effective formation of graphene nuclei, i.e., graphene nucleation density is maximum in C_2H_2 in contrast to CH_4 and CF_4. The relative generation of carbon atoms on the catalyst surface (i.e., N_c) plays an imperative role in this regard. This is because, more carbon atoms on the catalyst-substrate surface give rise to faster saturation of the surface with carbon atoms and eventually increases the cluster formation, which in turn, enables faster nucleation of more graphene nuclei on the catalyst surface. Consequently, diffusion and coalescence of graphene nuclei/islands become more effective and more graphene islands of large size are formed, and therefore, the number density of graphene sheets increases. Accordingly, using Fig. 3 it can be concluded that the graphene sheet number density is highest in C_2H_2 contained gas mixture followed by CH_4 and then by CF_4. The similar observation has been reported by Shiji *et al.*[16] (cf. Fig. 3 of Shiji *et al.*[16]).

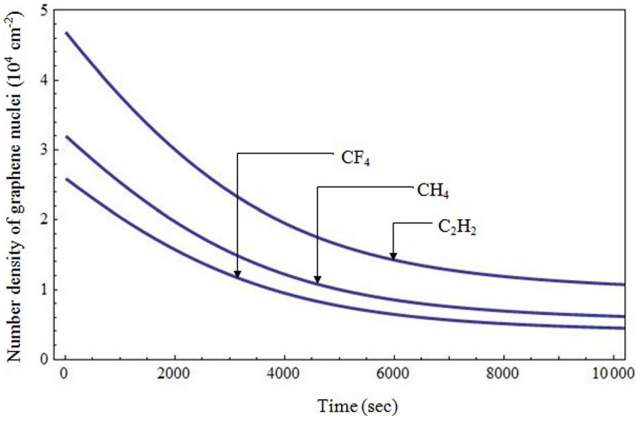


FIG. 6. The time evolution of graphene nuclei density for different carbon precursor gases C_2H_2, CH_4, and CF_4. The other parameters are given in the text.

Using the earlier results, we can conclude that geometry (i.e., height and thickness) of the graphene sheet alters considerably upon its PECVD synthesis using different gas mixtures. The field enhancement factor $\beta(\approx \frac{h_{gn}}{t_{gn}}$, where h_{gn} is the height and t_{gn} is the thickness of GS),[49,50] characteristic of electron field emission from the graphene sheet, is primarily dependent on the geometry of the graphene sheet. Thus, using the above results, we can interpret the field emission characteristics of the graphene sheet grown using different gas mixtures via PECVD process. In order to estimate β, we calculated the saturated value of thickness (i.e., 6.15 nm, 4.2 nm, and 4.95 nm) and height (i.e., 704 nm, 382 nm, and 277.54 nm) of the graphene sheet corresponding to C_2H_2, CH_4, and CF_4 containing gas mixtures, respectively. It can be seen from Fig. 7 that graphene sheet grown using C_2H_2 as a carbon source gas has highest field enhancement factor which is followed by CH_4 and then by CF_4. Hence, C_2H_2 containing gas mixture improves the field emission properties in contrast to CH_4 and CF_4. Although the graphene sheets grown using CH_4 and CF_4 are relatively thinner, C_2H_2 grown graphene sheet exhibits high field emission characteristics due to the fact that graphene sheets grown in this case are relatively much higher in height compared to CH_4 and CF_4. The above observations have been validated by Zhu *et al.*[10] and Saito.[51]

IV. CONCLUSION

A theoretical model based on surface deposition processes and plasma kinetics has been developed with the aim to elucidate the key processes in the PECVD growth of the graphene sheet and to study the effect of different gas mixtures, particularly different carbon source gases (i.e., C_2H_2, CH_4, and CF_4) on the nucleation and growth of the graphene nanosheet. It is observed that carbon generation on the catalyst surface is highest for C_2H_2 and least for CF_4, as a result of which the height, thickness, and number density of graphene nanosheets are maximum for C_2H_2. However, the thickness of the graphene nanosheet is least for CH_4 due to higher generation of etching species, i.e., hydrogen on the

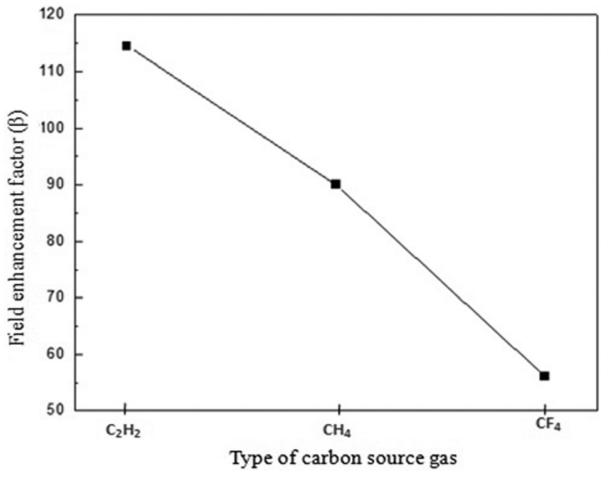


FIG. 7. The field enhancement factor of the graphene sheet for different carbon precursor gases C_2H_2, CH_4, and CF_4.

catalyst surface. From these results, we have also estimated the consequence of these carbon source gases on the field emission characteristics of the graphene sheet. It is found that the field enhancement factor $\beta(\approx \frac{h_{gn}}{t_{gn}}$, where h_{gn} is the height and t_{gn} is the thickness of GS$)$ is highest for C_2H_2 followed by CH_4 and CF_4. The present work would be significant for the development of efficient graphene based field emitters.

ACKNOWLEDGMENTS

This work has been partially supported by the Science and Engineering Research Board (SERB)–Department of Science and Technology (DST), Government of India.

[1]K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, Science **306**, 666 (2004).
[2]S. Gilje, S. Han, M. Wang, K. L. Wang, and R. B. Kaner, Nano Lett. **7**, 3394 (2007).
[3]M. Sprinkle, M. Ruan, Y. Hu, J. Hankinson, M. R. Roy, B. Zhang, X. Wu, C. Berger, and W. A. de Heer, Nat. Nanotechnol. **5**, 727 (2010).
[4]S. Stankovich, D. A. Dikin, R. D. Piner, K. A. Kohlhaas, A. Kleinhammes, Y. Jia, Y. Wu, S. B. T. Nguyen, and R. S. Ruoff, Carbon **45**, 1558 (2007).
[5]K. S. Subrahmanyam, L. S. Panchakarla, A. Govindaraj, and C. N. R. Rao, J. Phys. Chem. C **113**, 4257 (2009).
[6]I. Vlassiouk, S. Smirnov, S. P. Surwade, M. Regmi, N. Srivastava, R. Feenstra, G. Eres, C. Parish, N. Lavrik, P. Datskos, S. Dai, and P. Fulvio, J. Phys. Chem. C **117**, 18919 (2013).
[7]L. Jiang, T. Yang, F. Liu, J. Dong, Z. I. Yao, C. Shen, S. Deng, N. Xu, Y. Liu, and H.-J. Gao, Adv. Mater. **25**, 250 (2013).
[8]J. J. Wang, M. Y. Zhu, R. A. Outlaw, X. Zhao, D. M. Manos, B. C. Holloway, and V. P. Mammana, Appl. Phys. Lett. **85**, 1265 (2004).
[9]A. Malesevic, R. Kemps, A. Vanhulsel, M. P. Chowdhury, A. Volodin, and C. V. Haesendonck, J. Appl. Phys. **104**, 084301 (2008).
[10]M. Y. Zhu, R. A. Outlaw, M. B. Hansen, H. J. Chen, and D. M. Manos, Carbon **49**, 2526 (2011).
[11]Z. Bo, S. Mao, Z. J. Han, K. Cen, J. Chen, and K. Ostrikov, Chem. Soc. Rev. **44**, 2108 (2015).
[12]S. C. Sharma and N. Gupta, Phys. Plasmas **22**, 123517 (2015).
[13]L. V. Nang and E.-T. Kim, J. Electrochem. Soc. **159**, K93 (2012).
[14]W. Takeuchi, H. Sasaki, S. Kato, S. Takashima, M. Hiramatsu, and M. Hori, J. Appl. Phys. **105**, 113305 (2009).
[15]M. Cai, R. A. Outlaw, R. A. Quinlan, D. Premathilake, S. M. Butler, and J. R. Miller, ACS Nano **8**, 5873 (2014).
[16]K. Shiji, M. Hiramatsu, A. Enomoto, M. Nakamura, H. Amano, and M. Hori, Diamond Relat. Mater. **14**, 831 (2005).
[17]S. Kondo, S. Kawai, W. Takeuchi, K. Yamakawa, S. Den, H. Kano, M. Hiramatsu, and M. Hori, J. Appl. Phys. **106**, 094302 (2009).
[18]D. H. Seo, A. E. Rider, Z. J. Han, S. Kumar, and K. Ostrikov, Adv. Mater **25**, 5638 (2013).
[19]M. Y. Zhu, J. J. Wang, B. C. Holloway, R. A. Outlaw, X. Zhao, K. Hou, V. Shutthanandan, and D. M. Manos, Carbon **45**, 2229 (2007).
[20]K. Ostrikov, H.-J. Yoona, A. E. Rider, and S. V. Vladimirov, Plasma Processes Polym. **4**, 27 (2007).
[21]Z. Bo, Y. Yang, J. Chen, K. Yu, J. Yan, and K. Cen, Nanoscale **5**, 5180 (2013).
[22]A. N. Obraztsov, A. A. Zolotukhin, A. O. Ustinov, A. P. Volkov, Y. Svirko, and K. Jefimovs, Diamond Relat. Mater. **12**, 917 (2003).
[23]J. Zhao, M. Shaygan, J. Eckert, M. Meyyappan, and M. H. Rummeli, Nano Lett. **14**, 3064 (2014).
[24]K. Teii, S. Shimada, M. Nakashima, and T. H. Chuang, J. Appl. Phys. **106**, 084303 (2009).
[25]M. Hiramatsu and M. Hori, Jpn. J. Appl. Phys., Part 1 **45**, 5522 (2006).
[26]N. Gupta, S. C. Sharma, and R. Sharma, Phys. Plasmas **23**, 083509 (2016).
[27]D. J. Dagel, C. M. Mallouris, and J. R. Doyle, J. Appl. Phys. **79**(11), 8735 (1996).
[28]I. B. Denysenko, S. Xu, J. D. Long, P. P. Rutkevych, N. A. Azarenkov, and K. Ostrikov, J. Appl. Phys. **95**, 2713 (2004).
[29]C. Deschenaux, A. Affolter, D. Magni, C. Hollenstein, and P. Fayet, J. Phys. D: Appl. Phys. **32**, 1876 (1999).
[30]M. Mao and A. Bogaerts, J. Phys. D: Appl. Phys. **43**, 205201 (2010).
[31]J. Benedikt, J. Phys. D: Appl. Phys. **43**, 043001 (2010).
[32]J. R. Doyle, J. Appl. Phys. **82**, 4763 (1997).
[33]K. Miyata, M. Hori, and T. Goto, J. Vac. Sci. Technol., A **14**, 2343 (1996).
[34]H. Mehdipour and K. Ostrikov, ACS Nano **6**, 10276 (2012).
[35]B. B. Wang, K. Zheng, Q. J. Cheng, and K. Ostrikov, Appl. Surf. Sci. **325**, 251 (2015).
[36]H. Mehdipour, K. Ostrikov, and A. E. Rider, Nanotechnology **21**, 455605 (2010).
[37]M. A. Lieberman and A. J. Lichtenberg, *Principles of Plasma Discharges and Materials Processing* (Wiley Interscience Publication, USA, 1994).
[38]M. Chhowalla, K. B. K. Teo, C. Ducati, N. L. Rupesinghe, G. A. J. Amaratunga, A. C. Ferrari, D. Roy, J. Robertson, and W. I. Milne, J. Appl. Phys. **90**, 5308 (2001).
[39]R. Gupta, S. C. Sharma, and R. Sharma, Plasma Sources Sci. Technol. **26**, 024006 (2017).
[40]Z. Marvi, S. Xu, G. Foroutan, and K. Ostrikov, Phys. Plasmas **22**, 013504 (2015).
[41]N. V. Mantzaris, E. Gogolides, A. G. Boudouvis, A. Rhallabi, and G. Turban, J. Appl. Phys. **79**, 3718 (1996).
[42]B. B. Wang, K. Ostrikov, T. van der Laan, K. Zheng, J. J. Wang, Y. P. Yan, and X. J. Quan, J. Mater. Chem. C **1**, 7703 (2013).
[43]D. H. Seo, S. Kumar, A. K. Rider, Z. Han, and K. Ostrikov, Opt. Mater. Express **2**, 700 (2012).
[44]M. S. Sodha, S. Misra, S. K. Misra, and S. Srivastava, J. Appl. Phys. **107**, 103307 (2010).
[45]O. A. Louchev, C. Dussarat, and Y. Sato, J. Appl. Phys. **86**, 1736 (1999).
[46]I. Denysenko and N. A. Azarenkov, J. Phys. D: Appl. Phys. **44**, 174031 (2011).
[47]H. Maeda, M. Irie, T. Hino, K. Kusakabe, and S. Morooka, Diamond Relat. Mater. **3**, 1072 (1994).
[48]I. Schmidt, F. Hentschel, and C. Benndorf, Diamond Relat. Mater. **5**, 1318 (1996).
[49]S. Watcharotone, R. S. Ruoff, and F. H. Read, Phys. Procedia **1**, 71 (2008).
[50]R. Miller, Y. Y. Lau, and J. H. Booske, Appl. Phys. Lett. **91**, 074105 (2007).
[51]*Carbon Nanotube and Related Field Emitters: Fundamentals and Applications*, edited by Y. Saito (John Wiley & Sons, 2010).

Jestr

Journal of Engineering Science and Technology Review 10 (3) (2017) 115-122

Research Article

JOURNAL OF
Engineering Science and Technology Review

www.jestr.org

Efficient Color Cast Correction Based on Fuzzy Logic

Om Prakash Verma[1] and Nitin Sharma[2,*]

[1] *Department of Computer Science Engineering, Delhi Technological University, Delhi, India*
[2]*Department of Electronic and Communication Engineering, Maharaja Agrasen Institute of Technology, Rohini, Delhi, India*

Received 31 October 2016; Accepted 6 July 2017

Abstract

Color cast defect detection of an image is challenging and of great importance in the field of image quality evaluation. This paper presents a simple and less complex method to improve the visibility in color cast images. An objective measure called color cast could be independent of the illumination condition. There are different assumptions present in the literature to provide an estimate of the Red cast, Blue cast and Green cast present in the image. These assumptions sometimes failed to eliminate the effect of color cast under varying illumination condition. The proposed method performs well under these conditions which may involve two steps. Firstly it performs the color cast detection and thereafter, a power law transformation is introduced to achieve the true color distribution of the image. In our approach a new transformation method which corrects the color cast as well as improves the quality of image. The uncertainty in the color distribution due to the illumination is removed on the basis of fuzzy Gaussian membership functions and further, improves by using weighing distribution. The performance analysis of the proposed approach is carried out by using two objective measures the image distance metric criteria of CIELAB color space and the image quality assessment function. The effectiveness of the proposed approach is illustrated by minimizing the image distance between the reference image and corrected image. The computed higher value of the image quality assessment function obtained in the proposed algorithm application on images, also justifies that the proposed algorithm results the better quality of the image in comparison to existing methods.

Keywords: Color, Fuzzy Logic, Correction

1. Introduction

The color cast quality metric of images is an important feature in the application area like digital camera variants, biomedical and other practical problems field. The main cause of change in color information of the image is the variation in source illumination. The atmospheric conditions are random in nature. Thus, an image captured with camera at variant weather conditions like underwater images results the deviation in the color quality, visibility, contrast and brightness. The human visual system (HVS) have inherent ability to adjust the said deviation to a certain limit to minimize the impact of different light conditions on the color and hence can perceived the true color. The image acquisition devices at the front end of monitoring system, however, do not have this automatic adjustment ability. Thus, it is an essential objective to perform color correction in order to reproduce the same color image formation in the presence of different light characteristics.

There are some automated tools are also present to correct the color cast like adobe Photoshop, Picasa etc. But the automation doesn't work as desired. It is difficult to control the effect of color cast due to the illumination conditions. Many researchers worked on color cast problem [1-6]. The existing techniques used for color correction are max-RGB [3, 4], Grey World assumption [5], Grey Edge [6], Shades of Grey [7] etc. It may not provide a better solution in terms of removing a particular cast like red, green or blue. Although these techniques can be effective to solve the problem of color cast. Single parameter cannot control the color correction at different level. These techniques may be extended by using the combination with different weights [8]. The weights are applied with the combination of different algorithm. The computation of weights is complex and computationally expensive. These weights are dependent on the classification of images. This classification involved high dimension problem. Jonathan et al. [9] proposed the selection of white patch, the gray world and the max RGB color correction algorithms on dark images in accordance with the actual content of an image but not much effective for the color cast images. Since all algorithm developed have certain assumptions and hence, no algorithm is to be called as a universally true algorithm. In the weighting approach, finding suitable weights for combination is a difficult and error prone task. In other word, it is important to combine the algorithms that have good performance on the input image.

More recently, for color image enhancement [11], O. P. Verma et al. [12] proposed to learn the parameters of membership function used for fuzzification and Rajni et al. [13] proposed the color correction for underwater images. Since the systems based on fuzzy logic are resembled with human decision. In this work, we have proposed a fuzzy based color cast correction algorithm in RGB model that perform an efficient color cast detection and correction in compare to the state of art methods. Firstly, the color cast detection using fuzzy inference system (FIS) is evaluated and thereafter, a new power law transformation is introduced to achieve true color distribution of the image. The uncertainty in the color distribution due to the illumination is removed on the basis of fuzzy Gaussian membership functions and further, improves by using

*E-mail address: sharmaisnitin@gmail.com
ISSN: 1791-2377 © 2017 Eastern Macedonia and Thrace Institute of Technology. All rights reserved.
doi:10.25103/jestr.103.16

weighing distribution. The color distributions are balanced by defining the fuzzy rule based system. By controlling the weights depending on regional properties, we investigate that a visually better result is obtained. The basic idea is to give larger weights to the pixels in noticeable regions where the color cast is high, so that well captured regions are kept intact and others are changed according to the manipulated regions.

The paper is as organized as follow. The existing algorithms for color cast correction are described in Section 2. Section 3 discusses the fuzzy logic and fuzzy rules. Later in Section 4, the proposed algorithm is applied for color correction together with state of the art comparisons and provides significant results. Finally, conclusions drawn from the results obtained are mentioned in Section 5.

2. Color Cast Correction Algorithm

In this section, we discussed the estimation of illumination of different color constancy methods are discussed. The limitation of gray world based approach merits and limitations are also discussed.

2.1 Estimation of illumination based algorithms

Image is formed by multiplication of illumination component and reflectance component. To estimate the illuminated component which depends on the illuminant spectral power distribution P(λ) and camera sensitivity c(λ). The image $I_k(x,y)$ is given by

$$I_k(x,y)=\int_{\text{visible range}} P(\lambda)c(\lambda)\,d\lambda \tag{1}$$

For Lambertian surface, the image $I_{,}(x,y)=(R,G,B)$ are dependent on the light source $e(\lambda)$, the surface reflectance $S(x,y,\lambda)$ and the camera sensitivity functions $c(\lambda)$ where λ is the wavelength. Thus, the image $I_k(x,y)$ is given by,

$$I_k(x,y)=\int_{\text{visible range}} S(x,y,\lambda)e(\lambda)c(\lambda)\,d\lambda \tag{2}$$

As the integral term involves two unknown terms e(λ) and S(x, y, λ), thus, integral becomes an under constrained problem, that means integral cannot be solved without any further assumptions. Many color constancy methods have been proposed to estimate the illuminant. The grey world algorithm is just based on the grey world assumption, i.e. the average reflectance in a scene is achromatic. Finlayson et al. [16] proposed a more general color constancy algorithm called Shades of grey (SoG), which can include grey world and Max RGB based on the Minkowski-norm. The general estimation approach can be expressed as

$$\left(\int\left|\frac{\partial^n I^\sigma(x)}{\partial x^n}\right|^p dx\right)^{\frac{1}{p}} = Ae^{n,p,\sigma} \tag{3}$$

Where, *n* is the order of spatial derivative, *p* is the Minkowski norm which measures the relative weights and *σ* is the parameter for smoothing the image. When n is zero, *σ* is zero and p=1 then the expression (3) represents the gray world assumptions, otherwise if it has nonzero finite value then expression represents the shades of gray. If p becomes infinite then expression is of a max-RGB.

2.2. Gray World Based Assumption Algorithm

The assumption made in the gray world algorithm is that the given color image has a sufficient amount of color variations. In the present application we have considered Red (R) Green (G) and Blue (B) color model of color space. The mean value of each R, G, and B component of the image should have a common mean intensity of the image. In a real world, this assumption for images may hold true. Color balancing algorithms can also create images to have a uniform gray value for its R, G, and B channel. If the image is captured by a digital camera under a particular lighting condition, the effect of the special lighting cast can be removed by taking the mean value of each channel. As a result, it estimates the color of the image linearly which is much closer to the original scene. In terms of implementation, the mean value $\overline{R}$, $\overline{G}$ and $\overline{B}$ of red, green and blue channel respectively, of a color image of size M x N, can be calculated as

$$\overline{R}=\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}R(i,j) \tag{4}$$

$$\overline{G}=\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}G(i,j) \tag{5}$$

$$\overline{B}=\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}B(i,j) \tag{6}$$

Where, i and j denotes the indices of pixel position. The cast arises here is due to non-linear behaviour of the different illumination conditions. This problem cannot completely solve with the gray world assumption.

The mean intensity of an image is given by

$$\overline{W}=(\overline{R}+\overline{G}+\overline{B})/3 \tag{7}$$

3. Proposed Algorithm

At the time of image acquisition, each camera uses its own RGB space that is converted to a standard RGB space later after achieving color constancy, this shows that the transformation between R/G/B spaces, which requires calibrating the imaging device. Therefore, it will inevitably fail on those images from any unknown camera. The proposed approach first detects the presence of dominance of color in the given image without any pre-assumptions and then corrects the color cast in the image in the same space. We first detect the color cast of the image depending on the color dominance presence. Most of the algorithms used for finding the cast are depends on constant value as given in equation (3).The constant value based on the assumption. The effect of color cast is uncertain which is perceived by the human eye. The vagueness in perception is increased under varying illumination condition. Here, we reduced the vagueness of perception by using Fuzzy rules. These rules calculate the color cast on the basis of intensity value. Thereafter, remove the color cast in such a way that the true distribution of colors present in the image does not get altered. Further, we applied a correction factor which behaves non-linearly for obtained the color corrected enhanced image. The correction factors have been formed with the combination of base and the weighting distribution. The weighting distribution is introduced to reduce the effect of color cast. Finally, the color corrected enhanced image is obtained.

3.1 Color Cast Detection

In order to detect the dominant color, first we need to calculate the normalized value of R, G, and B channel from the given input color cast image henceforth referred to as "test image". Then the mean intensity of each color channel $\bar{R}$, $\bar{G}$ and $\bar{B}$ is calculated. In this work, the fuzzy approach is included for the decision of dominance of color presence. Each color channel is fuzzified by using Gaussian membership function. We use Gaussian membership function as shown in Fig.1 to avoid boundary value and represent vague or linguistic terms. Gaussian membership function is given by:

$$\mu_A(x)=\exp\left(\frac{(c-x)^2}{2\sigma^2}\right) \quad (8)$$

Here, c and σ are centre and width of the fuzzy set A and $x \in A$

The maximum number of pixels belongs to either higher or lower side of membership function, the Gaussian fuzzy sets where, $x \in \{0,1\}$ and standard deviation σare 0.19667. The fuzzy rules of the proposed Mamdani fuzzy inference system are defined in Table 1.

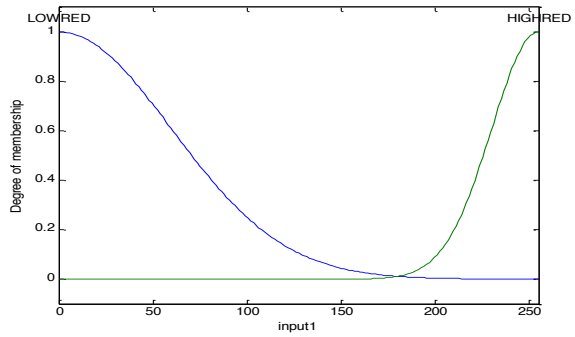


Fig. 1. Gaussian membership function curve for the LOW and the HIGH intensity values

Table 1. Color cast determination fuzzy rules

S.No.	R(i,j)	G(i,j)	B(i,j)	Color cast
1	HIGH	LOW	LOW	Red
2	LOW	HIGH	LOW	Green
3	LOW	LOW	HIGH	Blue
4	HIGH	HIGH	LOW	Red and Green
5	LOW	HIGH	HIGH	Green and Blue
6	HIGH	LOW	HIGH	Red and Blue

Generally, the intensity levels are not well spread with the correction of color cast. The pixel intensity as shown in table1 defines the belonging of the cast. The cast of each pixel evaluates by using membership function. The higher the pixel intensity greater the membership value and vice versa. After evaluation of color cast, we find the mean value of each channel as defined in equation (4), equation (5) and equation (6). Now depending on the region, a proposed base value is defined in Table 2.

Table 2. Proposed base value for the different region of casts

S.No.	Region	Base
1	Region 1	$\bar{R}$
2	Region 2	$\bar{G}$
3	Region 3	$\bar{B}$
4	Region 4	$\frac{\bar{R}+\bar{G}}{2}$
5	Region 5	$\frac{\bar{G}+\bar{B}}{2}$
6	Region 6	$\frac{\bar{R}+\bar{B}}{2}$

Accordingly, the color cast of the image is obtained as an output after the defuzzification and aggregation process using the above defined fuzzy rules. The output membership function is shown in Fig. 2. The rules depend on the degree of membership function. The overall cast evaluates the rule under membership function. The higher number of pixels decides the cast of the image.

Color Cast Correction- After the finalization of cast type, color cast of image is corrected by applying correction factor α_r, α_g and α_b on red, green and blue channel respectively are obtained from equation given by,

$$\alpha_r=\frac{\bar{R}}{\text{base}},\ \alpha_g=\frac{\bar{G}}{\text{base}} \text{ and } \alpha_b=\frac{\bar{B}}{\text{base}} \quad (9)$$

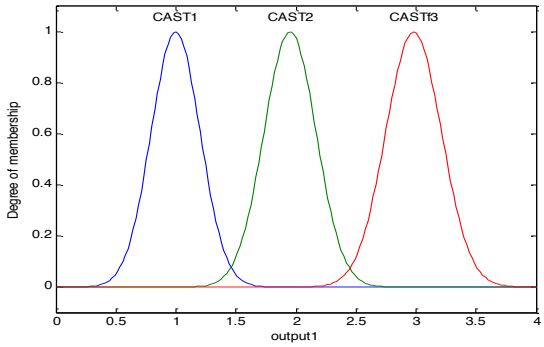


Fig. 2. Gaussian membership function for output color cast- Red cast, Green cast and Blue cast

These factors are depending on the base of the channel. This base has the value of color cast obtained in the first operation of proposed approach that is the color cast detection. This color cast effect is corrected by using the equation (10).

$$\tilde{R}(i,j)=R(i,j)^{\blacklozenge_r/w_r},\ \tilde{G}(i,j)=G(i,j)^{\blacklozenge_g/w_g}$$
$$\text{and } \tilde{B}(i,j)=B(i,j)^{\blacklozenge_b/w_b} \quad (10)$$

Where, w_r, w_g, and w_b are the color correction factor that reduced the effect of color cast. The power law coefficient is determined by the dominant color detected by fuzzy color correction factor to the weighted factor given in equation (11). Whereas, correction factor is obtained by computing the ratio of the mean value of intensity of color channels present on base depend in the actual image. The value of w_1, w_2, and w_3 is chosen as per the type of cast. If the resulted image is Red cast then $w_r=\bar{w}$ and $w_g=1$ and $w_b=1$, image is Green cast then $w_r=1$, $w_g=\bar{w}$ and $w_b=1$. Similarly in case of blue cast $w_r=1$, $w_g=1$ and $w_b=\bar{w}$. This weighing factor $\bar{w}$ is determined from equation (7). This factor gives mean intensity of the test image. Mathematically the weighing factor is calculated as

$$if\ \alpha_i=1,\ then\ w_i=\bar{W}\ and\ w_j=1, j\neq i.\ where\ i,j\in\{r,g,b\}$$

Finally, the corrected pixel values $\tilde{R}(i,j)$, $\tilde{G}(i,j)$ and $\tilde{B}(i,j)$ of each color channel is obtained. This law maintains the correction

with wide intensity range. We calculate the correction factor as in the gray world based correction but apply directly on each intensity value non- linearly. Hence the color cast corrected image is obtained.

Proposed Approach

The algorithm is outlined as follows:-

1. Obtain the RGB test image
2. Normalize each channel
3. Obtain the mean value of each channel
4. Determine the cast type by passing the mean value of each channel as argument to the Mamdani fuzzy Inference system designed. Thus, de-fuzzify the image into red cast, green cast and blue cast.
5. Define the base value as per the cast type resulted from step 4
6. Obtain the value of color correction factor for each channel α_r α_g nd α_b respectively using equation (9)
7. Finally obtain the corrected value of each channel i.e. $\tilde{R}(i,j)$., $\tilde{G}(i,j)$., $\tilde{B}(i,j)$ .using equation (10)

4. Performance Measures:-

The two image quality performance measures are considered to evaluate the effectiveness of the proposed approach. These performance measures are comprehensive image quality assessment model and CIELAB metric.

Comprehensive Image Quality Assessment Model:-The color quality of the image is measured subjectively as well as objectively. The subjective evaluation is based on the human perception of the image. It may vary from individual to individuals. The objective evaluation is more reliable, consistent and replicable especially when the goal is to compare with the other existing approaches. The comprehensive image quality parameters are proposed in the literature [14].A comprehensive image quality assessment model, called as the best quality assessment function for the color image was established. The quality parameters like luminance, contrast and entropy are quality evaluator and are described below Average Luminance

The average luminance of an image $\overline{L}$ is given by:-

$$\overline{L}=\frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}L(i,j) \tag{11}$$

Where, M and N are the number of rows and columns of an image respectively.

Average Information Entropy

In color images, information entropy evaluates for each channel red, green and blue is given by:-

$$H=-\sum_{i=0}^{255}p(i)\log_2 p(i) \tag{12}$$

where, $p(i)$ is the probability distribution of the pixel information for three bit image. Thereafter, take the average information of three channels ($\overline{H}$).

Average Contrast:

The color contrast can be calculated for each RGB channel are given by,

$$\vec{C_c}(i,j)=\nabla R(i,j)\hat{x}+\nabla G(i,j)\hat{y}+\nabla B(i,j)\hat{z} \tag{13}$$

Where, ∇R , ∇G and ∇B are the gradient of red, green and blue channel respectively. The total contrast computes by taking the average of these channels.

$$C_c=\left\|\vec{C}_C\right\|=\sqrt{\frac{\nabla^2 R+\nabla^2 G+\nabla^2 B}{3}} \tag{14}$$

Table 3. Comparison of luminance, entropy, average contrast and CAF value for two different images using proposed, grey world shades of gray and Max-RGB methods

			Average Luminance $\overline{L}$	Average Entropy $\overline{H}$	Average Contrast $\overline{C}$	CAF
Two men	Red Cast	Test Image	105.874	7.6176	117.6348	14.3537
		Proposed Approach	121.4628	7.8592	128.8760	22.8940
		Grey world	114.6916	7.4689	126.3866	18.2284
		Max-RGB	105.8476	7.5950	116.4403	14.2747
		Shades of Grey	106.1014	7.7031	123.5935	14.8013
	Green Cast	Test Image	154.3369	7.7214	107.404	12.2340
		Proposed Approach	139.3876	7.7641	113.2544	18.8837
		Grey world	147.5456	7.6409	103.9105	14.6033
		Max-RGB	154.2861	7.7219	107.5911	12.2576
		Shades of Grey	150.8933	7.7248	106.0571	13.4952
	Blue Cast	Test Image	95.2479	7.6093	124.4972	10.5967
		Proposed Approach	122.7724	7.8828	127.7406	23.6610
		Grey world	116.2059	7.4491	137.1300	19.2994
		Max-RGB	95.3044	7.5772	121.0267	10.4963
		Shades of Grey	101.4161	7.6467	130.5941	13.0091
Color checker	Red Cast	Test Image	117.0651	6.9913	44.2898	18.0358
		Proposed Approach	136.8766	7.3210	51.4229	19.6046
		Grey world	126.2622	7.1793	53.9270	19.4550

		Max-RGB	116.9710	7.0403	46.3699	18.3716
		Shades of Grey	117.4282	7.1331	49.4301	18.9138
	Green Cast	Test Image	139.2496	7.0017	42.9251	17.92191
		Proposed Approach	132.1887	7.3055	49.7295	19.4002
		Grey world	124.4725	7.1858	48.8869	19.0008
		Max-RGB	136.9524	7.0310	43.6200	18.0696
		Shades of Grey	125.7044	7.1179	45.1199	18.4477
	Blue Cast	Test Image	101.8550	7.1648	47.7581	18.8349
		Proposed Approach	132.1369	7.3283	51.2518	19.6078
		Grey world	130.3048	6.1334	45.3594	15.9172
		Max-RGB	125.5105	6.1117	36.1218	14.9832
		Shades of Grey	122.2703	6.1470	39.4600	15.4065

Table 4. Comparison of ΔE CIELAB metric for Proposed Approach, grey world, max-RGB and shades of grey

		ΔE_{rc} (test)	ΔE_{re} (proposed Approach)	ΔE_{rg} (grey)	ΔE_{rm} (max-RGB)	ΔE_{rc} (shades of grey)
Two men Image	Red Cast	0.1147	0.0738	0.0922	0.1146	0.0774
	Green Cast	0.0968	0.0577	0.0854	0.0983	0.0896
	Blue Cast	0.2241	0.0479	0.1329	0.2227	0.1607
Color checker Image	Red Cast	0.1996	0.0637	0.1041	0.1992	0.1318
	Green Cast	0.1414	0.0602	0.0724	0.1412	0.1030
	Blue Cast	0.2571	0.1019	0.1425	0.2572	0.1308

The average color contrast for each channel of an RGB image are given by

$$\overline{C} = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1} C_C(i,j) \quad (15)$$

Where, $\overline{C}$ represents average contrast of the image. The color image quality assessment function defined in [14].

$$CAF = \overline{H}^{\alpha} + \overline{C}^{\beta} + \overline{N} \quad (16)$$

The higher CAF value gives better image quality. This value is computed by adding the parameters like luminance, entropy, contrast and noise.

CIELAB Metric: -This metric estimates accuracy of the color reproduction in comparison to the original when analysed by a human observer [15]. This metric is suitable for measuring color difference of large uniform color targets. CIELAB is based on one channel for Luminance (L) and two color channels (a and b). The a-axis starts from green (-a) to red (+a) and the b-axis starts from blue (-b) to yellow (+b). The Luminance (L) starts from the bottom to the top of the three-dimensional model. The ΔE_c and ΔE_e metric are given by

$$\Delta E_c = \sqrt{\Delta L_c^2 + \Delta a_c^2 + \Delta b_c^2} \quad (17)$$

where,

$$\Delta L_c = L_{cast} - L_{original}$$
$$\Delta a_c = a_{cast} - a_{original}$$
$$\Delta b_c = b_{cast} - b_{original}$$

$$\Delta E_e = \sqrt{\Delta L_e^{\,2} + \Delta a_e^{\,2} + \Delta b_e^{\,2}} \quad (18)$$

where,

$$\Delta L_e = L_{enhanced} - L_{original}$$
$$\Delta a_e = a_{enhanced} - a_{original}$$
$$\Delta b_e = b_{enhanced} - b_{original}$$

Parameters ΔL_c, Δa_c and Δb_c is the difference in the cast and test image coordinates of L, a, and b of CIELAB color space and ΔL_e, Δa_e and Δb_e are the coordinates L, The 'a' and 'b' of enhanced image in CIELAB color space. ΔE is the Euclidean distance for measuring the difference between colours. Smaller value of ΔE indicates that the enhanced test image is closer to the reference image.



(a)



Fig. 3. Two-men image (a) Reference, (b) Red cast (c) Green cast, (d) Blue cast, (e) proposed corrected Red cast (f) proposed corrected Green cast (g) proposed corrected Blue cast (h) Grey world corrected Red cast (i) Grey world corrected Green cast (j) Grey world corrected Blue cast (k) max-RGB corrected Red cast (l) max-RGB corrected Green cast Two-men image (m) max-RGB corrected Blue cast, (n) shades of grey corrected Red cast, (o) shades of grey approach corrected Green cast (p) shades of grey approach corrected Blue cast

5. Results and Discussions

In this section experimental results are presented to prove the effectiveness of the proposed color correction algorithm on the natural images under red, green and blue cast. The proposed approach has been successfully implemented on Intel core i3 at 2.40 GHz using MATLAB version 2014. All images used in experiment are cast using paintshop pro software and evaluate the luminance, entropy, average contrast and CAF values. The color corrected images are shown in Fig. 3 and Fig. 4.The reference image is the original image without any bad illumination effect. The distorted images are obtained by adding color cast to them. The distorted images are corrected by some well-known methods in literature for color correction: max-RGB, Grey World assumption, Grey Edge, Shades of Gray and the proposed corrected approach. The image quality parameters: luminance, entropy, average contrast and CAF value for two different images are summarized in Table-3. The results show a good match with human ground truth with proposed technique. The higher entropy value is obtained as compared with existing methods. An additional performance measure ΔE had been taken in which image distances are calculated using equation (17), equation (18) and are summarized in Table-4. It is observed that the distance is significantly reduced with the proposed approach. The reference two men image is shown in Fig. 3(a). The red, green and blue cast images are shown in Fig.3 (b), (c) and (d) respectively having entropy of 7.6176, 7.7214 and 7.6093 respectively, are efficiently corrected with proposed technique. The images shown in Fig.3 (e), (f) and (g) are the resulted color cast corrected images using proposed technique.

It is clearly observed that proposed technique removes the color cast correctly and it is confirmed from the metric results

obtained that is the entropy value of the corrected images obtained are 7.8592, 7.7614 and 7.8828 respectively. The overall contrast values of the test images are 14.3537, 12.2340 and 10.5967.And these values are significantly enhanced in the corrected images to 22.8940, 18.8837 and 23.6610 respectively. The color of the over coat in two men image under different cast are appear slightly black but after applying the proposed method the color of over coat becomes truly black as it appears in the reference image. The same result can be inferred by employing the proposed technique over the color Color checker image. Images resulted after processing through the proposed technique are shown in Fig. 4 (e), (f) and (g),



Fig.4-Color checker image (a) Reference, (b) Red cast, (c) Green cast, (d) Blue cast, (e) proposed corrected Red cast (f) proposed corrected Green cast (g) proposed corrected Blue cast (h) Gray corrected Red cast (i) Gray corrected Green cast (j) Gray corrected Blue cast (k) max-RGB corrected Red cast (l) max-RGB corrected Green cast (m) max-RGB corrected Blue cast, (n) shades of grey corrected Red cast, (o) shades of grey approach corrected Green cast (p) shades of grey approach corrected Blue cast

It is clearly visible that proposed technique effectively remove the color cast and retain the originality of the image. Statistically obtained color corrective measures like entropy, contrast for corrected Color checker images are 7.3210, 7.3055, 7.3283 and 19.6046, 19.4008, 19.6078 respectively. The obtained CAF values of proposed algorithm are highest as compared to all other existing methods for all images.

6. Conclusions

In this paper, the Fuzzy based approach is used for detecting and removing all types of color cast. This algorithm detects the actual cast on the image in an effective manner without assuming it like Reddish, Greenish and Bluish. The uncertainty in the cast further removed completely by using the fuzzy logic. Thus, we recover true color distribution of the image and develop algorithm which is robust under different lighting conditions. As a result, the image obtained with high entropy, luminance and contrast. In addition to that the image obtained with high intensity value. The effectiveness of the proposed algorithm is shown by measuring the quality assessment function and CIELAB metric. These performance measures are compared with the existing color cast algorithm. The maximum value of the function is achieved by using proposed approach which shows a better quality of the image. It was inferred from the decrement nature of distance measured through CIELAB metric that the resulted image approaches towards the reference image.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence

References

1. Land, E.H., McCann, J.J., Lightness and retinex theory. J. Opt. Soc. Amer. 1974, 61 (1), 1–11.
2. Maloney, L.T., Wandell, B.A. Color constancy: a method for recovering surface spectral reflectance. J. Opt. Soc.Amer.A3., 1986, 3 (1), 29–33.
3. E. H. Land, "The retinex theory of color vision," Scientific American, vol. 237, no. 6, pp. 108–128, December 1977. 1, 2
4. B. Funt and L. Shi, "The rehabilitation of maxRGB," in Color and Imaging Conf., 2010, no. 1, pp. 256–259, Soc. Imag. Sci. Technology.
5. G. Buchsbaum, "A spatial processor model for object color perception," Journal of the Franklin Institute, vol. 310, no. 1, pp. 1–26, July 1980. 1, 2
6. G. D. Finlayson and E. Trezzi, "Shades of gray and color constancy," in Proc. IS&T/SID Color Imaging conference, 2004, pp. 37–41.
7. J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge based color constancy," IEEE Transactions on Image Processing, vol. 16, no. 9, pp. 2207–2214, 2007.
8. A. Gijsenij and T. Gevers Color Constancy Using Natural Statistics and Scene Semantics IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 33, No. 4, 2011
9. J. Negrete, R. E. Sanchez-Yanez, "Automatic selection of color constancy algorithms for dark image enhancement by fuzzy rule-based reasoning" Applied Soft Computing, Volume 28, March 2015, pp. 1–10.
10. L. A. Zadeh,"Fuzzy Sets, Information and Control", 1965, pp. 338-353.
11. Rizzi. A, Gatta. C, Marini. D, "A new algorithm for unsupervised global and local color correction", Pattern Recognition Letters, Vol 24, No. 11, pp.1663-1677, July. 2003.
12. Om Prakash Verma, Puneet Kumar, Madasu Hanmandlu, Sidharth Chhabra, High dynamic range optimal fuzzy color image enhancement using Artificial Ant Colony System, Applied Soft Computing, Volume 12,Issue 1, pp. 394-404 January 2012.
13. Sethi, R., Sreedevi, I., Verma, O. P., & Jain, V. (2015, December). An optimal underwater image enhancement based on fuzzy gray world algorithm and Bacterial Foraging algorithm, IEEE Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG),pp. 1-4 December,2015.
14. Z. X. Xie, Z. F. Wang, "Color Image Quality Assessment Based on Image Quality Parameters Perceived by Human Vision System", International Conference on Multimedia Technology (ICMT), Ningbo 29-31 Oct. 2010.
15. X. M. Zhang, B. A. Wandell, "A spatial extension to CIELAB for digital color image reproduction," Symposium Proceedings, 1996.
16. G.D. Finlayson and E. Trezzi, "Shades of gray and color constancy", Proc. of IS&T/SID Color Imaging Conference, Scottsdale, Arizona, USA, pp.37–41, 2004.

MR ADITYA JAIN (Orcid ID : 0000-0003-2509-4608)

DR. AMRISH PANWAR (Orcid ID : 0000-0001-5076-8593)

Article type : Article

Corresponding author mail id: panwaramar@gmail.com

Enhanced structural, dielectric, ferroelectric and piezoelectric properties of (1-*x*)$Ba_{0.9}Sr_{0.1}TiO_3$ – (*x*)$Ba_{0.7}Ca_{0.3}TiO_3$ ceramics derived using mechano-chemical activation technique

Aditya Jain[a], Amrish K. Panwar[*,a], Rakesh Saroha[a], A. K. Jha[b]

[a]*Department of Applied Physics, Delhi Technological University (Formerly Delhi College of Engineering), Bawana Road, Delhi-110042, India.*

[b]*Department of Applied Science, A.I.A.C.T.R. Geeta Colony, Delhi-110031, India.*

Abstract

The effect of $Ba_{0.7}Ca_{0.3}TiO_3$ (BCT) substitution on the structural, dielectric, ferroelectric and piezoelectric properties of mechano-chemically synthesized lead-free (1-*x*)$Ba_{0.9}Sr_{0.1}TiO_3$-(*x*)$Ba_{0.7}Ca_{0.3}TiO_3$ (*x* = 0.0, 0.10, 0.20, 0.35 and 0.50) ceramics have been investigated. X-ray diffraction patterns confirms the formation of tetragonal phase with *P4mm* space group. The results indicate a strong influence of BCT substitution on the structural and electrical properties of $Ba_{0.9}Sr_{0.1}TiO_3$ (BST) ceramic. BST-BCT ceramic for *x* = 0.35 have shown high dielectric constant ε_m ~ 21800, high remnant polarization P_r ~ 10.16 μC/cm^2, large piezoelectric charge constant D_{33} ~ 293 pC/N, large piezoelectric voltage constant g_{33} ~ 5.80 mV.m/N and highest dielectric breakdown strength E_{bd} ~ 224 kV/cm among the five synthesized samples. The correlation between structural,

* Corresponding author: Amrish. K. Panwar, E-mail: panwaramar@gmail.com, Tel: +91-9891778307 (Mobile).

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/jace.15061

This article is protected by copyright. All rights reserved.

Accepted Article

dielectric, ferroelectric and piezoelectric properties of the BST ceramic with increasing BCT content have been systematically described on the basis of domain wall motion and grain size effect.

Keywords: Dielectric; Microstructure; Ferroelectric; Piezoelectric; Energy storage.

1. Introduction

Barium titanate is a lead-free ferroelectric material, possesses high dielectric constant and low dielectric loss, which makes it useful for applications such as multi-layered ceramic capacitors and random access memory (RAM) devices [1]. Ferroelectric and piezoelectric ceramics have been extensively investigated and they are regarded as primary functional materials for various electronic, electrical and microelectronic component applications. With the recent advances in pulsed power technology, high energy density capacitors have been employed in pulse power equipments due to their better discharge efficiency and high energy storage capability [2]. Various lead-free $BaTiO_3$ based composites, such as $BaTiO_3$-$Bi(Ni_xTi_{1-x})O_3$ [3], $BaTiO_3$-$Bi(Zn_xTi_{1-x})O_3$ [4], $BaTiO_3$-$BiScO_3$ [5], etc. have been investigated and were found to exhibit thermally stable high dielectric constant over a broad temperature range. Alternatively, various other lead-free systems like $(Bi_{0.5}Na_{0.5})TiO_3$, $Bi_{0.5}(Na_{0.8}K_{0.2})_{0.5}TiO_3$, $(Na_{0.5}K_{0.5})NbO_3$, etc. were also explored for multifunctional device applications and these materials were found to exhibit better piezoelectric properties. Although, these material suffer reduced densification due to the presence of highly volatile elements like Na, Bi and K, which may have degraded their physical properties [6].

Currently, research is focused on the investigation of lead-free $BaTiO_3$ based materials which is attributed to their improved structural, dielectric, ferroelectric, pyroelectric, piezoelectric and multiferroic properties [7]. **Barium strontium titanate ($Ba_{1-x}Sr_xTiO_3$) composites are the most widely investigated materials until now due to their large tunability and extremely low dielectric loss around room temperature. In recent years, $Ba_{1-x}Sr_xTiO_3$ ceramics have long been investigated for applications in electronics and material science field. BST solid solutions need comparatively small DC bias voltage to achieve large tunability. Further, $Ba_{1-x}Ca_xTiO_3$ also possess interesting dielectric, ferroelectric and piezoelectric properties. It has been reported that up to 15 mol% of Ca^{2+} substitution at Ba^{2+} site, the ferroelectric to paraelectric phase transition temperature were found to be fairly stable [8]. Beyond the solid solubility limit of x = 23 mol%**

This article is protected by copyright. All rights reserved.

Accepted Article

in $Ba_{1-x}Ca_xTiO_3$ lattice, the tetragonal and orthorhombic phases coexists which results in improved electrical properties due to interaction of tetragonal and orthorhombic phases [9]. Therefore, combining the characteristics of both BST and BCT may result in improvement in their dielectric, ferroelectric and energy storage properties.

The aim of this work is to improve the electrical properties $BaTiO_3$ based ceramics, its versatility and to explore new functionalities. Our emphasis in the present work is to investigate the beneficial effect of $Ba_{0.7}Ca_{0.3}TiO_3$ addition on the structural and electrical properties. Our selection for the parent composition is justified by high dielectric constant, low dielectric loss, good ferroelectric properties and large piezoelectric constant. Further, In our enduring research for noble materials with overall superior dielectric, ferroelectric and piezoelectric properties for multi-layered ceramic capacitor, energy storage capacitors, etc.; herein we report the synthesis of (1-*x*)$Ba_{0.9}Sr_{0.1}TiO_3$-(*x*)$Ba_{0.7}Ca_{0.3}TiO_3$ (*x* = 0.0, 0.10, 0.20, 0.35 and 0.50) and their structural, dielectric, ferroelectric and piezoelectric properties.

2. Experimental

(1-*x*)$Ba_{0.9}Sr_{0.1}TiO_3$-*x*$Ba_{0.7}Ca_{0.3}TiO_3$ (BST-BCT) ceramics for *x* = 0.0, 0.10, 0.20, 0.35 and 0.50 were prepared using mechano-chemical activation technique. BST-BCT was prepared by taking AR grade (Sigma-Aldrich make) $BaCO_3$ (99.99%), TiO_2 (99.9%), $SrCO_3$ (99.98%) and $CaCO_3$ (99.0%) as the starting precursor reagents. All the raw precursor reagent powders were taken in required stoichiometric quantity and the powders were grinded in *Retsch PM100* planetary ball mill in air atmosphere at 440 rpm for a period of 25 h. A cylindrical Zirconia vial with 45 mm inner diameter, 50 ml capacity and 150 Zirconia balls with 5 mm diameter were used as milling medium. A ball to powder ratio of 12:1 was kept inside the ball milling vial. After ball milling, BST-BCT solid solutions were calcined at 1050 °C for 4 h to make a desired phase. The resulting powders were then mixed with 2.5 wt% polyvinyl alcohol (PVA) and pressed into cylindrical pellets of approximately 12 mm diameter and thickness of 1.20 mm by applying a force of 70 kN using a hydraulic press machine. The pellets of the composite samples were then sintered at 1250 °C with heating rate of 6 °C/min for 4 h in air atmosphere. Hereafter, these specimens were identified as BSC0, BSC10, BSC20, BSC35 and BSC50 for *x* = 0.0, 0.10, 0.20, 0.35 and 0.50, respectively.

This article is protected by copyright. All rights reserved.

Accepted Article

The constituent phase and crystal structure of the synthesized BSC*x* specimens were examined by *Bruker D8 Advance* X-ray diffractometer having $CuK_{\alpha 1}$ radiation with operational wavelength of 1.541 Å. Lattice parameters and theoretical densities were evaluated using the XRD data with the help of *X'Pert Highscore Plus* software followed by Rietveld analysis. The microstructural characteristics and surface morphology of high temperature sintered pellet samples were examined using *Hitachi S-3700N* scanning electron microscope. Grain size distribution analysis has been carried out using *ImageJ* software. TEM/HRTEM and SAED analysis have been carried out using TEM (FEISTWIN PF30) at 300 kV by using copper grids (CF200-Cu). The sintered pellets of (1-*x*)BST-(*x*)BCT solid solutions were polished to a thickness of nearly 1.2 mm and slightly viscous silver paste was coated on both sides of the pellets to make electrodes, following this, the pellet samples were heated at 160 °C for 1 h. These pellets were then used to carry out dielectric measurement by an *Agilent E4980A* LCR meter at multiple frequencies (100 Hz, 1 kHz, 10 kHz, 50 kHz and 100 kHz). Temperature dependent dielectric measurement was performed over the wide temperature range of 25 to 200 °C. The ferroelectric hysteresis loops of the composites were recorded using a Sawyer-Tower based P-E loop equipment (*M/s Marine India, New Delhi*). **Electrical poling of the samples has been carried out using a DC poling equipment (*M/s Marine India, New Delhi*). The samples were poled by keeping the sample in a silicon oil bath for 2 h at *E* ~ 15 kV/mm at an optimized temperature of 70 °C.**

3. Results & discussion

3.1. X-ray diffraction (XRD)

The combined X-ray diffraction patterns of (1-*x*)BST-(*x*)BCT (*x* = 0.0, 0.10, 0.20, 0.35 and 0.50) samples are presented in Fig. 1(a1). All the BST-BCT samples prepared through mechano-chemical activation process exhibits a tetragonal structure without any traces of impurities. The sharp and intense diffraction peaks of BST-BCT samples indicate the crystalline nature of ceramics. Barium titanate (JCPDS file no. 00-005-0626) was taken as the reference XRD pattern, and the Pseudo-Voigt function has been used to refine the diffraction peaks. The smooth and well-fitted refined profile of BST-BCT samples with its experimental profile clearly suggest the validity of the refined model as well as Rietveld parameters. The observed and calculated d values of all the reflections (diffraction

This article is protected by copyright. All rights reserved.

Accepted Article

lines) of BST-BCT samples indicate that the basic crystal structure (i.e. tetragonal) has not changed as *x* varied from 0.0 to 0.50. However, as *x* increases, a slight increase in *c/a* ratio has been observed which suggest an increase in tetragonality with the addition of BCT. The diffraction peak shown in Fig. 1(a2) indicates that as *x* increases, the diffraction pattern shifts towards higher angle side. Further, the lattice parameters '*a*' and '*c*' were also found to decrease with BCT concentration which may be attributed to lower ionic radii of Ca^{2+} (1.34 Å) ion as compared to Sr^{2+} (1.44 Å) and Ba^{2+} (1.61 Å). The peak splitting around $2\theta = 45°$ in the XRD patterns of all the five samples also confirms that it exhibits a tetragonal structure with *P4mm* space group. Table 1 lists the Rietveld refinement structural parameters for the five synthesized samples. The unit cell volume of (1-*x*)BST-(*x*)BCT is observed to decrease as *x* increases, which results in lattice shrinkage and consequently may produce micro-strain in the structure.

3.2. Microstructural studies

The effect of $Ba_{0.7}Ca_{0.3}TiO_3$ addition in the $Ba_{0.9}Sr_{0.1}TiO_3$ ceramic is correlated with the microstructural formation, grain growth behavior and density of the resulting composites. The SEM micrograph shown in Fig. 2(a-e) demonstrate well developed polygonal grain type morphology with clearly distinguishable grain boundaries. The density of the BST-BCT composites were found to increase with BCT substitution and BSC35 shows the highest density among the five synthesized specimens. However, the addition of BCT in BST lattice leads to an increase in grain size, which suggests that BCT addition could assist the grain growth of the ceramics. **The increase in grain size with BCT substitution may be due to the fact that the grains of BCT may have higher probability of growing as compared to BST and therefore bigger grain grows at the cost of smaller grain. Moreover, in case of substitution/doping, it is hard to obtain an identical morphology and grain size due to the considerable difference in the sintering conditions and the thermal expansion coefficients of the two compounds [10, 11].** In higher BCT content ceramics, it may happen that the smaller grains merge into each other and grow into bigger ones, resulting in an increase in the fraction of larger grains in the entire microstructure. The average grain size has been calculated using ImageJ software and it is found to be approximately ~274, 286, 309, 335 and 323 nm for BSC0, BSC10, BSC20, BSC35 and BSC50, respectively. In earlier reports, it has been observed

This article is protected by copyright. All rights reserved.

Accepted Article

Accepted Article

that the grain size of ceramics is reliant on its calcination and sintering temperature as well as soaking period at the peak temperature [12]. These two aspects have a foremost effect on the microstructural properties of the final sintered specimen and therefore, they eventually affect the dielectric and ferroelectric properties of the ceramics. In the present work, a better morphology combined with small grain size is obtained as compared to the earlier reports on $BaTiO_3$ based ceramics [7, 13]. **The density of the samples is measured for the pellet samples by the Archimedes method and the corresponding values are shown in Table 2. The density of the samples increases with increasing BCT substitution upto x = 0.35 and then decreases on further increase in BCT content. The values were within the range of 84-95 % of the theoretical density of $BaTiO_3$ (6.05 gm/cm3). Further, to clearly understand the grain growth of materials, the microstructure and density values were reported for BSC35 sample before PVA mixing, after PVA mixing and after 4 h of sintering. Fig. S1 in supplemental materials shows the above stated results which is available online.**

TEM, HRTEM and SAED studies of BSC35 have been performed as it shows overall superior electrical properties among the five synthesized samples. The average particle size is measured to be approximately ~145 nm for BSC35 powder as observed from TEM analysis (Fig. 3(a)). The perfectly parallely arranged lattice fringes do indicate the formation of perfectly crystalline material. HRTEM images can provide valuable evidences regarding the crystal phase and anisotropy in morphology of the material under study. In HRTEM images, it is considered that the orientation of the number of stacking layers is associated with the shape and atomic arrangement of unit cells. The HRTEM image of BSC35 particle (Fig. 3(b)) shows the lattice fringes corresponding to plane (101) with an inter-fringe distance of 1.77 Å. Fig. 3(c) depicts the selected area (electron) diffraction (SAED) pattern and it confirms the polycrystalline nature of BSC35 sample. The SAED pattern illustrates bright and dark spots in a circular pattern that may be assigned to different planes in the tetragonal crystal. Furthermore, both the HRTEM as well as SAED pattern recorded for the selected sample indicate that there are no noteworthy structural defects in the observed BSC35 sample.

This article is protected by copyright. All rights reserved.

3.3. Dielectric studies

Figs. 4(a-j) shows the temperature dependence of the dielectric constant, ε and dielectric loss at various frequencies over the temperature range from room temperature to 200 ºC for all the synthesized samples. The dielectric constant of the samples were found to increase as x increases, becomes maximum at x = 0.35 and then decreases on further increase in x. **$Ba_{0.7}Ca_{0.3}TiO_3$ addition leads to compositional fluctuation, structural reorder as well as rearrangement of cations in one or more crystallographic sites of the structure. All three parameters result in change in overall polarization inside the material and this leads to variation in dielectric properties of the material. Further, BCT addition in BST also results in microscopic inhomogeneity in the overall compound leading to a little change in the phase transition temperature of the material.** The dielectric constant at Curie temperature for the samples BSC0, BSC10, BSC20, BSC35 and BSC50 at 1 kHz were found to be 15947, 16344, 17247, 21792 and 16687, respectively. BSC35 exhibits the highest dielectric constant which may be due to larger grain size as well as the density of the material. The dielectric properties of the material are significantly affected by the grain size of the material [14]. The high dielectric constant value for BSC35 may be attributed to the fact that the increase in grain size results in a decrease in the number of insulating grain boundary layers. The fewer number of grain boundaries corresponds to lesser trapping of space charge and thus domain wall pinning does not contribute to dielectric constant resulting in high dielectric constant [15]. Further, a ferroelectric single crystal contains multiple ferroelectric domains separated by an interface known as domain walls. These domain walls act as transition between different dipole alignments [16]. Also, the addition of BCT into BST lattice results in a small increase in transition temperature. The slight shift in Curie temperature towards high-temperature side may be caused by the occurrence of ion disorder and internal stress developed inside the material [7]. The dielectric constant value of all the BST-BCT ceramic samples increases with rise in temperature and close to Curie temperature it shows a strong anomaly in dielectric constant value, which is a characteristic of ferroelectric behavior. This behavior can be described on the basis of increase in domain wall mobility with rise in temperature [17].

Moreover, the dielectric constant of the samples displays an inverse dependence on frequency, which is the inherent characteristics of most of the dielectric and ferroelectric materials [18]. The

This article is protected by copyright. All rights reserved.

Accepted Article

Maxwell-Wagner multi-layer model for space charge or interfacial polarization can explain the behavior of dielectric constant with variation in frequency [19]. This model demonstrates that any dielectric material contains a large number of grains that are separated by thin insulating grain boundaries. On the application of an external electric field, the positive and negative charge carriers are relocated inside the grain and they get accumulated at the edges of the grain. This accumulation of electric charges at boundaries causes a large polarization resulting in high dielectric constant [20]. At lower frequency of operation, all four kinds of polarizations namely ionic, electronic, dipolar and interfacial polarization contribute to the total polarization of the material and this result in an enhancement in dielectric constant value. At higher operating frequency, some of the polarizations cease to contribute, and therefore total polarization inside the material consist of only interfacial and dipolar polarization and hence at higher frequency dielectric constant decreases [18]. Similar behavior has been reported in many ferroelectric materials [21].

It can be observed that the dielectric loss of BST-BCT samples increases as the temperature increases. This increase in dielectric loss value at high temperature may be due to (i) increased mobility of ions, (ii) the imperfection in the material and (iii) shallow contribution of thermally activated charge carriers inside the material which would hinder the contribution of dipolar polarization [22]. Further, the dielectric loss of the materials were found to decrease with BCT substitution; becomes minimum for BSC35 and increases again on increasing the BCT content. The decrease in dielectric loss with increasing BCT content may be attributed to reduction in leakage current that rely on the crystal structure impurities like vacancies, dislocations and ionic charge imbalance [23].

3.4. Ferroelectric studies

Fig. 5 shows electric field induced polarization loops recorded at 10 Hz test frequency. The *P-E* loop measurement have been carried out at three different temperatures by increasing the electric field up to ~ 20 kV/cm. The insets of Figs. 5(a), 6(b) and 6(c) shows the variation of remnant polarization ($2P_r$) and ratio of remnant and saturation polarization (P_r/P_s) with increasing BCT content. Well saturated P-E loops confirm the ferroelectric nature of all the five samples. The shape of hysteresis *P-E* loops are associated with the crystalline structure of the material under study. The room temperature

This article is protected by copyright. All rights reserved.

Accepted Article

saturation polarization ($2P_s$), remnant polarization ($2P_r$) and coercive field ($2E_c$) values for the BST-BCT samples are shown in Table 2. It is observed that the remnant polarization first decreases up to x = 0.10 and then increases with the increase in x content and becomes maximum for x = 0.35. A higher remnant polarization combined with high coercive field is achieved for BSC35 sample that may be attributed to higher grain size of the composite observed in SEM micrograph in Fig 3(d). Further, the high remnant polarization for the BSC35 sample is possibly due to low defect density, compositional homogeneity, uniform distribution of grains, nonexistence of domain-wall pinning centers, etc. Further, the sharp nature of *P-E* hysteresis loop confirms the high resistance and low loss nature of BSC*x* ceramics. Larger grain size may contain larger individual ferroelectric domains resulting in lesser number of grain boundary regions, which are known to be low permittivity regions with poor ferroelectricity [24]. Moreover, domain walls among larger grains are easier to switch under externally applied electric field, which results in higher remnant polarization. The addition of BCT in BST lattice leads to an improvement in the shape of ferroelectric hysteresis loop, which may be ascribed to reduction in leakage current due to decrease in number of oxygen vacancies [25]. The ferroelectricity of the samples were found to decrease at higher temperatures. The decrease in remnant polarization at higher temperatures may be attributed to decrease in ferroelectric and increase in paraelectric phase of the material. At all three temperatures, a steady increase in coercive field of the BSC*x* samples is observed with the subsequent increase of BCT content. The occurrence of high coercive field with an increase in BCT content makes the samples ferroelectrically hard. Further, the coercive field (E_c) obtained from *P-E* hysteresis loop measurement has been used to determine the poling electric field to reorient the ferroelectric domains of BSC*x* ceramics.

3.5. Breakdown strength measurement

The dielectric breakdown strength (E_{bd}) of all the synthesized specimens have been measured and shown in Fig. 6. **Dielectric breakdown strength (E_{bd}) measurements of BST-BCT samples were performed on the silver electroded pellet by immersing the specimen in a silicon oil bath onto a sheet of highly conductive copper connected to high voltage DC supply. Some external parameters such as frequency and ramp rate of the externally applied electric field also affect the dielectric breakdown strength. In the present study, the frequency is set to 50 Hz and the**

This article is protected by copyright. All rights reserved.

Accepted Article

ramp rate of applied electric field were kept constant at 1 kV/s. The E_{bd} value is known to be dependent on the grain size, electrode geometry, density, area and thickness of the specimen under study. Moreover, homogeneous microstructure and high density are beneficial in improving dielectric breakdown strength [26]. Table 2 list the dielectric breakdown strength of all the five synthesized samples. The E_{bd} is found to be maximum for BSC35, which may be ascribed to uniform grain size and high density of the material. In a ceramic material, under the application of externally applied electric field, as the number of weakly related ions with higher porosity and lower density rises, the space displacement polarization increases. This increase in space displacement polarization forces the micropores inside the material to offer very small resistance to electrical breakdown and depreciate the E_{bd} value [27]. Thickness dependent dielectric study has also been carried out which is available online in supplemental materials.

3.6. Piezoelectric studies

The electromechanical response as a function of BCT content for the five synthesized samples was recorded. Fig. 7 depicts the variation of piezoelectric constant (d_{33}) and piezoelectric voltage coefficient (g_{33}) as a function of BCT content. In the process of electrical poling of ceramics when an external electric field is applied, materials with arbitrarily distributed dipoles become polarized in the same direction as that of the applied electric field. Poling plays an essential role in determining the piezoelectric properties of any polycrystalline ferroelectric material. Further, some other parameters such as poling time, poling voltage and poling temperature affect the piezoelectric properties and so they need to be optimized. In the present study, all the samples were poled for 2 h at $E \sim 15$ kV/mm at an optimized poling temperature of $T_P \sim 70$ ºC, at which superior piezoelectric properties were observed. Poling of ceramics at an appropriate temperature helps in accumulation of charge carriers around the grain boundaries introducing an internal electric field that can increase the electromechanical response of the poled materials [28]. The magnitude of longitudinal piezoelectric constant is found to be maximum for BSC35 and decreases on further increase of BCT concentration. The d_{33} values of each of the five synthesized BST-BCT samples are shown in Table 2. It is believed that higher value of BSC35 may be attributed to better crystallinity as well as an optimum combination of grain size and density of the material. All these piezoelectric constants are affected by

This article is protected by copyright. All rights reserved.

Accepted Article

ferroelectric domain structure. It has been reported that piezoelectric constant decreases with decrease in grain size [29]. Furthermore, previous investigations on piezoelectric materials suggested a parabolic relation between grain and domain sizes [30]. The piezoelectric voltage constant, g_{33} has been calculated by dividing the piezoelectric constant (d_{33}) of the material by its permittivity, ε. Finally, it can be established that BSC35 exhibits the optimum ferroelectric and piezoelectric properties among the five synthesized samples. Moreover, the piezoelectric coefficients, d_{33} and g_{33} obtained in the present study are comparable to several lead-based ceramics reported in the literature [31, 32].

4. Conclusion

Mechano-chemical activation technique was employed to synthesize BCT substituted BST crystalline powders. The XRD patterns and Rietveld refinement of BST-BCT samples confirms a tetragonal phase with *P4mm (99)* space group for all the synthesized samples. There was no noteworthy change in phase and crystal structure of samples with an increase in BCT content. Microstructural (grain size & relative density), dielectric, ferroelectric and piezoelectric properties were found to improve with BCT addition. Overall, BSC35 exhibits better microstructural as well as electrical properties among the five synthesized samples. The optimized poling conditions have resulted in high piezoelectric charge constant d_{33} ~ 293 pC/N and large piezoelectric voltage constant g_{33} ~ 5.82 mV.m/N for BSC35 sample. Finally, it can be concluded that the BST-BCT materials are easy and inexpensive to design. In addition, improved microstructural and electrical properties indicate that the mechano-chemical synthesis approach represents a better way to prepare materials for application in multi-layer ceramic capacitors and low loss dielectric materials.

5. Acknowledgements

Authors gratefully acknowledge the support of DST-SERB (Grant no. PS/039/2011), India to successfully complete this research.

This article is protected by copyright. All rights reserved.

Accepted Article

Accepted Article

References

1. Zimmermann F, Voigts M, Menesklou W, et al. $Ba_{0.6}Sr_{0.4}TiO_3$ and $BaZr_{0.3}Ti_{0.7}O_3$ thick films as tunable microwave dielectrics. *J. Eur. Ceram. Soc.* 2004;24:1729-1733.
2. Xiao S, Xiu S, Zhang W, et al. Effects of $Ba_xSr_{1-x}TiO_3$ ceramics additives on structure and energy storage properties of $Ba_{0.4}Sr_{0.6}TiO_3$–BaO–B_2O_3–Al_2O_3–SiO_2 glass-ceramic. *J. Alloys Compd.* 2016;675:15-21.
3. Fujii I, Nakashima K, Kumada N, et al. Structural, dielectric, and piezoelectric properties of $BaTiO_3$-$Bi(Ni_{1/2}Ti_{1/2})O_3$ ceramics. *J. Ceram. Soc. Jpn.* 2012;120:30-34.
4. Raengthon N and Cann DP. Dielectric relaxation in $BaTiO_3$–$Bi(Zn_{1/2}Ti_{1/2})O_3$ ceramics. *J. Am. Ceram. Soc.* 2012;95:1604-1612.
5. Ogihara H, Randall CA, and Trolier- McKinstry S. Weakly coupled relaxor behavior of $BaTiO_3$–$BiScO_3$ ceramics. *J. Am. Ceram. Soc.* 2009;92:110-118.
6. Puli VS, Pradhan DK, Riggs BC, et al. Synthesis and characterization of lead-free ternary component BST–BCT–BZT ceramic capacitors. *J. Adv. Dielectr.* 2014;04:1450014.
7. Chao X, Wang J, Liang P, et al. Phase transition and improved electrical performance of $Ba_{0.85}Ca_{0.15}Zr_{0.1}Ti_{0.9}O_3$–$Ca_{0.28}Ba_{0.72}Nb_2O_6$ ceramics with high Curie temperature. *Mater. Des.* 2016;89:465-469.
8. McQUARRIE M and BEHNKE FW. Structural and dielectric studies in the system $(Ba,Ca)(Ti,Zr)O_3$. *J. Am. Ceram. Soc.* 1954;37:539-543.
9. Souza AE, Teixeira SR, Morilla-Santos C, et al. Photoluminescence activity of $Ba_{1-x}Ca_xTiO_3$: dependence on particle size and morphology. *J. Mater. Chem. C.* 2014;2:7056-7070.
10. Puli VS, Kumar A, Chrisey DB, et al. Barium zirconate-titanate/barium calcium-titanate ceramics via sol–gel process: novel high-energy-density capacitors. *J. Phys. D: Appl. Phys.* 2011;44:395403.
11. Rai AK, Rao K, Kumar LV, et al. Synthesis and characterization of ultra fine barium calcium titanate, barium strontium titanate and $Ba_{1-2x}Ca_xSr_xTiO_3$ (x= 0.05, 0.10). *J. Alloys Compd.* 2009;475:316-320.
12. Herabut A and Safari A. Processing and electromechanical properties of $(Bi_{0.5}Na_{0.5})_{(1-1.5x)}La_xTiO_3$ ceramics. *J. Am. Ceram. Soc.* 1997;80:2954-2958.
13. Zimmermann F, Voigts M, Menesklou W, et al. $Ba_{0.6}Sr_{0.4}TiO_3$ and $BaZr_{0.3}Ti_{0.7}O_3$ thick films as tunable microwave dielectrics. *J. Eur. Ceram. Soc.* 2004;24:1729-1733.
14. Jain A, Panwar AK, and Jha A. Effect of ZnO doping on structural, dielectric, ferroelectric and piezoelectric properties of $BaZr_{0.1}Ti_{0.9}O_3$ ceramics. *Ceram. Int.* 2017;43:1948-1955.
15. Jain A, Panwar AK, and Jha AK. Influence of milling duration on microstructural, electrical, ferroelectric and piezoelectric properties of $Ba_{0.9}Sr_{0.1}Zr_{0.04}Ti_{0.96}O_3$ ceramic. *Ceram. Int.* 2016;42:18771-18778.
16. Mudinepalli VR, Feng L, Lin W-C, et al. Effect of grain size on dielectric and ferroelectric properties of nanostructured $Ba_{0.8}Sr_{0.2}TiO_3$ ceramics. *J. Adv. Ceram.* 2015;4:46-53.
17. Jain A, Panwar AK, and Jha AK. Structural, dielectric and ferroelectric studies of $Ba_{1-x}Sr_xTiO_3$ ceramics prepared by mechanochemical activation technique. *J. Mater. Sci. - Mater. Electron.* 2016;27:9911-9919.
18. Alkathy MS, Gayam R, and Raju KCJ. Effect of sintering temperature on structural and dielectric properties of Bi and Li co-substituted barium titanate ceramic. *Ceram. Int.* 2016;42:15432-15441.

This article is protected by copyright. All rights reserved.

Accepted Article

19. Montanari G and Morshuis P. Space charge phenomenology in polymeric insulating materials. *IEEE Trans. Dielectr. Electr. Insul.* 2005;12:754-767.
20. Liu G, Fan H, Xu J, et al. Colossal permittivity and impedance analysis of niobium and aluminum co-doped TiO_2 ceramics. *RSC Adv.* 2016;6:48708-48714.
21. Chou X, Zhai J, and Yao X. Relaxor Behavior and Dielectric Properties of $MgTiO_3$-Doped $BaZr_{0.35}Ti_{0.65}O_3$ Composite Ceramics for Tunable Applications. *J. Am. Ceram. Soc.* 2007;90:2799-2801.
22. Ang C, Yu Z, and Cross L. Oxygen-vacancy-related low-frequency dielectric relaxation and electrical conduction in Bi:$SrTiO_3$. *Phys. Rev. B: Condens. Matter*. 2000;62:228.
23. Verma A, Goel T, Mendiratta R, et al. Dielectric properties of NiZn ferrites prepared by the citrate precursor method. *Mater. Sci. Eng., B* 1999;60:156-162.
24. Ganguly M, Rout S, Ahn C, et al. Structural, electrical and optical properties of $Ba(Ti_{1-x}Yb_{4x/3})O_3$ ceramics. *Ceram. Int.* 2013;39:9511-9524.
25. Ganduglia-Pirovano MV, Hofmann A, and Sauer J. Oxygen vacancies in transition metal and rare earth oxides: Current state of understanding and remaining challenges. *Surf. Sci. Rep.* 2007;62:219-270.
26. Lei W, Yan Y-Y, Wang X-H, et al. Improving the breakdown strength of $(Mg_{0.9}Zn_{0.1})_2(Ti_{1-x}Mn_x)O_4$ ceramics with low dielectric loss. *Ceram. Int.* 2015;41:521-525.
27. Lei W, Ang R, Wang X-C, et al. Phase evolution and near-zero shrinkage in BaAl2Si2O8 low-permittivity microwave dielectric ceramics. *Mater. Res. Bull.* 2014;50:235-239.
28. Wu J, Xiao D, Wu W, et al. Composition and poling condition-induced electrical behavior of $(Ba_{0.85}Ca_{0.15})(Ti_{1-x}Zr_x)O_3$ lead-free piezoelectric ceramics. *J. Eur. Ceram. Soc.* 2012;32:891-898.
29. Randall CA, Kim N, Kucera JP, et al. Intrinsic and extrinsic size effects in fine- grained morphotropic- phase- boundary lead zirconate titanate ceramics. *J. Am. Ceram. Soc.* 1998;81:677-688.
30. Cao W and Randall CA. Grain size and domain size relations in bulk ceramic ferroelectric materials. *J. Phys. Chem. Solids*. 1996;57:1499-1505.
31. Li T and Long X. Piezo-/dielectric properties of perovskite-structure high-temperature relaxor ferroelectrics: The $Pb(Lu_{1/2}Nb_{1/2})O_3$–$Pb(Zn_{1/3}Nb_{2/3})O_3$–$PbTiO_3$ ternary ceramics. *Mater. Res. Bull.* 2014;51:251-257.
32. Xing Z, Wang H, Xu B, et al. Structural integrity and ferroelectric–piezoelectric properties of $PbTiO_3$ coating prepared via supersonic plasma spraying. *Mater. Des.* 2014;62:57-63.

This article is protected by copyright. All rights reserved.

Table 1. Comparison of Rietveld refined parameters for BST-BCT specimens.

Sample	a,b (Å)	c (Å)	c/a	Cell volume ($Å^3$)	Structure	χ^2	R_{exp}	R_f	R_{Bragg}
BSC0	**3.9822**	**4.0127**	**1.0076**	**63.63**	Tetragonal	**2.71**	**1.02**	**6.19**	**5.12**
BSC10	**3.9787**	**4.0082**	**1.0074**	**63.45**	Tetragonal	**3.48**	**3.11**	**5.88**	**3.87**
BSC20	**3.9765**	**4.0063**	**1.0074**	**63.34**	Tetragonal	**3.02**	**3.94**	**6.93**	**6.41**
BSC35	**3.9708**	**4.0046**	**1.0085**	**63.14**	Tetragonal	**2.98**	**3.71**	**6.87**	**1.74**
BSC50	**3.9658**	**4.0014**	**1.0082**	**62.93**	Tetragonal	**3.23**	**3.58**	**6.95**	**4.35**

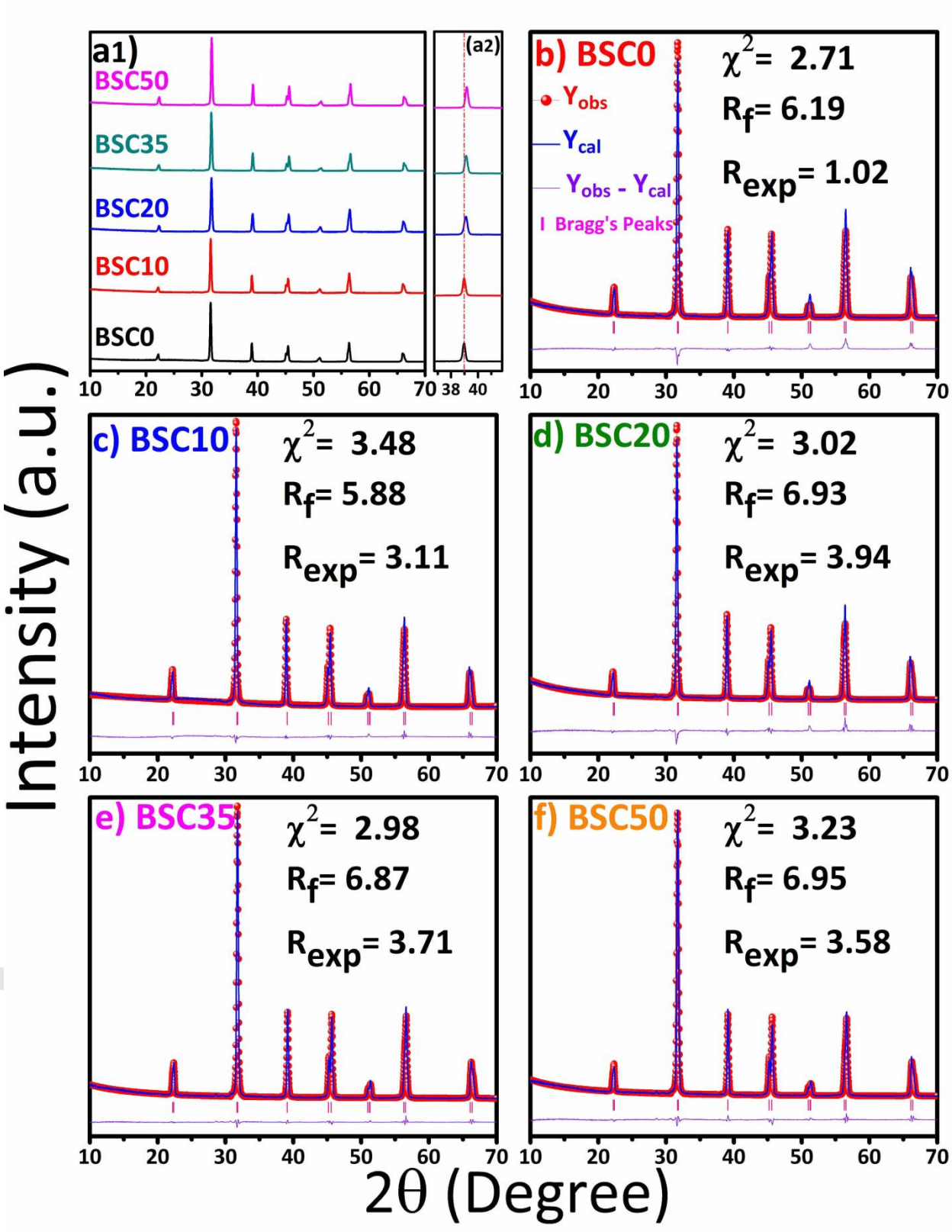
Table 2. Room temperature saturation polarization (P_s), remnant polarization ($2P_r$), coercive field ($2E_c$), density, dielectric breakdown strength (E_{bd}) and piezoelectric constant (d_{33}) for BSC0, BSC10, BSC20, BSC35 and BSC50.

Sample	$2P_s$ ($\mu C/cm^2$)	$2P_r$ ($\mu C/cm^2$)	$2E_c$ (kV/cm)	Density (gm/cm^3)	E_{bd} (kV/cm)	d_{33} (pC/N)
BSC0	23.48	9.17	4.54	5.082	183	213
BSC10	21.55	4.56	4.74	5.293	154	256
BSC20	18.85	4.69	5.15	5.421	173	288
BSC35	21.39	10.16	5.65	5.705	224	293
BSC50	21.60	9.50	9.53	5.477	181	241

Accepted Article

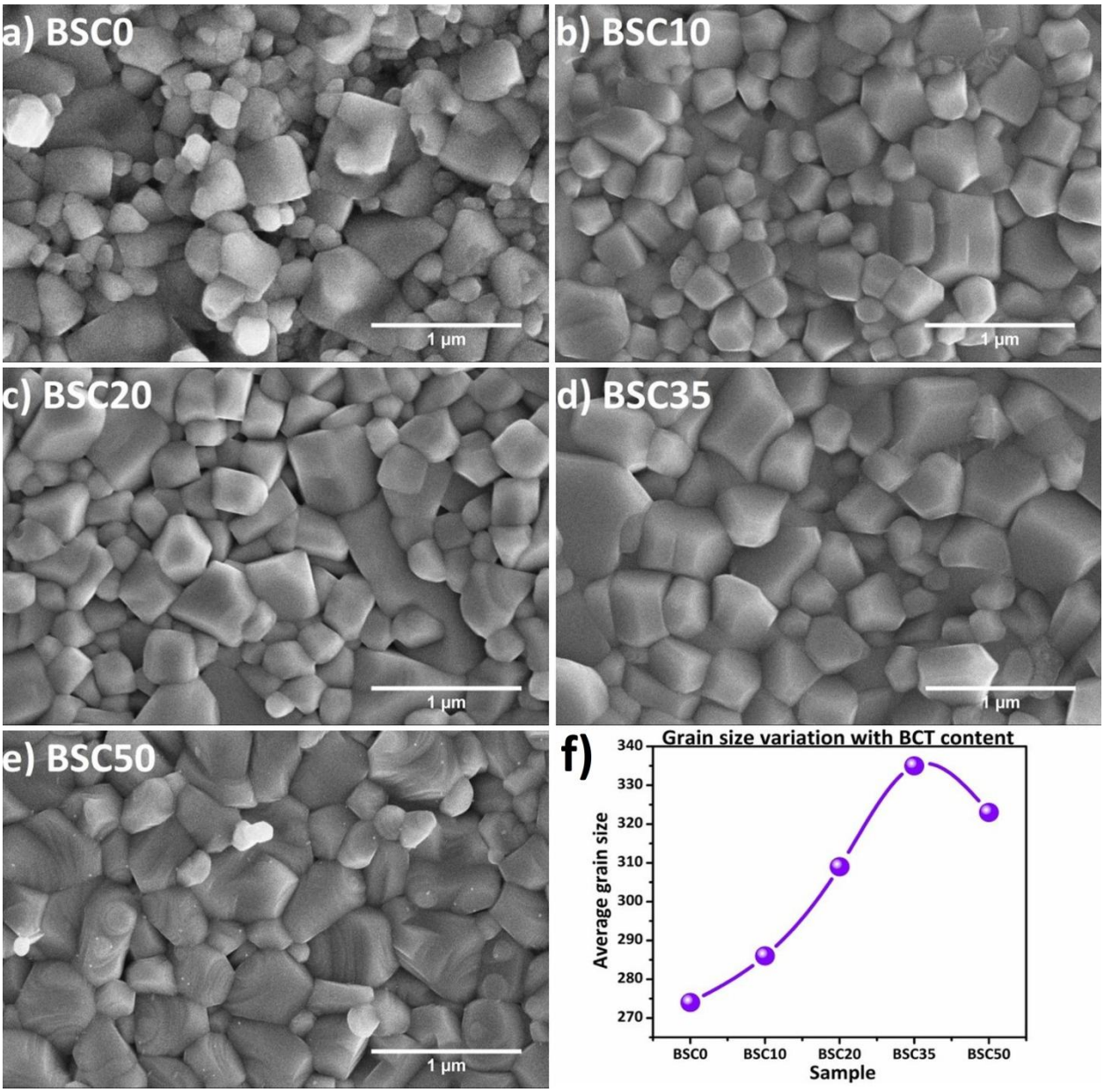
This article is protected by copyright. All rights reserved.

Accepted Article



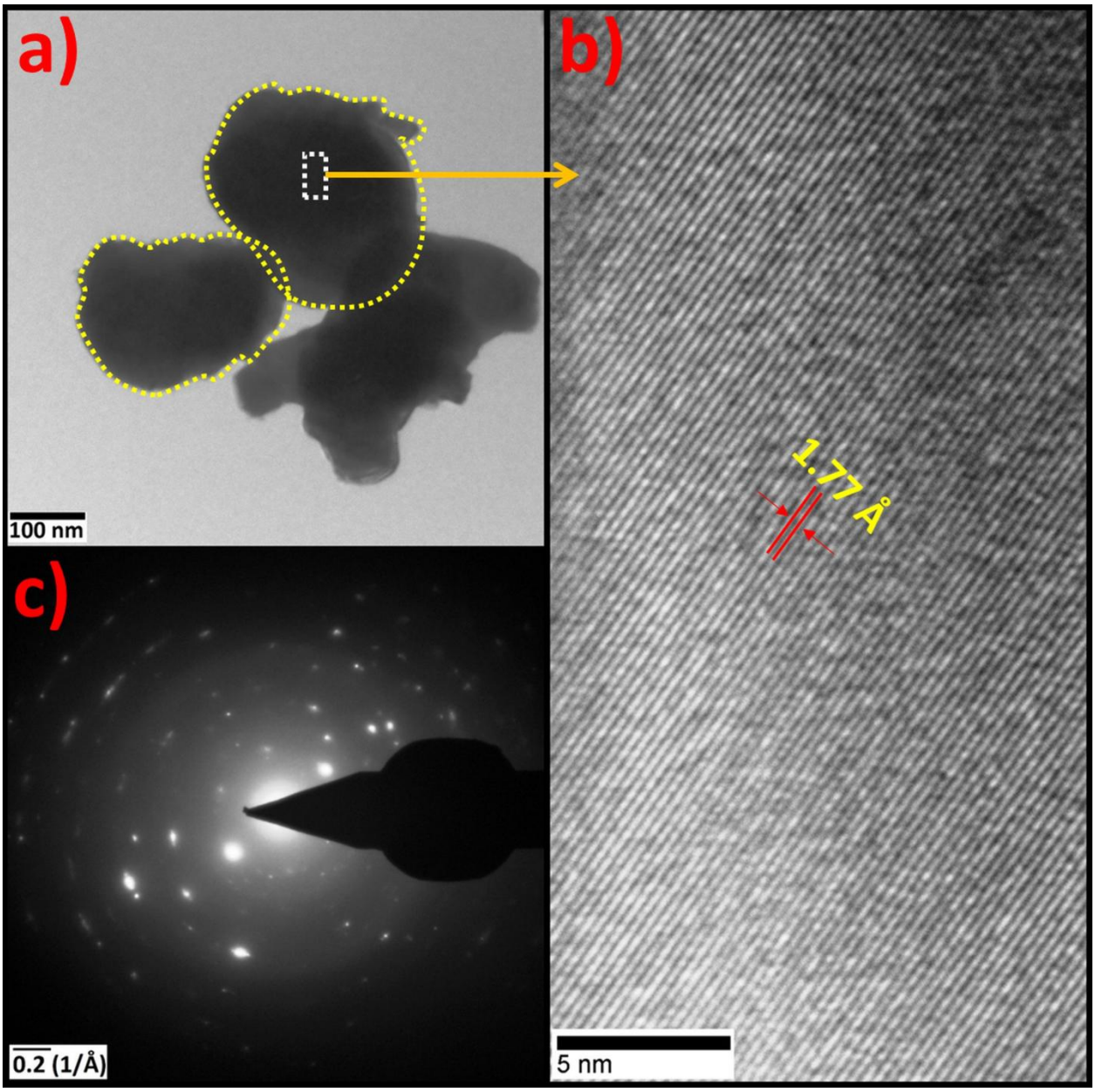
This article is protected by copyright. All rights reserved.

Accepted Article



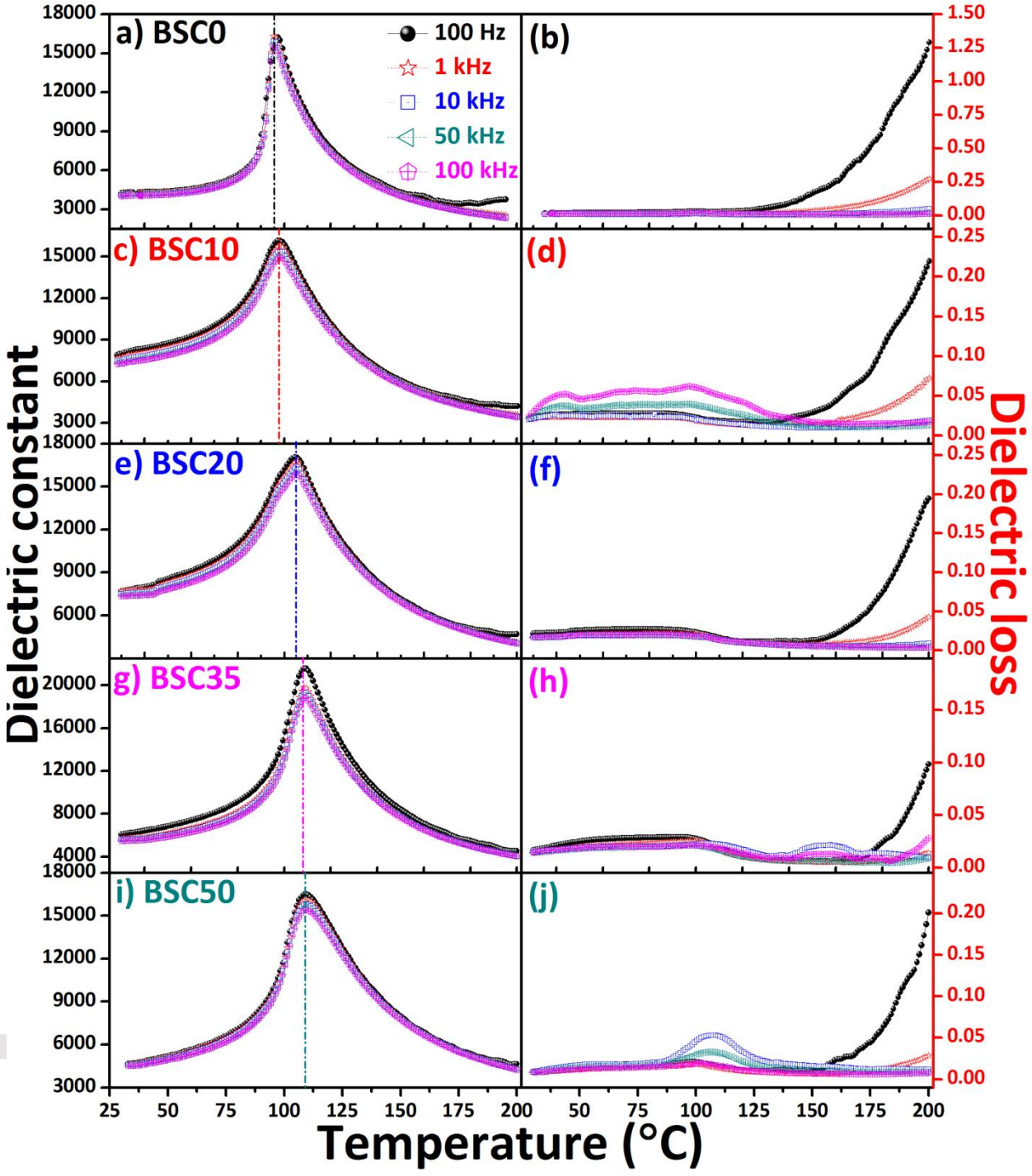
This article is protected by copyright. All rights reserved.

Accepted Article



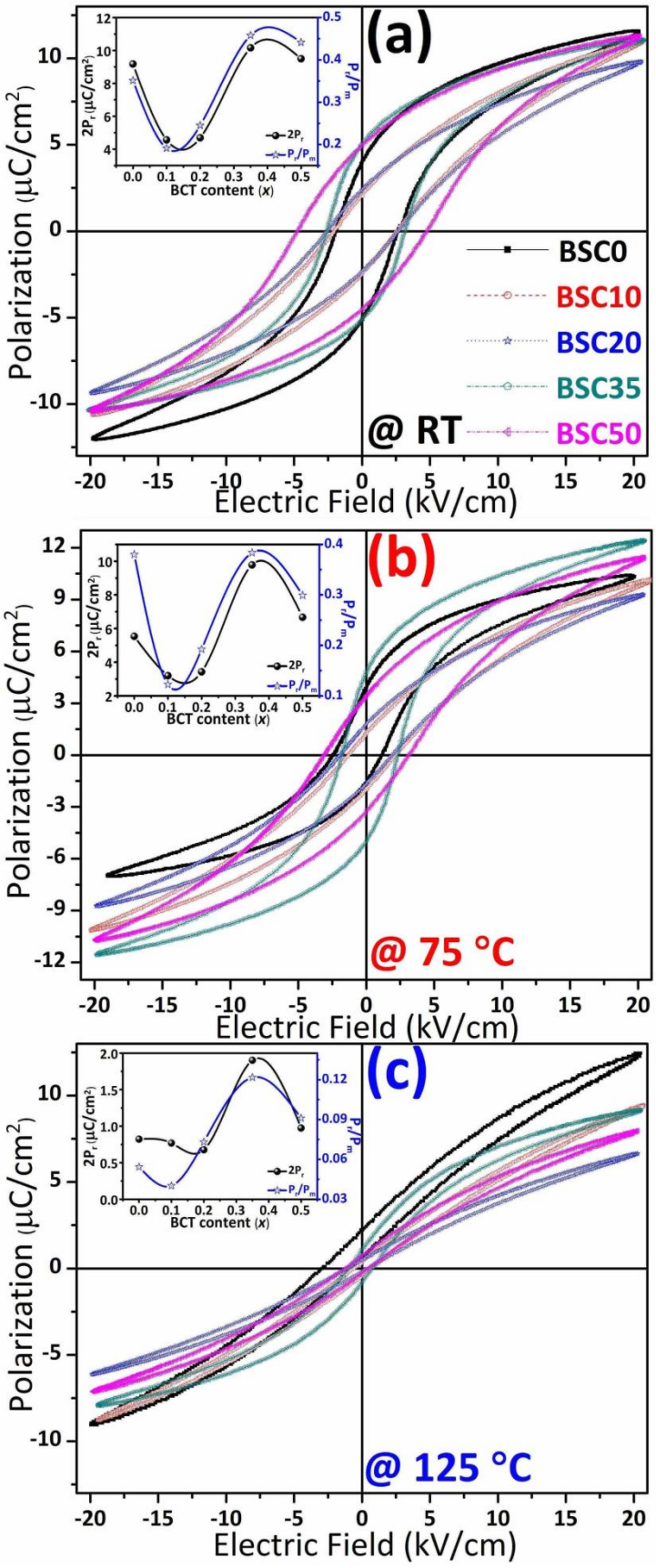
This article is protected by copyright. All rights reserved.

Accepted Article



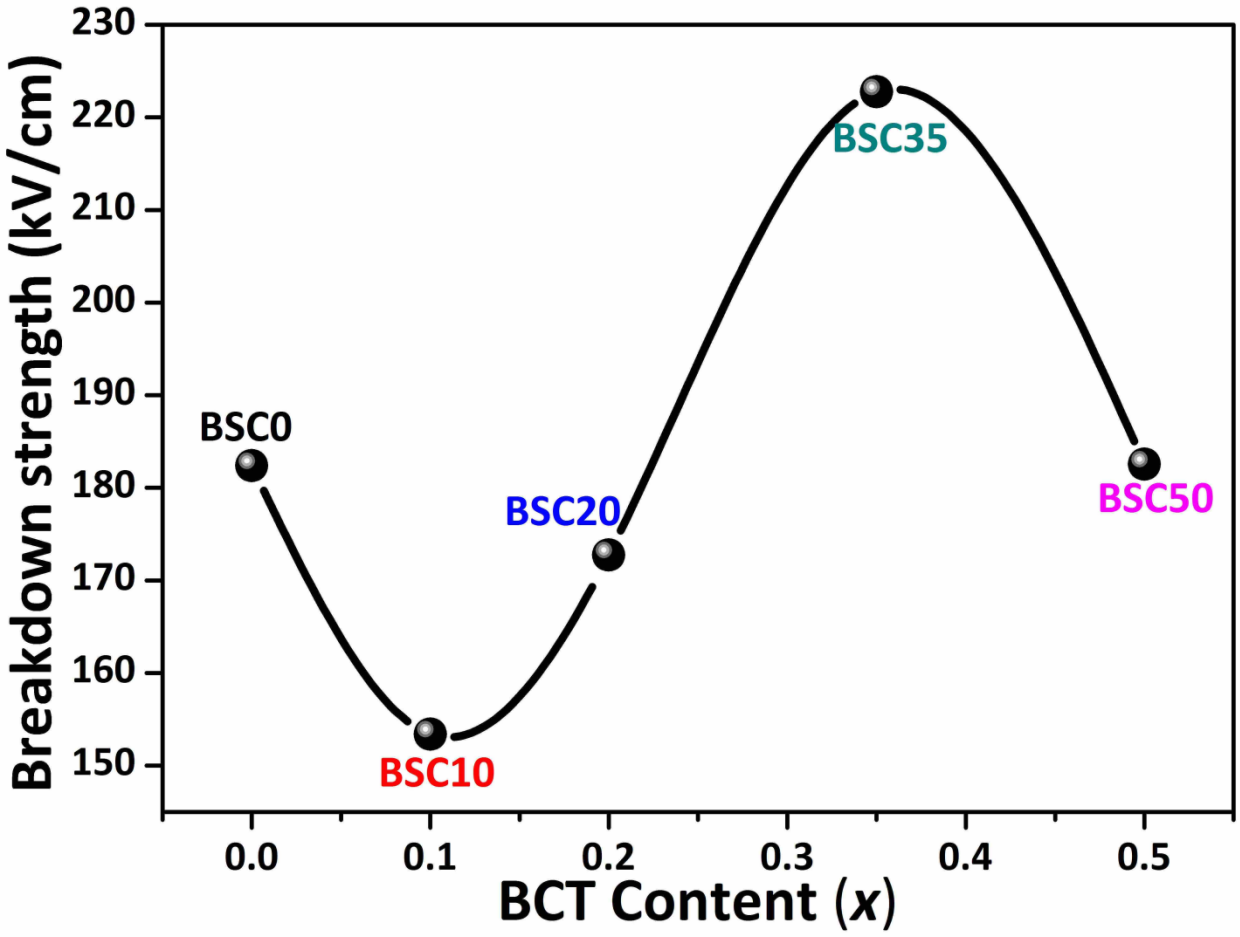
This article is protected by copyright. All rights reserved.

Accepted Article



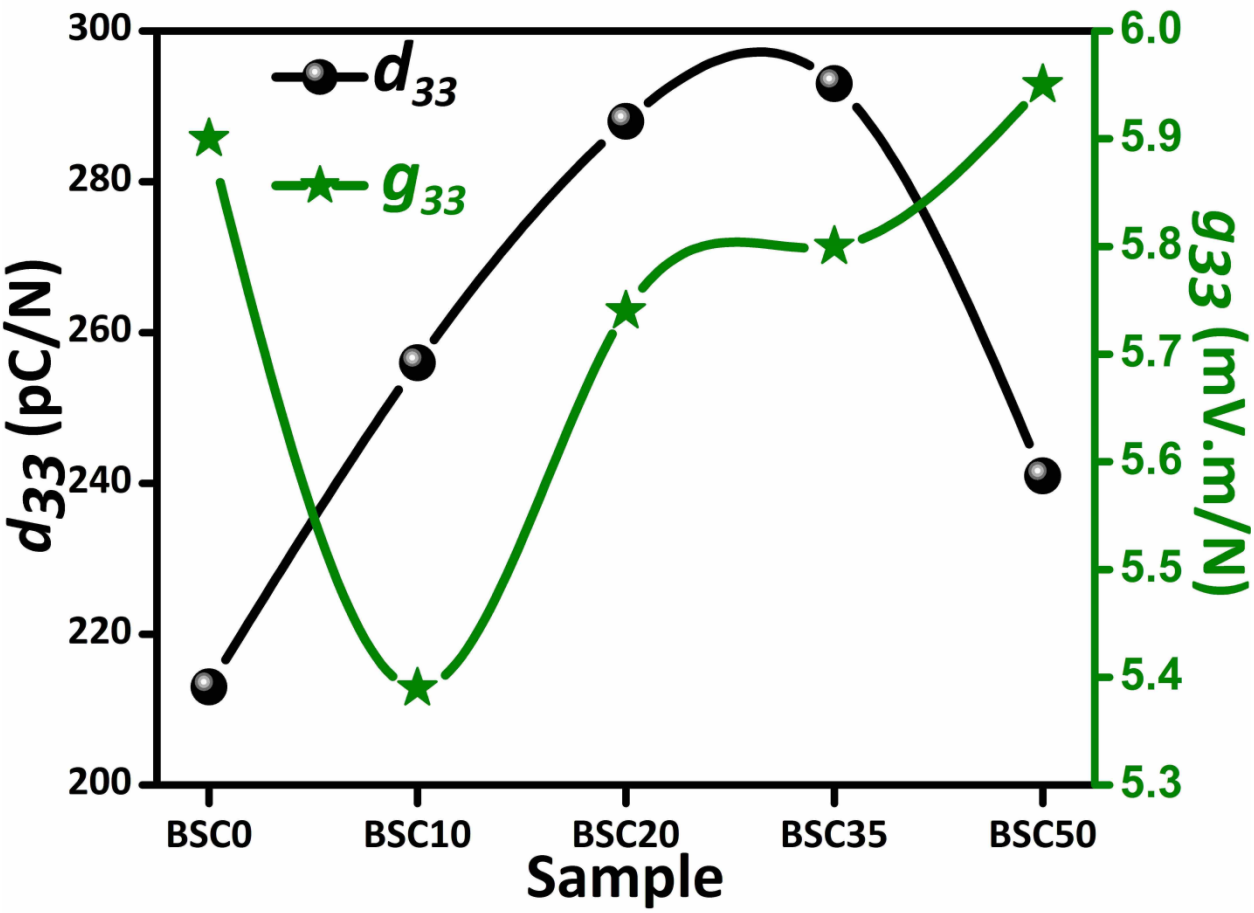
This article is protected by copyright. All rights reserved.

Accepted Article



This article is protected by copyright. All rights reserved.

Accepted Article



This article is protected by copyright. All rights reserved.

Article

Vision
21(3) 1–9
© 2017 MDI
SAGE Publications
sagepub.in/home.nav
DOI: 10.1177/0972262917716763
http://vision.sagepub.com

Evaluating the Feedback Effects of Brand Extension on Parent Brand Equity: A Study on Indian FMCG Industry

Richa Joshi[1]
Rajan Yadav[2]

Abstract

Brand extension is an important strategy to utilize the credibility of the brand and to minimize the advertising costs. The strategy is used in various industries these days but fast moving consumer goods (FMCG) is such an industry which lie closer to a consumer, hence the immediate effect of any strategy can be properly viewed in this industry. The study is an attempt to provide a framework to examine the effects of brand extension strategy on parent brand equity. It is based on two real FMCG brands of Indian market and their brand extensions. Two frameworks are shown based on the two brands Saffola and Fortune with a sample size of 285 and 278 respondents, respectively. Structural equation modelling is used to analyze the effectiveness of both the frameworks. The findings indicate that brand extensions do affect parent brand equity. Therefore, extensions should be introduced in such a manner that they help to strengthen parent brand equity.

Key Words

Brand, Brand Extension, Feedback Effects, Parent Brand Equity

Introduction

Several multinationals are entering into new markets, and they are leveraging their existing equity by practicing brand extension. Hence, it is important to study the factors which affect brand extension (Kaur & Pandit, 2015). Although extensions take advantage from parent brand's existing brand equity, the extensions also impact the equity of parent brand which is termed as reciprocal effects (Ahluwalia & Gurhan-Canli, 2000; Dwivedi & Merrilees, 2013). Successful brands are the assets of the company. They help in creating value for the company. Firms have different strategies for extending brands. The first strategy is to employ individual brand names to different products without any association to the parent company, for example, Procter & Gamble and its various brands. The second strategy is to use an umbrella brand name for every product, for example, Tata Tea, Tata Steel, etc. The third strategy is to use the combination of both, a sub-brand category which has parent brand and the individual brand names, such as Cadbury Oreo and Cadbury Dairy Milk (Pitta & Katsanis, 1995). The need for brand extensions was explained by Tauber (1988). According to Tauber, way back in the 1960s it was believed that to grow a company has to market new products otherwise the company will eventually die. Product life cycle concept was believed by most marketers. By 1980s, a study by Tauber for 7,000 items in 10-year period resulted in the success of only 93 items; hence, new product development was questionable. In 1980s, brand leverage was the strategy where an existing brand name was used for new product category so as to minimize the risk. Also, brand leverage or extension helped in improving the sales of the parent brand. Brand extensions are identified along with another category of extensions called line extensions although there is a slight difference between the two. Brand extension is the use of well-known brand name

[1] Research Scholar, Delhi School of Management, Delhi Technological University, Shahbad Daulatpur, Bawana Road, Delhi, India.
[2] Associate Professor and Head,Delhi School of Management, Delhi Technological University, Shahbad Daulatpur, Bawana Road, Delhi, India.

Corresponding author:
Richa Joshi, Research Scholar, Delhi School of Management, Delhi Technological University, Shahbad Daulatpur Bawana Road, Delhi - 110042, India.
E-mail: Richardias86@gmail.com

to penetrate a new product class (Aaker & Keller, 1990). And a line extension is the use of a recognized brand name for a new offering in the same product class (Reddy, Holak, & Bhat, 1994).

Brand extension has been studied extensively from the last few years. The brand extension studies can be divided into two broad categories: the first category refers to the studies related to the factors affecting brand extension. And the second category refers to the studies which evaluate the effect of brand extension on parent brand which is called as feedback effects on parent brand (Dwivedi & Merrilees, 2013). The current study is the combination of both the categories. It is an attempt to investigate the factors which affect brand extension and the impact of brand extension on parent brand equity. The framework is formulated keeping in view both the aspects of the research carried out till now. So that a new approach to study extensions and branding can be presented.

An extensive literature review has shown that the previous studies were focused on hypothetical brand extensions. Many brand extension studies use fictitious brands (Dens & Pelsmacker, 2010). Therefore, the actual behaviour of customers is not revealed. Also, the Indian context of brand extension has not been studied rigorously. The models and framework proposed in developed markets cannot be replicated in the Indian context. The study is an attempt to reduce the shortcomings and research gaps in the existing literature. The current study is underlining the feedback effects of brand extension on parent brand equity along with the effect of various factors on brand extension are being analyzed simultaneously. The study has identified the importance of brand extension and its feedback effects in the context of Indian brands and Indian population. The study is divided into four parts: first part comprises the literature review followed by research methodology in the second part. In the third segment of study, analysis and hypothesis testing are carried out. In the last part of the study, discussion and the implications of study along with its limitations and future research are covered.

Literature Review and Proposed Framework

Brand extension feedback effects are important while evaluating the effectiveness of brand extension. The reason is that the favourable evaluation of extension is supposed to be crucial in developing equity of a brand (Pitta & Katsani, 1995). Recent studies have focused on determining the usefulness of extension for the established parent brand. Feedback effects of successful extension products can also be negative when the perceived quality of the extension does not succeed in meeting the high-quality level of the parent brand (Volckner, Sattler, & Kaufmann, 2008). In another study (Zimmer & Bhat, 2004), it is shown that the quality of extension did not diminish the parent brand image, it either improved or left it unchanged. The associative-network memory theory is important to understand brand extension evaluation and their feedback effects. Knowledge of brand in the consumer memory is held as information nodes. Once an extension is introduced, the information already held in nodes is transferred to the extension. If consumers are satisfied or dissatisfied, it flows back to information nodes (Keller, 2007). Figure 1 represents the proposed framework of the study.

Consumer Innovativeness

The consumer responds differently to new products (Gatingnon & Robertson, 1985). Consumer innovativeness is defined as a person's tendency to adopt with new ideas or products comparatively earlier than the other people (Klink & Athaide, 2009). Consumer innovativeness is the inherent willingness to innovate which develops market and the product. According to diffusion of innovation theory, innovator is the first individual to adopt an innovation. Consumer innovators find comfort in taking the risk (Rogers, 1983). Every consumer, to some extent, is an innovator; all of us adopt some new ideas or objects in our lives (Hirschman, 1980). Brand extension is usually associated with a relatively new product. Therefore, innovators tend to purchase extension products. Hence, it can be postulated that

H1 Consumer innovativeness has significant positive effect on brand extension evaluation.

Perceived Risk

Bauer (1960) for the first time proposed that consumer behaviour is to be viewed as risk taking. The risk is also differentiated by several other parameters (Dowling & Staelin, 1994). The type of risk first reflects the person's perception of risk which is inbuilt in buying any particular product in a specific product category; this predisposition is referred to as product-category risk (PCR). The second type of risk is the perceived risk associated with the choice of particular product in a product class. Hence, some products in a product class may be considered safe while others may be perceived as risky. It is called product-specific risk (PSR). Entry in a risky product category can be made with a well-known brand name. The brand can serve as risk reliever and increase the acceptance of new extension (Hem, Chernatony, & Iversen, 2003). Established brand helps to alleviate the risk and the extension product is benefitted by such consumer perception. Hence, it can be postulated that

H2 Higher the perceived risk associated with the extension category more positive and significant will be the brand extension evaluation.

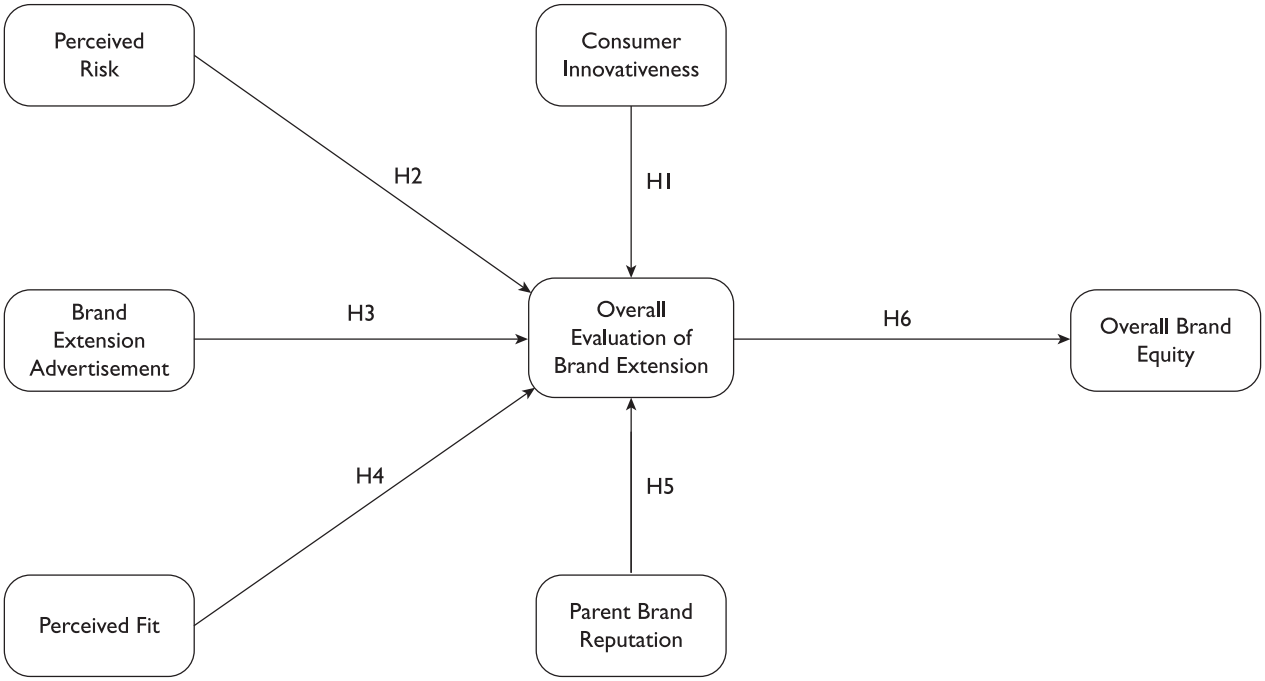


Figure 1. Proposed Conceptual Framework

Source: Authors' own.

Brand Extension Advertisement

The theory of the hierarchy of effects is useful for studying the influence of advertising on the consumer behaviour towards brand extensions. According to this theory, the first aim of advertising is not to sell but to enlarge brand beliefs by articulating the existence of brand (Martinez, Montaner, & Pina, 2009). Advertising of extension product makes consumer aware of the performance of the company and therefore increases the quality perceptions of the parent brand. Advertising of extension product strengthens the parent brand image (Volckner et al., 2008). Advertising is responsible for extension success (Taylor & Bearden, 2003). Thus, it can be postulated that

H3 Brand extension advertisement has significant positive impact on the overall evaluation of brand extension.

Perceived Fit

The perceived fit is the amount of congruence or similarity between the extension and the parent brand (Buil, Chernatony, & Hem, 2009). Consumer reactions to brand extensions involve a classification process in which new product is assessed according to the appropriateness of its membership in a category that already has a set of products which have some identifiable label it is termed as perceived fit. The perceived fit is defined with the help of two concepts: product similarity and concept consistency (Park, Milberg, & Lawson, 1991). Product similarity is referred to as the extent to which consumers perceive the extension as similar to the other products associated with the brand in terms of need they accomplish and situations in which they are used (Smith & Park, 1992). Concept consistency relies on the extension product's ability to hold the brand concept. An extension with good fit fortifies parent brand equity magnitude such as brand image (Zimmer & Bhat, 2004). Therefore, it can be postulated that

H4 Perceived fit has significant positive impact on the overall evaluation of brand extension.

Parent Brand Reputation

The effect of brand extension strategies on parent brand equity is expected to depend on the reputation of the parent brand. Brand extensions of the high equity brands benefit of a more positive attitude. Strong brand enhances the chances of a positive consumer attitude towards brand extensions. Strong brands enjoy high awareness level and clearly defined associations which are transmitted to the extension (Buil et al., 2009). When a new brand is launched consumers do not have prior experience with that brand; therefore, consumers rely greatly on cues, such as brand reputation. Brands with higher alleged reputation provide consumers with greater risk relief (Hem et al., 2003).

Hence, it can be postulated that

H5 Parent brand reputation has significant positive impact on brand extension.

Effect of Brand Extension on Overall Brand Equity

Effect of brand extension on brand image is important when developing strategies. Attitude towards the brand extension changes the consumers associations about the core brand (Martinez & Chernatony, 2004). A new brand extension gives new information to consumers about the parent brand which impacts the existing brand schema. If consumers believe a brand offers high quality, on the whole attitude towards parent brand is improved (Dens & Pelsmacker, 2010). The evaluation of the parent brand will be positive if consumers hold a positive attitude towards the extensions. If consumers possess negative attitude towards the extension, this will convey to the parent brand (Dwivedi, Merrilees, & Sweeney, 2010).

Hence, it can be postulated that

H6 Brand extension has significant positive impact on brand equity of parent brand.

Research Methodology

Research Design

A cross-sectional survey using online and offline methods was administered. In offline survey, respondents were provided with the hard copy of questionnaire of the brand which they were using or have used. In the case of the usage of both the brands, both questionnaires were provided. The online survey was conducted by designing links for both questionnaires with the help of Google Forms. The links of the questionnaires were circulated through e-mails, Facebook and LinkedIn. Again in the online survey also the customers were advised to fill the questionnaires of the brand which they were using and for usage of both brands they were advised to visit the links of both questionnaires and fill them.

Respondents

In view of the lack of formal sampling frame in such surveys, the studies used non-random convenience sampling technique. To ensure the true representation of the population under study, the survey was conducted in all parts of the Delhi among housewives, school teachers and postgraduate students in Delhi NCR. The sample size is 285 respondents for the brand Saffola and 278 respondents for brand Fortune.

Stimulus Selection

Five FMCG brands were chosen on the basis of their visibility in electronic and print media advertisements for their recent brand extensions. Several pre-tests were carried out before finalizing two brands and their respective extensions for the main study. The brands which were chosen for the research for stimulus development were Fair & Lovely (Hindustan Unilever), Saffola (Marico), Fortune (Adani Wilmar Limited), Nivea (Beiersdorf Global AG) and Duracell (Berkshire Hathaway). Their respective brand extensions were Fair & Lovely Face Wash, Saffola Oats, Fortune Rice, Nivea Deo and Duracell Powermat. The pre-tests were divided into a two-step process of determining the brand associations of the respondents with the five chosen brands and the selection of two brands out of five brands.

Area of the Study

The pre-tests were conducted in various regions of Delhi, that is, East Delhi, South Delhi, North Delhi, West Delhi and Central Delhi. In pre-test one, respondents were asked to write down the associations and other thoughts which they could recollect after hearing the name of the brand. Table 1 depicts the associations which the respondents could link with a particular brand.

Pre-test two was conducted with the help of focus group technique. Respondents in this type of research are chosen on the basis of their familiarity with the topic of research, they are in the age group prescribed by the study and have similar socio-cultural characteristics and are able to express their views to the interviewer (Rabiee, 2004). The results of the pre-tests demonstrated that Saffola and Fortune were chosen by respondents in three regions of Delhi (East Delhi, West Delhi and central Delhi) (East Delhi, South Delhi and North Delhi). Other brands, such as fair & lovely, were selected in two regions, and Nivea and Duracell are chosen in one region by the respondents. As a result of all the five pre-tests, two brands are chosen for the study (the brands with maximum frequency of appearance in focus groups):

1. Saffola
2. Fortune

Instrument and Measures

Given the cultural difference, buying habits and nature of Indian consumer necessary modifications and change have

Table 1. Summary of Brand Associations Collected from Questionnaire

Brand	Brand Associations
Fair & Lovely	Bleaching ingredients, cosmetic cream, Fair, whiteness cream, beautiful
Saffola	Health, heart, cholesterol free, less calorie, yellow
Fortune	Refined oil, light, easy availability, for low medium class people
Nivea	Cream, blue colour, tender feel, winters
Duracell	Long lasting battery, keep running, rabbit, powerful battery

Source: Authors' own.

been made in the scales which were available through literature review. Likert seven-point scale is used to frame the questionnaire (ranging from strongly disagree [1] to strongly agree [7]). The scale has been developed with the help of extensive literature review. Four items for the construct Overall Brand Equity were taken from Yoo and Donthu (2001). Out of five items for the construct Overall Evaluation of Brand Extensions, four were taken from Dwivedi et al. (2010) and one was taken from Aaker and Keller (1990). For measuring consumer innovativeness, four items were taken from Klink and Athaide (2009) and two items were taken from Hem et al. (2003). Three items for the construct parent brand reputation were taken from Hem et al. (2003) and three items were taken from Dwivedi et al. (2010). For measuring perceived fit, two items were taken from Hem et al. (2003) and one each from Dwivedi et al. (2010), Aaker and Keller (1990) and Martinez et al. (2009). Perceived risk is measured with the help of five items adapted from Hem et al. (2003). For measuring Brand Extension Advertising impact, two items were adapted from Volcker et al. (2008) and two were taken from Kaur and Pandit (2015). Table 2 shows the scales which are used in the questionnaire.

Reliability and Validity

Reliability

Reliability is the ability of a measuring instrument to give accurate and constant results. The reliability for proposed model 1 was assessed using Cronbach's α. The values of Cronbach's α for the constructs of model 1 depicting the

Table 2. Scales Used in the Questionnaire

Construct	Items
Overall Brand Equity	• It makes sense to buy Saffola/Fortune instead of any other brands. • Even if another brand has the same features as Saffola/Fortune, I buy Saffola/Fortune. • If there is another brand as good as Saffola/Fortune, I still buy Saffola/Fortune. • If another brand is not different from Saffola/Fortune in any way, it is smarter to purchase Saffola/Fortune.
Overall Evaluation of Brand Extensions	• Overall, I am very positive to extension Saffola/Fortune oats. • I am inclined towards extension Saffola/Fortune oats. • According to me extension Saffola/Fortune oats/rice is great. • I admire extension Saffola/Fortune oats/rice a lot. • I feel good about Saffola/Fortune oats/rice.
Consumer Innovativeness	• Overall, I like buying the most recent oats/rice. • If I needed to purchase oats/rice I would buy latest one available. • When I see one new brand of oats/rice product in the store, I often buy it because it is new. • I like to purchase the latest oats/rice before others do. • I like to experience novelty and change in my daily routine. • I like surprises.
Parent Brand Reputation	• Altogether, I am very positive to brand Saffola/Fortune. • Altogether, I am very satisfied with brand Saffola/Fortune. • Altogether, I associate positive things with brand Saffola/Fortune. • The SAFFOLA/FORTUNE brand is interesting. • According to me Saffola/Fortune brand is different from other competing brands. • I feel that I have a clear reason to buy Saffola/Fortune instead of others.
Perceived Fit	• The extension Saffola/Fortune oats/rice makes sense. • Saffola/Fortune brand has skills to launch the brand extension. • How similar or dissimilar are 'new product' to the products usually offered by Saffola/Fortune? • How similar is the position of Saffola/Fortune user with extension Saffola/Fortune oats/rice? • Think about brand Saffola/Fortune how similar is the competence for making original brand and extension Saffola/Fortune oats?
Perceived Risk	• When I'm in front of oats/rice section, I always feel rather unsure about what to pick. • When you buy oats/rice it's easy to make a wrong choice. • It's difficult to know that which oats/rice is the best option in the market. • I should be annoyed with myself if it turned out I'd made the wrong choice. • It's not so dangerous to make a wrong choice of oats/rice.
Brand Extension Advertisements	• Extension product is well supported in advertisements. • Extension product advertisements are memorable. • Frequency of extension product advertisement is 3–4 times a day. • Advertisements of extension product have strong appeal.

Source: Authors' own.

study of brand Saffola were 0.871, 0.899, 0.847, 0.862, 0.873, 0.891 and 0.808.

For proposed model 2, the values of Cronbach's α for the constructs model 2, that is, the study of brand Fortune were 0.868, 0.896, 0.845, 0.861 and 0.875, 0.891 and 0.803. Reliability estimate of 0.7 or higher denotes that the scale is reliable. Reliability between 0.6 and 0.7 is considered acceptable (Hair et al.,2013). Reliability for all constructs is in the suitable range for the models of both brands.

Validity

Construct validity indicates the degree to which the constructs used for the study actually measures the projected performance in comparison to the proposed measurement standards.

The composite reliability values for all the constructs in the model 1 created using brand Saffola were all greater than 0.80, the values exceeded the suggested values of 0.7 (Hair et al., 2013); thus all lie in acceptable range. The AVE scores is greater than 0.50; this establishes convergent validity (values of AVE are > 0.50 exceeding the recommended level, Hair et al., 2013). Discriminant validity is measured by the values of AVE, MSV and ASV. The values of AVE > MSV & ASV indicate that discriminant validity is also established. Table 3 depicts the values of reliability and validity of constructs for model 1 (Saffola).

The composite reliability values for all the constructs in model 2 created using brand Fortune were all greater than 0.80. The AVE scores is greater than 0.50. Both the values of CR and AVE exceeded the suggested values of 0.7 and 0.5 (Hair et al., 2013). Discriminant validity is measured using the values of AVE and MSV, ASV. The values of AVE > MSV and ASV indicate that discriminant validity is also established. Table 4 depicts the values of reliability and validity of constructs for model 2 (Fortune).

Measurement Model Analysis

Model 1

A confirmatory factor analysis was run in AMOS 22.0. The GFI of 0.848 was achieved for the brand Saffola; the values close to 1 indicate good fit (Byrne, 2010); CFI of 0.937 was obtained; CFI value more than 0.90 indicates that model fits well (Hair et al., 2013). TLI of 0.93 was obtained. The model with values that approach 1 suggests a better fit (Hair et al., 2013) was achieved for the study of brand Saffola which was acceptable. Indices of the badness of fit, namely, RMSEA, for brand Saffola was 0.047, and the cut-off value between 0.03 and 0.08 is advisable (Hair et al., 2013).

Table 3. Reliability and Validity of Constructs

	Model 1			
	CR	AVE	MSV	ASV
BEA	0.874	0.634	0.346	0.085
CIN	0.874	0.539	0.360	0.183
PBR	0.900	0.600	0.430	0.250
PR	0.863	0.559	0.346	0.072
PF	0.847	0.526	0.487	0.267
OBE	0.814	0.524	0.419	0.196
OEBE	0.891	0.622	0.487	0.256

Source: Authors' own.

Table 4. Reliability and Validity of Constructs

	Model 2			
	CR	AVE	MSV	ASV
BEA	0.876	0.639	0.343	0.089
CIN	0.871	0.534	0.354	0.176
PBR	0.897	0.592	0.424	0.246
PR	0.862	0.556	0.343	0.073
PF	0.845	0.522	0.472	0.263
OBE	0.809	0.516	0.424	0.196
OEBE	0.891	0.621	0.472	0.257

Source: Authors' own.

Notes: Indices: CR, composite reliability; AVE, average variance extracted; MSV, maximum shared variance; ASV, average shared variance; BEA, brand extension advertisement; CIN, consumer innovativeness; PBR, parent brand reputation; PR, perceived risk; PF, perceived fit; OBE, overall brand equity and OEBE, overall evaluation of brand extension.

Model 2

In the study of brand Fortune GFI of 0.845 was achieved; the values close to 1 indicate good fit (Byrne, 2010); CFI of 0.934 was obtained; CFI value more than 0.90 indicates that model fits well (Hair et al., 2013) and TLI of 0.927 was obtained for brand Fortune. The model with values that approach 1 suggests a better fit (Hair et al., 2013). The value of RMSEA was 0.048, and the cut-off value between 0.03 and 0.08 is advisable (Hair et al., 2013).

Structural Model Analysis

Model 1

The structural model analysis elicits the GFI 0.844 of CFI of 0.929 and TLI of 0.922 for the study of Saffola brand. All the values of model 1, that is, Saffola, indicate the acceptable fit.

Model 2

The structural model analysis elicits the GFI 0.841 of CFI of 0.926 and TLI of 0.919 for the study of Fortune brand. All the values of model 2, that is, Fortune, indicate the acceptable fit.

Hypothesis Testing

An examination of significance levels reveals that all six hypotheses are significant. Table 5 reports the significance

Table 5. Structural Results

Model 1		
Hypothesis	Sig. Level	Hypothesis Test Results
OEBE ◄—— CIN	0.000	Accepted
OEBE ◄—— PR	0.008	Accepted
OEBE ◄—— BEA	0.006	Accepted
OEBE ◄—— PF	0.000	Accepted
OEBE ◄—— PBR	0.000	Accepted
OBE ◄—— OEBE	0.000	Accepted

Source: Authors' own.

levels for the brand Saffola. It suggests that consumer innovativeness has significant impact on overall evaluation of brand extension ($p < 0.01$), thus supporting H1. Perceived risk also has significant impact on overall evaluation of brand extension ($p < 0.05$); thus H2 is supported. Brand extension advertisement has significant impact on overall evaluation of brand extension ($p < 0.05$); H3 is hence supported. Perceived fit also has significant impact on overall evaluation of brand extension ($p < 0.01$), thus supporting H4. Parent brand reputation has exerted a significant impact on overall evaluation of brand extension ($p < 0.01$); H5 is also supported. Overall evaluation of brand extension also had a significant impact on overall brand equity ($p < 0.01$); therefore H6 is also supported.

Table 6 reports the significance levels for the brand Fortune. It suggests that consumer innovativeness has a significant impact on the overall evaluation of brand extension ($p < 0.01$), hence supporting H1. Perceived risk also has a significant impact on the overall evaluation of brand extension ($p < 0.05$); thus H2 is supported. Brand extension advertisement has a significant impact on the overall evaluation of brand extension ($p < 0.05$); H3 is hence supported. Perceived fit also has a significant impact on the overall evaluation of brand extension ($p < 0.01$), thus supporting H4. Parent brand reputation has exerted a significant impact on the overall evaluation of brand extension ($p < 0.01$); H5 is also supported. Overall evaluation of brand extension also had a significant impact on overall brand equity ($p < 0.01$); therefore H6 is also supported.

Table 6. Structural Results

Model 2		
Hypothesis	Sig. Level	Hypothesis Test Results
OEBE ◄—— CIN	0.000	Accepted
OEBE ◄—— PR	0.016	Accepted
OEBE ◄—— BEA	0.009	Accepted
OEBE ◄—— PF	0.000	Accepted
OEBE ◄—— PBR	0.000	Accepted
OBE ◄—— OEBE	0.000	Accepted

Source: Authors' own.

Discussion and Implications

The present study penetrates the brand extension feedback effects in much deeper context. The literature present till date had although presented several frameworks which were empirically tested, but the current research has undergone a further exhaustive examination of factors affecting brand extension. Brand extension advertisements emerged out to be one of the important factors that affect brand extension. In most of the studies, the brand extension strategy is underlined and emphasized (Hem et al., 2003; Kaur & Pandit, 2015; Afzal, 2013; Tauber, 1988). The finding is consistent with the literature (Kaur & Pandit, 2015). Consumer innovativeness emerged out to be one of the important factors which affect brand extensions. More innovative consumers evaluate brand extension positively (Hem et al., 2003). Perceived risk has also been established as one of the important factors which affect the brand extension. There is always some risk associated with new product and this reinforces the perception of brands as risk relievers. And such benefit is shared by the extension product (Hem et al., 2003). Advertising is a manner to promote the new product; however, extension product advertising serves the purpose of reinforcement. If a brand has good image, the advertising campaigns of the extension reminds consumer about the brand association related to that brand (Martinez et al., 2009). The findings of the article reveal that brand extension advertisements do affect brand extension. Perceived fit has a positive effect on brand extensions and it has been shown with the help of two FMCG brands. This is consistent with the literature (Aaker & Keller, 1990; Buil et al., 2009; Kaur & Pandit, 2014; Martinez & Chernatony, 2004; Park et al., 1991) and services (Dwivedi & Merrilees, 2013; Hem et al., 2003). It has been a governing force in explaining success of extensions in the literature of feedback effects (Dwivedi et al., 2010). Another important factor observed is parent brand reputation. It can be seen that stronger brands provide more influence for extension than weaker brands (Hem et al., 2003). The findings of the study are coherent with the literature. The feedback effects on parent brand are not studied on real brand extensions in the Indian context. Therefore, the study is one of its kinds which mark the empirical testing of a framework which not only displays the brand extension aspect but its further impact on parent brand equity is also demonstrated. Though the results are found significant in Indian context, it shows the positive impact of brand extension on parent brand equity (Dwivedi & Merrilees, 2013).

The first major contribution is that the study moves beyond the hypothetical extensions to the real ones. Also, the extensions covered are FMCG products which exhibit high-frequency purchases. The consumer behaviour can be studied much easily in such kind of products since they are purchased continuously and the element of biases is reduced. The second major contribution is the group of

respondent belonged to different socio-economic status. The group is heterogeneous in nature, covering a diverse age group, education and profile. Therefore, results have practical significance. Managers can utilize the results in constructing the strategies for the success of brand extensions. Brand extensions should be introduced by evaluating the factors which affect them. The impact of brand extension on brand equity shows that if extensions are strategically positioned, they can enhance parent brand equity. The portion of parent brand equity which was capitalized by brand extensions is regained rather amplified by successful extensions.

Limitations and Future Research

The current study identifies few inherent limitations. First, the study is based on only two FMCG brands of the Indian market. The FMCG industry in India is a big industry with thousands of brands. Therefore, the study of two brands limits the generalizability of framework. The second limitation is that the current study is carried out in Delhi which abridges the applicability of results universally. Future studies can be done covering the metropolitan cities of India. Along with the Indian view of brand extension, cross-cultural studies can be carried out in Asian, European and American perspective regarding the spill over effects of brand extension. Third, the current study has not considered the fact that a parent brand can also be affected by several other brand extensions which are present in the market. For example, the other extensions of Saffola are Saffola Muesli and Saffola Salt plus. Similarly, Fortune also has other extensions, such as Fortune besan, Fortune pulses and Fortune soya chunks, which are not covered in the present study.

References

Aaker, D. A., & Keller, K. L. (1990). Consumer evaluations of brand extensions. *Journal of Marketing*, *54*(1), 27–41.

Afzal, S. (2013). Factors influencing successful brand extension into related and unrelated product categories. *Asian Economic and Financial Review*, *3*(2), 216–226.

Ahluwalia, R., & Gurhan-Canli, Z. (2000). The effects of extensions on the family brand name: An accessibility-diagnosticity perspective. *Journal of Consumer Research*, *27*(3), 371–381.

Bauer, R. A. (1960). Consumer behavior as risk taking, dynamic marketing for changing world. In Robet S. Hancock (Ed.), *Dynamic marketing for a changing world* (pp. 389–398). Chicago, IL: American Marketing Association.

Buil, I., Chernatony, L. de, & Hem, L. E. (2009). Brand extension strategies: Perceived fit, brand type, and culture influences. *European Journal of Marketing*, *43*(11/12), 1300–1324.

Byrne, B. M. (2010). *Structural equation modeling with Amos: Basic concepts, applications, and programming* (Multivariate Applications Series, 2nd ed.). New York, NY: Routledge.

Dens, N., & Pelsmacker, P. De. (2010). Attitude toward the extension and parent brand in response to extension advertising. *Journal of Business Research*, *63*(11), 1237–1244.

Dowling, G. R., & Staelin, R. (1994). A model of perceived risk and intended risk handling activity. *Journal of Consumer Research*, *21*(June) 119–134.

Dwivedi, A., Merrilees, B., & Sweeney, A. (2010). Brand extension feedback effects: A holistic framework. *Journal of Brand Management*, *17*(5), 328–342.

Dwivedi, A., & Merrilees, B. (2013a). Brand extension feedback effects: Towards a mediated framework. *Journal of Consumer Marketing*, *30*(5), 450–461.

———. (2013b). Brand-extension feedback effects: An Asian branding perspective. *Asia Pacific Journal of Marketing and Logistics*, *25*(2), 321–340.

Gatingnon, H., & Robertson, T. S. (1985). A propositional inventory for new diffusion research. *Journal of Consumer Research*, *11*(4), 849–867.

Gurhan Canli, Z., & Maheswaran, D. (1998). The effects of extensions on brand name and enhancement. *Journal of Marketing Research, 35*(November), 464–473.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2013). *Multivariate data analysis* (9th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Hem, L., Chernatony, L. D., & Iversen, N. M. (2003). Factors influencing successful brand extensions. *Journal of Marketing Management*, *19*(7/8), 781–806.

Hem, L., Iversen, N. M., & Olsen, L. E. (2014). Category characteristics effects on brand extension attitudes: A research note. *Journal of Business Research, 67*(8), 1589–1594.

Hirschman, E. C. (1980). Innovativeness, novelty seeking, and consumer creativity. *Journal of Consumer Research, 7*(3), 283–295.

Kaur, H., & Pandit, A. (2014). Consumer evaluation of brand extension: Empirical generalization and comparative analysis. *Journal of Empirical Generalizations in Marketing Science*, *15*(1), 37–48.

Kaur, H., & Pandit, A. (2015). Modelling consumer evaluation of brand extensions: Empirical evidence from India. *Vision, 19*(1), 37–48.

Keller, K. L. (1993). Conceptualizing, measuring and managing customer based brand equity. *Journal of Marketing*, *57*(1), 1–22.

———. (2007). *Strategic brand management: Building measuring and managing brand equity*. Upper Saddle River, NJ: Prentice Hall.

Klink, R. R., & Athaide, G. A. (2009). Consumer innovativeness and the use of new versus extended brand names for new products. *Journal of Product Innovation & Management*, *27*(1), 23–32.

Martinez, E., & Chernatony, L. de. (2004). The effect of brand extension strategies on brand image. *Journal of Consumer Marketing*, *21*(1), 39–50.

Martinez, E., Montaner, T., & Pina, J. (2009). Brand extension feedback: The role of advertising. *Journal of Business Research*, *62*(3), 305–313.

Park, C. W., Milberg, S., & Lawson, R. (1991). Evaluation of brand extensions: The role of product feature similarity and brand concept consistency. *Journal of Consumer Research*, *18*(2), 185–193.

Pitta, D. A., & Katsanis, L. P. (1995). Understanding brand equity for successful brand extension. *Journal of Consumer Marketing*, *12*(4), 51–64.

Rabiee, F. (2004). Focus-group interview and data analysis. *Proceedings of the Nutrition Society*, *63*(4), 655–660.
Reddy, S. K., Holak, S. L., & Bhat, S. (1994). To extend or not to extend: Success determinants of line extensions. *Journal of Marketing Research*, *31*(2), 243–262.
Rogers, E. M. (1983). *Diffusion of innovations* (3rd ed.). The Free Press. London: Collier Macmillan.
Smith, D. C., & Park, C. W. (1992). The effects of brand extensions on market share and advertising efficiency. *Journal of Marketing Research*, *29*(3), 296–313.
Tauber, E. M. (1988). Brand leverage: Strategy for growth in a cost control world. *Journal of Advertising Research*, *28*(August), 26–30.
Taylor, V. A., & Bearden, W. O. (2003). Ad spending on brand extensions success. *Brand Management*, *11*(1), 63–74.
Volckner, F., Sattler, H., & Kaufmann, G. (2008). Image feedback effects of brand extensions: Evidence from a longitudinal field study. *Marketing Letters*, *19*(2), 109–124.
Yoo, B., & Donthu, N. (2001). Developing and validating a multidimensional consumer-based brand equity scale. *Journal of Business Research*, *52*(1), 1–14.
Yoo, B., Donthu, N., & Lee, S. (2000). An examination of selected marketing mix elements and brand equity. *Journal of the Academy of Marketing Science, 28*(2), 195–211.
Zimmer, M. R., & Bhat, S. (2004). The reciprocal effects of extension quality and fit on parent brand attitude. *Journal of Product and Brand Management*, *13*(1), 37–46.

About the Authors

Richa Joshi (richardias86@gmail.com) is a Research scholar in Delhi School of Management, Delhi Technological University and she is also working as an Assistant Professor in Management with Gitarattan International Business School, an institute affiliated with GGSIP University. The area of her research is Branding. She has 8 years of experience of teaching in various undergraduate and postgraduate courses of GGSIP University. Ms Richa has done her graduation from Delhi University and Postgraduation from GGSIP University with the qualification of UGC-NET in Management.

Rajan Yadav (raj_yadav1974@yahoo.co.in) is an Associate Professor and head in Delhi School of Management, Delhi Technological University. His interest area is marketing management. He graduated from Delhi University and obtained his Master's Degree in Business Management from the MD University. He was awarded PhD in Organized Retail and also qualified National Eligibility Test (NET) conducted by UGC. His research interests include behavioural dimensions of technology diffusion, online social networks, retail management and student community -related social marketing issues. He has more than 15 years of postgraduate teaching experience.

Feasible star – delta and delta – star transformations for reliability networks

V.C. PRASAD

Department of Electrical Engineering, Faculty of Engineering, Dayalbagh Educational Institute, Dayalbagh, Agra 282005, India . prasadvishdelhi@gmail.com (Retired from Indian Institute of Technology, New Delhi, 110016, India).

Summary: Consider the problem of transforming a star (delta) into an equivalent delta (star) in a reliability network with imperfect undirected edges and perfect vertices. It is believed that such transformations are not possible in general if the probabilities of the elements of the given star / delta are rational numbers. Contrary to this, it is shown here that star – delta and delta – star transformations are possible under certain conditions. Further the probability of success of an element of an equivalent star (delta) is shown to be equal to the probability of failure of the corresponding element of the given delta (star).

Key words:
star – delta transformation, delta – star transformation, series parallel reliability networks

I. INTRODUCTION

Consider the problem of evaluating the reliability of a network with respect to a pair of terminals, s (source/input) and t (load / output). All elements are bidirectional and imperfect and the probabilities of success/failure of elements are known. All nodes are perfect. We wish to calculate the probability of success of the output i.e., there is connectivity between input and output. This problem is important for most systems including many electronic, electrical and power networks [1], [2]. It is well known that the evaluation of reliability of such networks is a NP hard problem [14]. On the other hand evaluation of reliability of series parallel networks is polynomial [14]. Therefore series parallel techniques [1] – [14] are important for reliability analysis. Star (Delta) — Delta (Star) transformations help us to transform some networks into series parallel networks. Further they help us to reduce computation even when series parallel techniques do not work on the entire network. Thus star ↔ delta transformations are important. They attracted the attention of many researchers in the last few decades [2] – [11]. Wang and Sun [10] studied this problem extensively and gave conditions under which star — delta and delta — star transformations are possible. But they argued that these conditions are never satisfied in practice except in special cases in which there are flow restrictions as in Gupta and Sharma [7] and Prasad [11]. Thus at present it is widely believed that star — delta and delta — star transformations are not possible in real life networks. Contrary to this it is shown in this paper that these transformations are possible. Therefore Wang and Sun's conditions are useful in practice. They are not just theoretical. Their conditions are greatly simplified to achieve this. It is further shown that the probability of success of an element of an equivalent delta (star) is same as the probability of failure of the corresponding element of the given star (delta).

II. STAR — DELTA, DELTA — STAR TRANSFORMATIONS

The star and delta networks under consideration are shown in Figure 1. Throughout this paper we assume that no element of star or delta has a probability zero or one. There is no loss of generality in this assumption because transformation of a star or delta with zero or one probability for one or more elements is trivial [10].

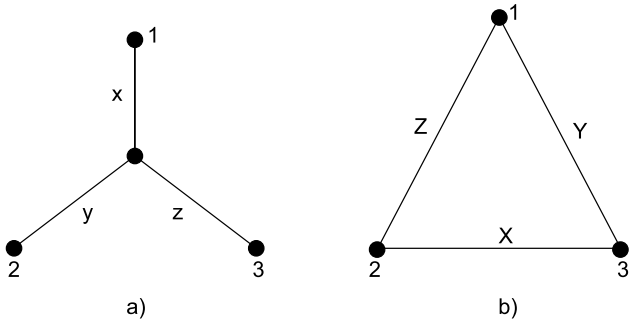


Fig. 1 (a) Star network, (b) Delta network.

2.1. Star — Delta transformation

Given the star of Figure 1 (a) with probabilities of success p_x, p_y and p_z, we want to determine p_X, p_Y and p_Z, the probabilities of success of the elements of the equivalent delta. Prasad [11] studied this problem when there are restrictions in flows. Wang and Sun [10] studied the general case. They partitioned the connection states into four classes for each node, derived equations, simplified them and got the following expressions.

$$p_X = \frac{S_3}{K + S_3} \tag{1}$$

$$p_Y = \frac{S_5}{K + S_5} \tag{2}$$

$$p_Z = \frac{S_6}{K + S_6} \tag{3}$$

where

$$K = S_0 + S_1 + S_2 + S_4 \tag{4}$$

S_0, S_1, S_2, S_3, S_4, S_5, S_6, S_7 are the eight possible states of star. S_0 is $q_x\, q_y\, q_z$. It corresponds to 000. S_1 is $q_x\, q_y\, p_z$. It corresponds to 001. Other states are also understood similarly. Wang and Sun [10] showed that equations (1), (2) and (3) can be used provided the probabilities p_x, p_y and p_z of the elements of star satisfy the equation

$$K^2 = (K + S_3)(K + S_5)(K + S_6) \tag{5}$$

i.e., star to delta transformation is possible if equation (5) holds.

We will first simplify this equation so that it is easy to use. It can be written as: $K^3 + K^2(S_3 + S_5 + S_6 - 1) + + K(S_3S_5 + S_5S_6 + S_3S_6) + S_3S_5S_6 = 0$.

But $K = S_0 + S_1 + S_2 + S_4 = 1 - (S_3 + S_5 + S_6 + S_7)$ i.e., $S_3 + S_5 + S_6 - 1 = -(K + S_7)$. Substitute into the above equation and simplify. This gives:

$$S_7K^2 - BK - C = 0 \tag{6}$$

where $B = S_3S_5 + S_5S_6 + S_3S_6$ and $C = S_3S_5S_6$.

$B = (q_x\, p_y\, p_z)(p_x\, q_y\, p_z) + (p_x\, q_y\, p_z)(p_x\, p_y\, q_z) + (q_x\, p_y\, p_z)(p_x\, p_y\, q_z) = p_x\, p_y\, p_z(q_x\, q_y\, p_z + p_x\, q_y\, q_z + q_x\, p_y\, q_z) = = S_7(S_1 + S_4 + S_2)$.

Similarly $C = (q_x\, p_y\, p_z)(p_x\, q_y\, p_z)(p_x\, p_y\, q_z) = S_7^2 S_0$

Substitute into equation (6) and cancel S_7 throughout. This gives: $K^2 - K(S_1 + S_2 + S_4) - S_7S_0 = 0$ i.e., $K[K - (S_1 + S_2 + S_4)] - S_7S_0 = 0$. This implies that $KS_0 - S_7S_0 = 0$. Thus equivalent delta exists if the condition $K = S_7$ is satisfied.

But $K = (S_0 + S_1 + S_2 + S_3) - S_3 + S_4 = q_x - q_x\, p_y\, p_z + + p_x\, q_y\, q_z$. $K = S_7$ means that $q_x + p_x\, q_y\, qz = p_y\, p_z$. This gives:

$$q_x + q_y + q_z = 1 + q_x\, q_y\, q_z \tag{7}$$

This equation is a simplified form of equation (5).

$q_x = \frac{2}{7},\ q_y = \frac{1}{2},\ q_z = \frac{1}{4}$ satisfies this equation. Thus a star with these probabilities has an equivalent delta showing that Wang and Sun's condition for star delta transformation is useful in practice. Further Theorem 2 of Wang and Sun which says that star to delta transformation is not possible for rational probabilities of star is wrong. Substituting these probabilities into equations (1), (2) and (3), the probabilities of equivalent delta are $p_X = \frac{2}{7},\ p_Y = \frac{1}{2}$ and $p_Z = \frac{1}{4}$.

Some irrational numbers also satisfy equation (7).

$q_x = \frac{1}{\sqrt{2}},\ q_y = \frac{\sqrt{2}}{7},\ q_z = \frac{7\sqrt{2} - 9}{6\sqrt{2}}$ is an example.

Equations (1), (2) and (3) help us to calculate the probabilities of the elements of delta from those of star. But they can be simplified as follows.

Theorem 1: $p_X = q_x$ if equation (7) is satisfied.

Proof: We want to prove that $\frac{S_3}{K + S_3} = q_x$ i.e., $K + S_3 = p_y\, p_z$ as $S_3 = q_x\, p_y\, p_z$. But $K + S_3 = S_0 + S_1 + S_2 + S_3 + S_4 = q_x + p_x\, q_y\, q_z$ as $q_x = S_0 + S_1 + S_2 + S_3$. $p_y\, p_z = = (1 - q_y)(1 - q_z) = 1 - q_y - q_z + q_y\, q_z$. So we want $q_x + q_y + q_z + p_x\, q_y\, q_z = 1 + q_y\, q_z$. Using equation (7), L.H.S $= 1 + q_x\, q_y\, q_z + p_x\, q_y\, q_z$ = R.H.S. Hence the result.

Theorem 2: $p_Y = q_y$ if equation (7) is satisfied.

Theorem 3: $p_Z = q_z$ if equation (7) is satisfied.

The proofs of Theorems 2 and 3 are similar to Theorem 1.

2.2 Delta – Star Transformation:

Given the delta of Figure 1(b) with probabilities p_X, p_Y, p_Z, we want to determine the probabilities p_x, p_y, p_z of the equivalent star. Gupta and Sharma [7] studied this problem when there are flow restrictions. Wang and Sun [10] considered the general case and derived the following expressions.

$$p_x = \frac{N}{N + D_4} \tag{8}$$

$$p_y = \frac{N}{N + D_2} \tag{9}$$

$$p_z = \frac{N}{N + D_1} \tag{10}$$

where $N = D_3 + D_5 + D_6 + D_7$ (11)

and D_0, D_1, D_2, D_3, D_4, D_5, D_6, D_7 are states of delta. D_0 is $q_X\, q_Y\, q_Z$. It corresponds to 000. D_2 is $q_X\, p_Y\, q_Z$ corresponding to 010. Other states are also understood in a similar way. Wang and Sun [10] showed that equivalent star exists if

$$N^2 = (N + D_1)(N + D_2)(N + D_4) \tag{12}$$

Simplifying this we get

$$D_0N^2 - BN - C = 0 \tag{13}$$

where $B = D_1D_2 + D_2D_4 + D_1D_4$ and $C = D_1D_2D_4$. It can be shown that $B = D_0(D_3 + D_6 + D_5)$ and $C = D_0^2 D_7$. Substitute in equation (13) and simplify. This gives $N = D_0$. But $N = (D_4 + D_5 + D_6 + D_7) - D_4 + D_3 = p_X - - p_X\, q_Y\, q_Z + q_X\, p_Y\, p_Z$. Substitute for N and D_0 and simplify this equation. We can then show that the required condition for the existence of an equivalent star is

$$p_X + p_Y + p_Z = 1 + p_X p_Y p_Z \qquad (14)$$

This equation is a simplified form of equation (12). $p_X = \frac{2}{7}$, $p_Y = \frac{1}{2}$ and $p_Z = \frac{1}{4}$ satisfy this equation. Thus a delta with these probabilities has an equivalent star showing that Wang and Sun's condition for delta star transformation is useful in practice. Further Theorem 3 of Wang and Sun which says that delta to star transformation is not possible for rational probabilities of delta is wrong. Substituting these probabilities into equations (8), (9) and (10), the probabilities of equivalent star are $p_x = \frac{5}{7}$, $p_y = \frac{1}{2}$, and $p_z = \frac{3}{4}$.

$p_X = \frac{1}{\sqrt{2}}$, $p_Y = \frac{\sqrt{2}}{7}$, $p_Z = \frac{7\sqrt{2}-9}{6\sqrt{2}}$ are an example of irrational numbers which satisfy equation (14).

Equations (8), (9) and (10) can be used to calculate the probabilities of the elements of star from those of delta. But these expressions can be simplified as shown below.

Theorem 4: $p_x = q_X$ if equation (14) is satisfied.
Theorem 5: $p_y = q_Y$ if equation (14) is satisfied.

Proof: $p_y = q_Y$ if $\frac{D_2}{N + D_2} = p_Y$.

So we want to prove that $N + D_2 = q_X q_Z$ as $D_2 = q_X p_Y q_Z$. $N + D_2 = D_3 + D_5 + D_6 + D_7 + D_2 = p_Y + p_X q_Y p_Z$ where $D_2 + D_3 + D_6 + D_7 = p_Y$. $q_X q_Z = (1 - p_X)(1 - p_Z) = 1 - p_X - p_Z + p_X p_Z$. Substituting we want to prove that $p_X + p_Y + p_Z + p_X q_Y p_Z = 1 + p_X p_Z$.

Using equation (14), L.H.S. = $1 + p_X p_Y p_Z + p_X q_Y p_Z$ = R.H.S. Hence the result.

Theorem 6: $p_z = q_Z$ if equation (14) is satisfied.

Proofs of Theorems 4 and 6 are similar to Theorem 5. It is clear from Theorems (1) to (6) that equations (7) and (14) imply each other. Let S_D (S_Y) be the set of all deltas (stars) satisfying equations (14), (7). Let S'_Y (S'_D) be the set of all equivalent stars (deltas) obtained using the star – delta and delta – star transformations described above. Then $S_D = S'_D$ and $S_Y = S'_Y$ because equations (7) and (14) are equivalent.

The following example illustrates the use of star ↔ delta transformations.

Example: Consider the IEEE 5 Bus power network of Figure 2 (a). It has generators at buses I and II. Loads are present at buses II, III, IV, V. Seven transmission lines numbered 3 to 9 connect the various buses as shown in the figure. The generators and the transmission lines are not perfect i.e., they are subject to failure. We wish to know how these imperfections affect the availability of power at the loads. To do this we must first know the imperfections of these elements in terms of probability of success / failure. Billington and Allan [1] suggested

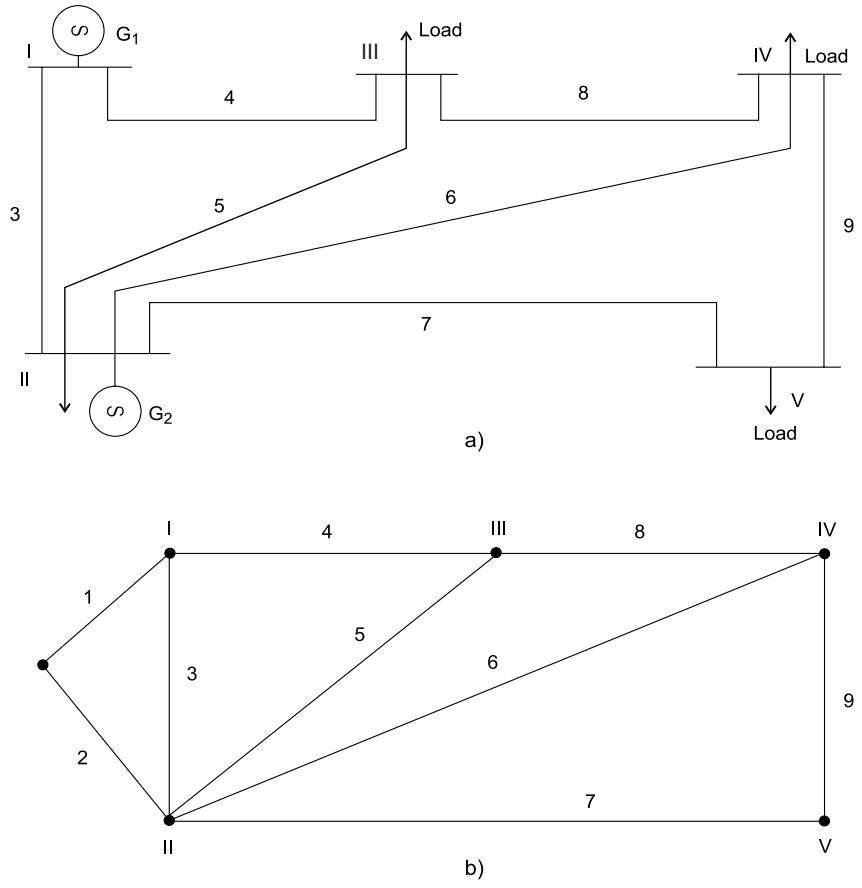


Fig. 2. (a) IEEE 5 Bus network, (b) its Graph.

concepts like failure rate, repair rate, repair time etc to calculate the probability of a component being in up (success) or down (failure) state. Using such ideas let the probabilities of the various elements of the 5 bus system be $p_1 = p_2 = 3/4$, $p_3 = 7/15$, $p_4 = 3/4$, $p_5 = p_6 = p_7 = p_8 = p_9 = 7/8$ where p_i (q_i) denotes probability of success (failure) of i^{th} element. The graph of the system is shown in Figure 2 (b).

Edges 3 to 9 correspond to the transmission lines. Edges 1 and 2 correspond to generators G_1 and G_2. The probabilities of these edges are the probabilities of the corresponding elements of the system. Node s is the source vertex which combines both generators into a single source. Therefore there will be power to a load point if there is at least one functional path from s to it. All vertices of this graph are perfect and all edges are imperfect. We wish to determine the probability of availability of power to load point 3. The graph is not series parallel. Therefore series parallel techniques cannot be applied. However it can be converted into a series parallel network by transforming the star (1, 3, 4) into a delta. Prasad's [11] star – delta transformation cannot be used because there are no restrictions in the direction of flows in the elements of star. q_1, q_2, q_3 satisfy equation (7). So this star can be transformed into a delta using the results of this paper. Let 1', 3', 4' be the corresponding edges of the delta. Edges 1', 3', 4' are between nodes (II, III), (III, s) and (II, s) respectively. Let edge 10 (11) denote the parallel combination of edges 2 and 4' (5 and 1'). The resulting graph is shown in Figure 3. Let edge 14 denote the result of the series parallel combinations of edges 7, 9, 6, 8 and 11. 10 and 14 are in series and the combination is in parallel with 3'. Combining them we get the probability of failure of load point 3. Calculations of these probabilities are shown below. From the theorems

$$p_{1'} = q_1 = \frac{1}{4},\ p_{3'} = q_3 = \frac{8}{15},\ p_{4'} = q_4 = \frac{1}{4},\ q_{10} = q_2\, q_{4'} = \frac{3}{16},$$
$$q_{11} = q_{1'}\, q_5 = \frac{3}{32},\ p_{14} = \frac{128021}{131072} \cong 0.977.$$

Failure probability of load point III = $q_{III} = q_3$, $(1 - p_{10}\, p_{14}) = 0.096$. Direct calculations also give the same result.

This clearly shows that a star can be transformed into a delta in real life networks. Therefore Wang and Sun's conditions are useful in practice. They are not just theoretical.

Consider the problem of determining the probability of supplying power to load point IV. Referring to Figure 3, this requires transforming the star (3', 8, 11) into delta. This star does not satisfy equation (7). So it cannot be transformed into a delta and the network cannot be reduced to a single element using series parallel techniques. But comparing Figure 2 (b) and Figure 3 we notice that Figure 3 has fewer paths and fewer elementary events to load point IV. Therefore calculation of probability for load point IV using conventional methods is more easily done with Figure 3 than Figure 2 (b). Thus star delta transformation is useful even when the network is not reducible to a single element.

Next consider the problem of transforming a star with

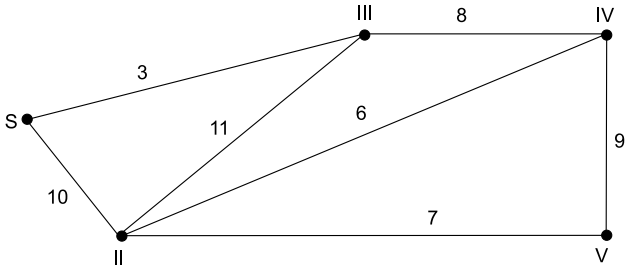


Fig. 3. Simplified graph of Figure 2 (b).

imperfect centre vertex 'd' into a delta. Wang and Sun showed that this is possible if

$$M^2 = (M + Y_3)(M + Y_5)(M + Y_6) \qquad (15)$$

where $Y_i = p_d S_i$, i = 0, 1, 2 ,…..,7, p_d is the probability of success of the centre vertex d and $M = (Y_0 + Y_1 + Y_2 + Y_4) + q_d$. Following the approach presented here, we can simplify equation (15) to

$$M(Y_0 + q_d) = Y_7 Y_0 \qquad (16)$$

When $p_d = 1$ this reduces to $K = S_7$. Equation (16) can be written as

$$S_0 p_d^2 (K - S_7) + S_0 p_d q_d + q_d^2 = 0$$

If $K = S_7$, $S_0 = -\, q_d/p_d$ which is not possible. This proves that a star with an imperfect centre vertex has no equivalent delta if equation (7) is satisfied.

III. CONCLUSIONS

Star delta and delta star transformations are shown to be feasible in reliability networks. Well known conditions of the literature are applicable. The probabilities of the edges of the equivalent star/delta are trivially calculated. It is in fact much simpler than calculating the equivalent resistances of the corresponding star delta problem of electrical circuits. Star ↔ delta transformations are useful whether series parallel techniques work on the entire network or not.

REFERENCES

[1] R. Billington and R.N. Allan, *Reliability evaluation of Power systems*, 2nd ed. New York: Plenum Press, 1996.

[2] M. Ramamoorthy and Balgopal, "Block diagram approach to power system reliability," *IEEE Trans. Power App. Syst*., vol. 89, no. 5/6, pp. 802–811, May/Jun. 1970.

[3] S.K. Banerjee and K. Rajamani, "Closed form solutions for delta-star and star-delta conversions of reliability networks," *IEEE Trans. Reliab*., vol. R-25, no. 2, pp. 118–119, Jun. 1976.

[4] C. Singh and M.D. Kankam,"Comment on: Closed form solutions for delta-star and star-delta conversions of reliability networks," *IEEE Trans. Reliab*., vol. R-25, no. 2, pp. 336–339, Jun. 1976.

[5] A. Rosenthal and D. Frisque, "Transformations for simplifying network reliability calculations," *Networks*, vol. 7, no. 2, pp. 97–111, Jun. 1977.

[6] A. Rosenthal, "Note on: Closed form solutions for delta-star and star-delta conversions of reliability networks," *IEEE Trans. Reliab*., vol. R-27, no. 2, pp. 110–111, Jun. 1978.

[7] H. Gupta and J. Sharma, "A delta-star transformation approach for reliability evaluation," *IEEE Trans. Reliab*., vol. R-27, no. 3, pp. 212–214, Aug. 1978.
[8] D.L. Grosh, "Comments on Delta-Star problem," *IEEE Trans. Reliab*., vol. R-32, no. 4, pp. 391–394, Oct. 1983.
[9] L. Traldi, "On the star delta transformation in network reliability," *Networks*, vol. 23, no. 3, pp. 151–157, May 1993.
[10] S.D. Wang and C.H. Sun, "Transformations of star-delta and delta-star reliability networks," *IEEE Trans. Reliab*., vol. 45, no. 1, pp. 120–126, Mar. 1996.
[11] V.C. Prasad, "Transformation of a star to delta for network reliability calculations," *Asian J. of Mathematics and Computer Research*, vol. 9, no. 1, pp. 46–56, 2016.
[12] G. Levitin, A. Lisnianski, H. Ben-Haim and D. Elmakis, "Redundancy optimization for series-parallel multi-state systems," *IEEE Trans. Reliab*., vol. 47, no. 2, pp. 165–172, Jun. 1998.
[13] R. Bris, E. Chatelet and F. Yalaoui, "New method to minimize the preventive maintenance cost of series parallel systems," *Reliability Engineering and System Safety*, vol. 82, no. 3, pp. 247–255, Dec. 2003.
[14] J.S. Provan, "The complexity of reliability computations in planar and acyclic graphs," *SIAM Journal on Computing*, vol. 15, no. 3, pp. 694–702, 1986.



Vishnuvajjula Charan Prasad
Bachelor's degree from Indian Institute of Science, Bangalore, Master's and Ph.D from Indian Institute of Technology, Delhi. Asst. Research Engineer during 1978–1979 at University of California, Berkeley; 1974–1978: lecturer Delhi college of Engineering. During 1978–1983: Reader Dayalbagh Educational Institute, Dayalbagh, Agra. Since 1983 he was with Indian Institute of Technology Delhi in various capacities till he voluntarily retired in 2006 as a Professor of Electrical Engineering. He worked as a Professor in a Malaysian University till 2012. Since then he is with Dayalbagh Educational Institute as an honorary Professor. His research interests are in nonlinear networks, analog fault diagnosis, network reliability, graph algorithms and binary decision diagrams and their applications.

Generating tunable THz radiation using rippled density plasma driven by density modulated relativistic electron beam

Pratibha Malik, Suresh C. Sharma, and Rinku Sharma

Citation: Physics of Plasmas **24**, 073101 (2017); doi: 10.1063/1.4990075
View online: http://dx.doi.org/10.1063/1.4990075
View Table of Contents: http://aip.scitation.org/toc/php/24/7
Published by the American Institute of Physics



Generating tunable THz radiation using rippled density plasma driven by density modulated relativistic electron beam

Pratibha Malik, Suresh C. Sharma, and Rinku Sharma
Department of Applied Physics, Delhi Technological University (DTU), Shahbad Daulatpur, Bawana Road, Delhi 110042, India

(Received 13 January 2017; accepted 9 June 2017; published online 28 June 2017)

The generation of Terahertz (THz) radiation by a density modulated relativistic electron beam (REB) using rippled density plasma, oriented at a suitable angle along the direction of radiation wave, is being investigated in this paper. The non-linear interaction of density modulated REB with ripple density plasma modifies the dispersion relation of the radiation wave co-propagating with the beam wave. Using fluid equations model, it is found that the requisite ripple wavelength decreases as the ripple angle increases and becomes steeper for resonant THz radiation emission. Thus, the radiation wavelength in terahertz range can be tuned by varying the ripple wavelength and beam energy. In addition, it is investigated that the growth rate of THz radiation emission scales as the one-third power of beam current, two-third power of ripple plasma density, and one-third power of modulation index. The output power and efficiency of THz radiation emission depend on the modulation index and reach the largest value when modulated beam velocity is comparable with the phase velocity of the wave as the modulation index approaches unity. *Published by AIP Publishing*. [http://dx.doi.org/10.1063/1.4990075]

I. INTRODUCTION

The development of terahertz (THz) radiation sources has been of growing interest due to its potential applications in medical imaging, detection of explosives and hazardous chemicals, and identification of different semiconducting structures.[1–3] Recent innovations in THz spectroscopy, remote sensing, and particularly in defence security have shown the remarkable new ideas in the field of science and technology. The generation of THz pulses has been demonstrated experimentally by the non-linear interaction of super luminous laser pulse with large band gap semiconductors and electro-optic crystals such as ZnSe, ZnTe, GaSe, $LiNbO_3$ semiconductors, and photoconductive antennas.[4–6] While several conventional techniques have also employed plasma as a nonlinear medium to generate THz radiation pulses, another scheme has been proposed to generate the THz radiation by synchrotron/transition radiation from bunched electron beams[7] and modulated relativistic electron beam,[8] which provides a better alternative to generate high THz energy in comparison with the laser induced THz radiation energy. However, the impression of plasma wiggler/undulator with its application to a FEL scheme to produce coherent radiation of wavelengths ranging from tens to hundreds nm has been first proposed and shown as rational.[9,10] In addition, a few other techniques have been investigated for the generation of THz radiation such as beating of two lasers in a periodic density plasma,[11] modulated electron beam with semiconducting planar structure,[12] and interaction of un-modulated electron beam in a ripple density plasma.[13]

Weiss *et al.*[14] have demonstrated the continuous increment in THz power from semiconductors like GaAs and InAs with the increase in the magnetic field. Liu and Tripathi[15] have proposed the application of a static magnetic field to provide the tunability of the frequency and a density ripple to control the angular orientation of the emitted radiation. The THz pulse generation based on tunnel ionization of a gas jet,[16] laser modulated electron beam,[17] laser accelerated electron bunches,[18] slow wave structure,[19,20] and cosh-Gaussian lasers rippled density plasma[21] has been reported earlier in different works. Nevertheless, the propagation of coherent THz pulses of electron bunch radiation sources has been focused to generate the ability of tens of mJ/pulse,[22–24] which shows the various orders of magnitude beyond that of the conventional THz sources.

However, Hamster *et al.*[25] have reported the generation of THz pulses from the interaction of intense laser pulse with plasma. In this case, a laser pulse with the power of 10^{12} W has been focused on both the gas and solid target for a short duration of 0.1 ps. Antonsen, Jr., *et al.*[26] have employed a scheme to generate a phase matched THz radiation pulses in a corrugated plasma channel, by the generation of ponderomotive force from laser. In recent years, the generation of coherent radiation has become a surgical field of pursuit by using short bunches of electrons and can be utilized as a high resolution monitor for detecting the bunching form.[27] More recently, Zhang *et al.*[28] have studied the generation of terahertz radiation by beating of the two Langmuir waves in a warm and collisional plasma as a function of critical angle.

In this paper, we present the generation of terahertz radiation by a density modulated relativistic electron beam (REB) in a rippled density plasma. The schematic representation of the model is shown in Fig. 1. The pre-modulated relativistic electron beam plays a significant role in the enhancement of growth rate, efficiency, and output power of the THz radiation emission. The physical mechanism of the process can be understood as follows: whenever a density modulated relativistic electron beam is injected into a ripple density plasma, the beam electrons are wiggled and emit

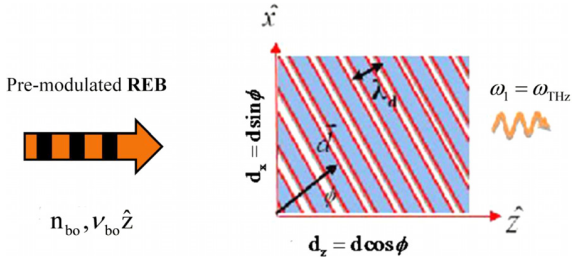


FIG. 1. Schematic diagram of the generation of THz radiation by a pre-modulated relativistic electron beam (REB) in a rippled density plasma.

light (THz) of ripple wavelength. The emitted photons (THz radiation) move slightly faster than the beam electrons and interact with them in each wiggler period and produce the beam mode (sideband). The THz radiation wave imparts its oscillatory velocity to plasma electrons and beam electrons, and the THz mode $(\omega_1, k\hat{z})$ couples with beam mode $(\omega_1, k\hat{z} + \vec{d})$. While the oscillatory velocity of plasma electrons due to beam mode interacts with density ripple and produces a nonlinear current density at $(\omega_1, \vec{k}_s)$. This is further responsible to drive the THz radiation signal. This derived THz signal distributes its energy to the plasma electrons and beam electrons and interacts with ripple density to generate the current which amplifies the beam mode. This feedback process is responsible for the amplification in the beam mode and electromagnetic wave, which continues with time at the cost of beam energy (kinetic energy) till the system goes into saturation.

The organization of the paper is as follows: In Sec. II, we discuss the basic mathematical model equations for the response of beam and plasma electrons, respectively. In Sec. III, we examine the instability of the THz mode, which leads to the amplification in THz unstable mode leading to efficiency enhancements. Later, in Sec. IV, the results are explained with physical phenomenon of the process and concluded in Sec. V.

II. MODEL EQUATIONS FOR THz RADIATION AND BEAM MODE WAVE

The generation and amplification of the THz radiation pulses by a density modulated electron beam in under-dense rippled plasma, free from the collisions, are studied [cf. Fig. 1]. The condition for under-dense plasma is expressed as $n_{po} < n_c$, where n_{po} is the density of electrons at equilibrium and n_c is the critical density of plasma. Under this condition, the wave vector of the radiation wave will be no longer imaginary as the radiation frequency scales as exp(-ikz).

We consider a density modulated relativistic electron beam of velocity $v_{bo}\hat{z}$, and it is being launched into a rippled density plasma of wave vector $\vec{d}$. The interaction of density modulated electron beam with ripple density plasma acts as a wiggler, which provides an opportunity for generating the THz radiation of frequency ω_1 with a wave vector $k\hat{z}$ and accompained with sideband wave of frequency ω_1 and wave vector $\vec{k}_s (= k\hat{z} + \vec{d})$.

The density modulated electron beam $n_b (= n_{bo} + n_{bm} \exp\{-i\omega_o(t - z/v_{bo})\})$ is propagating in a static rippled density plasma of wave vector $\vec{d} (= d_x\hat{x} + d_z\hat{z})$, making a small angle ϕ along with the direction of the wave. The plasma density is expressed as $n_p = n_{po} + n_1 (= n_d \exp[i(\vec{d}.\vec{r})])$, where n_{po} is the density at equilibrium and n_d is the amplitude of ripples. The relativistic electron beam parameters are expressed as $n_{bm} = \Delta n_{bo}$ and $\omega_0 = k_{oz} v_{bo}$, where Δ is the modulation index (its value ranges from 0 to 1), ω_0 and $k_{0z} (\sim \frac{\omega_0}{v_{bo}})$ are the frequency and wave vector of the modulated beam, respectively.

The electric and magnetic fields associated with the THz radiation wave are expressed as $\vec{E}_T = A_1 \exp[-i(\omega_1 t - kz)]$ and $\vec{B}_T = \frac{c(\nabla \times \vec{E}_T)}{i\omega_1}$. However, the radiation wave is accompanied with a sideband wave subjected to an electric field $\vec{E}_S = A_2 \exp[-i(\omega_1 t - \vec{k}_s.\vec{r})]$ with wave vector $\vec{k}_s = k\hat{z} + \vec{d}$ due to the rippled density plasma.

The radiation frequency of the THz wave can be written as $\omega_1 = 2\gamma_0^2 d_z v_{bo}$, where v_{bo} represents the beam velocity and γ_0 is the relativistic gamma factor. The response of beam electrons in the presence of radiation wave and sideband wave is governed by the equation of motion and continuity equation

$$\frac{\partial}{\partial t}(\gamma_b \vec{v}_b) + (\vec{v}_b.\nabla)(\gamma_b \vec{v}_b) = -\frac{e}{m}\left(\vec{E}_T + \vec{E}_S + \frac{1}{c}\vec{v}_b \times \vec{B}_T\right), \tag{1}$$

$$\frac{\partial n_b}{\partial t} + \nabla.(n_b \vec{v}_b) = 0. \tag{2}$$

A. Excitation and dispersion relation of THz radiation wave

The electric fields $\vec{E}_T$ and $\vec{E}_S$ are associated with the THz radiation wave and sideband wave, respectively. The relativistic gamma factor is given as $\gamma_0 = (1 - \frac{v_{bo}^2}{c^2})^{-1/2}$ and $\gamma_b = \gamma_o + \frac{v_{bo}}{c^2}[\gamma_o^3 v_{bz}^T + \gamma_o^3 v_{bz}^S]$, where -e and m_e are the electronic charge and mass, respectively. Therefore, considering only first order perturbation terms for beam velocity and density as

$$\vec{v}_b = v_{bo}\hat{z} + \vec{v}_{b1}^T \exp[-i(\omega_1 t - kz)] + \vec{v}_{b1}^S \exp\left[-i(\omega_1 t - \vec{k}_s.\vec{r})\right], \tag{3}$$

$$n_b = n_{bo} + n_{bm} \exp[-i\omega_o(t - z/v_{bo})] + n_{b1}^T \exp[-i(\omega_1 t - kz)] + n_{b1}^S \exp\left[-i(\omega_1 t - \vec{k}_s.\vec{r})\right]. \tag{4}$$

On linearizing Eqs. (1) and (2), we obtain the perturbed velocity and density of the THz radiation wave as

$$v_{b1}^T = \frac{eE_{Tx}}{im_e\omega_1\gamma_0}\hat{x} + \frac{eE_{Tz}}{im_e(\omega_1 - kv_{bo})\gamma_0^3}\hat{z}, \tag{5}$$

$$n_{b1}^T = \frac{n_{bo}}{(\omega_1 - kv_{bo})}\left[(1 + \Delta)\left(\vec{k}.\vec{v}_b^T\right) + \Delta k_{oz} v_{bz}^T\right]. \tag{6}$$

Similarly, the perturbed velocity and density for the sideband wave are expressed as

$$v_{b1}^{s} = \left[\frac{eE_{sx}}{im_e\omega_1\gamma_0} + \frac{ek_{sx}v_{bo}E_{sz}}{im_e\omega_1(\omega_1 - k_{sz}v_{bo})\gamma_0}\right]\hat{x} + \frac{eE_{sz}}{im_e(\omega_1 - k_{sz}v_{bo})\gamma_0^3}\hat{z}, \quad (7)$$

$$n_{b1}^{S} = \frac{n_{bo}}{(\omega_1 - k_{sz}v_{bo})}\left[(1+\Delta)\left(\vec{k}_s.\vec{v}_b^{s}\right) + \Delta k_{oz}v_{bz}^{s}\right]. \quad (8)$$

The response of ripple density plasma electrons to the perturbation of the THz radiation wave and sideband wave is also governed by the equation of motion and continuity, which yields the perturbed velocity and density as

$$\vec{v}_p = \vec{v}_p^{T} + \vec{v}_p^{S}, \quad \gamma_o = 1, \quad \vec{v}_b = 0, \quad (9)$$

$$n_p = n_{po} + n_d \exp\left[i(\vec{d}.\vec{r})\right] + n_p^{T} + n_p^{S}, \quad (10)$$

where n_p^T and n_p^S are the perturbed densities for THz radiation and sideband waves, respectively, and d represents the periodicity $(2\pi/\lambda_d)$ of ripple density plasma.

The perturbed velocities for THz wave and sideband wave are

$$\vec{v}_p^{T} = \frac{e\vec{E}_T}{im_e\omega_1}, \quad \vec{v}_p^{S} = \frac{e\vec{E}_s}{im_e\omega_1}. \quad (11)$$

However, the equation of continuity for the plasma electrons yields

$$n_p^{T} = \frac{\vec{k}}{\omega_1}.\left(n_{po}\vec{v}_p^{T}\right) + \frac{1}{2}\frac{\vec{k}}{\omega_1}.\left(n_d^{*}\vec{v}_p^{S}\right), \quad (12)$$

$$n_p^{S} = \frac{\vec{k}_s}{\omega_1}.\left(n_{po}\vec{v}_p^{S}\right) + \frac{1}{2}\frac{\vec{k}_s}{\omega_1}.\left(n_d\vec{v}_p^{T}\right). \quad (13)$$

Therefore, the non-linear current density for driving the THz radiation wave and sideband wave can be written as

$$\vec{J}_{Nl} = -n_{tot}e\vec{v}_{tot}, \quad (14)$$

where $n_{tot} = n_p + n_b$ and $\vec{v}_{tot} = \vec{v}_p + \vec{v}_b$. The nonlinear current density for THz radiation wave and sideband wave can be written as

$$\vec{J}_{Nl(\omega_1,k)} = -e\left[n_p\vec{v}_p^{T} + \frac{1}{2}n_d^{*}\vec{v}_p^{S} + (1+\Delta)n_{bo}\vec{v}_b^{T} + n_b^{T}\vec{v}_b^{o}\right], \quad (15)$$

$$\vec{J}_{Nl(\omega_1,k_s)} = -e\left[n_p\vec{v}_p^{S} + \frac{1}{2}n_d\vec{v}_p^{T} + (1+\Delta)n_{bo}\vec{v}_b^{S} + n_b^{S}\vec{v}_b^{o}\right]. \quad (16)$$

The general wave equation can be modified as

$$\nabla\left(\nabla.\vec{E}_I\right) - (\nabla.\nabla)\vec{E}_I = -\frac{4\pi}{c^2}\frac{\partial\vec{J}_I}{\partial t} - \frac{1}{c^2}\frac{\partial^2\vec{E}_I}{\partial t^2}. \quad (17)$$

Here, we get THz radiation wave and the side band wave, correspondingly, after putting $I = T$ and $I = S$, and replacing the terms $\nabla = i\vec{k}$ and $\nabla = i\vec{k}_s$ with $\frac{\partial}{\partial t} = -i\omega_1$ in Eq. (17), as

$$\underline{\underline{A}}.\vec{E}_T = -\frac{n_d\omega_{pd}^2}{2n_{po}c^2}\vec{E}_s, \quad (18)$$

$$\underline{\underline{B}}.\vec{E}_s = -\frac{n_d\omega_{pd}^2}{2n_{po}c^2}\vec{E}_T, \quad (19)$$

where $\underline{\underline{A}} = \left[\left(k^2 - \frac{\omega_1^2-\omega_{pe}^2}{c^2}\right)\underline{\underline{I}} - \vec{k}.\vec{k} + \frac{\underline{\underline{\rho}}^T}{c^2}\right]$, $\rho_{xx}^T = \omega_{pd}^2 + \frac{(1+\Delta)}{\gamma_0}\omega_{pb}^2$, $\rho_{zz}^T = \omega_{pd}^2 + \frac{\omega_{pb}^2\omega_1}{\gamma_o^3(\omega_1 - kv_{bo})^2}[(1+\Delta)\omega_1 + \Delta k_{oz}v_{bo}]$, and I is a Unitary matrix of order 3, $\omega_{pd}\left(=\sqrt{\frac{4\pi n_d e^2}{m_e}}\right)$, $\omega_{pb}\left(=\sqrt{\frac{4\pi n_{bo}e^2}{m_e}}\right)$, and $\omega_{pe}\left(=\sqrt{\frac{4\pi n_{po}e^2}{m_e}}\right)$ are the ripple density plasma frequency, beam plasma frequency, and electron plasma frequency, respectively.

The parameters take the form as $\underline{\underline{B}} = \left[\left(k_s^2 - \frac{\omega_1^2-\omega_{pe}^2}{c^2}\right)\underline{\underline{I}} - \vec{k}_s.\vec{k}_s + \frac{\underline{\underline{\rho}}^S}{c^2}\right]$, $\rho_{xx}^S = \omega_{pd}^2 + \frac{(1+\Delta)\omega_{pb}^2}{\gamma_o}$,

$$\rho_{zz}^S = \omega_{pd}^2 + \frac{(1+\Delta)\omega_1\omega_{pb}^2}{\gamma_o^3(\omega_1 - k_{sz}v_{bo})} + \frac{\omega_1\omega_{pb}^2v_{bo}}{\gamma_o(\omega_1 - k_{sz}v_{bo})^2} \times \left[(1+\Delta)\frac{k_{sz}}{\gamma_o^2} + (1+\Delta)\frac{k_{sx}^2v_{bo}}{\omega_1} + \Delta\frac{k_{oz}}{\gamma_o^2}\right],$$

and the other parameters are same in Eq. (19) as earlier. Therefore, re-arranging the terms in Eqs. (18) and (19), we obtain

$$\underline{\underline{A}}.\vec{E}_T + m\vec{E}_s = 0, \quad (20)$$

$$m\vec{E}_T + \underline{\underline{B}}.\vec{E}_s = 0, \quad (21)$$

where $m = \frac{n_d\omega_{pd}^2}{n_{po}2c^2}$, and the solutions obtained by Eqs. (20) and (21) under the condition $n_{po} \gg n_d$ such that the ratio of $\frac{n_d}{n_{po}}$ becomes too small to be considered, we obtain the tensor form of Eqs. (20) and (21) as

$$\begin{bmatrix} \underline{\underline{A}} & \frac{\omega_{pd}^2}{2c^2} \\ \frac{\omega_{pd}^2}{2c^2} & \underline{\underline{B}} \end{bmatrix} = 0, \quad (22)$$

$$\left(\underline{\underline{B}} - \frac{\omega_{pd}^4}{4c^4}\underline{\underline{A}}^{-1}\right) = 0. \quad (23)$$

Equation (23) is the dispersion relation of the THz electromagnetic wave. The solution of which gives the growth rate of the radiation wave.

III. INSTABILITY ANALYSIS

In order to obtain the growth rate of unstable mode of THz radiation wave, first, we calculate $\underline{\underline{A}}^{-1}(= Adj\underline{\underline{A}}/|\underline{\underline{A}}|)$, and considering only those terms, which contains $(\omega_1 - k_{sz}v_{bo})^2$ in the denominator of Eq. (23), we obtain

$$\underline{\underline{A}}^{-1} = \begin{bmatrix} \rho_{11} & 0 & \rho_{13} \\ 0 & \rho_{23} & 0 \\ \rho_{31} & 0 & \rho_{33} \end{bmatrix},$$

where the parameters are defined as

$$\rho_{11} = k^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2} - k_z k_z + \frac{\rho_{zz}^T}{c^2}, \quad \rho_{23} = k^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2} - k_z k_z,$$

$$\rho_{13} = \frac{k_x k_z}{\left[\left(k^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2} - k_x k_x + \frac{\rho_{xx}^T}{c^2}\right)\left(k^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2} - k_z k_z + \frac{\rho_{zz}^T}{c^2}\right) - k_x^2 k_z^2\right]},$$

$$\rho_{31} = \frac{k_x k_z}{\left[\left(k^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2} - k_x k_x + \frac{\rho_{xx}^T}{c^2}\right)\left(k^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2} - k_z k_z + \frac{\rho_{zz}^T}{c^2}\right) - k_x^2 k_z^2\right]}, \quad \text{and}$$

$$\rho_{33} = \frac{\left(k^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2} - k_x k_x + \frac{\rho_{xx}^T}{c^2}\right)}{\left[\left(k^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2} - k_x k_x + \frac{\rho_{xx}^T}{c^2}\right)\left(k^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2} - k_z k_z + \frac{\rho_{zz}^T}{c^2}\right) - k_x^2 k_z^2\right]}.$$

On solving Eq. (23), we obtain

$$(\omega_1 - k_{sz} v_{bo})^2 \left(k^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2} - \frac{\omega_{pd}^4}{4c^4} C\right) - \frac{D}{\left(\omega_1^2 - \omega_{pe}^2\right)\left(k_s^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2}\right)} = 0, \tag{24}$$

where

$$C = \frac{\frac{\omega_1^2 - \omega_{pe}^2}{c^2} + k^2\left(\frac{k_{sz}^2 c^2}{\omega_1^2 - \omega_{pe}^2} - 1\right) - \left(k_s^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2}\right)}{\left(\frac{\omega_1^2 - \omega_{pe}^2}{c^2}\right)\left(k^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2}\right)} \tag{25}$$

and

$$D = \frac{\omega_1 \omega_{pb}^2 k_{sz} v_{bo}}{\gamma_o^3}\left[(1+\Delta)\left(1 + \frac{\gamma_o^2 v_{bo} k_{sx}^2}{\omega_1 k_{sz}}\right) + \Delta \frac{k_{oz}}{k_{sz}}\right] \times \left[\left(k_{sz}^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2}\right)\left(k^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2}\right) - \frac{n_d^2 \omega_{pe}^4}{4c^4 n_{po}^2}\right]. \tag{26}$$

Equation (24) gives the modified dispersion relation of the THz radiation wave. The solution of Eq. (24) can be obtained under two conditions, one when the constants C and D are neglected and the other when constants C and D are considered. The consideration and non-consideration of constants C and D provide two distinct solutions of Eq. (24). When C and D are neglected in Eq. (24), we obtain the uncoupled modes of THz radiation and sideband waves, which physically represents a condition of non-growth of the radiation wave in the system. When C and D are considered for the solution of Eq. (24), we obtain the growth rate of the instability of THz radiation wave and sideband wave, which represents that the pre-modulated electron beam imparting its energy to the system ultimately results in a significant enhancement of THz radiation and sideband waves.

Now, simplifying Eq. (24) when the constants C and D are neglected, we obtain a dispersion relation for the uncoupled modes as

$$(\omega_1 - k_{sz} v_{bo})^2 \left(k^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2}\right) = 0. \tag{27}$$

The first solution of Eq. (27), at $(\omega_1 - k_{sz} v_{bo})^2 = 0$, i.e., $\omega_1 = \omega_{1r} = k_{sz} v_{bo}$, gives beam mode and second solution at $\left(k^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2}\right) = 0$ is obtained, as $\omega_1 = \omega_{1r} = (k^2 c^2 + \omega_{pe}^2)^{1/2}$ gives radiation mode. By solving Eq. (27), simultaneously, we obtain the relation for the ripple density wavelength, which varies with the radiation frequency as

$$\lambda_d = \frac{4\pi c\left(\gamma_o^2 - 1\right)\omega_1}{\omega_1^2 + \gamma_o^2 \omega_{pe}^2} \cos\phi, \tag{28}$$

where ω_1 is the angular frequency of the radiation wave and ϕ is the ripple angle. Equation (28) gives a qualitative idea of the radiation between angular frequency and ripple angle ϕ.

However, the inclusion of C and D, the solution of Eq. (24) is given as

$$(\omega_1 - k_{sz} v_{bo})^2 \left[k^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2} + \frac{k_{sx}^2 - \frac{\left(\omega_1^2 - \omega_{pe}^2\right)}{c^2}}{k_s^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2}} \times \frac{c^2}{\left(\omega_1^2 - \omega_{pe}^2\right)} \frac{n_d^2 \omega_{pe}^4}{4c^4 n_{po}^2}\right] = \frac{n_d^2 \omega_{pe}^4 \omega_{pb}^2 \omega_1 k_{sz} v_{bo}}{4c^4 n_{po}^2 \gamma_o^3 \left(\omega_1^2 - \omega_{pe}^2\right)\left(k_s^2 - \frac{\omega_1^2 - \omega_{pe}^2}{c^2}\right)} \times \left[(1+\Delta)\left(1 + \frac{\gamma_0^2 v_{bo} k_{sx}^2}{\omega_1 k_{sz}}\right) + \Delta \frac{k_{oz}}{k_{sz}}\right]. \tag{29}$$

Considering the terms $\omega_1 = k_{sz}v_{bo} + \delta = \omega_1 = \omega_r + \delta$, where δ is the small frequency mismatch. Therefore, we acquire the relations as $\omega_1 = (\omega_{pe}^2 + k^2c^2)^{1/2} + \delta$ or $\omega_1^2 - (\omega_{pe}^2 + k^2c^2) = 2\delta\sqrt{\omega_{pe}^2 + k^2c^2}$, and following the sequence, we obtain the value of δ from Eq. (29) as

$$\delta^3 = \frac{\omega_{pb}^2\omega_{pd}^4k_{sx}^2}{8c^2\gamma_o^3(\omega_{pe}^2 + k^2c^2)^{1/2}(k_s^2 - k^2)^2} \times \left[(1+\Delta)\left(1 + \gamma_o^2\frac{k_{sx}^2}{k_{sz}^2}\right) + \Delta\frac{k_{oz}}{k_{sz}}\right]. \quad (30)$$

Therefore, the growth rate (i.e., the imaginary part of)

$$\gamma = Im\left[\frac{\omega_{pb}^2\omega_{pd}^4k_{sx}^2}{8c^2\gamma_o^3(\omega_{pe}^2 + k^2c^2)^{1/2}(k_s^2 - k^2)^2} \times \left((1+\Delta)\left(1 + \gamma_o^2\frac{k_{sx}^2}{k_{sz}^2}\right) + \Delta\frac{k_{oz}}{k_{sz}}\right)\right]^{\frac{1}{3}} e^{\frac{i2n\pi}{3}}, \quad n = 0, 1, 2,$$

$$\gamma = \frac{\sqrt{3}}{4}\left[\frac{\omega_{pb}^2\omega_{pd}^4d_x^2}{\gamma_o^3c^2(\omega_{pe}^2 + k^2c^2)^{1/2}(k_s^2 - k^2)^2} \times \left((1+\Delta)\left(1 + \gamma_o^2\frac{d_x^2}{k_{sz}^2}\right) + \Delta\frac{k_{oz}}{k_{sz}}\right)\right]^{\frac{1}{3}}. \quad (31)$$

In Eq. (31), we can say that the growth rate γ (in s^{-1}) of the tunable unstable THz radiation wave scales as one-third power of the beam current I_b and two-third power of the ripple density n_d. In addition, the growth rate of the unstable wave increases with the density of beam electrons and modulation index Δ.

The power density of the THz radiation emission of the wave can be defined as[29]

$$P = \left(\omega_1\frac{\partial\varepsilon_0}{\partial\omega_1}\right)\frac{A_1^2}{8\pi}v_{te}A_b, \quad (32)$$

where the values of A_1, v_{te}, and A_b are the electric field amplitude, thermal velocity of the plasma electrons, and beam area, respectively. The symbol ε_0 is the dielectric permittivity of the plasma. This can be expressed by the relation $\vec{D} = \varepsilon_0\vec{E}$, when $\vec{P}$ (polarization) $\ll \vec{E}$ and $\varepsilon_0 \simeq 1$ for the plasma.

From Eq. (32), we can see that the power density of the radiation increases with thermal velocity of the plasma electrons. In addition, the power density also increases linearly with the amplitude of radiation wave.

The efficiency of the THz radiation emission can also be written as[29]

$$\eta = \frac{\gamma_o^3\gamma}{(\gamma_o - 1)\omega_1}$$

$$\text{or} \quad \eta = \frac{\sqrt{3}}{4}\frac{\gamma_o^3}{(\gamma_o - 1)\omega_1}\left[\frac{\omega_{pb}^2\omega_{pd}^4d_x^2}{\gamma_o^3c^2(\omega_{pe}^2 + k^2c^2)^{1/2}(k_s^2 - k^2)^2} \times \left((1+\Delta)\left(1 + \gamma_o^2\frac{d_x^2}{k_{sz}^2}\right) + \Delta\frac{k_{oz}}{k_{sz}}\right)\right]^{\frac{1}{3}}. \quad (33)$$

Using Eq. (33), we observe that the efficiency η of the THz radiation emission depends on the growth rate γ (in s^{-1}) of the unstable mode and modulation index Δ.

IV. RESULTS AND DISCUSSION

In this work, we study the generation of THz radiation emission by pre-modulated relativistic electron beam in a rippled density plasma. The calculations have been performed to investigate the effect of plasma density, beam density, modulation index, and radiation wave frequency on the growth rate, power, and efficiency of the THz radiation emission of the wave. The methodology prepared is in agreement with the experimental and analytical results obtained by Refs. 11, 26, and 30–34. The typical parameters used to study numerically the behaviour of the THz radiation wave in a rippled density plasma are electronic mass $m_e = 9.1 \times 10^{-28}$ g, electronic charge $e = 4.8 \times 10^{-10}$ statC, electric field amplitude $A_1 = 95$ Statvolt/cm, electron plasma density n_{po} varies as 10^{13}–10^{16} cm^{-3}, initial beam modulated frequency ω_0 $(= 9.3 \times 10^{12} radsec^{-1})$, beam radius $r_b = 0.7$ cm, and the modulation index Δ values from 0 to 1 in the steps of 0.1 and the ripple density amplitude is being expressed by the relation $n_d = 0.6n_{po}$. Using Eq. (32), we have estimated the THz emission power $P \approx 750$ W for the following parameters: plasma temperature $T = 1$ eV, frequency $f = 3$ THz with other parameters same as above.

Using Eq. (28), we have plotted in Fig. 2 the variation of wavelength λ_d(in μm) of rippled plasma with the normalized radiation frequency ω_1/ω_{pe} for different values of ripple angles $\phi = 10^\circ, 20^\circ, 30^\circ$, and 40°. From Fig. 2, it is observed that for all the values of ripple angles the ripple wavelength increases and reaches maximum and then falls off rapidly as radiation frequency increases. In addition, it is also observed that the ripple wavelength decreases as the ripple angle increases. Thus, one may conclude that the steep ripples would be rather useful for the excitation of the resonant THz radiation. It is also evident from Fig. 2 that the closely spaced density ripples can offer more strong THz radiation emission. This result is in compliance with the experimental observation obtained by Antonsen, Jr., *et al.*,[26] in which they

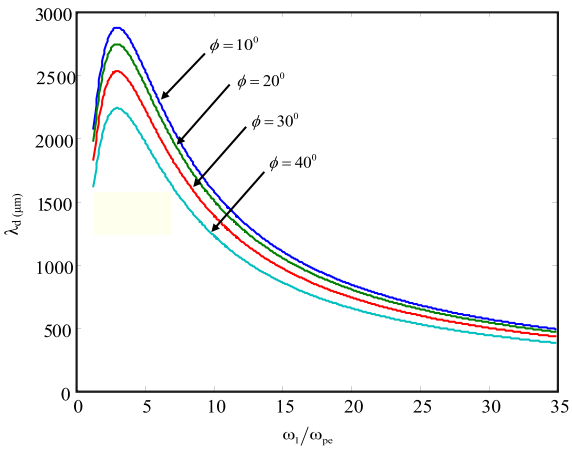


FIG. 2. Ripple plasma wavelength λ_d as a function of the normalized radiation frequency ω_1/ω_{pe} for different ripple orientations, i.e., $\phi = 10^\circ, 20^\circ, 30^\circ, 40^\circ$ and for $E_b = 1.0$ MeV and $n_{po} = 3 \times 10^{13} cm^{-3}$.

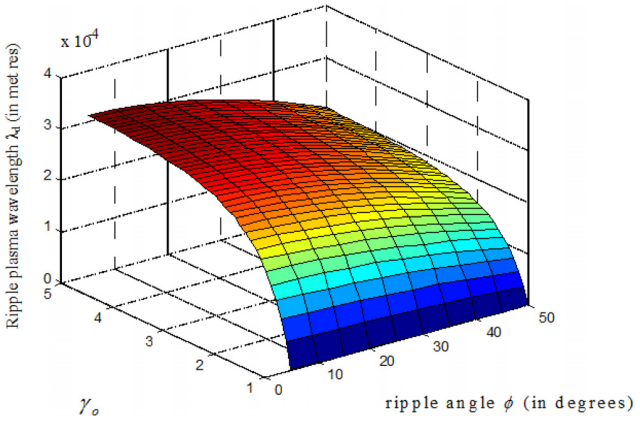


FIG. 3. Variation of ripple plasma wavelength $\lambda_d(\text{in m})$ with ripple angle ϕ (in deg) and beam energy γ_o at $\omega_1 = 2.2 \times 10^{12}$rad sec^{-1} and other variables are same as in Fig. 2.

have suggested that the delta function type corrugated plasma density structure is more encouraging for better phase matching condition and maximum transfer of energy to the THz radiation wave excited by laser pulses.

Besides this, the maxima of the curves at different ripple angles give the cut-off normalized radiation frequency $\omega_1/\omega_{pe} = 3.058$, i.e., plasma density $n_{po} = 3 \times 10^{13}\text{cm}^{-3}$ with beam energy $E_b = 1.0\,\text{MeV}$. The cut-off frequency demarcates the two regions in plasma. However, Fig. 3 depicts the variation of ripple wavelength with ripple angle and beam energy at radiation frequency with other parameters same as in Fig. 2. Applying Eq. (28), it can be seen from Fig. 3 that for a particular value of ripple angle, first, ripple wavelength increases as beam energy increases and then it starts decreasing as ripple angle increases further.

Using Eq. (31), in Fig. 4, we have plotted the growth rate γ (in s^{-1}) of the radiation wave with modulation index Δ for different beam densities, e.g., $n_{bo} = 0.5 \times 10^7$ cm^{-3}, 1.0×10^8 cm^{-3}, 0.5×10^9 cm^{-3}, and $n_{bo} = 1.5 \times 10^9$ cm^{-3} for ripple angle $\phi = 10^\circ$ and beam energy $E_b = 1.0\,\text{MeV}$. From Fig. 4, it can be seen that the growth rate of the THz radiation mode increases with the beam density and reaches the largest value when the beam velocity is comparable with the phase velocity of the radiation wave, i.e., $\omega_0 \sim \omega_1$ and $k_{oz} \sim k$, i.e., the modulation frequency and wave vector are comparable with the radiation frequency and wave vector. The variation of the growth rate with beam densities results in Fig. 4 is consistent with the experimental observations of Namiot and Shchurova.[30] In this case, they have observed that an increment in the beam electrons results in more transfer of energy to the THz electromagnetic wave and this leads to the amplification of the wave in a corrugated surface.

Figure 5 represents the variation of the growth rate γ (in s^{-1}) of unstable radiation wave with modulation index Δ, using Eq. (31), for different plasma densities, e.g., $n_{po} = 1.5 \times 10^{13}$ cm^{-3}, 4.5×10^{13} cm^{-3}, 1.0×10^{14} cm^{-3}, and $n_{po} = 1.0 \times 10^{15}$ cm^{-3}. From Fig. 5, it can be seen that the growth rate of the radiation wave as a function of modulation index increases with plasma density. In addition, the growth rate of the radiation wave also increases with the modulation index. Zhang[31] has proposed a scheme for the amplification of the radiation wave using relativistic electron beam in sinusoidal ripple boundary with planar electrostatic system and found that our theoretical results are in good agreement with the existing results.

In Fig. 6, we have plotted the growth rate γ (in s^{-1}) of the unstable mode with the modulation index Δ for different beam energies E_b (=0.5 MeV, 1.0 MeV, 1.5 MeV, and 2.0 MeV) and for plasma density $n_{po} = 2.0 \times 10^{14}$ cm^{-3}, radiation frequency $f = 4$ THz, and $n_{b0} = 10^{-5} n_{po}$. It can be seen from Fig. 6 that the growth rate of the unstable radiation wave increases with the modulation index and decreases with beam energy. We know that a wave can grow when the beam electrons impart their energy to the radiation wave and this can happen only if the beam velocity exceeds the phase velocity of the radiation wave.

Figure 7 displays the variation of efficiency η of the THz radiation emission with the modulation index Δ for different radiation frequencies, e.g., f = 3 THz, 4 THz, 5 THz, and 6 THz and for beam energy $E_b = 1.0\,\text{MeV}$. The

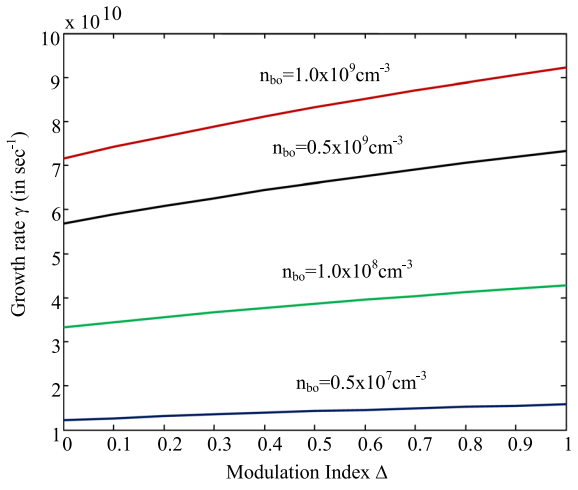


FIG. 4. Variation of the growth rate γ (in s^{-1}) with modulation index Δ for different beam densities, i.e., n_{bo}=0.5×10^7cm^{-3}, 1.0×10^8 cm^{-3}, 0.5×10^9 cm^{-3} and $n_{bo} = 1.5 \times 10^9$ cm^{-3} and for $E_b = 1.0\,\text{MeV}$ and $n_{b0} = 10^{-5} n_{po}$.

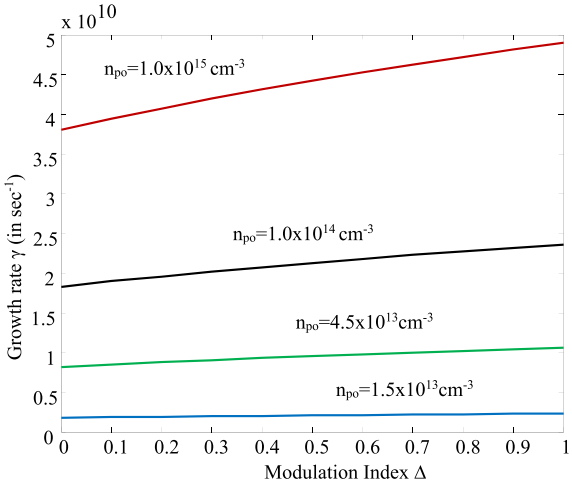


FIG. 5. Graph for the variation of the growth rate γ (in s^{-1}) with modulation index Δ for different plasma densities, i.e., $n_{po} = 1.5 \times 10^{13}$ cm^{-3}, 4.5×10^{13} cm^{-3}, 1.0×10^{14} cm^{-3}, and $n_{po} = 1.0 \times 10^{15}$cm^{-3} and for beam energy E_b = 1.0 MeV and $n_{b0} = 10^{-5} n_{po}$.

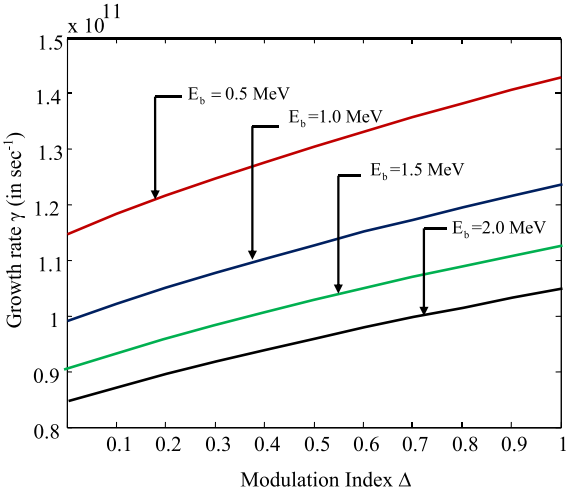


FIG. 6. Variation of the growth rate γ (in s^{-1}) with modulation index Δ for different beam energies: $E_b = 0.5$ MeV, 1.0 MeV, 1.5 MeV, and 2.0 MeV with plasma density $n_{po}=2.0 \times 10^{14}$ cm^{-3} and for f = 4 THz and $n_{b0}= 10^{-5} n_{po}$.

efficiency of THz radiation emission of the wave as a function of modulation index decreases with radiation frequency. This is because for a particular value of beam energy, as the frequency of radiation wave increases, the condition for the growth of the wave is no longer satisfied. In this case, the phase velocity of the wave (energy of radiation wave) exceeds the beam velocity (or beam energy) and the radiation wave attenuates, thereby leading to the decrement in the efficiency of the radiation emission. In Fig. 7, we also observe that the higher efficiency can be achieved as radiation frequency becomes small, which is higher as compared with efficiency obtained by beating of two lasers in periodic density plasma.[11,32–34] It implies that our physical model concept is quite efficient to get the THz radiation emission.

In Fig. 8, we see the variation of efficiency of the radiation wave with ripple angle ϕ, using Eq. (33) for the parameters—$E_b = 1.53$ MeV, $n_{po} = 1.5 \times 10^{14} cm^{-3}$, and $\Delta = 0.7$. In Fig. 8, we observe that the efficiency of the radiation wave increases as ripple angle increases which proves that

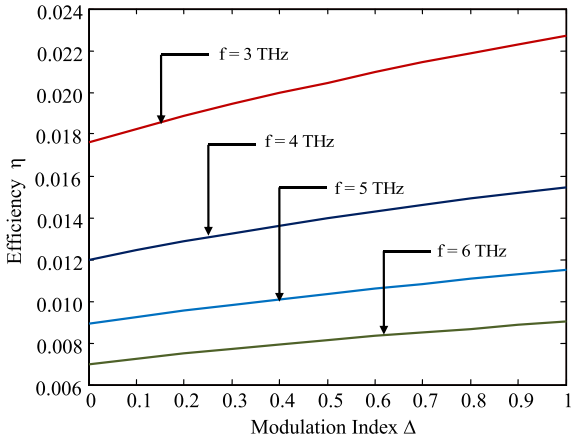


FIG. 7. A plot of efficiency η of the THz radiation emission with modulation index Δ for E_b =0.5 MeV and for radiation frequencies, f = 3, 4, and 5 THz; plasma density $n_{po} = 4.5 \times 10^{13}$ cm^{-3}.

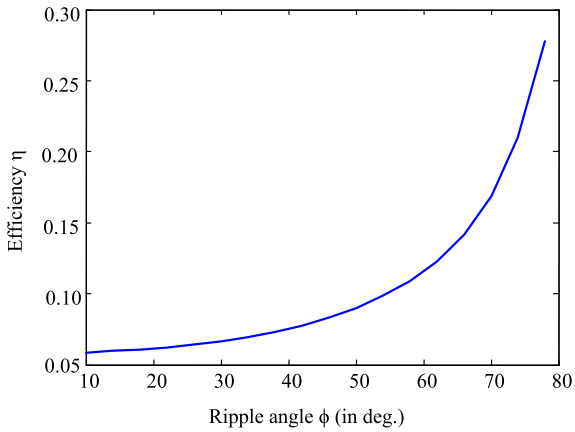


FIG. 8. Variation of the efficiency η of THz radiation emission with ripple angle ϕ (in deg) for beam energy $E_b = 1.53$ MeV, $\Delta = 0.7$, $n_{po} = 1.5 \times 10^{14}$ cm^{-3}, and radiation frequency f = 2.9 THz.

the closer ripples provide better efficiency of the THz radiation emission.

V. CONCLUSION

This work explores the possibility for the generation of THz radiation with substantial advantage of pre-modulated relativistic electron beam in a rippled density plasma or wiggler in Free electron Laser (FEL) scheme. This scheme suggests that the tunable and coherent THz radiation wavelengths depend on rippled density plasma and energy of the beam electrons. In addition, the effect of plasma density, beam density, beam energy, and the modulation index on the growth rate of radiation wave is studied. The variation of the efficiency of THz emission with frequency of radiation wave has also been studied in this investigation. It is found that the plasma and beam densities enhance the growth rate of radiation wave. Moreover, the radiation wave attenuates with the increase in the beam energies. The frequency of radiation wave is seen to reduce the efficiency of the THz radiation. An important observation of this work is the pre-modulation of the electron beam that leads to a significant enhancement in the growth rate, efficiency, and power density of the tunable THz radiation wave and reaches the largest value (i.e., satisfies the Cerenkov resonance condition) as the modulation index approaches unity. This analysis also reveals that the amplification and generation mechanism of the THz wave pumped by the rippled density plasma are favourable for mildly relativistic electron beam. The results of this work may be employed in the practical applications of THz radiation emission. To the best of the authors' knowledge, no experimental work has been reported so far on THz radiation generation by a density modulated electron beam in a rippled density plasma.

[1]M. Tonouchi, Nat. Photonics **1**, 97 (2007).
[2]B. Ferguson and X. C. Zhang, Nat. Mater. **1**, 26 (2002).
[3]E. Pickwell and V. P. Wallace, J. Phys. D **39**, 301 (2006).
[4]J. Faure, J. V. Tilborg, R. A. Kanidal, and W. P. Leemans, Opt. Quantum Electron. **36**, 681 (2004).
[5]D. Hashimshony, A. Zigler, and K. Papadopoulos, Phys. Rev. Lett. **86**, 2806 (2001).
[6]J. F. Holzman and A. Y. Elezzabi, Appl. Phys. Lett. **83**, 2967 (2003).

[7]J. M. Byrd, Z. Hao, M. C. Martin, D. S. Robin, F. Sannibale, R. W. Schoenlein, A. A. Zholents, and M. S. Zolotorev, Phys. Rev. Lett. **96**, 164801 (2006).
[8]Y. Shen, X. Yang, G. L. Carr, Y. Hidaka, J. B. Murphy, and X. Wang, Phys. Rev. Lett. **107**, 204801 (2011).
[9]C. Joshi, T. Katsouleas, J. M. Dawson, Y. T. Yan, and J. M. Slater, IEEE Trans. J. Quantum Electron. **23**, 1571 (1987).
[10]R. Fedele, G. Miano, and V. G. Vaccaro, Phys. Scr. T **30**, 192 (1990).
[11]A. K. Malik, H. K. Malik, and Y. Nishida, Phys. Lett. A **375**, 1191 (2011).
[12]S. C. Sharma and P. Malik, Phys. Plasmas **22**, 043301 (2015).
[13]V. B. Pathak, D. Dahiya, and V. K. Tripathi, J. Appl. Phys. **105**, 013315 (2009).
[14]C. Weiss, R. Wallenstein, and R. Beigang, Appl. Phys. Lett. **77**, 4160 (2000).
[15]C. S. Liu and V. K. Tripathi, J. Appl. Phys. **105**, 013313 (2009).
[16]A. K. Malik, H. K. Malik, and S. Kawata, J. Appl. Phys. **107**, 113105 (2010).
[17]D. Xiang and G. Stupakov, Phys. Rev. Spec. Top. Accel. Beams **12**, 080701 (2009).
[18]W. P. Leemans, J. V. Tilborg, J. Faure, C. G. R. Geddes, C. Toth, C. B. Schroeder, E. Esarey, G. Fubiani, and D. Dugan, Phys. Plasmas **11**, 2899 (2004).
[19]B. D. Layer, A. York, T. M. Antonsen, S. Varma, Y. H. Chen, Y. Leng, and H. M. Milchberg, Phys. Rev. Lett. **99**, 035001 (2007).
[20]S. A. Maier, S. R. Andrews, L. M. Moreno, and F. J. G. Vidal, Phys. Rev. Lett. **97**, 176805 (2006).
[21]M. Singh, R. K. Singh, and R. P. Sharma, Europhys. Lett. **104**, 35002 (2013).
[22]E. Budiarto, J. Margolies, S. Jeong, J. Son, and J. Bokor, IEEE J. Quantum Electron. **32**, 1839 (1996).
[23]H. Parchamay, N. Yugami, T. Muraki, H. Noda, and Y. Nishid, Jpn. J. Appl. Phys. **44**, 5 (2005).
[24]X. G. Dong, Z. M. Sheng, H. C. Wu, W. M. Wang, and J. Zhang, Phys. Rev. E **79**, 046411 (2009).
[25]H. Hamster, A. Sullivan, S. Gordon, W. White, and R. W. Falcone, Phys. Rev. Lett. **71**, 2725 (1993).
[26]T. M. Antonsen, Jr., J. Palastro, and M. Milchberg, Phys. Plasmas **14**, 033107 (2007).
[27]M. Nejati, C. Javaherian, B. Shokri, and B. Jazi, Phys. Plasmas **16**, 022108 (2009).
[28]X. B. Zhang, X. Qiao, L. H. Cheng, R. A. Tang, A. X. Zhang, and J. K. Xuea, Phys. Plasmas **22**, 094502 (2015).
[29]C. S. Liu and V. K. Tripathi, *Interaction of Electromagnetic Waves With Electron Beams and Plasmas* (World Scientific Publishing Co. Pvt. Ltd., Singapore, 1994).
[30]V. A. Namiot and L. Y. Shchurova, Phys. Lett. A **375**, 2759 (2011).
[31]S. C. Zhang, Phys. Plasmas **16**, 093107 (2009).
[32]A. K. Malik and K. P. Singh, Laser Part. Beams **33**, 519 (2015).
[33]A. K. Malik, K. P. Singh, and V. Sajal, Phys. Plasmas **21**, 073104 (2014).
[34]A. K. Malik and H. K. Malik, IEEE J. Quantum Electron. **49**, 232 (2013).

Biocatalysis and Agricultural Biotechnology 11 (2017) 248–251

Contents lists available at ScienceDirect

Biocatalysis and Agricultural Biotechnology

journal homepage: www.elsevier.com/locate/bab

Immobilization studies of cellulase on three engineered polymer surfaces

Rajesh Ahirwar[a,b], Jai G. Sharma[c], Pradip Nahar[a], Saroj Kumar[a,c,*]

[a] CSIR-Institute of Genomics and Integrative Biology, Delhi, India
[b] Department of Biochemistry, ICMR-National Institute for Research in Environmental Health, Bhopal, India
[c] Department of Biotechnology, Delhi Technological University, Delhi, India

ARTICLE INFO

Keywords:
Immobilization
Cellulase
Photolinker
Inert surface
Enzymatic catalysis

ABSTRACT

Enzymes are valuable micro-machines that catalyze a variety of making and breaking reactions. It is well known that immobilization confer stability to enzymes, however, it came at a cost of calculative optimization, support selection and choice of linking method. Desirable properties like hydrophilicity, biocompatibility, resistance to microbial attacks and easy access at low cost has enabled agarose, sepharose, silica, celite, glass and certain plastics as frequently used supports for enzyme immobilization. Herein, we have compared three common plastic polymers- polystyrene, polypropylene and polyethylene for chemical immobilization and subsequent biotransformation using the cellulase enzyme. The results show that the immobilized cellulase on polystyrene, polypropylene and polyethylene beads retain almost two-fold residual activity compared to free enzyme after three-hour incubation at 65 °C. Also, the immobilized cellulase retained significant activity after multiple uses and upon ambient storage for 35 days. Overall, improved catalysis was observed for cellulase upon immobilization, especially to polystyrene surface.

1. Introduction

Cellulases, the multicomponent enzymes (endocellulases, cellobiohydrolase, and β-glucosidase) produced primarily by fungi, bacteria, and protozoans catalyze the hydrolysis of the β-(1-4) linkages in cellulose to convert it to soluble sugars (Zhang and Zhang, 2013). Cellulases covers a large sphere of industrial applications including textile, detergent, brewery and wine, food, animal feed, leather, pulp, and paper, as well as in research and development (Pazarlioğlu et al., 2005; Filos et al., 2006; Juturu and Wu, 2014). The cellulosic materials serve as a substrate for the production of chemicals, and fuel ethanol by enzymatic conversion (Wayman et al., 1992; Philippidis et al., 1993; Fujita et al., 2002). However, industrial setups relying on the use of free cellulases can suffer an unfavorable economic because of the high requirement, low recycling, and susceptibility to inactivation of the enzyme. The immobilization, on the other hand, can overcome some of these shortcomings for applications such as the production of sugar from cellulosic materials (Tébéka et al., 2009). The quest to reduce the cost while enhancing activity and stability of enzyme to a wide range of pH and temperature has resulted in various innovative and effective physical and chemical methods for immobilization of the cellulases (Hirsh et al., 2010). Usually, covalent methods are preferred for immobilization as it restricts the loss (leaching) of enzyme by making a stable covalent bond with the support. Available information suggests the use of various materials such as liposome, PIII-treated polystyrene, Si wafers, and polyvinyl alcohol membrane for covalent immobilization of cellulases (Wu et al., 2005; Li et al., 2007; Tébéka et al., 2009; Hirsh et al., 2010). The choice of chemical linkers for immobilization varies depending on the linking chemistry. Several studies have used photolabile linkers for covalent immobilization of a variety of biomolecules (Liu and Yan, 2010). Photolabile linkers allow easy and efficient cleavage reactions that can be performed directly into solvents suitable for many biological analyses. Particularly, 1-fluoro-2-nitro-4-azidobenzene (FNAB) is used efficaciously for surface engineering of various polymers for immobilization and bioconjugation of proteins and enzymes (Kumar et al., 2016). For example, biomolecular surfaces constructed using FNAB were used in applications like diagnostic assays (e.g. ELISA), aptamer screening (SELEX), microarray fabrication, and to construct conducting polymers and biosensors (Bora et al., 2002; Deepa et al., 2009; Ahirwar et al., 2014, 2016). Considering the diverse applications of FNAB, the present communication has tried to explore the utility of FNAB constructs in biocatalysis of cellulose. The immobilized cellulase on polystyrene, polypropylene and polyethylene polymer beads was analysed for its relative and residual activity, stability and reusability.

2. Materials and methods

The plastic polymer beads of polystyrene (spherical; diameter =

* Correspondence to: Young Scientist, Delhi Technological University, Delhi 110007, India.
E-mail address: sarojkumar@dtu.ac.in (S. Kumar).

http://dx.doi.org/10.1016/j.bcab.2017.07.014
Received 19 May 2017; Received in revised form 20 July 2017; Accepted 21 July 2017
Available online 25 July 2017
1878-8181/ © 2017 Elsevier Ltd. All rights reserved.

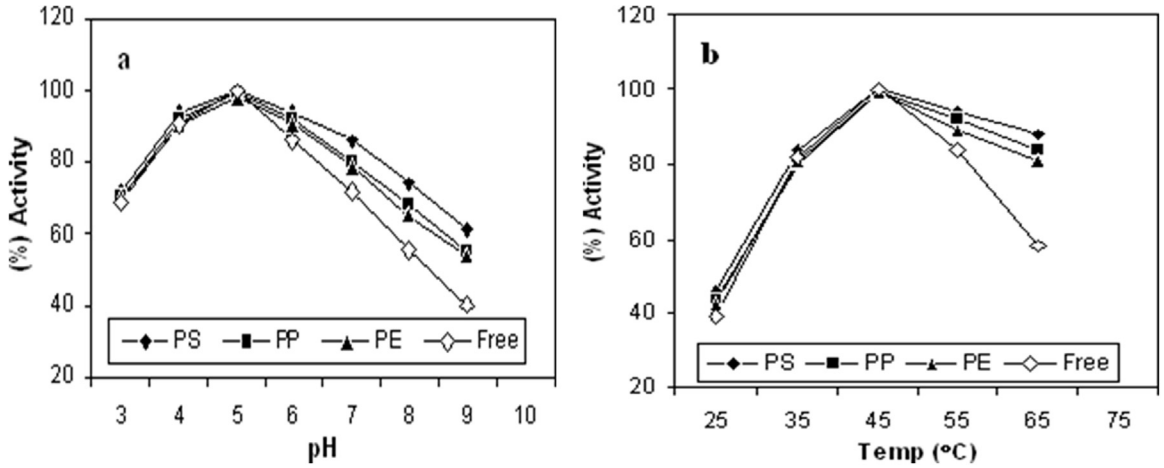


Fig. 1. The effect of pH (a) and temperature (b) on the catalytic activity of immobilized and free cellulase.

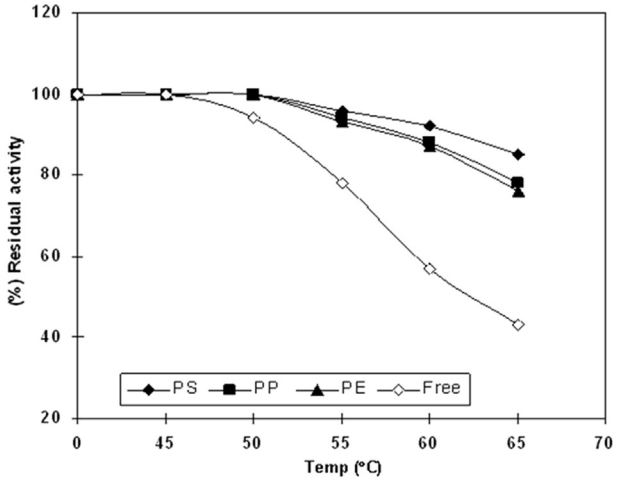


Fig. 2. Thermal stability of immobilized and free enzyme.

3.30 ± 0.36 mm), polypropylene (diameter = 3.40 ± 0.30 mm) and polyethylene (diameter = 3.43 ± 0.55 mm) were prepared for enzyme immobilization by spreading the ethanolic solution of FNAB (300 mg FNAB in 7.5 ml of ethanol) onto respective category of beads, followed by the evaporation of solvent under dark conditions. After complete evaporation of methanol, the beads were irradiated with UV radiations (365 nm) for 15 min in a UV Stratalinker. Afterwards, cellulase was immobilized to the activated polymer beads by incubating 1 g of each polymer beads with cellulase solution (1 mg Cellulase from *Aspergillus niger* in 1 ml of 0.1 M phosphate buffer saline, pH 7.2) for 45 min at 50 °C. The efficacy of FNAB for activation of polymer beads and subsequent immobilization of cellulase was checked by enzymatic assay of cellulase (Ghose, 1987).

The effect of pH on the activity of immobilized enzyme was studied by incubating cellulase-immobilized polymer beads with 1% carboxymethyl cellulose (CMC) solution having different pH (sodium citrate buffer, pH 3–6; PBS, pH 7–9) at 45 °C for 60 min. The resulting reducing sugars were estimated on spectrophotometer at 546 nm using Dinitrosalicylic Acid (DNS) method (Ghose, 1987). Control experiments for comparative analysis were carried out with free cellulase.

Similarly, the effect of temperature on the activity of immobilized and free cellulase was studied by incubating the 1% CMC solution with respective enzyme preparations at 25, 35, 45, 55 and 65 °C respectively, for 60 min. Percentage activity was calculated with respect to the maximum activity of the enzyme. In another experiment, the thermal stability of immobilized and free cellulase was checked by incubating enzyme preparations at 45, 50, 55, 60 and 65 °C for 180 min, followed by assessment of the residual activity by analysing the degradation of 1% CMC solution. Initial activity (before incubation) of the untreated enzyme was taken as control. In subsequent experiments, the reusability of PS, PP and PE- immobilized cellulase was checked for 8

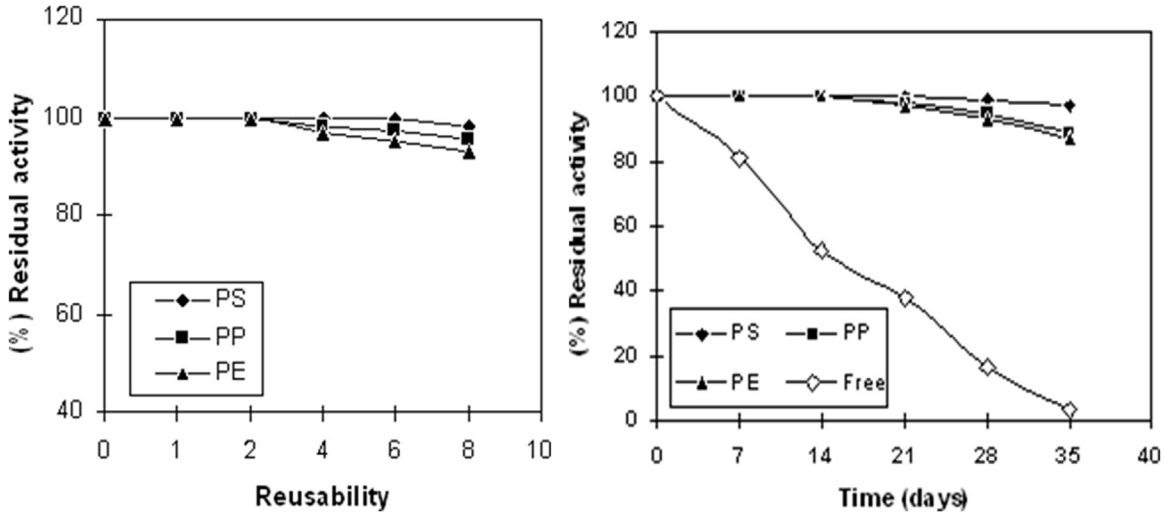


Fig. 3. (a) Reusability of immobilized enzyme. (b) Storage stability of immobilized and free enzyme.

subsequent uses. The first batch of repeated uses was started by adding 5 mg beads of each polymer carrying the immobilized cellulase to 10 ml of 1% CMC solution, followed by 60 min incubation at optimum pH and temperature. Residual activity after each repeated use was determined and expressed as percent to control activity. Storage stability of the immobilized and free cellulase was ascertained by incubating the enzyme preparations at ambient temperature (30 °C) for 7, 14, 21, 27 and 35 days respectively. The activity of the covalently immobilized enzyme was expressed as a percentage of its residual activity with respect to its initial activity.

3. Results and discussion

Restriction of biomolecules, especially the enzymes to a particular support make them robust and resistant to heterogeneous environment. It also prevents biomolecular leaching while allowing multiple re-use, convenient handling of enzyme preparations and an easy recovery of enzymes and products. The last two decades of enzymology has witnessed expand in the use of cellulases in food, agriculture, textile, paper and pulp, detergents, and biosensors industries. A further push to cellulase industry can be provided by developing innovative and cost-effective protocols for cellulase mediated biotransformation. Trying to meet these expectations, we used a photo-labile linker to covalently couple cellulase to polymer surfaces as the FNAB-mediated enzyme immobilization is a simple, cost-effective and highly reproducible procedure that involves reactive nitrene based chemical insertion reaction (Kumar et al., 2016). We used plastic polymer beads (spherical) of polystyrene (diameter = 3.30 ± 0.36 mm), polypropylene (diameter = 3.40 ± 0.30 mm) and polyethylene (diameter = 3.43 ± 0.55 mm) as solid supports because of their chemical inertness toward cellulase, easy availability and low cost of production.

Various reaction parameters were optimized for the immobilized cellulase on plastic polymer beads to assess its ability to degrade cellulosic material. The effect of pH on the activity of immobilized cellulase was studied over a wide range of pH 3.0–9.0. The maximum residual activity for both free and immobilized cellulase was observed at pH 5.0. However, the immobilized cellulase retained better activity at higher pH (Fig. 1a). At same time, the polystyrene-bound cellulase depicted best catalysis among the analysed supports. These observations imply that cellulase stabilizes upon immobilization against conformational disruptions at extreme pH.

Temperature is known to enhance enzyme-catalyzed reactions by allowing better kinetics resulting from ample collision between enzyme and ligands. Analysis of the effects of temperature on the activity of free and immobilized cellulase showed that both free and immobilized cellulase produce similar activity upto an optimum temperature of 45 °C, beyond which (e.g. 65 °C) the free enzyme loss its activity faster than the immobilized enzyme (Fig. 1b). It may be a result of shielding of amino groups at surface and active sites leading to lesser denaturation of enzyme, while the higher kinetic energy of substrate still facilitating higher substrate conversion.

Thermal stability of both immobilized and free cellulase was determined by measuring the residual activity after incubation at elevated temperatures for 3 h. At 65 °C for 180 min incubation, the residual activities of immobilized cellulase on PS, PP and PE were calculated as 85%, 78% and 76.9%, respectively (Fig. 2). On the other hand, the residual activity of free enzyme reduced below 45% after similar incubation at 65 °C for 180 min. Overall, though there was a decrease in activity of both free and immobilized cellulase with increase in incubation temperature, the covalently immobilized cellulase retained its activity for longer duration as compared to the free cellulase enzyme. Among the three tested plastic polymers, thermal stability of cellulase immobilized onto polystyrene was comparatively high than the polypropylene and polyethylene.

From an economic point of view, the reusability, operation stability and easy separation are considered important parameters of enzyme immobilization. Herein, the reusability of immobilized cellulase on PS, PP and PE beads was checked by repeated use for hydrolysis of CMC for 8 times. We observed that after 8 reuses, the immobilized cellulase on PS beads lost around 3% of its initial activity whereas those bound to PP and PE beads lost only approximately 7% of its initial activities during 8 times repeated use (Fig. 3a). This loss in activity of immobilized cellulase is comparatively lower than previous studies which observed a 15% drop in cellulase activity after 6 reuses (Li et al., 2007). The loss in activity can be attributed to inactivation of enzyme. Longer incubations are presumed to further decrease the residual activity of both immobilized and free cellulase because of the time-dependent denaturation from environmental conditions. Incubations requiring longer time should optimize conditions for better biotransformation with minimum enzyme and activity loss. Nevertheless, immobilized enzyme will always tend to retain its activity unlike the free enzyme. The storage stability of the immobilized cellulase was found better compared to the free enzyme. The free cellulase completely loss its activity when stored at 30 °C for 35 days, whereas the immobilized enzyme stored in same conditions retained 100% activity for 16 days (Fig. 3b). After 35 days of storage, the activity of immobilized cellulase onto PS, PP and PE decreased by 2.6%, 11% and 13% respectively. The improved storage stability of immobilized enzymes can be attributed to the reduction in the rate of denaturation of enzyme.

In conclusion, cellulase immobilized to PS, PP and PE beads shows better kinetics, reusability and storage stability. Further, PS-bound cellulose shows excellent operational reusability, pH, thermal and storage stability over the cellulase immobilized onto PP and PE. The PS beads can be used for any enzymatic conversion in a flow-through system such as packed bead reactor. Possibly, the PS beads can also be used as an ideal support for immobilization of other protein.

Acknowledgement

The authors thank funding agencies Council of Scientific and Industrial Research grant CARDIOMED (BSC0122) and SERB (SB/YS/LS-75/2013) Department of Science and Technology, Government of India, for providing the financial support to this work. S. K. thanks to SERB-DST, Government of India for the award of Fast Track Young Scientist.

References

Ahirwar, R., Tanwar, S., Parween, S., Kumar, A., Nahar, P., 2014. Image-based detection of oligonucleotides–a low cost alternative to spectrophotometric or fluorometric methods. Analyst 139 (9), 2186–2192.

Ahirwar, R., Vellarikkal, S.K., Sett, A., Sivasubbu, S., Scaria, V., Bora, U., Borthakur, B.B., Kataki, A.C., Sharma, J.D., Nahar, P., 2016. Aptamer-assisted detection of the altered expression of estrogen receptor alpha in human breast cancer. PloS One 11 (4), e0153001.

Bora, U., Chugh, L., Nahar, P., 2002. Covalent immobilization of proteins onto photo-activated polystyrene microtiter plates for enzyme-linked immunosorbent assay procedures. J. Immunol. Methods 268 (2), 171–177.

Deepa, M., Bhandari, S., Kant, R., 2009. A comparison of charge transport behavior in functionalized and non-functionalized poly 3,4-(ethylenedioxythiophene) films. Electrochim. Acta 54 (4), 1292–1303.

Filos, G., Tziala, T., Lagios, G., Vynios, D.H., 2006. Preparation of cross-linked cellulases and their application for the enzymatic production of glucose from municipal paper wastes. Prep. Biochem. Biotechnol. 36 (2), 111–125.

Fujita, Y., Takahashi, S., Ueda, M., Tanaka, A., Okada, H., Morikawa, Y., Kawaguchi, T., Arai, M., Fukuda, H., Kondo, A., 2002. Direct and efficient production of ethanol from cellulosic material with a yeast strain displaying cellulolytic enzymes. Appl. Environ. Microbiol. 68 (10), 5136–5141.

Ghose, T.K., 1987. Measurement of cellulase activities. Pure Appl. Chem. 59, 257.

Hirsh, S.L., Bilek, M.M.M., Nosworthy, N.J., Kondyurin, A., dos Remedios, C.G., McKenzie, D.R., 2010. A comparison of covalent immobilization and physical adsorption of a cellulase enzyme mixture. Langmuir 26 (17), 14380–14388.

Juturu, V., Wu, J.C., 2014. Microbial cellulases: engineering, production and applications. Renew. Sustain. Energy Rev. 33, 188–203.

Kumar, S., Kumar, D., Ahirwar, R., Nahar, P., 2016. Exploring the flexible chemistry of 4-fluoro-3-nitrophenyl azide for biomolecule immobilization and bioconjugation. Anal. Bioanal. Chem. 408 (25), 6945–6956.

Li, C., Yoshimoto, M., Fukunaga, K., Nakao, K., 2007. Characterization and immobilization of liposome-bound cellulase for hydrolysis of insoluble cellulose. Bioresour.

Technol. 98 (7), 1366–1372.
Liu, L.-H., Yan, M., 2010. Perfluorophenyl azides: new applications in surface functionalization and nanomaterial synthesis. Acc. Chem. Res. 43 (11), 1434–1443.
Pazarlioğlu, N.K., Sariişik, M., Telefoncu, A., 2005. Treating denim fabrics with immobilized commercial cellulases. Process Biochem. 40 (2), 767–771.
Philippidis, G.P., Smith, T.K., Wyman, C.E., 1993. Study of the enzymatic hydrolysis of cellulose for production of fuel ethanol by the simultaneous saccharification and fermentation process. Biotechnol. Bioeng. 41 (9), 846–853.
Tébéka, I.R.M., Silva, A.G.L., Petri, D.F.S., 2009. Hydrolytic activity of free and immobilized cellulase. Langmuir 25 (3), 1582–1587.
Wayman, M., Chen, S., Doan, K., 1992. Bioconversion of waste paper to ethanol. Process Biochem. 27 (4), 239–245.
Wu, L., Yuan, X., Sheng, J., 2005. Immobilization of cellulase in nanofibrous PVA membranes by electrospinning. J. Membr. Sci. 250 (1–2), 167–173.
Zhang, X.-Z., Zhang, Y.H.P., 2013. Cellulases: Characteristics, Sources, Production, and Applications, in Bioprocessing Technologies in Biorefinery for Sustainable Production of Fuels, Chemicals, and Polymers. John Wiley & Sons, Inc., Hoboken, NJ, USA.

Neurocomputing 000 (2017) 1–12

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Linguistic Lyapunov reinforcement learning control for robotic manipulators

Abhishek Kumar, Rajneesh Sharma*

Instrumentation and Control, Netaji Subhas Institute of Technology, Sector-3, Dwarka, Delhi 110085, India

ARTICLE INFO

Article history:
Received 15 January 2017
Revised 13 May 2017
Accepted 26 June 2017
Available online xxx

Communicated by Dr. Chenguang Yang

Keywords:
Reinforcement learning
Fuzzy Q learning
Linguistic Lyapunov RL
Two link robotic manipulator
SCARA

ABSTRACT

We propose a Lyapunov theory based linguistic reinforcement learning (RL) framework for stable tracking control of robotic manipulators. In particular, we employ Lyapunov theory to constrain fuzzy rule consequents for ensuring stability of the designed controller. Proposed fuzzy RL controller employs Lyapunov theory dictated rules for discovering an optimal yet stable control strategy for robotic manipulators. Furthermore, our proposed linguistic RL controller handles payload variations and external disturbances quite effectively. We validate linguistic Lyapunov RL controller on two benchmark control problems: (i) a standard two-link robotic arm manipulator, and (ii) a two link selective compliance assembly robotic arm (SCARA). Simulation results and comparison against (a) baseline fuzzy Q learning (FQL) controller, and (b) a recently proposed Lyapunov theory based Markov game controller showcases our controller's superior tracking performance and lower computational complexity. Furthermore, our controller exhibits high stability with disturbances and payload variations.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Controlling robot manipulators is an extremely challenging and complex control problem due to their highly nonlinear and coupled dynamics. Flawless modeling of robotic manipulators is an equally daunting task; forcing the designer to make idealizing assumptions with consequent modeling errors. This has prompted designers to go for model free approaches. RL has emerged as an alternative technique to design high performance controllers for robot manipulators [1]. However most of the RL approaches proposed so for do not guarantee any stability on the designed controller [2]. Lyapunov theory [3,4] offers a powerful platform for designing controllers that are guaranteed to be stable and can handle model uncertainties and parameter variations. Recently, in [3] authors have introduced stability in the RL paradigm by hybridizing Lyapunov theory into the RL action generation mechanism. The approach is excellent but has one drawback of higher computational complexity as one is required to solve one linear program (LP) per iteration to arrive at an optimal solution. Other approaches [5,6] that have tried to infuse stability in the RL domain via Lyapunov theory advocate an actor-critic (AC) formulation. In an AC formulation we, typically have two function approximators; one critic which approximates value function and an actor to approximate optimal policy. A major hurdle here is the simultaneous training of two networks and difficulty in convergence to the optimal solution. In contrast, we seek to guarantee stability of our controller by what may be called a "linguistic Lyapunov control".

In our approach, we use linguistic RL (wherein rule consequents are linguistic rather than crisp). These rules are Lyapunov constrained so as to impart stability to the designed controller. In our previous works, we have attempted at designing stable RL controllers by: (i) generating a hybrid controller by mixing Lyapunov theory based action with RL action [3] and (ii) curtailing the action set of the RL controller so as to satisfy Lyapunov conditions [4]. In this work, we take a conceptually different approach as: (a) we formulate "learning with words" or a linguistic RL approach where we frame linguistic fuzzy RL rules for learning optimal solution, and (b) we impose Lyapunov constraints on these linguistic RL rules or design a Lyapunov linguistic RL controller. This linguistic Lyapunov framework is then carefully fine-tuned for application on robotic manipulators.

Next, we look at some recently proposed soft computing approaches that have used Lyapunov theory for lending stability to the controller. In [7], authors have proposed neural networks based consensus control for multiple robotic manipulators systems and used Lyapunov theory to tune the NN weights for stability. In [8], Lyapunov theory has been used to design an adaptive neural controller for redundant robot manipulators constrained by mobile obstacles. Li et al. [9] propose an RL approach to coordinated manip-

* Corresponding author.
E-mail addresses: rajneesh496@gmail.com, rajneesh496@rediffmail.com (R. Sharma).

http://dx.doi.org/10.1016/j.neucom.2017.06.064
0925-2312/© 2017 Elsevier B.V. All rights reserved.

Please cite this article as: A. Kumar, R. Sharma, Linguistic Lyapunov reinforcement learning control for robotic manipulators, Neurocomputing (2017), http://dx.doi.org/10.1016/j.neucom.2017.06.064

ulation of multi robots to cope with unknown dynamics of robots and manipulated object. These approaches use NNs (at times two NNs in an actor critic RL formulation) which poses problems in terms of slow convergence and/ or divergence and high computational burden when the system complexity increases. In another approach [10], Lyapunov theory has been used to develop a concurrent learning implementation of model based RL for approximate optimal regulation. The approach relies on availability of a sufficiently accurate system model which is a restrictive pre condition. Our approach, in contrast, proposes a model free linguistic Lyapunov RL controller with no divergence issues and offers scalability to high dimensional non-linear systems.

In implementing RL control on continuous state action space problems such as the manipulator control, use of function approximators becomes a necessity to counter the "curse of dimensionality". We have used fuzzy inference system (FIS) as function approximator as in [3], but our approach has two distinct advantages vis a vis these approaches: (i) we have used only one function approximator unlike the AC based RL approaches [5,6] making it simple and convergence to the optimal solution is fast. In an AC RL setup simultaneous training of two function approximator is required, and (ii) it has lower computational complexity than Markov game based approach [1,3] where a linear program must be solved at each iteration to generate optimal solution. This makes our approach suitable for high dimensional problems as scalability is very easy.

In a recent survey [11], authors have detailed several neural networks (NN) and fuzzy logic based controllers. Some recent NN based RL controllers have also been described that employ two NNs in an actor critic (AC) configuration. However, there are two major concerns in an AC formulation: (i) Simultaneous training and ensuring convergence of actor and critic NNs is difficult and at times tricky, and (ii) severe computational burden due to large number of tunable parameters. In contrast, our proposed approach has fewer tunable parameters and employs a single function approximator leading to faster convergence.

An adaptive fuzzy controller for an exoskeleton robot is presented in [12] wherein fuzzy logic has been employed to approximate nonlinear dynamics of the manipulator. However, our approach is different from this approach [12] in two aspects: (i) in [12], fuzzy logic has been used to approximate the manipulator parameters whereas we use FIS as a function approximator for estimating the linguistic Q function, and (ii) Our approach is an RL approach wherein quadratic error signal is used to modify the q values for optimal trajectory tracking. Moreover our implementation is simpler as we use only one FIS with linguistic consequents.

We test the efficacy of our proposed approach on two benchmark robotic problems: (i) tracking control of a two link robotic manipulator, and (ii) control of a SCARA. The controller is subjected to disturbances and parameter variations as well. Performance comparison has been done against baseline fuzzy Q learning and Lyapunov theory based Markov game controller. Rest of the paper is organized as: theoretical background of reinforcement learning paradigm and fuzzy Q learning is given in Section 2. Section 3 details proposed approach and application of proposed scheme on two link manipulator and SCARA is given in Section 4. In Section 5, we give simulation results and comparison against fuzzy Q learning and Lyapunov theory based Markov game controller which is followed by conclusion and future scope in Section 6.

2. Reinforcement learning and fuzzy Q learning

Reinforcement learning is learning by interacting with the system one intends to control. Typical goal is to maximize an accumulated reward or minimize cost (obtained at each step/interaction). Basically, RL is a family of stochastic iterative algorithms for finding an optimal solution to a sequential decision making problem [13]. RL algorithm can be ; (i) model based wherein a model of the underlying system is learned and is used to generate an optimal solution, and (ii) model free, e.g., Q learning in which model of the plant/system is learned impromptu [13]. Model free RL techniques have gained importance as no model of the system is required and are simple to apply. Next, we describe Q learning (a model free RL approach).

2.1. Q learning

Q learning was proposed by Watkins [13] as a model free technique to optimize unknown systems. It models the sequential decision making task (optimization of controller) as a Markov Decision Process (MDP). MDP is a sequential decision making model in which at time instant k an agent (or controller) applies action $u \in U(s)$ in state $s \in S$ making the environment (or system) move to the next state $s' \in S$, emitting a reward (or cost) *c*. *U(s)* and *S* are set of all possible actions in state *s* and set of all possible states, respectively. This movement from *s* to *s'* is in accordance with state transition probability *p(s,u,s')*. Agent's objective is to optimize policy $\pi: \pi(s) \rightarrow u;\ u \in U(s)$, so that expected accumulated discounted cost is minimized [3]. In Q learning, we define a *Q*-value as the expected accumulated cost incurred by agent on taking action *u* in state *s* and then continuing on an optimal policy. *Q*-value, *Q(s,u)* is given by [3]

$$Q(s,u) = c(s,u) + \xi \sum_{s' \in S} p(s,u,s')V(s') \tag{1}$$

where $V(s') = \min_{u \in U(s')} Q^*(s',u)$ and $\xi \in (0,\ 1]$ is discount factor that gives relationship between present and future cost.

We can generalize (1) as update rule

$$Q(s^k,u^k) \leftarrow Q(s^k,u^k) + \lambda[c(s^k,u^k) + \xi \min_{u \in U(s')} Q(s^{k+1},u^k) - Q(s^k,u^k)] \tag{2}$$

where s^k and s^{k+1} denotes state at *k*th and $(k+1)$th iteration, u^k gives action taken at iteration *k*, $\lambda \in (0,\ 1]$ is learning rate parameter. Infinitely large visit to each state action pair and proper decrement in λ gives optimal *Q*-values i.e. Q^*.

2.2. Fuzzy Q learning

In Q learning, we need to store *Q* values for each visited state action pair; *Q* value storage becomes infeasible for large state spaces. This problem can be solved using generalization techniques, e.g., neural network (NN) and fuzzy logic have been used for generalization [1,3–6]. We use fuzzy inference system (FIS) to implement Q learning, also known as fuzzy Q learning. Rules formulated in fuzzy Q learning (FQL) have the form [3]:

$$\begin{aligned} R_n: \quad & \text{If } s_1^k \text{ is } L_1^n \text{ and } ..... \text{ and } s_m^k \text{ is } L_m^n \text{ then } u = u_1 \text{ with } q(n,1) \\ & \text{or } u = u_2 \text{ with } q(n,2) \\ & ..... \\ & \text{or } u = u_p \text{ with } q(n,p) \end{aligned} \tag{3}$$

In rule R_n linguistic term L_v^n for input variable s_v^k has membership function $\eta_{L_v^n}$. s^k denotes system state $\{s_1^k, s_2^k, ..., s_m^k\}$ at instant *k*, and acts as input to the FIS. Matching of rule premise to the input vector s^k is used to calculate truth-value of each rule $\mu(s^k)$: $[\mu_1(s^k)\,\mu_2(s^k)....\mu_N(s^k)]$ corresponding to *N* rules.

In *n*th rule (R_n) the agent selects minimizing action, i.e., the one with minimum $q(n,\ u_m)$ value from set $U = \{u_1, u_2, ..., u_p\}$:

$$u_n^* = \arg \min_{u_n \in U} q(n,u_n) \tag{4}$$

Please cite this article as: A. Kumar, R. Sharma, Linguistic Lyapunov reinforcement learning control for robotic manipulators, Neurocomputing (2017), http://dx.doi.org/10.1016/j.neucom.2017.06.064

Takagi Sugeno FIS is used to generate an aggregate action:

$$u_{fq}(s^k) = \frac{\sum_{n=1}^{N} \mu_n(s^k) u_n^*}{\sum_{n=1}^{N} \mu_n(s^k)} \tag{5}$$

Exploration exploitation policy (EEP) is adopted by an RL agent to look for better rewarding actions. Denoting EEP action by $u_n^\dagger$ and optimal action by u_n^* in nth rule we can define ε - *Greedy* (EEP) action as:

$$\varepsilon \text{ - } Greedy(u_n^*) : u_n^\dagger = \begin{cases} u_n^r \text{ with probability } \varepsilon \\ u_n^* \text{ with probability } 1-\varepsilon \end{cases} \tag{6}$$

where u_n^r is random action under rule n. The parameter ε is decreased from 0.4 to 0.01 over trials to make agent more exploratory in the beginning and optimal in later stages.

Q-value for defuzzified action $u_{fq}(s^k)$ at state s^k is given by

$$Q(s^k, u_{fq}(s^k)) = \frac{\sum_{n=1}^{N} \mu_n(s^k) q(n, u_n^\dagger)}{\sum_{n=1}^{N} \mu_n(s^k)} \tag{7}$$

Target value at state s^k is given by

$$V(s^k) = \frac{\sum_{n=1}^{N} \mu_n(s^k) q(n, u_n^*)}{\sum_{n=1}^{N} \mu_n(s^k)} \tag{8}$$

The system moves from state s^k to state s^{k+1} when action $u_{fq}(s^k)$ is applied on the system in state s^k incurring a cost c^k. Temporal difference (TD) error is calculated as:

$$\Delta Q = c^k + \lambda V(s^{k+1}) - Q(s^k, u_{fq}(s^k)) \tag{9}$$

Finally, Q values are updated as:

$$q(n, u^\dagger) \leftarrow q(n, u^\dagger) + \lambda \Delta Q \frac{\mu_n(s^k)}{\sum_{n=1}^{N} \mu_n(s^k)} \forall n \in N \tag{10}$$

where $0 < \lambda < 1$ denotes learning rate.

2.3. Lyapunov theorem

Lyapunov theorem has been used extensively to introduce stability in a nonlinear system control [3–6,14]. The trick is to select a Lyapunov candidate function (LCF) which is positive definite and decreasing or constant in nature. If function $P(s)$ has to be a LCF then it must satisfy following conditions:

$$\begin{aligned} &\text{L1. } P(0) = 0 \\ &\text{L2. } P(s) > 0, \ \forall s \neq 0 \\ &\text{L3. } \dot{P}(s) < 0, \ \forall s \neq 0 \end{aligned} \tag{11}$$

Here $s = 0$ is the equilibrium state. Selection of a befitting LCF is the most critical task in designing a stable controller. LCF is selected in such a way that LCF and/or its first derivative has some association with the control action and the control action is selected in such a manner that the LCF becomes positive definite and its first derivative becomes negative definite giving us a Lyapunov stable controller.

3. Linguistic Lyapunov reinforcement learning control for manipulators

3.1. Deriving linguistic Lyapunov rules

In trajectory tracking tasks, the manipulator has to follow a desired or target path θ_d. We assume that we have access to $\dot{\theta}_d$, $\ddot{\theta}_d$ and that these are bounded. Tracking error is given by $e = \theta_d - \theta$ and $\dot{e} = (\dot{\theta}_d - \dot{\theta})$ where $\theta_d, \theta, \dot{\theta}_d$ and $\dot{\theta}$ corresponds to desired position, actual position, desired velocity and actual velocity of system. Equilibrium state is denoted by $(e, \dot{e}) = (0, 0)$. We choose the Lyapunov candidate function (LCF) as $P = P(e, \dot{e}) = \frac{1}{2}(e^2 + \dot{e}^2)$. Clearly our choice of LCF satisfies L1 and L2 conditions (11) so only L3 needs to be satisfied. Now,

$$\dot{P} = e\dot{e} + \dot{e}\ddot{e} \tag{12}$$

For P to be an LCF:

$$e\dot{e} + \dot{e}\ddot{e} < 0 \tag{13}$$

We have no prior information about the system but we do know that $\ddot{e}$ depends directly on applied torque 'τ'. Assuming that we can measure position and velocity of the system (13) can be approximately written using the "computing with words" approach [14] as

$$e\dot{e} + \dot{e}\tau < 0 \tag{14}$$

From (14) it is implicit that if e and $\dot{e}$ have different sign then $\tau = 0$ or τ with an opposite sign to $\dot{e}$ will suffice to satisfy (14). Furthermore a) if e and $\dot{e}$ are positive then (14) will be satisfied if $\tau < -e$ and b) if e and $\dot{e}$ are negative then (14) will be satisfied if $\tau > -e$. Following the "computing with words" approach [14], we can say that for making $\tau > -e$ (for negative e) we can specify τ as "Positive Big", linguistically. Also for making $\tau < -e$ (for positive e) we can assign τ to the linguistic variable "Negative Big". This linguistic interpretation of Lyapunov condition (11) gives us following Lyapunov linguistic rules for manipulators ($e\dot{e} + \dot{e}\tau < 0$):

IF e is Pos AND $\dot{e}$ is Pos THEN τ is Neg_Big
IF e is Pos AND $\dot{e}$ is Neg THEN τ is Zero or Pos
IF e is Neg AND $\dot{e}$ is Pos THEN τ is Zero or Neg
IF e is Neg AND $\dot{e}$ is Neg THEN τ is Pos_Big (15)

here Pos, Neg, Neg_Big, Pos_Big and Zero are Linguistic variables representing Positive, Negative, Negative Big, Positive Big and zero, respectively. We have assigned these linguistic variables to e, $\dot{e}$ & τ to generate the above Lyapunov linguistic rules (15). Finally, we attempt to design what may be called "Linguistic Lyapunov RL Controller" by framing rules of the following form:

$$\begin{aligned} R_n: \quad &\text{If } s_1^k \text{ is } L_1^n \text{ and } \ldots\ldots \text{ and } s_m^k \text{ is } L_m^n \text{ then } U = U_{nL} \text{ and} \\ &u = u_{1L} \text{ with } q_{lL}(n,1) \\ &\text{or } u = u_{2L} \text{ with } q_{lL}(n,2) \\ &\ldots\ldots\ldots \\ &\text{or } u = u_{pL} \text{ with } q_{lL}(n,p) \end{aligned} \tag{16}$$

Here (u_{1L}, u_{2L}, ..., u_{pL}) are actions corresponding to linguistic Lyapunov action set U_{nL} at state s^k for rule R_n and q_{lL} is the q value corresponding to linguistic Lyapunov consequent.

Next, RL is used to select most optimal action in each rule for the linguistic Lyapunov rule base (16), which essentially selects an action from U_{lL} to make $\dot{P}$ as "negative" as possible. In other words, using RL we search for an optimal linguistic Lyapunov controller. Thus our proposed controller is a model free stable RL controller as it selects linguistic action from a Lyapunov dictated "term set".

3.1.1. Linguistic Lyapunov RL controller implementation

Input state vector $s^k = \{s_1^k, s_2^k, \ldots\ldots, s_m^k\}$ is matched with the rule antecedents (16) giving a rule firing strength vector $\mu(s^k)$: $[\mu_1(s^k)\,\mu_2(s^k)\ldots\mu_N(s^k)]$. For each rule, linguistic Lyapunov action set U_{lL} is estimated as described in earlier section. Optimal (or

Please cite this article as: A. Kumar, R. Sharma, Linguistic Lyapunov reinforcement learning control for robotic manipulators, Neurocomputing (2017), http://dx.doi.org/10.1016/j.neucom.2017.06.064

greedy action as per RL literature) is selected for this Lyapunov set as:

$$u_{nL}^{*} = \arg \min_{u_{nL} \in U_{nL}} q_{fL}(n, u_{nL}) \tag{17}$$

Where u_{nL} is the Lyapunov constrained action chosen from Lyapunov constrained linguistic action set U_{nL}. Aggregating these rule specific actions gives:

$$u_{lL}(s^k) = \frac{\sum_{n=1}^{N} \mu_n(s^k) u_{nL}^{*}}{\sum_{n=1}^{N} \mu_n(s^k)} \tag{18}$$

where $u_{lL}(s^k)$ is the linguistic Lyapunov action at state s^k.

We implement EEP as:

$$u_{nL}^{\dagger} = \varepsilon \text{ - } Greedy(u_{nL}^{*}) \tag{19}$$

Q-value for action $u_{lL}(s^k)$ at state s^k is given by

$$Q_{lL}(s^k, u_{lL}(s^k)) = \frac{\sum_{n=1}^{N} \mu_n(s^k) q_{lL}(n, u_{nL}^{\dagger})}{\sum_{n=1}^{N} \mu_n(s^k)} \tag{20}$$

Target value at state s^k is estimated as:

$$V_{lL}(s^k) = \frac{\sum_{n=1}^{N} \mu_n(s^k) q_{lL}(n, u_{nL}^{*})}{\sum_{n=1}^{N} \mu_n(s^k)} \tag{21}$$

The system emanates a cost signal c_L^k when transitioning from state s^k to s^{k+1} under action $u_{lL}(s^k)$. Temporal difference (TD) error is:

$$\Delta Q_{lL} = c_L^k + \lambda V_{lL}(s^{k+1}) - Q_{lL}(s^k, u_{lL}(s^k)) \tag{22}$$

Finally, we update q-values:

$$q_{lL}(n, u_{lL}^{\dagger}) \leftarrow q_{lL}(n, u_{lL}^{\dagger}) + \lambda \Delta Q_{lL} \frac{\mu_n(s^k)}{\sum_{n=1}^{N} \mu_n(s^k)} \forall n \in N \tag{23}$$

where $\lambda \in (0, 1)$ denote learning rate. This update makes q values converge, i.e., $\Delta Q_{lL} \to 0$ and the controller becomes a linguistic Lyapunov RL controller.

4. Linguistic Lyapunov controller formulation

4.1. Robotic manipulator modeling

Dynamic model of robot manipulator with i joints may be written as [1,3]:

$$M(\Theta)\ddot{\Theta} + C(\Theta, \dot{\Theta})\dot{\Theta} + g(\Theta) + f(\dot{\Theta}) + \tau_d = \tau \tag{24}$$

where Θ represents a vector of $i \times 1$ joint angles, τ and τ_d represents vector of $i \times 1$ input torques and external disturbance, respectively.$M(\Theta)$ is inertial matrix with matrix size $i \times i$, Centrifugal and Coriolis matrix is represented by $C(\Theta, \dot{\Theta})$, while $f(\dot{\Phi})$ and $g(\Phi)$ represent frictional force and gravitational force vector of size $i \times 1$, respectively.

We implement the proposed linguistic Lyapunov RL Controller as a decentralized controller, i.e., manipulator has a dedicated controller for each of its links. $\Theta(i)$ and $\dot{\Theta}(i)$ represent angle and velocity of ith link and is the input vector for the ith link. We fuzzify each input variable into 3 partitions giving 9 rules for individual link and overall 18 rules for the two link robot. We use a membership function as in [3]:

$$\mu_{lp}(x_j(i)) = e^{\frac{-\left(x_j(i) - x_j^{l_p}(i)\right)}{\sigma_j(i)}}; \; l_p = 1, 2, 3; \; j = 1, 2; \; i = 1, 2 \tag{25}$$

Here l_p denotes fuzzy label in ith joint for jth link, $((x_1(1) = \Theta(1), x_1(2) = \Theta(2), x_2(1) = \dot{\Theta}(1)$ and $x_2(2) = \dot{\Theta}(2))$.

Centers for these fuzzy labels have been derived empirically [1,3] and are defined as $x_j^{l_p}(i) = a_j(i) + b_j(i)(l_p - 1)$where $a_1(1) = a_1(2) = -2$, $a_2(1) = a_2(2) = -10$, $b_1(1) = b_1(2) = 2$, and $b_2(1) = b_2(2) = 10$ and widths of fuzzy label are given as $\sigma_1(1) = \sigma_1(2) = 1.5$ and $\sigma_2(1) = \sigma_2(2) = 8$. The tracking errors for links are: $e(1) = \Theta_d(1) - \Theta(1)$, $e(2) = \Theta_d(2) - \Theta(2)$ and cost function is given as $c^k = \frac{1}{2}((e(1))^2 + (e(2))^2 + (\dot{e}(1))^2 + (\dot{e}(2))^2)$. Desired trajectory for robot manipulator is given by $\Theta_d = [\Theta_d(1)\ \Theta_d(2)]$.

4.2. Linguistic Lyapunov robotic controller

Consider a function $P = \frac{1}{2}((e(1))^2 + (e(2))^2 + (\dot{e}(1))^2 + (\dot{e}(2))^2)$, then from (11) it is evident that condition L1 and L2 are satisfied for this function. For making P, an LCF, condition L3 has to be satisfied i.e.$\dot{P} < 0$

$$\begin{aligned}
&\Rightarrow e(1)\dot{e}(1) + e(2)\dot{e}(2) + \dot{e}(1)\ddot{e}(1) + \dot{e}(2)\ddot{e}(2) < 0 \\
&\Rightarrow e(1)\dot{e}(1) + e(2)\dot{e}(2) + \dot{e}(1)\left(\ddot{\Theta}_d(1) - \ddot{\Theta}(1)\right) \\
&\quad + \dot{e}(2)\left(\ddot{\Theta}_d(2) - \ddot{\Theta}(2)\right) < 0 \\
&\Rightarrow e(1)\dot{e}(1) + e(2)\dot{e}(2) + \dot{e}(1)\ddot{\Theta}_d(1) + \dot{e}(2)\ddot{\Theta}_d(2) \\
&\quad - \dot{e}(1)\ddot{\Theta}(1) - \dot{e}(2)\ddot{\Theta}(2) < 0
\end{aligned} \tag{26}$$

We don't know exact manipulator model but from the dynamics of robot manipulators we know that $\ddot{\Theta}(1)$ and $\ddot{\Theta}(2)$ depend directly on $\tau(1) - \tau(2)$ and $\tau(2) - \tau(1)$, respectively (simulation model and corresponding derivation given in Appendices A–C). Hence (26) gets modified to

$$\begin{aligned}
&e(1)\dot{e}(1) + e(2)\dot{e}(2) + \dot{e}(1)\ddot{\Theta}_d(1) + \dot{e}(2)\ddot{\Theta}_d(2) \\
&\quad - \dot{e}(1)(\tau(1) - \tau(2)) - \dot{e}(2)(\tau(2) - \tau(1)) < 0 \\
&\Rightarrow e(1)\dot{e}(1) + e(2)\dot{e}(2) + \dot{e}(1)\ddot{\Theta}_d(1) + \dot{e}(2)\ddot{\Theta}_d(2) \\
&\quad + \tau(1)\left(\dot{e}(2) - \dot{e}(1)\right) + \tau(2)\left(\dot{e}(1) - \dot{e}(2)\right) < 0
\end{aligned} \tag{27}$$

Considering $e(1)\dot{e}(1) + e(2)\dot{e}(2) + \dot{e}(1)\ddot{\Theta}_d(1) + \dot{e}(2)\ddot{\Theta}_d(2) = k1$ and $\tau(1)(\dot{e}(2) - \dot{e}(1)) + \tau(2)(\dot{e}(1) - \dot{e}(2)) = k2$, (27) reduces to

$$k1 + k2 < 0 \tag{28}$$

For satisfying (28) we formulate following rules:

IF $k1$ is Neg THEN $k2$ is Zero or Neg or Neg_Big
IF $k1$ is Pos THEN $k2$ is Neg_Big
IF $k1$ is Zero THEN $k2$ is Neg or Neg_Big (29)

Closer look on (29) shows that it is enough to choose $k2$ Neg_Big to satisfy (28). So to make P an LCF it is sufficient to ensure $k2$ is Neg_Big (irrespective of $k1$ value). In other words, $k2 = \tau(1)(\dot{e}(2) - \dot{e}(1)) + \tau(2)(\dot{e}(1) - \dot{e}(2))$ should be much less than zero, giving following linguistic rules: (Table 1).

We define two Lyapunov linguistic action sets: $U_{negL} = \{\text{Zero, Neg, Neg_Big}\}$ and $U_{posL} = \{\text{Zero, Pos, Pos_Big}\}$. Next, RL is employed to select one action from either U_{negL} or U_{posL} as the case may be, e.g., if $\dot{e}(1)$ is POS and $\dot{e}(2)$ is NEG then the Lyapunov linguistic controller must select an action from U_{posL} for $\tau(1)$ and U_{negL} for $\tau(2)$. This action set has been dictated by the Lyapunov condition (L3). Primary aim is to choose most optimal action from this Lyapunov linguistic action set or to choose an action that makes $\dot{P}$ as negative as possible. This combination of RL and Lyapunov theory based linguistic rules results in a self-learning stable controller. Complete linguistic Lyapunov controller is shown in Fig. 1.

Please cite this article as: A. Kumar, R. Sharma, Linguistic Lyapunov reinforcement learning control for robotic manipulators, Neurocomputing (2017), http://dx.doi.org/10.1016/j.neucom.2017.06.064

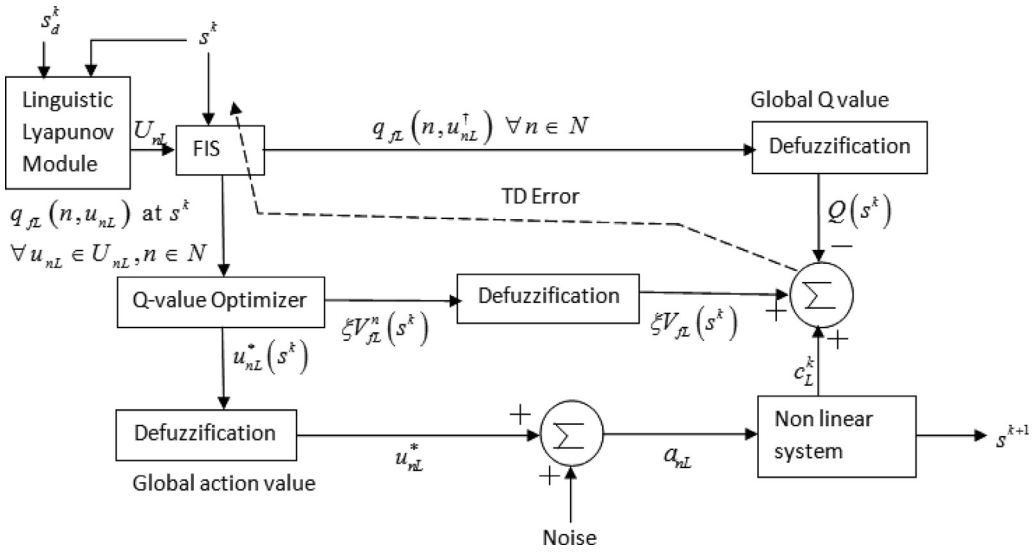


Fig. 1. Block diagram for linguistic Lyapunov reinforcement learning controller.

Table 1.
Linguistic rules for stable manipulator control.

$\dot{e}(1)$	$\dot{e}(2)$	$\tau(1)$	$\tau(2)$
Pos	Pos	for $\dot{e}(1) \geq \dot{e}(2) \Rightarrow$ Pos_Big	for $\dot{e}(1) \geq \dot{e}(2) \Rightarrow$Neg_Big
		for $\dot{e}(1) < \dot{e}(2) \Rightarrow$Neg_Big	for $\dot{e}(1) < \dot{e}(2) \Rightarrow$Pos_Big
Neg	Pos	Neg	Pos
Zero	Pos	Neg	Pos
Pos	Neg	Pos	Neg
Neg	Neg	for $\dot{e}(1) \geq \dot{e}(2) \Rightarrow$Pos_Big	for $\dot{e}(1) \geq \dot{e}(2) \Rightarrow$Neg_Big
		for $\dot{e}(1) < \dot{e}(2) \Rightarrow$Neg_Big	for $\dot{e}(1) \geq \dot{e}(2) \Rightarrow$Pos_Big
Zero	Neg	Pos	Neg
Pos	Zero	Pos	Neg
Neg	Zero	Neg	Pos
Zero	Zero	Zero	Zero

5. Simulation of linguistic Lyapunov RL control

We simulate the proposed controller on two robot manipulator tasks: (i) A two link manipulator, and (ii) SCARA. We have used two link manipulators in this study as they are computationally less expensive than six or more degree of freedom manipulators. However a two link manipulator suffices in terms of capturing all the nonlinearities, complexities and highly coupled nature of robot dynamics. We perform simulations on an Intel (R) Core i3, 2.27 GHz processor with 3 GB RAM using MATLAB®. Parameters used in the simulations are: learning rate parameter is 0.05, discount factor is 0.9, exploration is decreased gradually from 40% to 1% and sampling time is 0.02 s. Lyapunov linguistic action sets (empirical) used are U_{negL}= [−20 −10 0] Nm and U_{posL}= [20 10 0] N m. We have kept these parameters same for both the manipulator problems.

5.1. Two link robot arm manipulator

Simulation model and related details of two link robot arm manipulator (TLM) can be found in Appendix A. We simulate the manipulator for 20 s using 4th order Runge–Kutta method with sampling time 20 ms. Performance of our linguistic Lyapunov controller is compared against baseline FQL controller and a recently proposed Lyapunov theory based Markov game controller (LMGC), for two cases: (i) with external disturbance and fixed payload, i.e., control versus external disturbances and (ii) external disturbances with a varying payload, i.e., control against external disturbances with payload variations. External disturbance is introduced by adding a Gaussian distributed torque equal to ±20 % of the applied torque. Payload m_2 varies as [3]:

(a) $t \leq 4$ s, $m_2 = 1$ kg, (b) $4\text{ s} < t \leq 7$ s, $m_2 = 2.5$ kg,
(c) $7\text{ s} < t \leq 9$ s, $m_2 = 1$ kg, (d) $9\text{ s} < t \leq 12$ s, $m_2 = 4$ kg,
(e) $12\text{ s} < t \leq 15$ s, $m_2 = 1$ kg, (f) $15\text{ s} < t \leq 18$ s, $m_2 = 2$ kg, and
(g) $18\text{ s} < t \leq 20$ s, $m_2 = 1$ kg.

5.1.1. Two link robot arm manipulator results

5.1.1.1. Control with disturbances and payload changes. Figs. 2 and 3 show tracking errors for link1 and link 2 for the standard two link robot manipulator with disturbances and payload variations. Our linguistic Lyapunov RL controller (LLRLC) is found to have least tracking errors and handles disturbances quite effectively in comparison to the FQL controller (FQC) and Lyapunov theory based Markov game controller. Though LMGC gives lower tracking error for link 2 (i.e. for theta2), its superior tracking (for theta 2) is marred by spikes at several places which is undesirable. Our LLRLC does not give any spike/chatter during manipulator control.

Control torque comparison of controllers (LLRLC, FQC and LMGC) is shown in Figs. 4 and 5. Our LLRLC exerts least torque (link1) while LMGC exhibits spikes. Lower control torque requirement is an advantage of LLRLC. For link 2 (Fig. 5), we observe that LLRLC not only has least torque requirements but has least torque pulsations/chattering.

5.1.1.2. Results without payload variations. Figs. 6 and 7 show tracking errors for link 1 and link 2 of the standard two link robot manipulator with only disturbances (no payload variations). We observe that linguistic Lyapunov RL controller achieves lower tracking errors than the fuzzy Q-learning controller and Lyapunov theory based Markov game controller. Here again LMFC has comparable performance against LLRLC for link 1. From Fig. 7, we observe that LMGC exhibits torque pulsation or chattering at several points which is undesirable.

Control torque requirement for all the controllers are shown in Figs. 8 and 9. We observe that LLRLC exerts least torque amongst all controllers for achieving desired trajectory. Here again, we observe that LLRLC produces most smooth control torque (and almost chatter free) while LMGC has quite serious torque pulsation. For link 2 all controllers exhibit lot of chattering which is least in the case of LLRLC.

Please cite this article as: A. Kumar, R. Sharma, Linguistic Lyapunov reinforcement learning control for robotic manipulators, Neurocomputing (2017), http://dx.doi.org/10.1016/j.neucom.2017.06.064

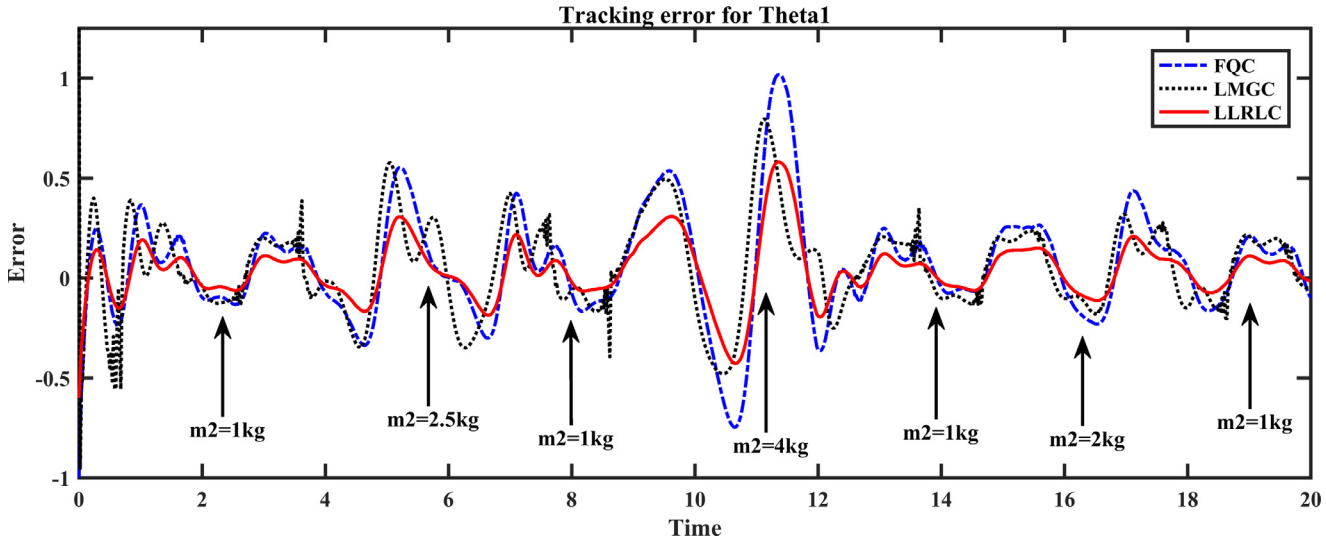


Fig. 2. Tracking error for theta1 with load variation.

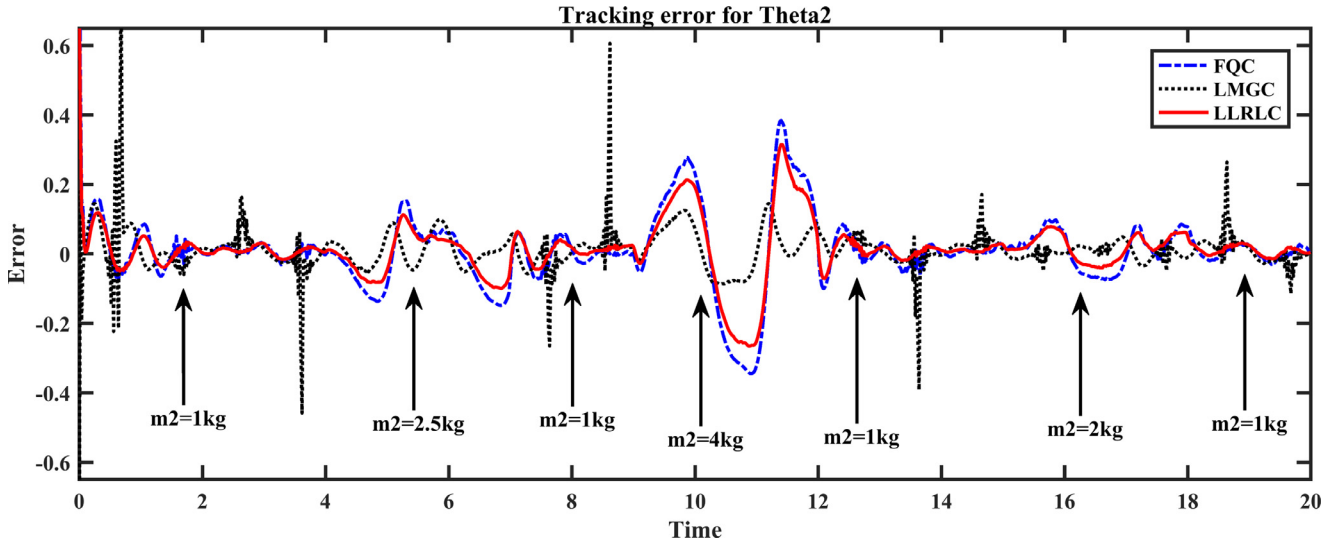


Fig. 3. Tracking error for theta2 with load variation.

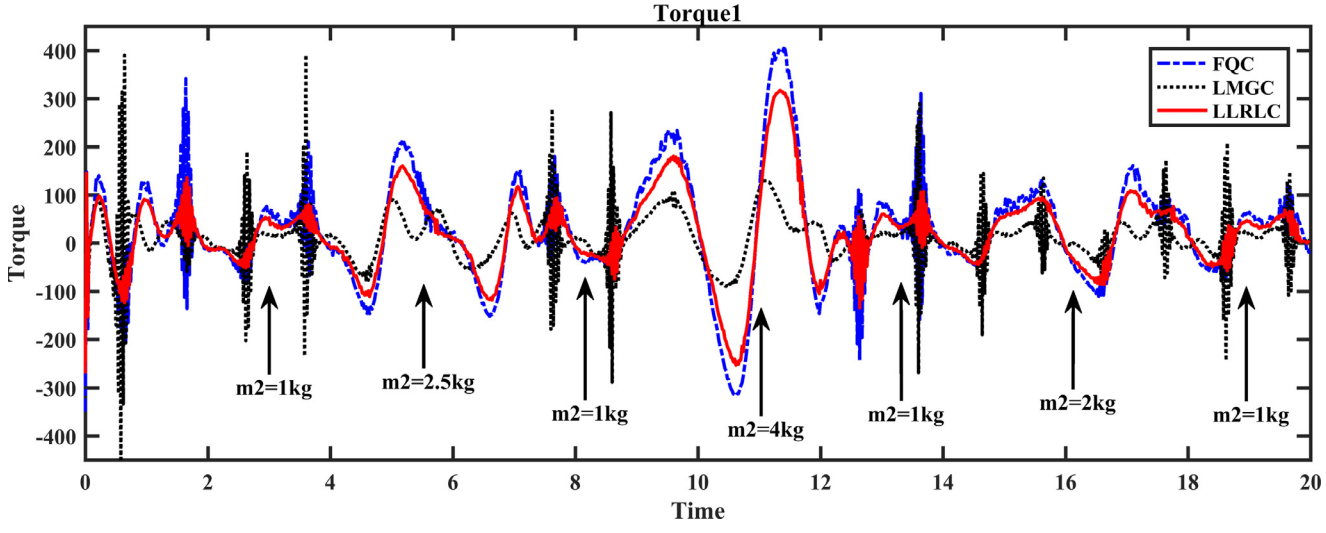


Fig. 4. Torque on link 1 with load variation.

Please cite this article as: A. Kumar, R. Sharma, Linguistic Lyapunov reinforcement learning control for robotic manipulators, Neurocomputing (2017), http://dx.doi.org/10.1016/j.neucom.2017.06.064

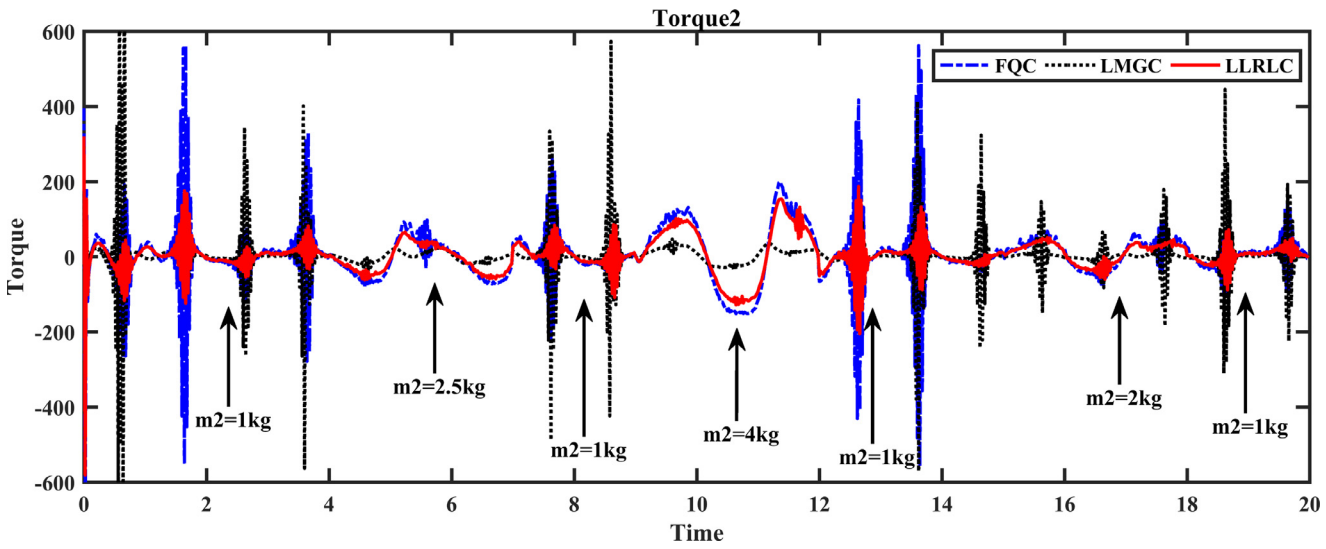


Fig. 5. Torque on link 2 with load variation.

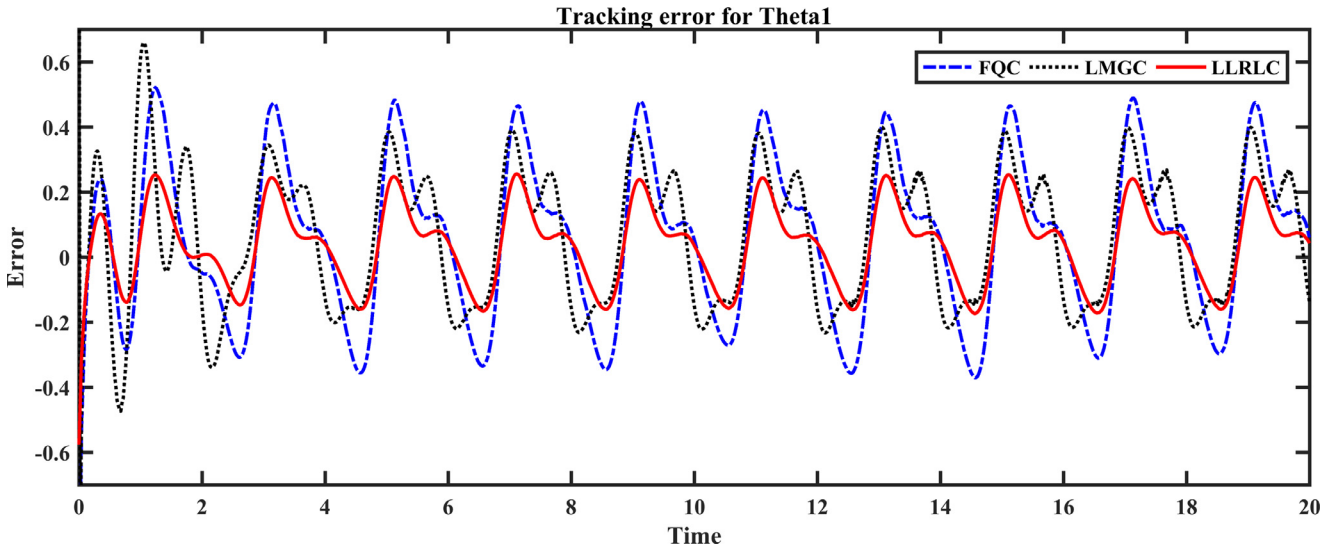


Fig. 6. Tracking error for theta1 without load variation.

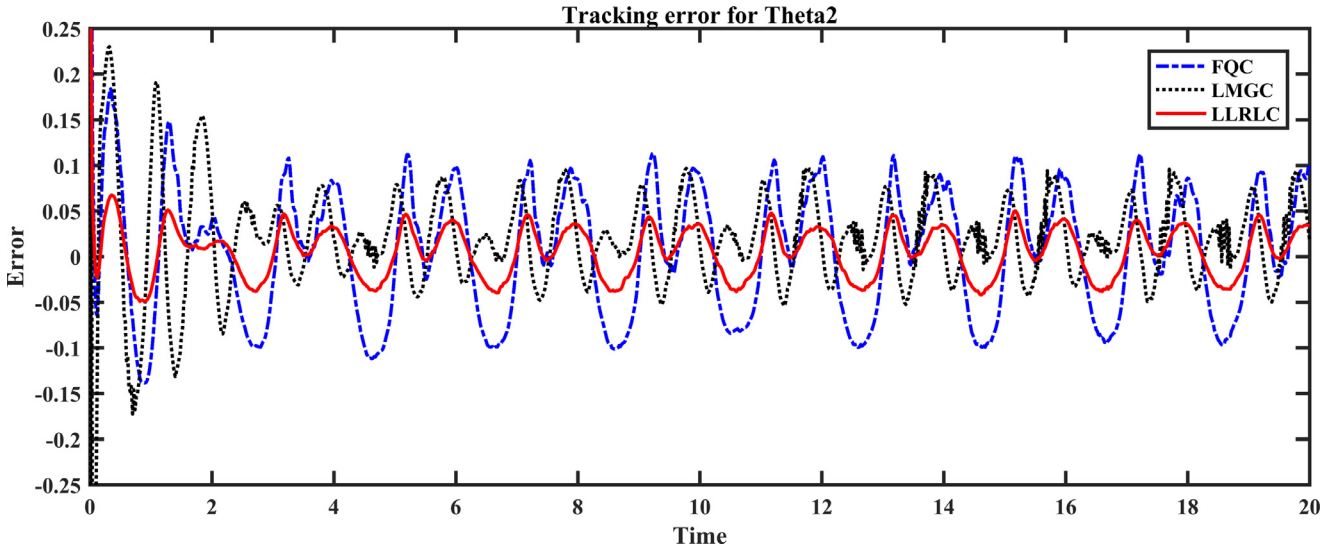


Fig. 7. Tracking error for theta2 without load variation.

Please cite this article as: A. Kumar, R. Sharma, Linguistic Lyapunov reinforcement learning control for robotic manipulators, Neurocomputing (2017), http://dx.doi.org/10.1016/j.neucom.2017.06.064

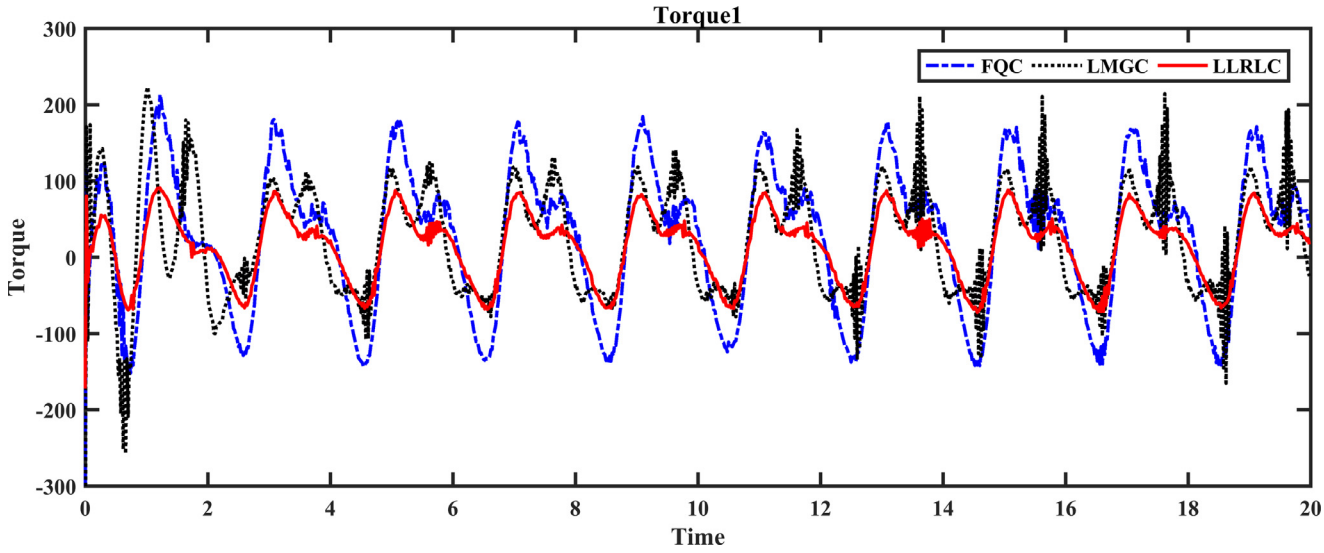


Fig. 8. Torque on link 1 without load variation.

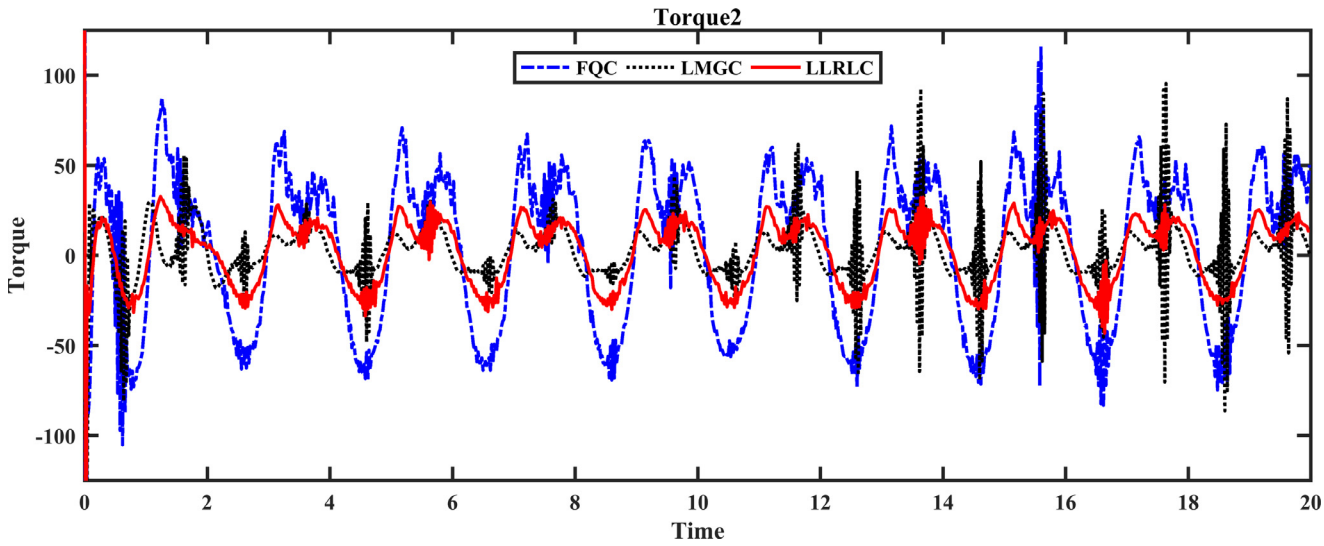


Fig. 9. Torque on link 2 without load variation.

These results show that our controller exhibits a stable and superior performance (in terms of tracking error and control torques) in comparison to the baseline FQL control and gives comparable (somewhat better) performance in comparison to LMGC. It is to be noted that FQC, LMGC and LLRLC are model free RL based controller requiring no a priori system information.

Table 2.
Computational complexity comparison of RL controllers.

	LMGC	FQC	LLRLC
TLM without load variation	165.32	15.02	14.84
TLM with load variation	184.52	17.95	14.99
SCARA without load variation	56.34	5.12	4.91

5.2. Selective compliance assembly robotic arm (SCARA) control

Appendix B gives simulation model and parameters. We simulate SCARA for 6 s using fourth-order Runge–Kutta method with fixed time step of 20 ms. All the controllers have been implemented with same simulation and learning parameters. Here we perform simulation for a pure tracking problem, i.e., with disturbances only. Fig. 10 shows tracking error for link1 wherein we observe that LLRLC achieves best tracking performance in least time, i.e., tracking error becomes quite low early on. From Fig. 11, we observe that LLRLC manages to achieve least tracking error.

For link1 (Fig. 12), we observe that LLRLC exhibits chattering higher than other controllers but has lower torques. For the SCARA (Fig. 13) torque pulsations are largest for LLRLC for link 2. Performance of LLRLC for SCARA is not on expected lines.

Computation time comparison of all the controllers has been used to compare computational complexity of the controllers. Table 2 gives simulation time (in seconds) for conducting 1000 trials for all controllers. It is evident from Table 2 that LLRLC takes less time as compared to FQC and LMGC showing it is computationally superior to both FQC and LMGC. Please note the stark difference in computational time of FQC and LLRLC from that of LMGC. This is because LMGC solves two linear programs; in each RL iteration.

From above results, we observe that LLRLC for SCARA is not as convincing as in the case of two link robotic arm. Lyapunov theory

Please cite this article as: A. Kumar, R. Sharma, Linguistic Lyapunov reinforcement learning control for robotic manipulators, Neurocomputing (2017), http://dx.doi.org/10.1016/j.neucom.2017.06.064

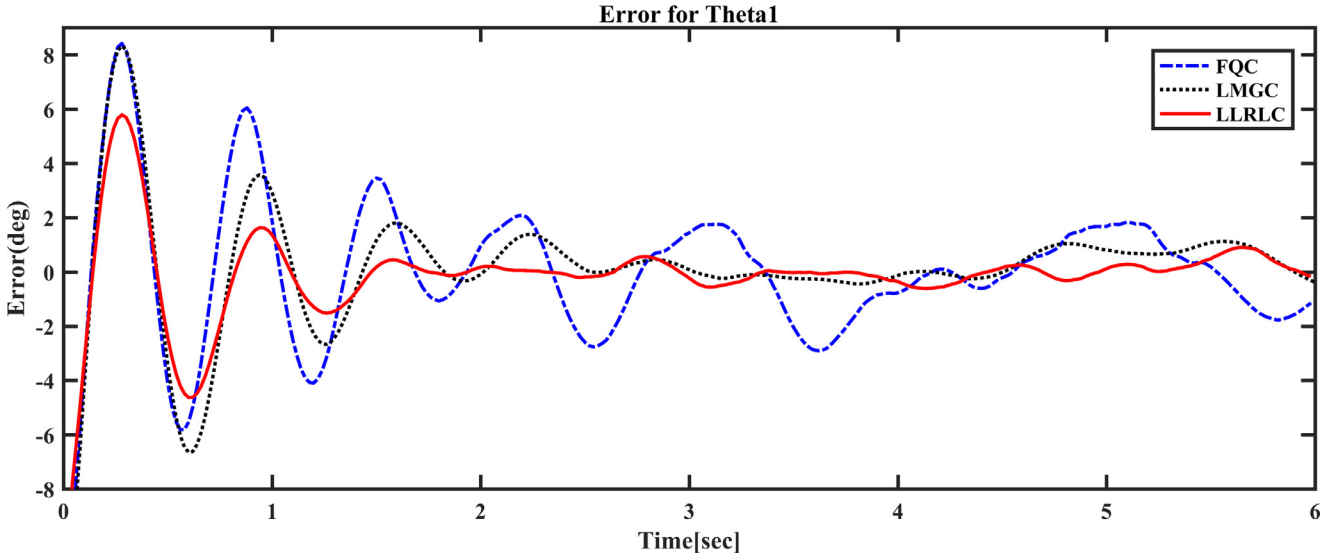


Fig. 10. Tracking error for theta1 for SCARA.

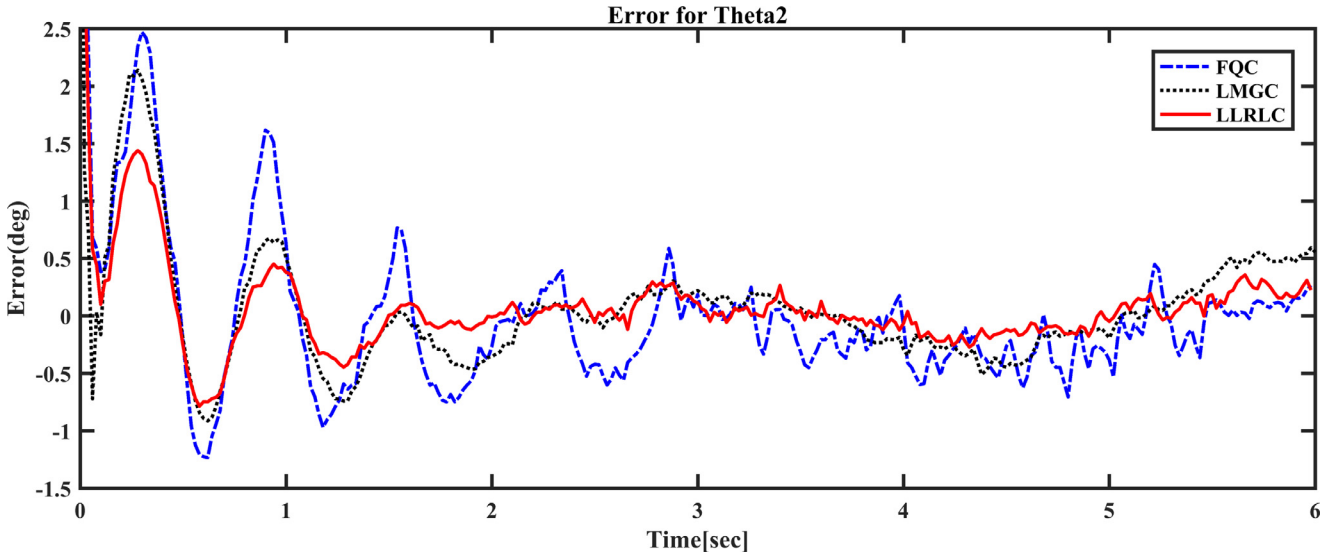


Fig. 11. Tracking error for theta2 for SCARA.

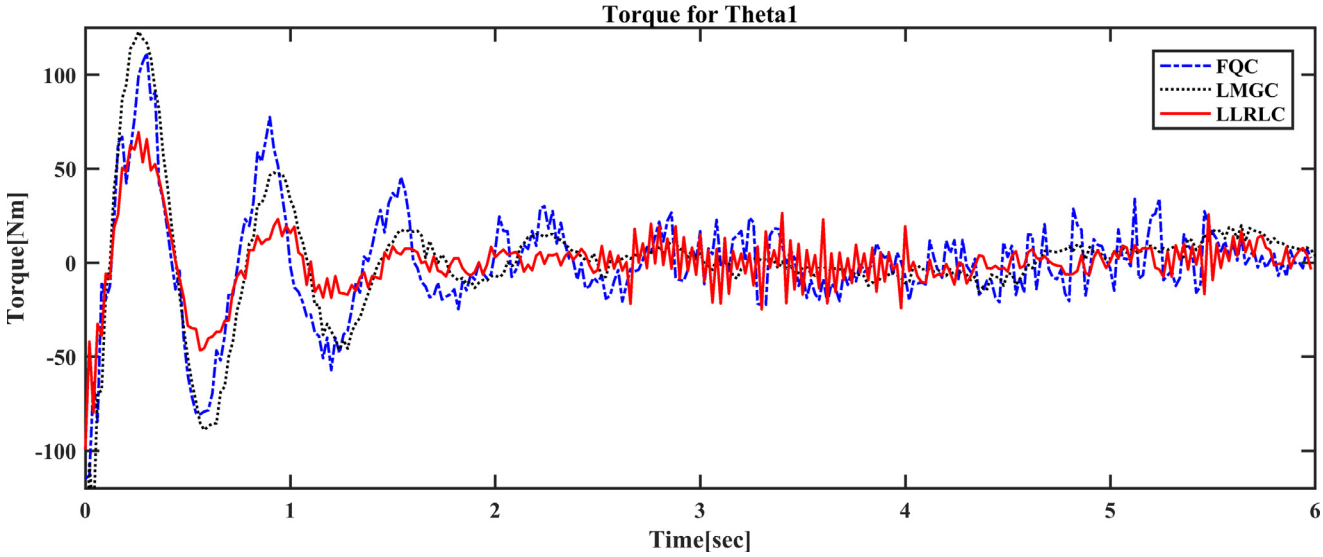


Fig. 12. Torque on link 1.

Please cite this article as: A. Kumar, R. Sharma, Linguistic Lyapunov reinforcement learning control for robotic manipulators, Neurocomputing (2017), http://dx.doi.org/10.1016/j.neucom.2017.06.064

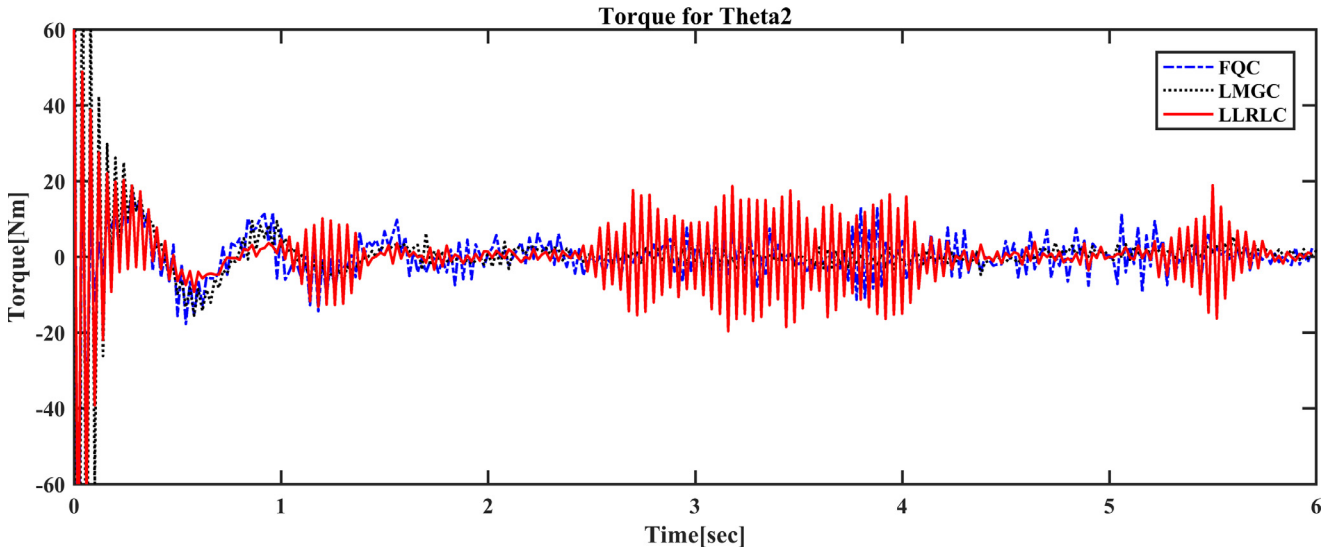


Fig. 13. Torque on link 2.

based FQL control is better than normal FQL control for both the systems in terms of handling disturbances and payload variation for path tracking problem and comparable (in terms of tracking performance) to LMGC. However Our LLRLC outperforms LMGC by a huge margin when it comes to computational burden. Thus even though LLRLC and LMGC draw their stability aspect from Lyapunov theory, LLRLC is better implementation wise.

6. Conclusions and scope for future work

We have proposed a linguistic Lyapunov fuzzy RL controller wherein Lyapunov conditions have been used to guarantee stability. To the best of our knowledge, ours is a first attempt at using Lyapunov dictated rules (rule consequents) in a linguistic RL setting for controlling robotic manipulators. Computational complexity of our proposed approach is at par, if not better, than the baseline FQC (without any stability guarantee). We implement our LLRLC approach on two benchmark manipulator problems: (i) Two link standard robotic arm and (ii) SCARA. Proposed LLRLC controller is found to have superior tracking performance, lower control torque requirement and is computationally less expensive to the FQL and Lyapunov theory based Markov game controllers. In future, we would implement a multi agent version of the proposed approach for applications involving coordination control of multi robot system [9]. Another future lead, which the author is currently working on, is to use LLRLC in conjunction with fuzzy Markov games [1] (or as standalone) to design robust, stable and adaptive controllers with applications in robotics, bio technology and power systems control [15].

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Two-link robot arm control

Fig. 14 gives simulation model for two link robot arm [3]. A two-link robot arm has all the nonlinear effects common to general robot manipulators and is easy to simulate.

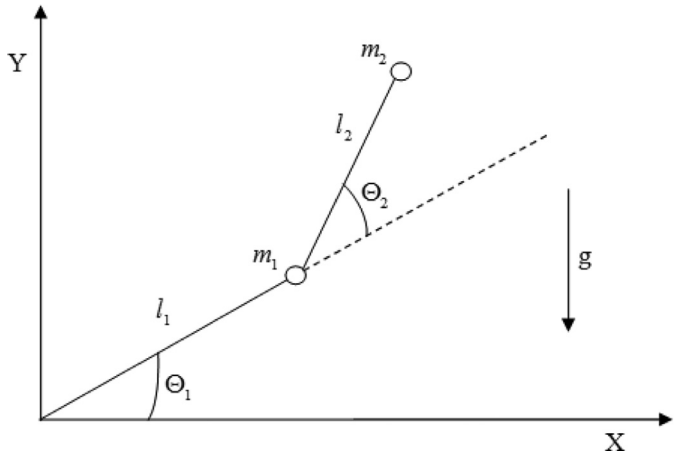


Fig. 14. Two link robot manipulator.

- Manipulator dynamics:

$$\begin{bmatrix} \alpha+\beta+2\eta\cos\Theta(2) & \beta+\eta\cos\Theta(2)\beta \\ \beta+\eta\cos\Theta(2) & \end{bmatrix}\begin{bmatrix} \ddot{\Theta}(1) \\ \ddot{\Theta}(2) \end{bmatrix} + \begin{bmatrix} -\eta\left(2\dot{\Theta}(1)\dot{\Theta}(2)+\dot{\Theta}^2(2)\right)\sin\Theta(2) \\ \eta\dot{\Theta}^2(1)\sin\Theta(2) \end{bmatrix} + \begin{bmatrix} \alpha\ell_1\cos\Theta(1)+\eta\ell_1\cos(\Theta(1)+\Theta(2)) \\ \eta\ell_1\cos(\Theta(1)+\Theta(2)) \end{bmatrix} = \begin{bmatrix} \tau(1) \\ \tau(2) \end{bmatrix} \quad (A.1)$$

$$\alpha=(m_1+m_2)a_1^2,\ \beta=m_2a_2^2,\ \eta=m_2a_1a_2,\ \ell_1=g/a_1 \quad (A.2)$$

and robot parameters are $a_1=a_2=1$ m, $m_1=m_2=1$ kg.
- Desired trajectory: $\Theta_d(1)=\sin(t)$, $\Theta_d(2)=\cos(t)$.
- System state-space is continuous and has four variables, i.e., $x(k)=(\Theta(1),\Theta(2),\dot{\Theta}(1),\dot{\Theta}(2))$.

Appendix B. SCARA robot control

Fig. 15 shows a 2 DOF SCARA with two parallel revolute joints and two links.

Dynamic model and parameters are taken from [3].

Please cite this article as: A. Kumar, R. Sharma, Linguistic Lyapunov reinforcement learning control for robotic manipulators, Neurocomputing (2017), http://dx.doi.org/10.1016/j.neucom.2017.06.064

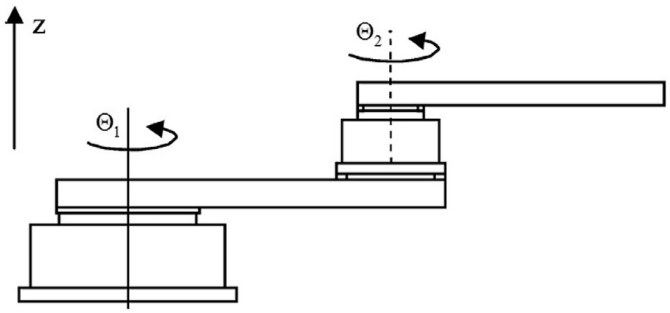


Fig. 15. 2 DOF SCARA.

- Dynamical model and parameters:

$$\begin{bmatrix} p_1 + p_2c_2 & p_3 + 0.5p_2c_2 \\ p_3 + 0.5p_2c_2 & p_3 \end{bmatrix} \begin{bmatrix} \ddot{\Theta}(1) \\ \ddot{\Theta}(2) \end{bmatrix} + \begin{bmatrix} \left(p_4 - p_2s_2\dot{\Theta}_2\right)\dot{\Theta}_2 - 0.5p_2s_2\dot{\Theta}_2 \\ 0.5p_2s_2\dot{\Theta}_1 + p_5\dot{\Theta}_2 \end{bmatrix} = \begin{bmatrix} \tau(1) \\ \tau(2) \end{bmatrix} \quad \text{(B.1)}$$

$$p_1 = 5.0,\ p_2 = 0.9,\ p_3 = 0.3,\ p_4 = 3.0,\ p_5 = 0.6,\ s_2 = \sin(\Theta(2)),\ c_2 = \cos(\Theta(2)) \quad \text{(B.2)}$$

Desired trajectory:

$$\Theta_d(1) = 0.5 + 0.2(\sin t + \sin 2t) rad$$
$$\Theta_d(2) = 0.13 - 0.1(\sin t + \sin 2t) rad$$

Control task is to apply a sequence of torques $\tau = [\tau_1 \tau_2]$ at each joint so that the joints follow a prescribed trajectory in joint space, i.e. $\Theta_d = [\Theta_d(1)\Theta_d(2)]$.

Appendix C. Torque and link angle derivation

I. For Two link robot manipulator

Taking $\alpha + \beta + 2\eta\cos\Theta(2) = k(1)$, $\beta + \eta\cos\Theta(2) = k(2)$, $-\eta(2\dot{\Theta}(1)\dot{\Theta}(2) + \dot{\Theta}^2(2))\sin\Theta(2) + \alpha\ell_1\cos\Theta(1) + \eta\ell_1\cos(\Theta(1) + \Theta(2)) = k(3)$ and $\eta\dot{\Theta}^2(1)\sin\Theta(2) + \eta\ell_1\cos(\Theta(1) + \Theta(2)) = k(4)$

(A.1) can be written as

$$\begin{bmatrix} k(1)k(2) \\ k(2)\beta \end{bmatrix} \begin{bmatrix} \ddot{\Theta}(1) \\ \ddot{\Theta}(2) \end{bmatrix} + \begin{bmatrix} k(3) \\ k(4) \end{bmatrix} = \begin{bmatrix} \tau(1) \\ \tau(2) \end{bmatrix}$$

where $k(1) > k(2) > \beta > 0$

$$\Rightarrow k(1)\ddot{\Theta}(1) + k(2)\ddot{\Theta}(2) + k(3) = \tau(1) \quad \text{(C.1)}$$

$$k(2)\ddot{\Theta}(1) + \beta\ddot{\Theta}(2) + k(4) = \tau(2) \quad \text{(C.2)}$$

$$\Rightarrow \ddot{\Theta}(1) = \frac{\tau(2) - \beta\ddot{\Theta}(2) - k(4)}{k(2)} \quad \text{(C.3)}$$

from C.3 and C.1, we have

$$k(1)\left(\frac{\tau(2) - \beta\ddot{\Theta}(2) - k(4)}{k(2)}\right) + k(2)\ddot{\Theta}(2) + k(3) = \tau(1)$$

$$\Rightarrow k(1)\tau(2) - k(1)\beta\ddot{\Theta}(2) - k(1)k(4) + k(2)^2\ddot{\Theta}(2) + k(2)k(3) = k(2\tau)(1)$$

$$\Rightarrow \left(k(2)^2 - k(1)\beta\right)\ddot{\Theta}(2) = k(2)\tau(1) - k(1)\tau(2) - k(2)k(3) + k(1)k(4)$$

Now, $\left(k(2)^2 - k(1)\beta\right) = (\beta + 2\eta\cos\Theta(2))^2 - (\alpha + \beta + 2\eta\cos\Theta(2))\beta$

$= \eta^2\cos^2\Theta(2) - \alpha\beta$

$= (m_2a_1a_2)^2\cos^2\Theta(2) - (m_1 + m_2)a_1^2m_2a_2^2$ (Using (A.2))

$= m_2a_1{}^2a_2{}^2\left(m_2\cos^2\Theta(2) - (m_1 + m_2)\right) < 0$

$\Rightarrow (-ve)\ddot{\Theta}(2) = k(2)\tau(1) - k(1)\tau(2) - k(2)k(3) + k(1)k(4)$

$\Rightarrow \ddot{\Theta}(2)$ depends directly on $\tau(2) - \tau(1)$

Also from C.2 we have

$$\ddot{\Theta}(2) = \frac{\tau(2) - k(2)\ddot{\Theta}(1) - k(4)}{\beta} \quad \text{(C.4)}$$

from C.1 and C.4

$$k(1)\ddot{\Theta}(1) + k(2)\left(\frac{\tau(2) - k(2)\ddot{\Theta}(1) - k(4)}{\beta}\right) + k(3) = \tau(1)$$

$$\Rightarrow (\beta k(1) - k(2)^2)\ddot{\Theta}(1) = \beta\tau(1) - k(2)\tau(2) + k(2)k(4) - k(3)\beta$$

As shown earlier $\beta k(1) - k(2)^2 > 0$

$$\Rightarrow (+ve)\ddot{\Theta}(1) = \beta\tau(1) - k(2)\tau(2) + k(2)k(4) - k(3)\beta$$

$\Rightarrow \ddot{\Theta}(1)$ depends directly on $\tau(1) - \tau(2)$

II. For SCARA:

Taking $p_1 + p_2c_2 = k(1)$, $p_3 + 0.5p_2c_2 = k(2)$, $p_3 = k(3)$ and $(p_4 - p_2s_2\dot{\Theta}_2)\dot{\Theta}_2 - 0.5p_2s_2\dot{\Theta}_2 = k(4)$, $0.5p_2s_2\dot{\Theta}_1 + p_5\dot{\Theta}_2 = k(5)$

Then (B.1) can be written as

$$\begin{bmatrix} k(1)k(2) \\ k(2)k(3) \end{bmatrix} \begin{bmatrix} \ddot{\Theta}(1) \\ \ddot{\Theta}(2) \end{bmatrix} + \begin{bmatrix} k(4) \\ k(5) \end{bmatrix} = \begin{bmatrix} \tau(1) \\ \tau(2) \end{bmatrix}$$

Next, using steps outlined for the two link robotic manipulator, we have

$$\left(k(2)^2 - k(1)k(3)\right)\ddot{\Theta}(2) = k(2)\tau(1) - k(1)\tau(2) - k(2)k(4) + k(1)k(5)$$

$$\text{Now } k(2)^2 - k(1)k(3) = (p_3 + 0.5p_2c_2)^2 - (p_1 + p_2c_2)p_3 = p_3(p_3 - p_1) + 0.25p_2{}^2c_2{}^2 = \text{Neg_big} + \text{Pos_small} < 0$$

$$\Rightarrow (-ve)\ddot{\Theta}(2) = k(2)\tau(1) - k(1)\tau(2) - k(2)k(4) + k(1)k(5)$$

$$\Rightarrow \ddot{\Theta}(2) = -k(2)\tau(1) + k(1)\tau(2) + k(2)k(4) - k(1)k(5)$$

$\Rightarrow \ddot{\Theta}(2)$ depends directly on $\tau(2) - \tau(1)$

Similarly $\ddot{\Theta}(1)$ depends directly on $\tau(1) - \tau(2)$

References

[1] R. Sharma, M. Gopal, A Markov game-adaptive fuzzy controller for robot manipulators, IEEE Trans. Fuzzy Syst. 16 (1) (2008) 171–186.
[2] M. Mohri, A. Rostamizadeh, A. Talwalkar, Foundations of Machine Learning, MIT Press, Cambridge, MA, USA, 2012.
[3] R. Sharma, Lyapunov theory based stable Markov game fuzzy control for non–linear systems, Eng. Appl. Artif. Intell. 55 (2016) 119–127.
[4] A. Kumar, R. Sharma, A stable Lyapunov constrained reinforcement learning based neural controller for nonlinear systems, in: Proceedings of International Conference on Computing, Communication and Automation (ICCCA), 2015, pp. 185–189.
[5] S. Bhasin, N. Sharma, P. Patre, W. Dixon, Asymptotic tracking by a reinforcement learning-based adaptive critic controller, J. Control Theory Appl. 9 (3) (2011) 400–409.
[6] L. Tang, Y.-J. Liu, S. Tong, Adaptive neural contro lusing reinforcement learning for a class of robot manipulator, Neural Comput. Appl. 25 (2013) 135–141.

Please cite this article as: A. Kumar, R. Sharma, Linguistic Lyapunov reinforcement learning control for robotic manipulators, Neurocomputing (2017), http://dx.doi.org/10.1016/j.neucom.2017.06.064

[7] D. Zhao, W. Ni, Q. Zhu, A framework of neural networks based consensus control for multiple robotic manipulators, Neurocomputing 140 (2014) 8–18.
[8] B. Daachi, T. Madani, A. Benallegue, Adaptive neural controller for redundant robot manipulators and collision avoidance with mobile obstacles, Neurocomputing 79 (2012) 50–60.
[9] Y. Li, L. Chen, K.P. Tee, Q. Li, Reinforcement learning control for coordinated manipulation of multi-robots, Neurocomputing 170 (2015) 168–175.
[10] R. Kamalapurkar, P. Walters, W.E. Dixon, Model-based reinforcement learning for approximate optimal regulation, Automatica 64 (2016) 94–104.
[11] Y. Jiang, C. Yang, H. Ma, A review of fuzzy logic and neural network based intelligent control design for discrete-time systems, Discrete Dyn. Nat. Soc. (2016), Hindawi, http://dx.doi.org/10.1155/2016/7217364.
[12] S. Qiu, Z. Li, W. He, L. Zhang, C. Yang, C. Su, Brain–machine interface and visual compressive sensing-based teleoperation control of an exoskeleton robot, IEEE Trans. Fuzzy Systems 25 (1) (2017) 58–69.
[13] R.S. Sutton, A.G. Barto, Reinforcement Learning: An Introduction, 2nd ed., The MIT Press, Cambridge, 2012 (in progress).
[14] M. Margaliot, G. Langholz, Fuzzy Lyapunov-based approach to the design of fuzzy controller, Fuzzy Sets Syst. 106 (1999) 49–59.
[15] D. Ernst, M. Glavic, L. Wehenkel, Power systems stability control: reinforcement learning framework, IEEE Trans. Power Syst. 19 (1) (2004) 427–435.



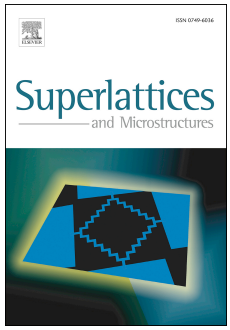
Abhishek Kumar did his B.Tech. in Electronics & Communication Engineering from BBDNITM, UPTU, Lucknow and M.E. in Power Engineering from Jadavpur University, Kolkata. He is doing his Ph.D. in Instrumentation & Control Engineering from Netaji Subhas Institute of Technology, Delhi University, Delhi.



Rajneesh Sharma did his B.E. in Electrical Engineering from Delhi College of Engineering, Delhi University and M.E. in Control and Instrumentation from Delhi College of Engineering, Delhi University. He did his Ph.D. in Intelligent Control of Non Linear Systems from Indian Institute of Technology, Delhi. Thereafter Dr. Sharma carried out Post-Doctoral Research at Institute for Systems and Robotics, IST, Lisbon, Portugal for one year. Dr. Sharma has worked in BHEL, Department of Telecommunications as IES and Indian Railways as IES officer prior to joining NSIT in 2001. He has published about 40 research papers in reputed refereed journals and conferences.

Please cite this article as: A. Kumar, R. Sharma, Linguistic Lyapunov reinforcement learning control for robotic manipulators, Neurocomputing (2017), http://dx.doi.org/10.1016/j.neucom.2017.06.064

Accepted Manuscript



$LiTaO_3$ based metamaterial perfect absorber for terahertz spectrum

Nishant Shankhwar, Yogita Kalra, Ravindra Kumar Sinha

PII: S0749-6036(17)31448-9

DOI: 10.1016/j.spmi.2017.07.034

Reference: YSPMI 5144

To appear in: *Superlattices and Microstructures*

Received Date: 13 June 2017

Accepted Date: 11 July 2017

Please cite this article as: N. Shankhwar, Y. Kalra, R.K. Sinha, $LiTaO_3$ based metamaterial perfect absorber for terahertz spectrum, *Superlattices and Microstructures* (2017), doi: 10.1016/j.spmi.2017.07.034.

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

$LiTaO_3$ based metamaterial perfect absorber for terahertz spectrum

Nishant Shankhwar[1], Yogita Kalra[1] and Ravindra Kumar Sinha[1,2]

[1] *TIFAC-Centre of Relevance and Excellence in Fiber Optics and Optical Communication, Department of Applied Physics, Delhi Technological University, Bawana Road, Delhi-110042, India*

[2] *CSIR-Central Scientific Instruments Organisation (CSIO), Chandigarh-160030, India*

Abstract

In this article a new design of metamaterial perfect absorber (MPA) has been proposed which comprises a periodic array of $LiTaO_3$ microcylinders grown over a silver substrate. The fundamental principle of the structure is the occurrence of Mie resonance in the cylinders due to high value of refractive index ($\sim$10) of $LiTaO_3$ in the wavelength range 65-100 μm. This remarkably high refractive index is a manifestation of polaritonic resonance of $LiTaO_3$ in the considered wavelength range. The proposed structure shows high absorption coefficient in the said wavelength range with a maximum of 99.9% around 88 μm.

Keywords: Terahertz, Metamaterial, Perfect absorber

1. Introduction

Mie resonance based metamaterials have been deeply explored in the past few years [1, 2]. Particles of high permittivity are suitably arranged to achieve desired functionality at the desired wavelength. Material of the particles is chosen depending upon the wavelength of operation [3, 4]. To achieve Mie resonance at telecom wavelength and in near infrared spectrum silicon is the most common choice [5, 6, 7, 8, 9, 10]. In terahertz spectrum, materials like lithium tantalate ($LiTaO_3$) [11, 12], silicon carbide (SiC) [13, 14],

Email address: dryogitakalragmail.com ()

polycrystalline titanium oxide (poly-TiO_2) [15], cuprous chloride (CuCl), are more suitable. Whereas particles made up of barium strontium titanate ($Ba_{0.5}Sr_{0.5}TiO_3$, BST) have been reported to show Mie resonance in microwave spectrum [1]. Dielectric metamaterials and metasurfaces designed and fabricated so far have a broad range of interesting applications including - negative refraction [12], controlling the wavefront [16, 17], electromagentic transparency [18, 19], electromagnetic cloaking [20, 21], perfect reflection [22, 8], perfect absorption [22, 23], subwavelength confinement [24], realization of zero index [7], etc. Moreover, highly directive all dielectric nanoantennas [6, 25, 26] including Yagi-Uda nanonantena, have aslo been reported in the past few years.

Out of all these potential applications, the property of perfect absorption has been addressed in this article. Metamaterial perfect absorbers (MPAs) have a large number of useful applications such as heating a bimaterial cantilever in MEMS for terahertz wave imaging [27], development of spatial light modulators [28], frequency selective filtering, improving the effciency of photovoltaics [29], etc. A large number of MPAs operating in microwave, visible and infrared spectrum have been reported in the past few years. Most of them are patterned metallic structures separated from a continuous metallic plane by a thin dielectric spacer layer [30, 31, 32, 33, 34]. Some other works have reported BST based dielectric MPAs whose fundamental principle of operation is occurrence of Mie resonance and they work in microwave spec-

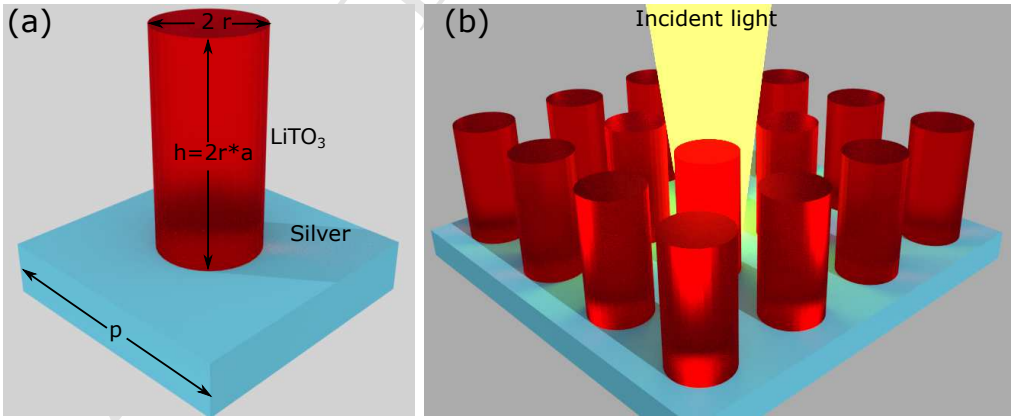


Figure 1: (a) Unit cell of the proposed metamaterial structure, (b) array of the $LiTaO_3$ cylinders over silver substrate.

trum [23]. In this article a $LiTaO_3$ based metamaterial structure has been proposed which shows high absorption in the range 65 μm - 100 μm (terahertz spectrum) with maximum absorption of 99.9% at 88 μm. The proposed structure is shown in figure 1. Figure 1(a) shows the unit cell of the MPA in which a cylinder has been shown resting on silver substrate. Figure 1(b) shows periodic array of the particles which behaves as a perfect absorber.

2. Empirical relation to determine resonant wavelength

The high absorbance of the proposed metamaterial structure is attributed to Mie resonance occurring in the particles. For which, a material having high value of refractive index in terahertz spectrum is needed. $LiTaO_3$ is one such material, whose refractive index is high ($\sim$10) for wavelength lying between 60 μm and 100μm . Moreover, its extinction coefficient is high as well which acts as an aid to the absorptive properties of the proposed structure. Analytical formulae for the Mie resonance in a dielectric sphere are well known [3, 4], but in this work cylindrical particles have been chosen for which no analytical formulae exist. Hence, a need was felt to establish an empirical relation which can express the resonant wavelength as a function of geometrical parameters of the cylinder within the above mentioned wavelength range. For this purpose, dependence of resonant wavelength on the radius and aspect ratio (ratio of height to diameter) was examined individually, by varying just one at a time and keeping the other a constant. Based on the obtained results, an empirical formula has been established to express the resonant wavelength as a function of radius and aspect ratio of the cylindrical particle, which is shown below as

$$\lambda_0 = -7.83erf(a)r + 16.18r + 56.97erf(a) \tag{1}$$

where λ_0 is the resonant wavelength, a is the aspect ratio, r is the radius and $erf()$ stands for error function. Figure 2(a) & (b) show the dependence of resonant wavelength on radius and aspect ratio of the cylinder respectively. It is visible that empirical formula fits well with the computaional data. It should be noted that, though the refractive index is not visible in the above formula, it is intrinsically included in it for material dispersion was included by means of a user define function during computation in COMSOL. The data for refractive index has been taken from Palik [35].

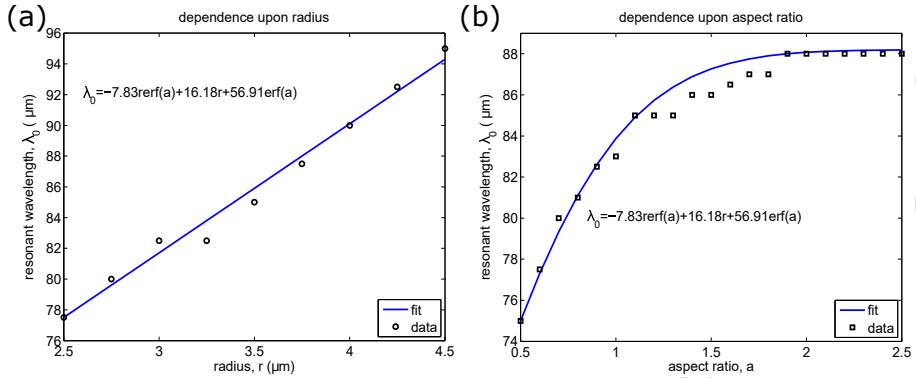


Figure 2: (a) Dependence of resonant wavelength at the radius and (b) dependence of the resonant wavelength on aspect ratio. The blue solid curve represents equation 1 and markers computational data

3. Design and numerical analysis

3.1. Cylinders

In this work, radius r and aspect ratio a have been chosen to be 3.75 μm and 2.0 respectively. These values have been such chosen that resonant wavelength lies between 80 and 90 μm. On substituting these values in equation 1, the corresponding resonant wavelength which is also the wavelength of maximum absorption, comes out to be 88 μm. This is confirmed by figure 3(a) in

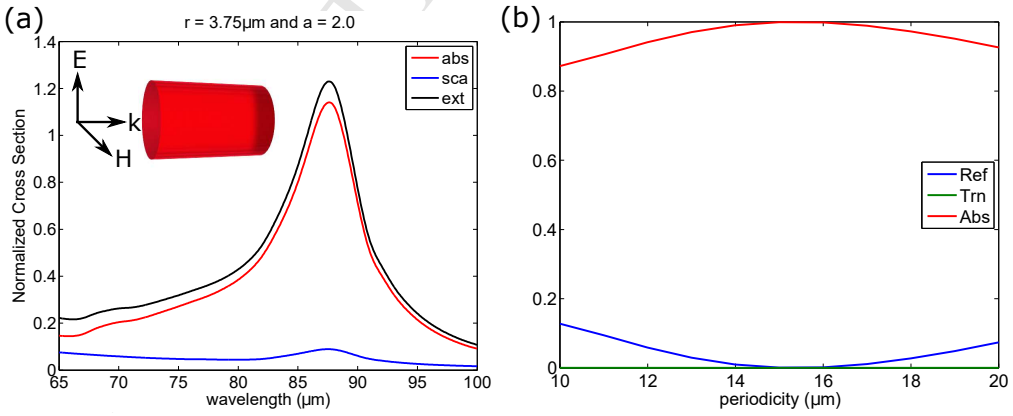


Figure 3: Optimization graph: variation of absorbance with periodicity of the array. Absorbance is maximum for periodicity $p = 15$ μm

ACCEPTED MANUSCRIPT

which extinction cross-section for a cylindrical particle of radius $r = 3.75\ \mu m$ and aspect ratio $a = 2.0$ is maximum at 88 μm. Furthermore, the structure has been optimized with respect to periodicity p in order to determine the value for which absorbance is maximum. The optimization graph has been shown in figure 3(b), from which it has been inferred that periodicity should be 15 μm.

3.2. Substrate

In case of perfect absorbers, role of the substrate if very crucial. It is a basic fact that absorbance A is given as $1 - R - T$ where R is the reflectance and T is the transmittance. Since the cylinders are made up a dielectric, a large amount of light is transmitted and only small amounts are reflected and absorbed as shown in figure 4(a). The small absorption peak seen in figure 4(a) is attributed to high extinction coefficient of $LiTaO_3$ resulting in absorption of light within the cylinders. The purpose of the substrate is to suppress transmittance to such negligible levels that $A \sim 1 - R$. To achieve this, the material of the substrate should be a metal and its thickness should be very large compared to the penetration depth of electromagnetic waves. Penetration depth δ of the incident electromagnetic wave of a particular frequency into a metal is given by [36],

$$\delta = \sqrt{\frac{2}{\omega\mu\sigma}} \tag{2}$$

where, ω is the frequency of electromagnetic wave, μ is the relative permeability and σ is the conductivity of the metal. For $\omega = 2.14 \times 10^{13}$ rad/s ($\equiv 88\ \mu m$ wavelength), $\mu = 1$ & $\sigma = 6.3 \times 10^7$ S/m, penetration depth $\delta = 0.385\ \mathring{A}$. In the proposed structure, the material of the substrate is silver and its thickness has been kept 1 μm which, compared to the penetration depth, is astronomical. Addition of an optically thick (thickness $\gg$ penetration depth) metallic substrate not just suppresses transmission but also enhances absorption, as shown in figure 4(b). Its should be noted that in place of silver, other metals like gold, copper and aluminium can also be used. Hence, if one intends to fabricate the proposed design, the metal which is compatible with the adopted fabrication technique should be chosen.

4. Results and Discussion

Scattering and absorption curves shown in figure 3(a) illustrate that absorption dominates scattering. When the field gets localized inside the cylin-

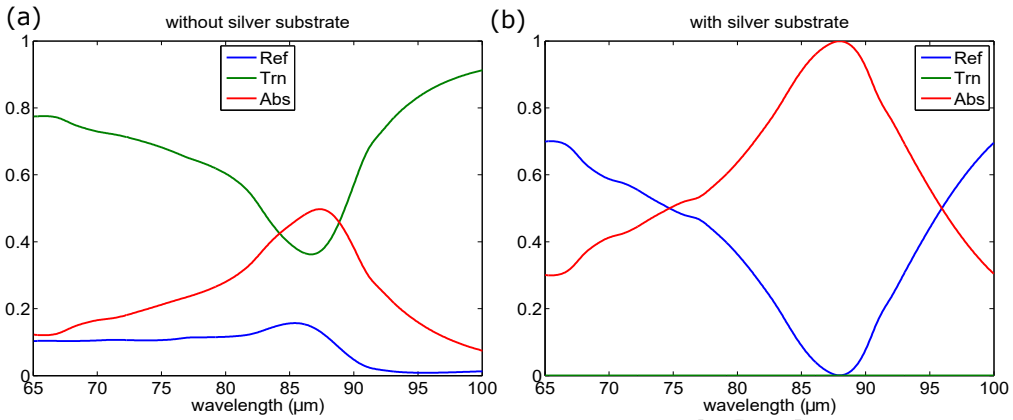


Figure 4: (a) Reflectance, transmittance and absorbance of an array of $LiTaO_3$ cylinders without including silver substrate. A small absorption peak shows that $LiTaO_3$ itself absorbs some amount of light in the absence of sliver substrate. (b) Reflectance, transmittance and absorbance of an array of $LiTaO_3$ cylinders with silver substrate. Metallic substrate suppressed transmission and enhanced absorption

der due to resonance, a significant amount of it gets absorbed by $LiTaO_3$ itself owing to high value of extinction coefficient of the material. To prove this, results of numerical analysis performed in the absence of the substrate have been plotted in figure 4(a), in which a small peak is visible in the absorption curve with $A \sim 0.5$ at 88 μm. However, 50% absorption it not good enough to entitle the structure as a perfect absorber. Reinforcement is provided by the metallic substrate which reduces transmission to almost zero and increases absorption upto 99.9% at 88 μm as shown in figure 4(b).

A perfect absorber is not just the one which absorbs the desired wavelength almost completely, but it should also be able to absorb the wavelength at large angles of incidence. To test this ability of the proposed structure absorbance has been calculated at different angles of incidence and results have been plotted in figure 5. It shows that the absorbance $A \geq 0.9$ upto 55^o angle of incidence. Hence, the structure qualifies to be considered as a broad-angle metamaterial perfect absorber.

Finally, the field and power distribution inside the metamaterial unit cell at resonance have been illustrated in figure 6. Figure 6(a) shows the distribution of y-component magnetic field and the arrows show rotating electric field vector which confirm the magnetic resonance occurring at 88 μm. Figure 6(b) shows power loss density inside the cylinder.

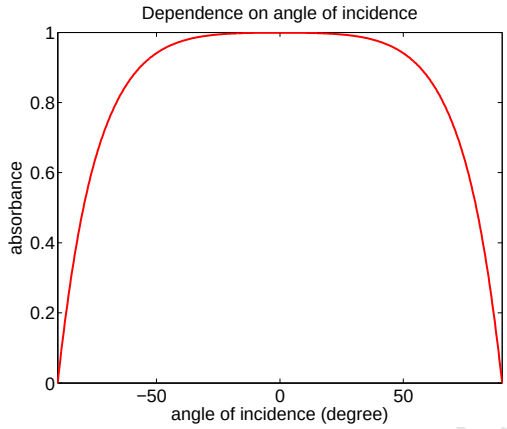


Figure 5: Variation of absorbance with respect to angle of incidence. Absorbance remains greater that 0.9 upto 55^o

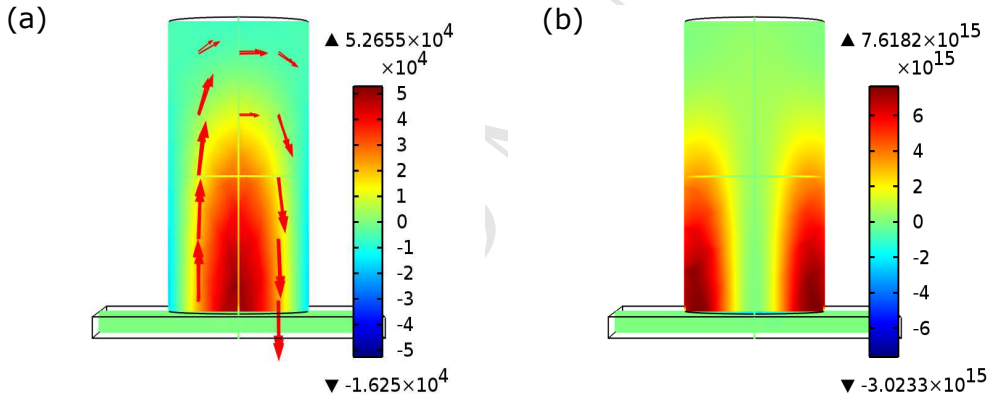


Figure 6: (a) Distribution y component of magnetic field (color contours) in A/m and rotating electric field (arrows) in xz plane; (b) Distribution of power loss density in W/m^3, at 88 μm

5. Conclusion

In light of the above analysis it can be concluded that the proposed structure is a broad-angle metamaterial perfect asbsorber which exhibit high absorbance in desired wavelength range for a broad range of angle of incidence. The MPA operates in terahertz spectrum and absorbs 99.9% of incident radiation at 88 μm ($\equiv$ 3.4 THz). Hence, on being integrated with a bimaterial cantilever focal plane array (FPA), it can be efficiently used for Terahertz

ACCEPTED MANUSCRIPT

wave imaging/sensing. Some other useful application include frequency selective filtering and spatial light modulation.

Acknowledgment

The authors gratefully acknowledge (i) the initiatives and support provided by establishment of the TIFAC-Centre of Relevance and Excellence in Fiber Optics and Optical Communication at Delhi Technological University (Formerly Delhi College of Engineering) Delhi, through the "Mission REACH" program of Technology Vision-2020 of the Government of India, (ii) support through DST-RMES (Indo-Russian) joint research project on "All-Dielectric, Plasmonic and Hybrid Photonic Nanostructures", (iii) support through DST-RFBR (Indo-Russian) joint research project entitled "From Plasmonic to Dielectric and Hybrid Nanoantenna: Novel approches to control electromagnetic waves and light". One of the authors (NS) acknowledges the support through Junior Research Fellowship awarded by University Grants Commision (India) under UGC-NET-JRF scheme. This work was supported by the Ministry of Education and Science of the Russian Federation (project 14.584.21.0009 with unique identificator RFMEFI58414X0009).

[1] Q. Zhao, L. Kang, B. Du, H. Zhao, Q. Xie, X. Huang, B. Li, J. Zhou, L. Li, Experimental demonstration of isotropic negative permeability in a three-dimensional dielectric composite, Physical review letters 101 (2) (2008) 027402.

[2] Q. Zhao, J. Zhou, F. Zhang, D. Lippens, Mie resonance-based dielectric metamaterials, Materials Today 12 (12) (2009) 60–69.

[3] L. Lewin, The electrical constants of a material loaded with spherical particles, Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering 94 (27) (1947) 65–68.

[4] C. F. Bohren, D. R. Huffman, Absorption and scattering of light by small particles, John Wiley & Sons, 2008.

[5] A. García-Etxarri, R. Gómez-Medina, L. S. Froufe-Pérez, C. López, L. Chantada, F. Scheffold, J. Aizpurua, M. Nieto-Vesperinas, J. J. Sáenz, Strong magnetic response of submicron silicon particles in the infrared, Optics express 19 (6) (2011) 4815–4826.

ACCEPTED MANUSCRIPT

[6] P. A. Belov, A. Krasnok, A. Miroshnichenko, C. R. Simovski, Y. Kivshar, All-dielectric nanoantennas, in: Integrated Photonics Research, Silicon and Nanophotonics, Optical Society of America, 2013, pp. IM1B–1.

[7] P. Moitra, Y. Yang, Z. Anderson, I. I. Kravchenko, D. P. Briggs, J. Valentine, Realization of an all-dielectric zero-index optical metamaterial, Nature Photonics 7 (10) (2013) 791–795.

[8] P. Moitra, B. A. Slovick, Z. Gang Yu, S. Krishnamurthy, J. Valentine, Experimental demonstration of a broadband all-dielectric metamaterial perfect reflector, Applied Physics Letters 104 (17) (2014) 171102.

[9] N. Shankhwar, R. K. Sinha, Y. Kalra, S. Makarov, A. Krasnok, P. Belov, High-quality laser cavity based on all-dielectric metasurfaces, Photonics and Nanostructures-Fundamentals and Applications 24 (2017) 18–23.

[10] I. Staude, J. Schilling, Metamaterial-inspired silicon nanophotonics, Nature Photonics 11 (5) (2017) 274–284.

[11] M. S. Wheeler, J. S. Aitchison, M. Mojahedi, Three-dimensional array of dielectric spheres with an isotropic negative permeability at infrared frequencies, Physical Review B 72 (19) (2005) 193103.

[12] M. S. Wheeler, J. S. Aitchison, M. Mojahedi, Coated nonmagnetic spheres with a negative index of refraction at infrared frequencies, Physical Review B 73 (4) (2006) 045105.

[13] J. A. Schuller, R. Zia, T. Taubner, M. L. Brongersma, Dielectric metamaterials based on electric and magnetic resonances of silicon carbide particles, Physical review letters 99 (10) (2007) 107401.

[14] W. Cai, U. K. Chettiar, A. V. Kildishev, V. M. Shalaev, Designs for optical cloaking with high-order transformations, Optics Express 16 (8) (2008) 5444–5452.

[15] K. Shibuya, K. Takano, N. Matsumoto, K. Izumi, H. Miyazaki, Y. Jimba, M. Hangyo, Terahertz metamaterials composed of tio2 cube arrays, in: Proceedings of the 2nd International Congress on Advanced Electromagnetic Materials in Microwaves and Optics, Pamplona, Spain, 2008.

ACCEPTED MANUSCRIPT

[16] M. I. Shalaev, J. Sun, A. Tsukernik, A. Pandey, K. Nikolskiy, N. M. Litchinitser, High-efficiency all-dielectric metasurfaces for ultracompact beam manipulation in transmission mode, Nano letters 15 (9) (2015) 6261–6266.

[17] K. E. Chong, I. Staude, A. James, J. Dominguez, S. Liu, S. Campione, G. S. Subramania, T. S. Luk, M. Decker, D. N. Neshev, et al., Polarization-independent silicon metadevices for efficient optical wavefront control, Nano letters 15 (8) (2015) 5369–5374.

[18] F. Zhang, Q. Zhao, J. Zhou, S. Wang, Polarization and incidence insensitive dielectric electromagnetically induced transparency metamaterial, Optics express 21 (17) (2013) 19675–19680.

[19] Y. Yang, I. I. Kravchenko, D. P. Briggs, J. Valentine, All-dielectric metasurface analogue of electromagnetically induced transparency, Nature communications 5.

[20] D. P. Gaillot, C. Croënne, D. Lippens, An all-dielectric route for terahertz cloaking, Optics express 16 (6) (2008) 3986–3992.

[21] J. Valentine, J. Li, T. Zentgraf, G. Bartal, X. Zhang, An optical cloak made of dielectrics, Nature materials 8 (7) (2009) 568–571.

[22] X. Liu, Q. Zhao, C. Lan, J. Zhou, Isotropic mie resonance-based metamaterial perfect absorber, Applied Physics Letters 103 (3) (2013) 031910.

[23] X. Liu, C. Lan, B. Li, Q. Zhao, J. Zhou, Dual band metamaterial perfect absorber based on artificial dielectric molecules, Scientific Reports 6.

[24] S. Jahani, Z. Jacob, Transparent subdiffraction optics: nanoscale light confinement without metal, Optica 1 (2) (2014) 96–100.

[25] Reena, Y. Kalra, A. Kumar, R. Sinha, Tunable unidirectional scattering of ellipsoidal single nanoparticle, Journal of Applied Physics 119 (24) (2016) 243102.

[26] I. Devi, R. Dalal, Y. Kalra, R. K. Sinha, Modeling and design of all-dielectric cylindrical nanoantennas, Journal of Nanophotonics 10 (4) (2016) 046011–046011.

ACCEPTED MANUSCRIPT

[27] F. Alves, D. Grbovic, B. Kearney, G. Karunasiri, Microelectromechanical systems bimaterial terahertz sensor with integrated metamaterial absorber, Optics letters 37 (11) (2012) 1886–1888.

[28] W. L. Chan, H.-T. Chen, A. J. Taylor, I. Brener, M. J. Cich, D. M. Mittleman, A spatial light modulator for terahertz beams, Applied Physics Letters 94 (21) (2009) 213511.

[29] C. Wu, B. Neuner III, J. John, A. Milder, B. Zollars, S. Savoy, G. Shvets, Metamaterial-based integrated plasmonic absorber/emitter for solar thermo-photovoltaic systems, Journal of Optics 14 (2) (2012) 024005.

[30] J. Hao, J. Wang, X. Liu, W. J. Padilla, L. Zhou, M. Qiu, High performance optical absorber based on a plasmonic metamaterial, Applied Physics Letters 96 (25) (2010) 251104.

[31] X. Liu, T. Starr, A. F. Starr, W. J. Padilla, Infrared spatial and frequency selective metamaterial with near-unity absorbance, Physical review letters 104 (20) (2010) 207403.

[32] K. Aydin, V. E. Ferry, R. M. Briggs, H. A. Atwater, Broadband polarization-independent resonant light absorption using ultrathin plasmonic super absorbers, Nature communications 2 (2011) 517.

[33] J. Hendrickson, J. Guo, B. Zhang, W. Buchwald, R. Soref, Wideband perfect light absorber at midwave infrared using multiplexed metal structures, Optics letters 37 (3) (2012) 371–373.

[34] C. Gong, M. Zhan, J. Yang, Z. Wang, H. Liu, Y. Zhao, W. Liu, Broadband terahertz metamaterial absorber based on sectional asymmetric structures, Scientific Reports 6.

[35] E. D. Palik, Handbook of optical constants of solids, Vol. 3, Academic press, 1998.

[36] A. Ghatak, Optics, Tata-McGraw Hill Publishing Company, 2003.

ACCEPTED MANUSCRIPT

Highlights:

- Fundamental principle is Mie resonance
- Absorbance is 99.9% at 88 μm
- Design is simpler than patterned metamaterial perfect absorbers
- Absorbance remains greater than 90% for upto 55º of angle of incidence.

ACCEPTED MANUSCRIPT

Global Journal of Enterprise Information System

DOI: 10.18311/gjeis/2017/15616

Location based Twitter Opinion Mining using Common-Sense Information

Amita Jain[1*] and Minni Jain[2]

[1]Department of Computer Science and Engineering, Ambedkar Institute of Advanced Communication Technologies and Research, Delhi, India; amita_jain_17@yahoo.com

[2]Department of Computer Science and Engineering, Delhi Technological University, Delhi, India

Abstract

Sentiment analysis research of public information from social networking sites has been increasing immensely in recent years. Data available at social networking sites is one of the most effective and accurate source to identify the public sentiment of any product/service. In this paper, we propose a novel localized opinion mining model based on common sense information extracted from ConceptNet ontology. The proposed methodology allows interpretation and utilization of data extracted from social media site "Twitter" to identify public opinions. This paper includes location specific, male- female specific and concept specific popularities of product. All extracted concepts are used to calculate senti_score and to build a machine learning model that classifies the user opinions as positive or negative.

Keywords: ConceptNet, Natural Language Processing, Sentiment Analysis, SentiWordNet

Paper Code: 15616; **Originality Test Ratio:** 08%; **Submission Online:** 11-Jan-2017; **Manuscript Accepted:** 15-Feb-2017; **Originality Check:** 18-Feb-2017; **Peer Reviewers Comment:** 25-Feb-2017; **Double Blind Reviewers Comment:** 12-April-2017; **Author Revert:** 03-May-2017; **Camera-Ready-Copy:** 10-May-2017

1. Introduction

Sentiment analysis of data collected from social networking sites is an effective mean of discovering public opinion. Today, many companies conduct paper or online surveys to collect customer opinions about their product/service. In paper based surveys very less customer base can be involved and there is no assurance of honest and complete opinions. Therefore paper based surveys are not an effective approach to collect product opinions from public. With the emergence of social networking applications and sites, people tend to express their opinion about the product on their twitter or Facebook profiles. Social media sites have a lot of data related to people's opinions for services/products they use. Hence mining opinions from social media sites is a much more advanced method for market analysis. A lot of work has been done on opinion mining from twitter, most of which focuses on people's sentiment towards various topics. But analyzing twitter data in this manner gives a much generalized idea. To make it more specific, opinion mining can be performed on social media data from explicit locations. Our approach is to find the product sentiments in specific locations. To determine the popularity of a given product in several locations, a large data set of Twitter is analyzed. Tweets are a reliable source of information mainly because people tweet about everything and anything they do including purchasing new products and reviewing them. This will allow companies to emphasis their marketing expenditures on areas where sentiment is low, while keeping minimum advertisement in areas of high popularity locations. This proposed research work deals with outcome prediction and explores localized outcomes based on common sense knowledge.

Common-sense knowledge represents basic understanding that people gain through experiences[1]. Commonsense, is required to properly identify sentiments in natural language text, for instance, the concept "small queue" is positive when referring to a post office, while the concept "small room" should be understood as negative in a hotel review. The concept-level sentiment analysis outperforms other existing research work because it preserves the semantic association with multiword expressions. For commonsense knowledge, in this paper ConceptNet ontology is used. An AI tool "NamSor" is used to exactly deduce gender from a person's user name. After completion of the analysis phase, experimental results are presented.

The rest of the paper is organized as follows. Section 2 describes prior research work that has been done to determine sentiment of twitter data. Section 3 discusses the entire procedure of sentiment analysis including data extraction & preprocessing, feature extraction and sentiscore calculation. The results of the experiment have been visualized graphically and also interpreted in section 4. Finally in section 5, discusses conclusion and future work respectively.

*Author for correspondence

2. Related Work

Various sentiment analysis methods have been proposed in literature to classify user opinions. However, most previous existing methods mainly rely on simple keywords spotting, syntactic information and POS tags. Mullen and Collier[2] proposed a method to expand the concept set based on Osgood's theory of semantic orientation. Turney's[3] described a semantic orientation method for supervised sentiment classification. Gelfand et al.[4] have developed a method based on sematic relation graph to extract concepts from a documents. They used the relationships between words extracted from a lexical database to form concepts. Agarwal et al.[5] proposed sentiment analysis with dependency- based semantic parsing. Hriday et al.[6] developed a location based methodology to identify sentiment of opinions from twitter. Nithish et al.[7] mainly focused on market reaction of smart phones using sentiment analysis. They intended to determine the explanation of factors that influences a product rating. DBSCAN: A clustering based data mining algorithm is proposed by Khanaferov et al.[8] to analyses tweets from same domain.

3. Proposed Methodology

Proposed approach for concept-level localized twitter opinion mining comprises three main steps 1) Data extraction and preprocessing 2) Concept extraction and 3) sentiment score calculation.

3.1 Data Extraction and Preprocessing

In the proposed work, tweets from twitter are used as a source of data. To extract an extensive amount of Tweets a public API of twitter is used. Each extracted tweet has several kind of information like tweet_id, text, username etc. Out of these extracted information only the tweet_id and text is useful for this work.

For this research, tweets are extracted from six major cities in the India. Due to the language constraints and data availability, the choice of cities is very limited. Six major cities selected for tweets collection and experiments are: New Delhi, Chennai, Mumbai, Bangalore, Pune, Surat. The longitude and latitude of each major city is used to define the city. Free map tools by Viklund are used to select estimated radius of coverage for each city[9]. The values of radius, latitude and longitude are assigned to the parameter locations in the query build for API.

For implementation and testing product iPhone7 is selected. Using the proposed approach it is possible to evaluate popularity of any product but with a condition that the product must have a good amount of data available on twitter. At the period of the research a large amount of tweets related to iPhone7 were available. So the tweets that contain the word "iPhone 7" are extracted. The most or least popular concepts of the iPhone7 are determined using some keywords to collect concept specific tweets. A case is 'iPhone7 battery'. This query will return all tweets containing the terms, iPhone 7 and battery together. Other keywords used are jet-black, 3.5 mm jack, camera, iOS, iTunes, screen, sound, and touch. From each the tweet text, location, and username are extracted.

3.2 Concept Extraction

In order to extract useful concepts, the Stanford Natural Language Processing tool by The Stanford NLP Group (SNLP Group 2015) is used[10,11]. This SNLP tool provides outputs that are grammatical relations between terms in a sentence. In this proposed work, to extract concepts from tweets, three main relations (also called dependencies) are used. These relations are **nsubj, dobj, amod**.

- The first relation is **nsubj**. "This relation is used to find dependencies between nouns and adjectives or verbs which are complementing the noun in a sentence". This relation provides an important knowledge that whether a sentence is complementing a noun or not. An instance of this **nsubj** relation is:

"My iPhone7 camera is amazing!" For this sentence, the concept **(camera, amazing)** is extracted.

- The second relation is **dobj**. "It is a direct object relation and used to find direct objects that a verb is referring to in a sentence". An example would be

"Hate the sound of iPhone7!" For this tweet the system extracts the concept **(Hate, sound)** and also **(Hate, iPhone7)**.

- The final relation is **amod**. "It is adjectival modifier relation and used to identify any adjectives used in a sentence to modify a noun phrase". For instance.

"Got my new Jet-black iPhone7, feeling wow!!" For this tweet, concept (**iPhone7, jet-black**) is extracted

3.3 Extracting Concepts from tweets using Common-Sense Knowledge

To extract more concepts related to product using commonsense knowledge 'ConceptNet' is used. "ConceptNet is a large semantic network consisting of large number of common-sense concepts" [12]. To mine several inferences from the text, commonsense knowledge available in ConceptNet can be used. It comprises vertices 'concepts' linked by edges 'relations between concepts'. Some of available relationships in the ConceptNet are: MadeOf, IsA, AtLocation, EffectOf, DesireOf, CapableOf etc[12]. Now to extend concept list, extracted concepts from previous steps using three dependencies relations are sent as query to ConceptNet. For instance, let the concept 'birthday party' is sent to ConceptNet as query, then output would be concepts such as buy gifts, cake. In ConceptNet, relations related to concepts are :"cake – AtLocation-- birthday party" and "buy gifts

– UsedFor—birthday party". These new set of concepts provide more knowledge. This combination of dependencies with commonsense knowledge offers a better understanding of text to the system[13]. The proposed system allows the machine to better understand the matter and the meaning conveyed by the natural language text.

3.4 Calculation of Sentiment Score of Tweets

To calculate the sentiment of extracted tweets a numeric metric is required. SentiWordNet (2015), is used to calculate overall sentiment of tweets[14]. SentiWordNet provides a score lies between −1 to 1, where higher value indicates the more positive sentiment and lower value indicates to lower sentiment. Score from SentiWordNet for each word in a tweet (containing at least one concept extracted on previous section) is obtained and then addition of all sentiment values to get a score for whole tweet. There is an important thing that SentiWordNet takes terms and their part of speech as input. The part of speech of the terms depends entirely on the sentence itself. So to map each term of sentence to its part of speech, a POS tagger is used.

The sentiment for each term is obtained using SentiWordNet and then the term scores are added to get final sentiment score of tweet, Eq. (1).

$$Score\ (location) = \frac{\sum_{i=1}^{n} Senti_{Score_i}}{n} \quad (1)$$

where, n is total number of tweets, Senti_Score i is SentiWordNet score for each tweet, Location j is refers to one particular city.

The scores calculated in this method do not follow any scale so it is necessary to normalize these final scores to get fixed sentiment scores for the tweets. To get information about users' gender from the tweets, a tool called "Nam-Sor" (2015) is used. It is an existing data mining tool[15]. The entire method is illustrated in Fig. 1.

4. Results and Discussion

Several comparisons are made to properly understand the variations in sentiments. Three comparisons are done to demonstrate the sentiment trends using proposed approach. These are:

1. *Countrywide Sentiment as Percentage:* Overall national sentiment of product.

2. *Countrywide concept Average Score:* This score includes concepts related to product for all the six cities. This score provides the overall view of sentiment towards the iPhone7 features.

3. *Gender Concept Average per city:* In this sentiment scores, all six cities are grouped on gender (male/female) basis for distinct concepts. It comprises all the mentioned variables that are gender, product features and specific location.

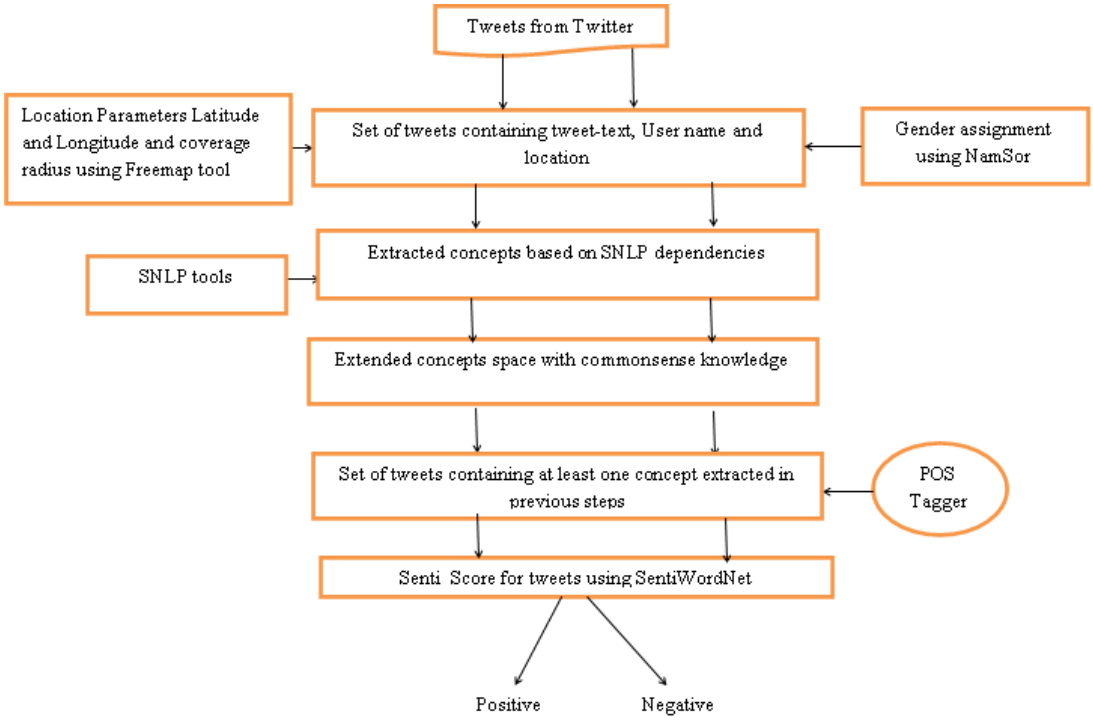


Figure 1. Flow diagram of proposed approach to location based sentiment analysis.

Originally 1040 tweets are collected using the twitter API. But after filtering only 552 are left for the further process. The tweets other than 552 are not valuable in for opinion mining. In this work, product search parameter and concept search parameter and location parameter are used so the count of tweets that obtained is lower than expectations.

For better comparability and easy understanding, all of the comparisons are explained using graphs. As shown in Fig. 2, 75% above users believed that iPhone7 is a worthy product. Usually excellent sentiments are more challenging to find because this requires huge number of tweets which cover words with very high positive SentiWordNet scores.

Next, the popularity of product iphone7 on the basis of its features is shown in Fig. 3. As per the results, feature iOS and Camera of the iPhone7 got highest positive sentiments. Result shows low positive sentiment of users about Jet-black color and battery of iPhone7. Now the final graph shown in fig. 4 includes results based on location, and gender (male/female). There are consistent sentiments for the both female and male with the touch and screen features. Observing Fig.4, the cities, New Delhi and Pune shows most positive sentiment scores with Mumbai following on close by cities Bangalore and Surat have moderate level in positive sentiments while the city Chennai has lower positive sentiments.

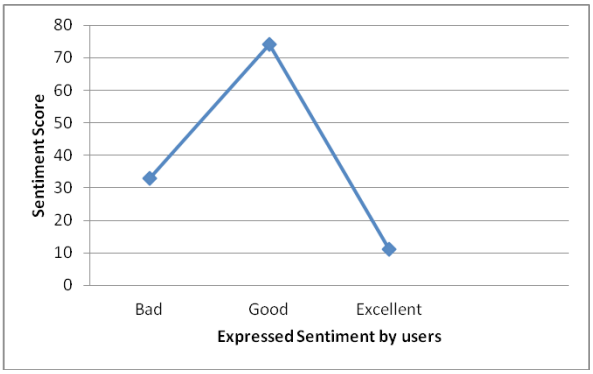


Figure 2. Overall countrywide sentiments as percentage.

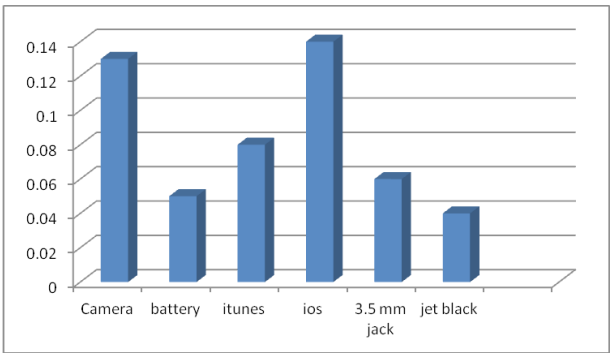


Figure 3. Countrywide concept average.

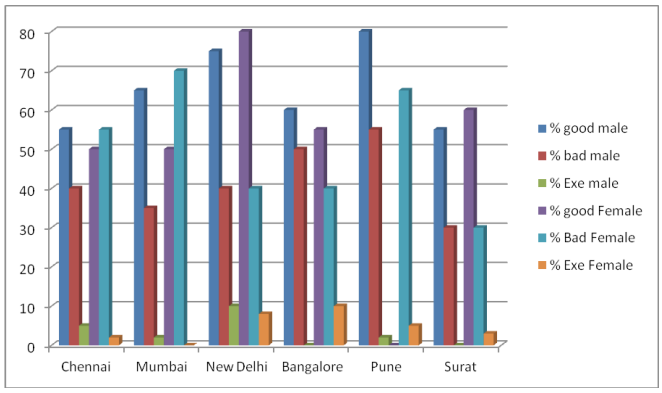


Figure 4. City based gender specific sentiment.

New Delhi shows maximum positive sentiment scores for both genders (female/male), followed closely by Pune and then by Mumbai. This is generally a good sign of diversity in these cities, meaning very diverse user sentiments for the iPhone7 are available. Here % bad, % good and % exe means percentage of marked as "bad", "good" and "excellent" respectively.

5. Conclusion and Future Work

In this research work, a methodology is proposed to identify the sentiment/opinion /popularity of any product/service across female and male users, in several locations. The proposed methodology is a generalized approach and can be used for tweets of any product and from any country with a condition of availability of product related tweets in a good amount. The tool named "NamSor" is used to classify gender of each tweet, SentiWordNet to calculate senti_Score and ConceptNet ontologies to extend concepts are used. Finally in result section, several comparisons are made to validate the accurateness of the proposed approach with the help of graphs. The analysis shows the positive and negative sentiment of male and female users from various locations, towards features of iphone7. It is also analyzed that adding of the concepts from commonsense knowledge increases the performance of proposed system. Now, the future work involves identifying the more relations to extract the concepts and to enhance the quality of concept. Ontologies other than ConceptNet can be explore to improve the concept mining method. To this end, fuzzy based algorithms van be used to enhance the effectiveness of the system in terms of performance.

6. References

1. Erik C, Amir H, Catheine H, Chris E. Common sense computing: From the society of mind to digital intuition and beyond. In: Lecture notes in computer science 5707, Springer. 2009; 252–9.

2. Mullen T, Collier N. Sentiment analysis using support vector machines with diverse information sources. In: EMNLP. 2004; 546–9. PMid:15256037
3. Turney PD. Thumbs u or thumbs down? Semantic orientation applied to unsupervised classification of reviews. ACL-2002. 2002; 417–24.
4. Gelfand B, Wulfekuler M, Punch WF. Automated concept extraction from plain text. In: AAAI workshop on text categorization. 1998; 13–7.
5. Agarwal B, Poria S, Mittal N, Gelbukh A, Hussain A. Concept-level sentiment analysis with dependency-based semantic parsing: a novel approach. Cognitive Computation. 2015; 7(4):487–99. https://doi.org/10.1007/s12559-014-9316-6
6. Hridoy SA, Ekram MT, Islam MS, Ahmed F, Rahman RM. Localized twitter opinion mining using sentiment analysis. Decision Analytics. 2015; 2(1):8. https://doi.org/10.1186/s40165-015-0016-4
7. Nithish R, Sabarish S, Abirami AM, Askarunisa A, Navaneeth Kishen M. An Ontology based Sentiment Analysis for mobile products using tweets. In Fifth International Conference on Advanced Computing. 2013. p. 242–7. https://doi.org/10.1109/icoac.2013.6921974
8. Khanaferov D, Luc C, Wang T. Social Network Data Mining Using Natural Language Processing and Density based Clustering. In IEEE International Conference on Semantic Computing (ICSC). 2014. p. 250–151. https://doi.org/10.1109/icsc.2014.48
9. Viklund A. Free Map Tools. Available from: http://www.freemap-tools.com/. Retrieved December 23, 2016.
10. SNLP Manua. Stanford Typed Dependencies Manual. Available from: http://nlp.stanford.edu/software/dependencies_manual.pdf.
11. SNLP Group. The Stanford NLP Group. Available from: http://nlp.stanford.edu/. Retrieved December 4, 2016.
12. Havasi C, Speer R, Alonso JB. Conceptnet 3; a flexible, multilingual semantic network for common sense knowledge, In: Recent advances in natural language processing. 2007; 27–9.
13. Wang Q-F, Cambria E, Liu C-L, Hussain A. Common Sense knowledge for handwritten Chinese text recognition. Cogn Compu. 2013; 5(2):234–42. https://doi.org/10.1007/s12559-012-9183-y
14. SentiWordNet. Available from: http://sentiwordnet.isti.cnr.it/. Retrieved January 30, 2017.
15. Namsor. Available from: https://github.com/namsor/namsor-api. Retrieved January 30, 2017

Citation:

Amita Jain and Minni Jain

"**Location based Twitter Opinion Mining using Common-Sense Information**",

Global Journal of Enterprise Information System. Volume-9, Issue-2, April-June, 2017. (http://informaticsjournals.com/index.php/gjeis)

Conflict of Interest:

Author of a Paper had no conflict neither financially nor academically.

Modeling the emission of high power terahertz radiation using Langmuir wave as a wiggler

Jyotsna Panwar and Suresh C. Sharma

Citation: Physics of Plasmas **24**, 083101 (2017); doi: 10.1063/1.4993611
View online: http://dx.doi.org/10.1063/1.4993611
View Table of Contents: http://aip.scitation.org/toc/php/24/8
Published by the American Institute of Physics

Articles you may be interested in

Two color laser driven THz generation in clustered plasma
Physics of Plasmas **24**, 073109 (2017); 10.1063/1.4991382

Generating tunable THz radiation using rippled density plasma driven by density modulated relativistic electron beam
Physics of Plasmas **24**, 073101 (2017); 10.1063/1.4990075

Terahertz radiation with high power and high efficiency in a magnetized plasma
Physics of Plasmas **24**, 073112 (2017); 10.1063/1.4991395

The pseudoforce approach to fully nonlinear plasma excitations
Physics of Plasmas **24**, 082302 (2017); 10.1063/1.4987016



PHYSICS OF PLASMAS **24**, 083101 (2017)

Modeling the emission of high power terahertz radiation using Langmuir wave as a wiggler

Jyotsna Panwar and Suresh C. Sharma
Department of Applied Physics, Delhi Technological University, Shahbad Daulatpur, Bawana Road, Delhi-110042, India

(Received 3 March 2017; accepted 29 June 2017; published online 18 July 2017)

The emission of high power terahertz (THz) radiation lying in the range of millimeter to submillimeter wavelengths has been studied analytically using the Langmuir wave as an electrostatic pump wave in the presence of static magnetic field for both finite and infinite geometries. The interaction of two laser beams with the relativistic electron beam leads to velocity modulation of the beam, which then translates into density modulation on traveling through the drift space. The premodulated beam on interacting with the pump wave acquires an oscillatory velocity that couples with the perturbed and modulated beam densities to result in nonlinear current density which helps in evaluating the growth rate and efficiency of the output THz radiation. The beam and plasma wave wiggler parameters are found to influence the growth rate and efficiency of the emitted THz radiation. *Published by AIP Publishing.*
[http://dx.doi.org/10.1063/1.4993611]

I. INTRODUCTION

Terahertz (THz) radiation poses a great scope of research in various potential fields[1–5] due to its high penetration and non-ionization characteristics. Several schemes have been proposed for the generation of these high power radiations including mainly optical rectification[6–8] in dielectrics, plasmas, and semiconductors, interaction of high power lasers with electro-optic crystals, or by employing prebunched electron beams[9–11] using free electron lasers.

The emission of terahertz radiation by optical rectification of amplitude modulated surface plasma, excited on the rippled metal surface by an obliquely incident laser beam, has been studied by Bhasin and Tripathi.[8] They concluded that the amplitude of the wave scales as the square root of the laser wavelength and laser spot size, and the efficiency of the wave are found to be high due to very weak linear damping.

Sharma and Malik[12] have observed analytically the effect of prebunched relativistic electron beam (REB) on the excitation of surface plasmons lying in the terahertz range using semiconductors and observed that the efficiency and the amplitude of the plasmons can be tuned by the beam parameters and the spacing of the semiconductor plates.

Welsh and Wynne[12] have analyzed the emission of radiation pulses by the incoherent optical rectification on metallic nanostructured surfaces via excitation of surface plasmons.

The effective schemes employed for the generation of THz radiation can be the interaction of relativistic electron beams or short wavelength lasers with the plasma medium. Plasma eigenmodes employed as an electrostatic wiggler help in obtaining radiation of shorter wavelength with a high wiggler strength.[13]

Singh *et al.*[14] have investigated the emission of terahertz radiation by the non-linear decay of upper hybrid wave in the presence of static magnetic field into a high frequency laser wave and terahertz wave of low frequency. They observed that the power and the growth rate of the terahertz radiation increases with the magnetic field.

Sharma *et al.*[15] have employed surface plasma as the wiggler in the presence of prebunched REB for the generation of radiation in the range of sub-millimeter wavelength. They concluded that the plasma density helps in improving the gain and efficiency of the device. The growth rate was found to increase with the plasma density and varied as one third power of the beam density.

Zolghadr *et al.*[16] have studied the effect of magnetized plasma on the output power and the saturation length using the whistler wave wiggler. They observed that the output power increases and the saturation length decreases with the plasma frequency. Further, the power decreases and the saturation length increases with the cyclotron frequency.

More recently, Hussain *et al.*[17] have investigated that the terahertz radiation emitted using the super Gaussian laser beams interaction with the plasma in the presence of static electric field is strongly affected by the plasma density, intensity of the laser beams, and strength of the externally applied field.

Radiation with frequencies up to 200 GHz generated from the Cerenkov interaction in the magnetized plasma has been experimentally detected by Yugami *et al.*[18] They found that the intensity of the radiation to be proportional to the magnetic field intensity and power as the square of the field strength. The results indicate that this is an efficient way to obtain tunable radiation with high power.

Plasma can withstand high fields due to being in the ionized state and hence overcome the limitation of material breakdown. The plasma medium has thus proved to be the most suited medium for the generation of high power radiation. The introduction of plasma in the free electron lasers helps in reducing the beam energy requirement and acts as a slow wave structure.

Tripathi and Liu[19] have addressed the effects on the radiation and the electron beam due to the plasma introduced in a magnetic wiggler. They observed that the plasma and the strong magnetic field reduces the electron beam energy requirement and enhances the growth rate. The depression in the electron density of plasma due to the beam acts as an optical guiding medium when the beam and plasma densities are comparable.

The emission of high power tunable radiation by the interaction of relativistic electron beam with a dusty plasma crystal has been studied by Mehdian *et al.*[20] They observed that the crystal as a wiggler provides high output frequency, power, and efficiency at comparatively low beam energy and quality.

Hedayati *et al.*[21] have investigated that the plasma medium in the presence of static magnetic field helps in obtaining radiation of shorter wavelength by reducing the wiggler period and lowering the electron beam energy. They also concluded that the electron bunching enhances at the wave frequency comparable to the plasma frequency.

The generation of terahertz radiation by employing rippled underdense plasma has been investigated by Pathak *et al.*[22] In this case, the frequency was observed to be tuned by the beam energy and ripple wavelength, and the parameters affecting the growth rate were plasma density, depth, and amplitude of density ripple.

Kumar and Tripathi[23] have employed the beating of two collinear laser pulses in a clustered gas to study the emission of radiation. In this case, it was concluded that the power of the THz radiation varies linearly as the square of the amplitude of laser, fourth power of the plasma frequency of cluster electrons, and inversely as the square of the laser frequency.

Gildenburg and Vvedenskii[24] have demonstrated the terahertz radiation generation by the linear conversion of an ultrashort laser pulse. The power of the radiation was obtained in the range of gigawatt using an optical intensity of about 10^{14}–10^{15} W cm^{-2}. The conversion efficiency was realized to be comparatively quite high when the length of the laser pulse was not much greater than the period of the optical field.

In the present paper, we have developed an analytical model to study the emission of terahertz radiation by the interaction of a density modulated electron beam with an electrostatic plasma pump wave wiggler. The mechanism for the analysis is as follows. The REB first gets energy modulated at the different frequency of the two laser beams exerting non-linear ponderomotive force on the electron beam in the modulator. The beam on travelling through the drift space gets density modulated. The density modulated beam interacts with the plasma region, where the Langmuir plasma wave is excited that acts as an electrostatic pump wave. The beam acquires an oscillatory velocity due to the pump wave. Further, the radiation wave and the pump lead to the perturbed beam density that couples with the oscillatory velocity to give nonlinear current density which behaves as an antenna for the emission of high power terahertz radiation. In Sec. II, we have shown the modulation of the REB and its interaction with the pump wave in infinite and finite geometries to study the behavior of high power output radiation. Section III includes the discussions for the obtained results and the investigations are concluded in Sec. IV.

II. PHYSICAL MODEL

In the present analysis, on taking into account the interaction of premodulated REB with the pump wave, we investigate the emission of THz radiation. Preliminarily, we investigate the electron beam modulation by considering an REB of density n_e^0 moving in the axial direction with initial velocity $v_e^0\,\hat{z}$ which is modulated by two lasers, laser L_a and laser L_b (cf. Fig. 1) polarized in the $\hat{x}$-direction and copropagating in the $\hat{z}$-direction with frequencies ω_a and ω_b, and wave numbers k_a and k_b, respectively. The electric and magnetic fields of the laser beams propagating in the modulator are considered as $\vec{E_{L_\xi}} = A_\xi e^{-i(\omega_\xi t - k_\xi z)}\hat{x}$ and $\vec{B_{L_\xi}} = \frac{c(\vec{k_\xi}\times\vec{E_{L_\xi}})}{\omega_\xi}\hat{y}$, where $\xi = a,b$ for lasers L_a and L_b, respectively.

The two laser beams interact with REB in the modulator region as a result of which the beam electrons acquire an oscillatory velocity $\vec{v_{0_\xi}} = \frac{e\vec{E_{L_\xi}}}{im_e\omega_\xi\gamma_0}$ (where $\xi = a,b$ for lasers L_a and L_b, respectively), where m_e is the mass of electron and $\gamma_0 = (1-(\frac{v_e^0}{c})^2)^{-1/2}$ is the relativistic gamma factor. Due to the interaction of oscillatory velocity of the beam and the magnetic fields of the laser beams, an axial nonlinear ponderomotive force $(\vec{f}_{pe})$[25] is exerted on the beam electrons

$$\vec{f}_{pe} = f_0\cos(\omega_e t - k_e z)\hat{z}, \tag{1}$$

where $f_0 = \frac{e^2k_eA_aA_b^*}{2im_e\gamma_0\omega_a\omega_b}$, $\omega_e(=\omega_a-\omega_b)$ is the frequency (lying in the terahertz range), $\vec{k_e}(=\vec{k_a}-\vec{k_b})$ is the wave number of the velocity modulated REB, and e is the electronic charge. The length of the overlap region (or modulator region) is $l_m(=\frac{2r_l}{\sin(\frac{\theta_l}{2})}\approx\frac{4r_l}{\theta_l}$, where θ_l is the small angle of interaction and r_l is the radius of laser spot). The angle of interaction is taken very small so as to have significant velocity modulation, and only the frequency components (ω_e) relevant for the generation of THz radiation are considered, whereas the other frequency components $(2\omega_a, 2\omega_b)$ are ignored. The electron beam response under the ponderomotive force is governed by the relativistic equation of motion

$$\frac{\partial}{\partial t}(\gamma\vec{v}) = \frac{\vec{f}_{pe}}{m_e}, \tag{2}$$

where $\vec{v} = \vec{v_e^0} + \vec{v_e^m}$, $\gamma = \gamma_0 + \frac{v_e^0 v_e^m \gamma_0^3}{c^2}$. On linearizing Eq. (2), the modulated velocity (v_e^m) attained by the beam is obtained as

$$v_e^m = \frac{f_0 l_m}{m_e\gamma_0^3 v_e^0}\frac{\sin(\theta_l)}{\theta_l}e^{-i(\omega_e t_m+\theta_l)} = -\alpha_m v_e^0\sin(\omega_e t_m+\theta_l), \tag{3}$$

where $\alpha_m = \frac{k_e l_m c^2}{2\gamma_0^4 v_e^{02}}\frac{e^2A_aA_b^*}{m_e^2\omega_a\omega_b c^2}\frac{\sin(\theta_l)}{\theta_l}$, t_m is the time at which beam arrives at the modulator, and $\frac{\sin(\theta_l)}{\theta_l}\approx 1$ for appreciable

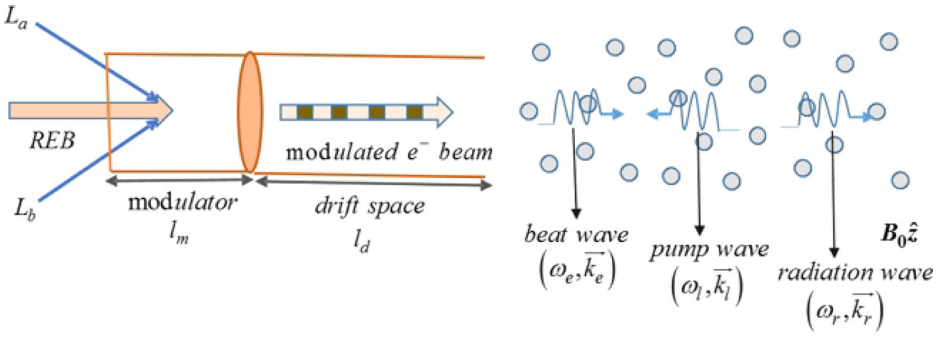


FIG. 1. Schematic of interaction of premodulated REB (by two laser beams) with the Langmuir plasma pump wave (in unbound plasma).

velocity modulation. The beam, thus, gets velocity modulated in the modulator region where the initial slow moving electrons (late electrons) start moving fast and fast moving electrons (early electrons) slow down in order to attain a common velocity.

The beam then passes through the drift space, where due to acceleration and deceleration, the late electrons catch the early electrons, leading to the density modulation of the electron beam. The density (n_e^m) of the modulated REB can be written as[24]

$$n_e^m = -2n_e^0 J_1\left(\frac{\omega_e l_d \alpha_m}{v_e^0}\right)\cos\left(\omega_e t_m - \varphi\right), \tag{4}$$

where $t_p\left(= t_m + \frac{l_m}{v_e^0} + \frac{l_d(1+\alpha_m \sin(\omega_e t_m + \theta_l))}{v_e^0}\right)$ is the time at which the beam electrons leave the drift space and enter the pump wave; l_d is the length of drift space.

In the next stage, we consider a system filled with an isothermal plasma of uniform density n_p^0 in which the premodulated REB excites an electrostatic plasma wave (pump wave) due to the Cerenkov interaction (i.e., $\omega_l \approx \vec{k}_l.\vec{v_e^0}$, where $(\omega_l, \vec{k}_l)$ is the frequency and the wave number of the pump wave, respectively) which acts as a wiggler in the present analysis.

The electrostatic potential of the plasma wave is defined as $\phi_l = A_l e^{-i(\omega_l t - k_l z)}$. The electric field of the pump wave $\vec{E}_l(= -\nabla\phi_l)$ is polarized in the $x - z$ plane.

The perturbed velocity ($\vec{v_{1l}}$) due to the pump wave, obtained using the linearized equation of motion, $m_e\left[\frac{\partial}{\partial t}\left(\gamma_0^3 \vec{v_{1l}}\right) + v_e^0 \frac{\partial}{\partial z}\left(\gamma_0^3 \vec{v_{1l}}\right)\right] = -e\vec{E}_l$, is given as

$$\vec{v_{1l}} = -\frac{e\vec{k}_l \phi_l}{m_e \gamma_0^3 \left(\omega_l - k_l v_e^0\right)}. \tag{5}$$

On using the equation of continuity, $\frac{\partial n_{1l}}{\partial t} + \nabla.\left(n_e^0 \vec{v_{1l}} + n_{1l}\vec{v_e^0}\right) = 0$, the perturbed beam density response for the beam electrons (n_{1l}) and the plasma electrons (n_{1p}) is obtained as

$$n_{1l} = -\frac{\left(n_e^0 + n_e^m\right) e k_l^2 \phi_l}{m_e \gamma_0^3 \left(\omega_l - k_l v_e^0\right)^2} \text{ and } n_{1p} = -\frac{n_p^0 e k_l^2 \phi_l}{m_e \omega_l^2}. \tag{6}$$

Poisson's equation, $\nabla^2\phi_l = 4\pi e(n_{1l} + n_{1p})$, can be simplified to $\left(\omega_l^2 - \omega_p^2\right)\left(\omega_l - k_l v_e^0\right)^2 = \frac{\omega_{pb}^2 \omega_l^2}{\gamma_0^3}$, where $\omega_{pb}\left[= \left(\frac{4\pi n_e^0 e^2}{m_e}\right)^{1/2}\right]$ is the beam plasma frequency, $\omega_p\left(= \left(\frac{4\pi n_p^0 e^2}{m_e}\right)^{1/2}\right)$ is the electron plasma frequency, and n_p^0 is the initial density of the plasma medium.

On using $\omega_l = \omega_p + \delta_l$ in the above equation, where δ_l is the small frequency mismatch, we obtain $\delta_l = \left[\frac{\omega_{pb}^2 \omega_p}{2\gamma_0^3}\right]^{1/3}\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right)$. The imaginary part of δ_l gives the growth rate as

$$\Gamma_p = \frac{\sqrt{3}}{2}\left(\frac{n_e^0 + n_e^m}{2n_p^0}\right)^{1/3}\frac{\omega_p}{\gamma_0}. \tag{7}$$

The motion of an electron moving along the axial direction is governed by the equation of motion as $\frac{d^2 z}{dt^2} = -\frac{eA_{lz}}{m_e\gamma_0} e^{-i(\omega_l t - k_l z)} = -\left(\omega_l^b\right)^2 z$, where $\omega_l^b\left[= \left(\frac{eA_{lz}k_l}{m_e\gamma_0}\right)^{1/2} = \Gamma_p\right]$ is the electron bounce frequency.

Therefore, the amplitude of the pump wave (A_{lz}) is given as

$$A_{lz} = \frac{3\omega_p^2 m_e}{4ek_l\gamma_0}\left(\frac{n_e^0 + n_e^m}{2n_p^0}\right)^{2/3}, \tag{8}$$

which can be defined as the function of beam parameters, that is, density of the modulated electron beam ($n_e^0 + n_e^m$), beam velocity $\left(\gamma_0 = \frac{1}{\sqrt{1 - \frac{v_e^{02}}{c^2}}}\right)$, and plasma density ($n_p^0$).

Now, the interaction of the premodulated REB with the excited pump wave in the presence of external static magnetic field B_0 along the z-axis is studied. On taking the beam response in the electron frame moving with velocity $v_e^0\hat{z}$, where the beam is treated as non-relativistic, we obtain the oscillatory velocity component as $\vec{v_e^{l'}} = \frac{e\left[i\omega_l'\left(\nabla'\phi_l'\right) - \left(\vec{\omega_{ce}} \times \nabla'\phi_l'\right)\right]}{m_e\left(\omega_l'^2 - \omega_{ce}^2\right)}$, where $'$ denotes the quantities in the electron frame and $\omega_{ce} = \left(\frac{eB_0}{m_e c}\right)$ is the electron cyclotron frequency. The perturbation induced due to an electromagnetic radiation wave is given as $\vec{E}_r = \vec{A}_r e^{-i(\omega_r t - k_r z)}$, where $(\omega_r, \vec{k}_r)$ is the frequency and the wave number of the radiation wave, respectively. The radiation wave and the pump wave simultaneously exert a nonlinear ponderomotive force[26–28] on the beam electrons, $\vec{F}_p' = e\nabla'\phi_p'$, where the ponderomotive potential (ϕ_p') can be simplified to

$$\phi'_p = -\frac{e\left[i\omega'_l(\nabla'\phi'^{*}_l) + \left(\vec{\omega'_{ce}} \times \nabla'\phi'^{*}_l\right)\right].\vec{E'_r}}{2im_e\omega'_r\left(\omega'^2_l - \omega'^2_{ce}\right)}. \tag{9}$$

The response of beam electrons in terms of velocity ($\vec{v'_e}$) and density (n'_e) due to the ponderomotive force and the self-consistent electric field is obtained as

$$\vec{v'_e} = -\frac{e\nabla'\left(\phi' + \phi'_p\right)}{im_e\omega'_e}, \tag{10}$$

$$n'_e = -\frac{k'^2_e e^3 n^{0'}_e\left[i\omega'_l(\nabla'\phi'_l) - \left(\omega'_{ce} \times \nabla'\phi'_l\right)\right]\left[1 + \frac{n^{m'}_e}{n^{0'}_e}\right]\left(\phi' + \phi'_p\right)}{m_e\omega'^2_e\left(\omega'^2_l - \omega'^2_{ce}\right)}. \tag{11}$$

The nonlinear current density ($\vec{J'_{NL}}$) being an important source for the emission of terahertz radiation arises by coupling of the perturbed beam density with the oscillatory beam velocity. Hence, the nonlinear current density at $(\omega_r, \vec{k_r})$ is obtained as

$$\vec{J'_{NL}} = -n'_e e\vec{v'_e} = \frac{k'^2_e e^3 A'_l(\phi' + \phi'_p)n^{0}_e\left[1 + \frac{n^{m'}_e}{n^{0'}_e}\right]\left[i\omega'_l(\nabla'\phi'_l) - \left(\omega'_{ce} \times \nabla'\phi'_l\right)\right]e^{-i\left(\omega'_r t - k'_r z\right)}}{im_e\omega'^2_e\left(\omega'^2_l - \omega'^2_{ce}\right)}, \tag{12}$$

where $\omega_r = \omega_e + \omega_l$ and $\vec{k_r} = \vec{k_e} + \vec{k_l}$.

Further to study the proficient generation of the high power THz radiation, the growth rate and efficiency radiation wave have been evaluated for the geometries with infinite[15] (local effects) and finite[28–30] (nonlocal effects) boundaries.

Case I: Effects of infinite boundary on the THz radiation wave.

We consider the excited pump wave in plasma in a very large dimensional space such that it can be considered as an infinite system. In order to study the generation of terahertz radiation with the local effects of the wiggler wave, the wave equation is employed as

$$-\nabla^2\vec{E_r} = \frac{4\pi i\omega_r}{c^2}\vec{J_{NL}} + \frac{i\omega_r}{c^2}\frac{\partial\vec{E_r}}{\partial t}. \tag{13}$$

The expression for non-linear current density that drives the radiation wave is substituted in the above wave equation after applying Lorentz transformation[31] on it in the lab frame to obtain

$$\left(\omega_r^2 - k_r^2c^2\right)\left(\omega_e - k_ev_e^0\right)^2 = \frac{\omega_{pb}^2e^2\left(k_e - \frac{\omega_ev_e^0}{c^2}\right)^2[i\gamma_0\omega_l(\nabla\phi_l) - (\omega_{ce} \times \nabla\phi_l)]^2\left(1 + \frac{n_e^m}{n_e^0}\right)\left(1 - \frac{\chi_e}{\varepsilon}\right)}{2m_e^2\gamma_0\left(\gamma_0^2\omega_l^2 - \omega_{ce}^2\right)^2}, \tag{14}$$

where χ_e is the beam susceptibility and $\varepsilon = 1 + \chi_e$. To simplify the above equation, we take $\omega_e = k_ev_e^0 + \delta_{ir}$ and $\omega_r = k_rc + \delta_{ir}$, where δ_{ir} is the small frequency mismatch

$$\delta_{ir} = \left(\frac{\omega_{pb}^2e^2\left(k_e - \frac{\omega_ev_e^0}{c^2}\right)^2[i\gamma_0\omega_l(\nabla\phi_l) - (\omega_{ce} \times \nabla\phi_l)]^2\left(1 + \frac{n_e^m}{n_e^0}\right)\left(1 - \frac{\chi_e}{\varepsilon}\right)}{4\omega_rm_e^2\gamma_0\left(\gamma_0^2\omega_l^2 - \omega_{ce}^2\right)^2}\right)^{\frac{1}{3}}e^{\frac{i2n\pi}{3}}.$$

The growth rate of the radiation wave (Γ_{ir}) is obtained by taking the imaginary part of δ_{ir} with n = 1.

$$\Gamma_{ir} = \mathrm{Im}\delta_{ir} = \frac{\sqrt{3}}{2}\left(\frac{\omega_{pb}^2e^2(k_e - \frac{\omega_ev_e^0}{c^2})^2A_l^2\left(1 + \frac{n_e^m}{n_e^0}\right)\left(1 - \frac{\chi_e}{\varepsilon}\right)}{4\omega_rm_e^2\gamma_0^3\omega_l^2\left(1 - \frac{\omega_{ce}^2}{\gamma^2\omega_l^2}\right)^2}\right)^{\frac{1}{3}}. \tag{15}$$

The efficiency of the radiation[32] ($\eta_{ir} = \frac{\Gamma_{ir}}{\omega_r}\frac{\gamma_0^3}{\sqrt{3}(\gamma_0 - 1)}$) can be evaluated as

$$\eta_{ir} = \frac{1}{2}\left(\frac{\omega_{pb}^2e^2\gamma_0^6\left(k_e - \frac{\omega_ev_e^0}{c^2}\right)^2A_l^2\left(1 + \frac{n_e^m}{n_e^0}\right)\left(1 - \frac{\chi_e}{\varepsilon}\right)}{4\omega_r^4m_e^2(\gamma_0 - 1)^3\omega_l^2\left(1 - \frac{\omega_{ce}^2}{\gamma_0^2\omega_l^2}\right)^2}\right)^{\frac{1}{3}}. \tag{16}$$

Case II: Effects of finite boundary on the THz radiation wave.

The radiation wave has been considered for the same with the finite geometry. In order to study the nonlocal theory of the pump wave, we consider a uniform cylindrical plasma column of radius a_0 as shown in Fig. 2. Using the expression of Lorentz transformed nonlinear current density in the wave equation, we obtain

$$\frac{\partial^2 E_r}{\partial r^2}+\frac{1}{r}\frac{\partial E_r}{\partial r}+k^2E_r=-\frac{\omega_{pb}^2e^2[i\gamma_0\omega_l(\nabla\phi_l)-(\omega_{ce}\times\nabla\phi_l)]^2\left(k_e-\frac{\omega_e v_e^0}{c^2}\right)\left(1+\frac{n_e^m}{n_e^0}\right)\left(1-\frac{\chi_e}{\varepsilon}\right)}{2c^2m_e^2\gamma_0\left(\omega_e-k_ev_e^0\right)^2\left(\gamma_0^2\omega_l^2-\omega_{ce}^2\right)^2}, \tag{17}$$

where $k^2=\frac{\omega_r^2}{c^2}\varepsilon_p-k_r^2$, and $\varepsilon_p(=1-\frac{\omega_p^2}{\omega_l^2})$ is the permittivity of the plasma medium.

In the absence of beam (i.e., $\omega_{pb}=0$), the equation reduces to

$$\frac{\partial^2 E_r}{\partial r^2}+\frac{1}{r}\frac{\partial E_r}{\partial r}+k^2E_r=0. \tag{18}$$

The solution of the above equation is of Bessel's form which can be given as $E_r=A_nJ_0(k_nr)$, where $J_0(k_nr)$ is the zeroth order Bessel function of first kind. At $r=a_0$, the electric field should vanish, i.e., $E_r=0$ and the solution reduces to $J_0(k_na_0)=0$, i.e., $k_n=\frac{X_n}{a_0}(n=1,2,3\ldots..)$, X_n are the zeros of the Bessel function and n is the mode number.

In the presence of beam (i.e., $\omega_{pb}\neq 0$), a series of orthogonal sets of wave functions can be used to express the solution of the wave function with $E_r=\sum_m A_mJ_0(k_mr)$. We substitute this expression in Eq. (17) and multiply both sides of the equation by $rJ_0(k_nr)$ and then integrate it over r from 0 to a_0. On retaining only the dominant mode (i.e., $m=n$), we obtain

$$\frac{\omega_r^2}{c^2}-k_{1r}^2=\frac{\omega_{pb}^2e^2\left(k_e-\frac{\omega_e v_e^0}{c^2}\right)^2[i\omega_l\gamma_0(\nabla\phi_l)-(\omega_{ce}\times\nabla\phi_l)]^2\left(1+\frac{n_e^m}{n_e^0}\right)\left(1-\frac{\chi_e}{\varepsilon}\right)\alpha_1}{2m_e^2c^2\gamma_0\left(\omega_e-k_ev_e^0\right)^2\left(\gamma_0^2\omega_l^2-\omega_{ce}^2\right)^2}, \tag{19}$$

where $k_{1r}^2=k_r^2+k_n^2$ and $\alpha_1=\frac{\int_0^{r_e}rJ_0(k_mr)J_0(k_nr)dr}{\int_0^{a_0}rJ_0(k_mr)J_0(k_nr)dr}$, where r_e is the radius of the electron beam. At $r_e=a_0$, $\alpha_1=1$.

Equation (19) can be simplified to

$$\left(\omega_r^2-k_{1r}^2c^2\right)\left(\omega_e-k_ev_e^0\right)^2=\frac{\omega_{pb}^2e^2\left(k_e-\frac{\omega_e v_e^0}{c^2}\right)^2[i\gamma_0\omega_l(\nabla\phi_l)-(\omega_{ce}\times\nabla\phi_l)]^2\left(1+\frac{n_e^m}{n_e^0}\right)\left(1-\frac{\chi_e}{\varepsilon}\right)\alpha_1}{2m_e^2\gamma_0\left(\gamma_0^2\omega_l^2-\omega_{ce}^2\right)^2}. \tag{20}$$

Here, $\omega_r=k_{1r}c$ agrees to the radiation mode and $\omega_e=k_ev_e^0$ corresponds to the beam mode. The growth rate (Γ_{fr}) is obtained on solving the above expression by using $\omega_r=k_{1r}+\delta_{fr}$ and $\omega_e=k_ev_e^0+\delta_{fr}$, where δ_{fr} is the small frequency mismatch and taking its imaginary part

$$\Gamma_{fr}=\frac{\sqrt{3}}{2}\left(\frac{\omega_{pb}^2e^2A_l^2\left(k_e-\frac{\omega_e v_e^0}{c^2}\right)^2\left(1+\frac{n_e^m}{n_e^0}\right)\left(1-\frac{\chi_e}{\varepsilon}\right)\alpha_1}{4\left[\left(\frac{X_n}{a_0}\right)^2+k_r^2\right]^{\frac{1}{2}}m_e^2c\gamma_0^3\omega_l^2\left(1-\frac{\omega_{ce}^2}{\gamma_0^2\omega_l^2}\right)^2}\right)^{\frac{1}{3}}. \tag{21}$$

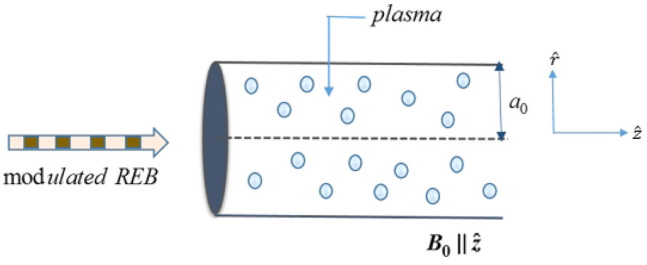


FIG. 2. Schematic of interaction of REB with the plasma pump wave in a cylindrical column (bound plasma).

In the case of finite geometry, the frequency of the output radiation depends on the modes and radius of the waveguide as

$$\omega_r=\left[\left(\frac{X_n}{a_0}\right)^2+k_r^2\right]^{\frac{1}{2}}c. \tag{22}$$

The efficiency of the output radiation ($\eta_{fr}=\frac{\Gamma_{fr}}{\omega_r}\frac{\gamma_0^3}{\sqrt{3}(\gamma_0-1)}$) can be expressed as

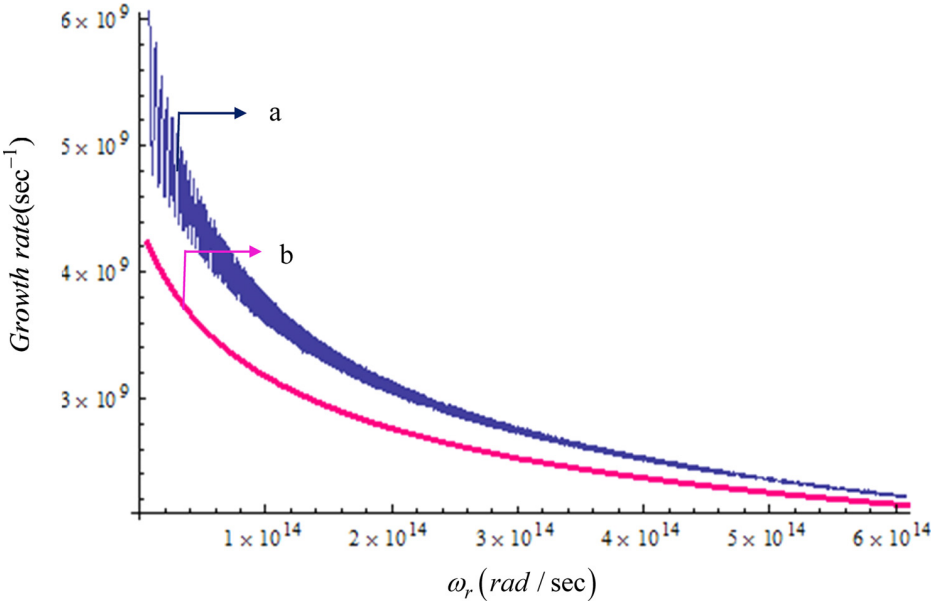


FIG. 3. Growth rate of the unstable wave as a function of radiation frequency for (a) modulated REB and (b) unmodulated REB.

$$\eta_{fr} = \frac{1}{2}\left(\frac{\omega_{pb}^2 e^2 A_l^2 \gamma_0^6 \left(k_e - \frac{\omega_e v_e^0}{c^2}\right)^2 \left(1 + \frac{n_e^m}{n_e^0}\right)\left(1 - \frac{\chi_e}{\varepsilon}\right)\alpha_1}{4\left[\left(\frac{X_n}{a_0}\right)^2 + k_r^2\right]^2 m_e^2 c^4 \omega_l^2 \left(1 - \frac{\omega_{ce}^2}{\gamma_0^2 \omega_l^2}\right)^2 (\gamma_0 - 1)^3}\right)^{\frac{1}{3}}. \tag{23}$$

For the infinite and finite geometries, the growth rate and efficiency of the output radiation is found to be influenced by the parameters of the beam (scales as one third power of the beam density) and the pump wave (scales as two third power of the electric field). For both the cases, at ω_{ce} is approximately equal to $\gamma_0\omega_l$, Cerenkov resonance occurs and the growth rate and efficiency of the THz radiation attain the maximum value. The frequency of the output radiation depends on the radius and modes of the cylindrical plasma column in the case of finite boundaries, whereas in the case of infinite, the dimensions of the system do not affect the radiation.

III. RESULTS AND DISCUSSION

The present analysis involves an analytical model for the study of the growth rate and efficiency of the terahertz radiation which is being radiated by the interaction of the laser premodulated relativistic electron beam with the Langmuir pump wave wiggler. The response of the beam and wiggler parameters on the output radiation has been studied by employing the following typical parameters, namely, electron beam velocity $v_e^0 = 0.95c$, relativistic gamma factor $\gamma_0 = 3.2$, beam energy $E_b = 1.12$ MeV, beam density $n_e^0 = 4 \times 10^{15}\text{cm}^{-3}$, beam current $I_e = 6.2 \times 10^{14}\text{gm}^{\frac{1}{2}}\text{cm}^{\frac{3}{2}}\text{s}^{-2}$, radius of the electron beam $r_e = 0.06$ cm, length $l_m = 0.75$ cm of modulator, drift space length $l_d = 1.5$ cm, pump wave frequency $\omega_l = 7 \times 10^{13}\text{rads}^{-1}$, and radius of the cylindrical waveguide $a_0 = 0.1$cm.

To investigate the effect of density modulation, the growth rate of the output radiation has been estimated in Fig. 3. From the figure, it can be seen that for both the finite and infinite geometries, the growth rate decreases with the increase of output radiation frequency. The maximum value of the growth rate achieved for an unmodulated beam is $4.2 \times 10^9\text{s}^{-1}$, whereas for a density modulated beam, the growth rate attains the maximum value of $6 \times 10^9\text{s}^{-1}$ which is nearly 43% higher than the unmodulated beam.

Figure 4 illustrates the influence of pump field and radiation frequency on the growth rate of the output radiation for different electron beam densities (i.e., $4 \times 10^{15}\text{cm}^{-3}$, $8 \times 10^{15}\text{cm}^{-3}$, and $12 \times 10^{15}\text{cm}^{-3}$) in the case of infinite geometry. The figure indicates that upon increasing the pump field and beam densities, the growth rate augments and scales as one-third of the beam density and two-third power of the pump field. The similar observation of increase of growth rate with beam density has been experimentally observed by Chang *et al.*[33] [cf. Fig. 4(b)]. Furthermore, the

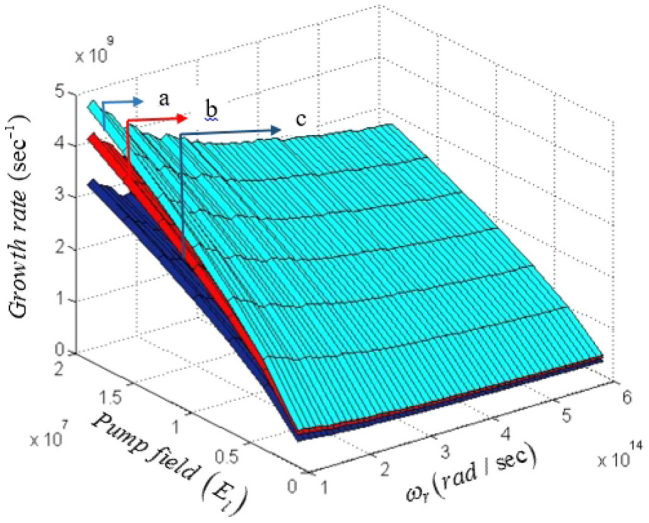


FIG. 4. Variation of growth rate with respect to the field amplitude of the pump wave and radiation frequency at different values of beam densities for (a) $4 \times 10^{15}\text{cm}^{-3}$, (b) $8 \times 10^{15}\text{cm}^{-3}$, and (c) $12 \times 10^{15}\text{cm}^{-3}$.

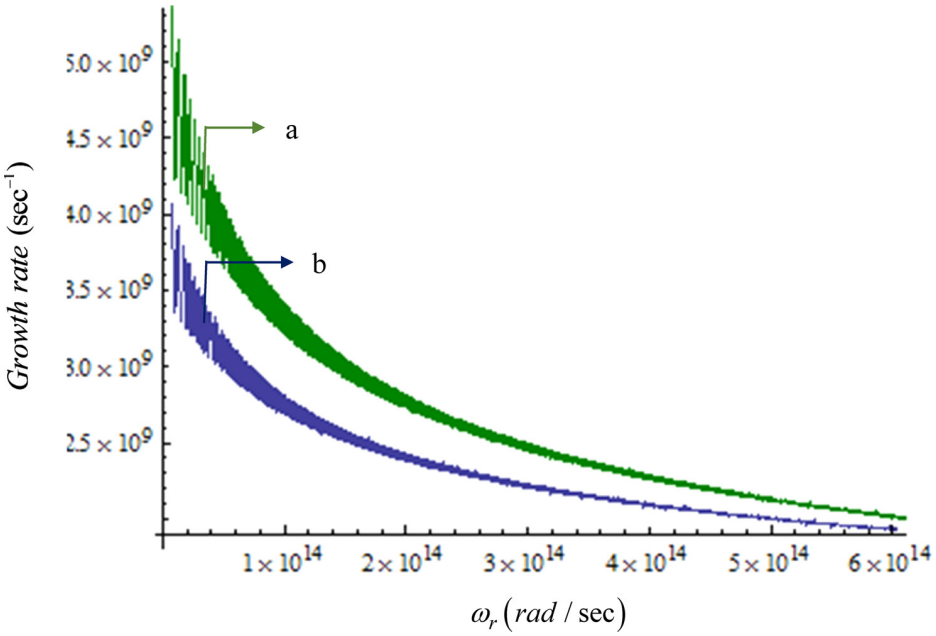


FIG. 5. Growth rate versus radiation frequency (a) in the presence of external magnetic field and (b) in the absence of magnetic field.

growth rate decreases with the increase of radiation frequency.

Figure 5 shows the growth rate variation as a function of radiation frequency with the effect of applied magnetic field. The figure shows that in the presence of magnetic field, the growth rate enhances to about 19%which is maximum and is achieved when the electron cyclotron frequency is comparable to the pump wave frequency.

In Fig. 6, we have illustrated the variation of the growth rate with the electron cyclotron frequency. Initially the growth rate increases with the magnetic field and then attains a maximum value at a point where the frequency of the pump wave becomes comparable to the electron cyclotron frequency and then starts decreasing with further increase of the magnetic field.

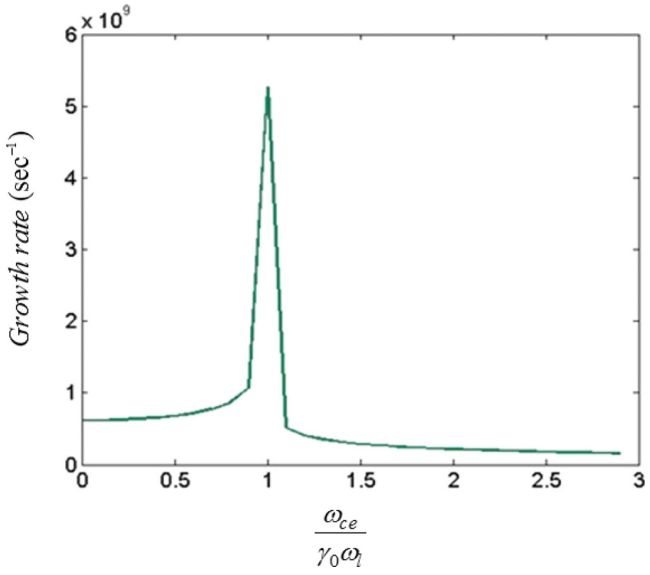


FIG. 6. Variation of growth rate as a function of normalized electron cyclotron frequency.

Using Eq. (22), the dependence of output radiation frequency on the radius and modes of the cylindrical plasma column has been shown in Fig. 7. The figure shows that the frequency of the radiation increases with the modes of the wave and decreases with the radius of the waveguide. This behavior is similar to the findings of Malik *et al.*[34] (cf. Fig. 6).

Figure 8 shows the variation of the growth rate with the radiation frequency for finite and infinite geometries. The decrease in growth is more in the case of finite geometry in comparison to the infinite case (the same behavior has been observed by Bhasin and Sharma[30]). As seen from Fig. 7 that the frequency of the radiation in the case of finite geometry depends on the parameters of the geometry. Hence, the growth is bounded by the dimensions of the cylindrical plasma column and increases with the increase in radius of the cylinder. Further, as the modes increases, the frequency increases and the growth rate decreases.

The efficiency of the emitted output radiation as a function of the radiation frequency has been plotted in Fig. 9.

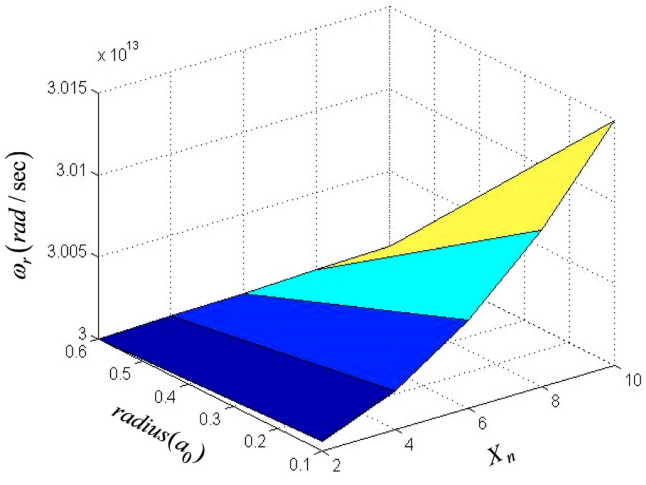


FIG. 7. Radiation frequency as a function of radius of the cylindrical plasma column and modes of the wave.

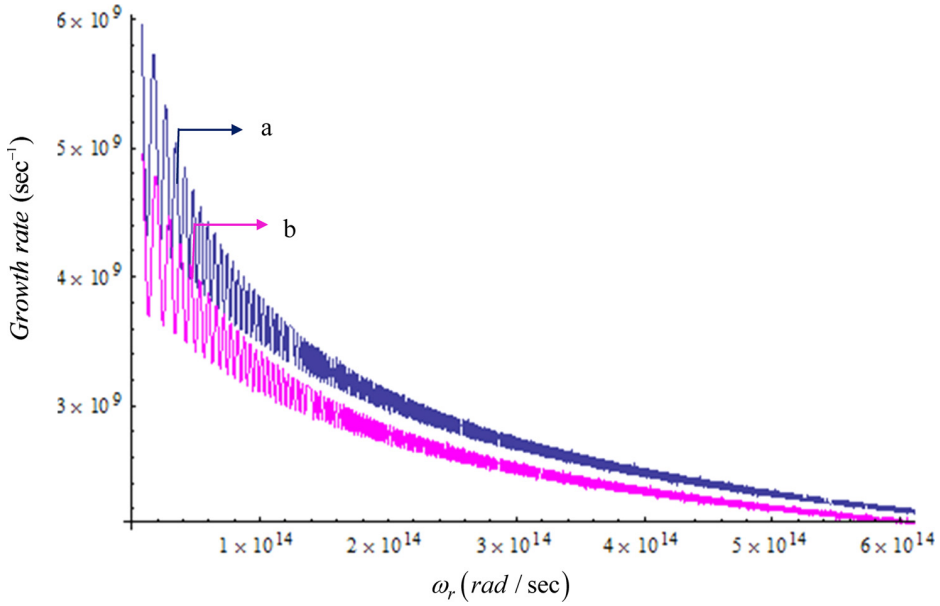


FIG. 8. Growth rate as a function of radiation frequency for (a) infinite geometry and (b) finite geometry.

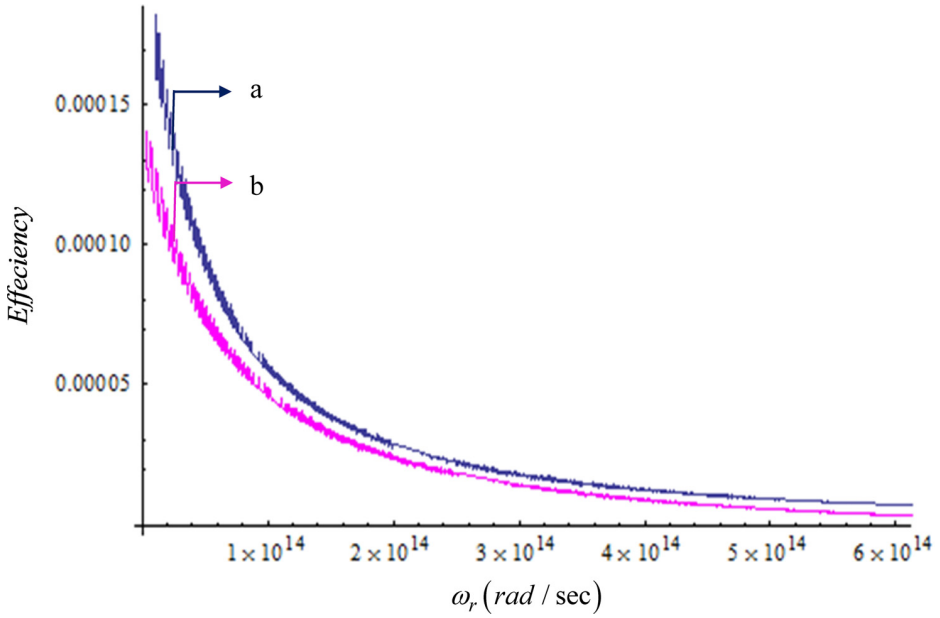


FIG. 9. Efficiency as a function of radiation frequency for (a) infinite geometry and (b) finite geometry.

The efficiency falls off with the frequency and is low for the cylindrical plasma column as compared to the unbound plasma.

IV. CONCLUSION

The emission of terahertz radiation has been studied with the effect of beam premodulation and in the presence of applied static magnetic field. The growth rate and efficiency of the radiation wave have been analyzed for both infinite and finite geometries. The growth rate decreases monotonically with the frequency of the output radiation. In the case of finite geometry, i.e., cylindrical plasma column, the growth rate decreases as the mode increases and is bounded by the dimensions of the plasma column; thus, the maximum value attained is less as compared to the infinite geometry. The modulation of the relativistic electron beam leads to an increase in the growth rate to about 43%. The growth of the radiation wave depends on the parameters of beam and the wiggler wave, and hence it can be improved by increasing the density of the electron beam and amplitude of the pump wave. The applied static magnetic field leads to the enhancement of the growth rate until the electron cyclotron frequency becomes comparable to the pump wave frequency. The scheme, thus, holds importance in obtaining tunable radiation with high power.

[1]J. F. Federici, B. Schulkin, F. Huang, D. Gary, R. Barat, F. Oliveira, and D. Zimdars, Semicond. Sci. Technol. **20**, S266 (2005).
[2]W. L. Chan, J. Deibel, and D. M. Mittleman, Rep. Prog. Phys. **70**, 1325 (2007).
[3]M. Tonouchi, Nat. Photonics **1**, 97 (2007).
[4]M.-A. Brun, F. Formanek, A. Yasuda, M. Sekine, N. Andoand, and Y. Eishii, Phys. Med. Biol. **55**, 4615 (2010).
[5]A. R. Orlando and G. P. Gallerano, J. Infrared Millimeter, Terahertz Waves **30**, 1308 (2009).
[6]G. H. Welsh and K. Wynne, Opt. Express **17**, 2470 (2009).

[7]Y.-S. Lee, T. Meade, V. Perlin, H. Winful, T. B. Norris, and A. Galvanauskas, Appl. Phys. Lett **76**, 2505 (2000).
[8]L. Bhasin and V. K. Tripathi, Phys. Plasmas **16**, 103105 (2009).
[9]S. C. Sharma, J. Panwar, and R. Sharma, Contrib. Plasma Phys. **57**, 167 (2017).
[10]J. G. Neumann, R. B. Fiorito, P. G. O'Shea, H. Loos, B. Sheehy, Y. Shen, and Z. Wu, J. Appl. Phys. **105**, 053304 (2009).
[11]A. Gover, Phys. Rev. Spec. Top.—Accel. Beams **8**, 030701 (2005).
[12]S. C. Sharma and P. Malik, Phys. Plasmas **22**, 043301 (2015).
[13]C. Joshi, T. Katsouleas, J. M. Dawson, Y. T. Yan, and J. M. Slater, IEEE J. Quantum Electron. **23**, 1571 (1987).
[14]M. Singh, S. Kumar, and R. P. Sharma, Phys. Plasmas **18**, 022304 (2011).
[15]S. C. Sharma, J. Sharma, A. Bhasin, and R. Walia, J. Plasma Phys. **78**, 635 (2012).
[16]S. H. Zolghadr, S. Jafari, and A. Raghavi, Phys. Plasmas **23**, 053104 (2016).
[17]S. Hussain, R. Kishor Singh, and R. P. Sharma, Phys. Plasmas **23**, 073120 (2016).
[18]N. Yugami, T. Higashiguchi, H. Gao, S. Sakai, K. Takahashi, H. Ito, and Y. Nishida, Phys. Rev. Lett. **89**, 065003 (2002).
[19]V. K. Tripathi and C. S. Liu, IEEE Trans. Plasma Sci. **18**, 466 (1990).
[20]H. Mehdian, S. Jafari, and A. Hasanbeigi, Plasma Phys. Controlled Fusion **52**, 055005 (2010).
[21]R. Hedayati, S. Jafari, and S. Batebi, Plasma Phys. Controlled Fusion **57**, 085007 (2015).
[22]V. Bandhu Pathak, D. Dahiya, and V. K. Tripathi, J. Appl. Phys. **105**, 013315 (2009).
[23]M. Kumar and V. K. Tripathi, Phys. Plasmas **18**, 053105 (2011).
[24]V. B. Gildenburg and N. V. Vvedenskii, Phys. Rev. Lett. **98**, 245002 (2007).
[25]M. Kumar and V. K. Tripathi, Phys. Plasmas **19**, 073109 (2012).
[26]W. L. Kruer, *The Physics of Laser Plasma Interactions* (Addison-Wesley Publishing, 1988).
[27]G. Basudev, *Basic Plasma Physics* (Alpha Science International, 2014).
[28]V. K. Tripathi and C. S. Liu, IEEE Trans. Plasma Sci. **17**, 583 (1989).
[29]D. Tripathi, R. Uma, and V. K. Tripathi, Phys. Plasmas **19**, 093105 (2012).
[30]A. Bhasin and S. C. Sharma, Phys. Plasmas **14**, 073102 (2007).
[31]J. D. Jackson, *Classical Electrodynamics* (John Wiley & Sons, 1962).
[32]C. S. Liu and V. K. Tripathi, *Interaction of Electromagnetic Waves with Electron Beams and Plasmas* (World Scientific Publishing Co. Pvt. Ltd., Singapore, 1994).
[33]R. P. H. Chang, Phys. Rev. Lett. **35**, 285 (1975).
[34]A. K. Malik, H. K. Malik, and S. Kawata, J. Appl. Phys. **107**, 113105 (2010).

ELK
Asia Pacific Journals

ELK Asia Pacific Journals – 978-93-85537-06-6
ARIMPIE -2017

Optimization of process parameters in Abrasive Flow Machining

Parvesh Ali
Deptt of Mechanical Engg
Delhi Technological University, Delhi-110042, India

R.S Walia
Deptt of Mechanical Engg
Delhi Technological University, Delhi-110042, India
waliaravinder@yahoo.com

Abhinav Sri
Deptt of Mechanical Engg
Guru Nanak Dev Engg. College, Ludhiana, India

Sehijpal Singh
Deptt of Mechanical Engg
Guru Nanak Dev Engg. College, Ludhiana, India

Abstract—In recent competitive environment there is a huge demand of products with high accuracy of dimensional tolerances with less defects and better life cycle. Both of these can be achieved by providing a better nano finishing technique. Abrasive Flow Machining is an unconventional process used for finishing of complex geometries, which uses a viscous media and pressurizes it to flow through the complex path. The abrasives in the media cause micro cutting action and enhance the surface quality of the product. This process is dependent upon various process parameters such as pressure, number of cycles, concentration of abrasives etc. So focusing on different types of abrasives and work pieces is also a point of concern for enhancing the surface quality of the product. This paper involves experimental observations and was further optimized to obtain better combinations of these parameters.

Keywords— Material Removal, Surface finish, abrasives

I. INTRODUCTION

In the present competitive environment there is need to develop the components with high dimensional accuracy, high finishing, better life cycle and economical. Abrasive Flow machining is an unconventional process used for finishing of the components with complex geometries, difficult to machine using any other conventional technique [1]. This process uses a non Newtonian fluid to carry the abrasives and also for finishing. Media used in this process has a good fluidity and is viscous in nature. It is mixed with abrasives and acts as a cutting tool during the finishing process. This process simultaneously performs finishing and deburring. The material removes in form of small chips. It provides better finishing in comparison to the machined and casted surface. AFM process enhances the engine performance, component life and also improves the fatigue strength of the product. This process has its applications in automobile industry, medical components and in aerospace industry. It provides finishing up to 0.05 micrometer [2]. It saves up to 90% of finishing time in comparison to other conventional finishing process [3]. AFM can be divided in three categories on the basis of their working and configuration i.e One way AFM [4], Two way AFM [5] and Orbital AFM [6]. One way AFM has the advantage of faster cycle processing, better control of media temperature and simple in construction while two way AFM is known for its better processing control and also it can finish both inner and outer diameter. Brar et.al. found that in Electrochemical assisted Abrasive Flow Machining material removal improves with increase in voltage [7]. Material removal depends on many parameters but it mainly depends upon the abrasive to media ratio among the various process parameters [8].

In Abrasive flow machining various studies has been conducted to know the after effect of the parameters on the system efficiency. The process parameters of this process may be extrusion pressure, number of cycles, amount of abrasives etc. This paper discusses the effects of some important parameters such as different types of abrasives and work pieces on the system efficiency. This information has an important role in enhancing the process performance. Some experiments were performed to get the best optimum condition by using Taguchi L_9 orthogonal Array.

Work piece : This experiment uses three types of work pieces i.e Aluminium, Mild steel and Cast Iron. The work piece has a standard dimension according to the Kohut [9].

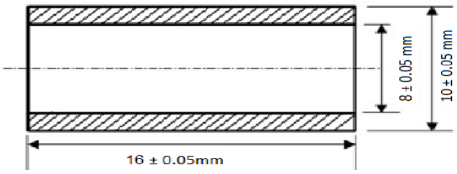


Fig 1. Test piece [9]

A. Abrasives**:**

Three types of abrasives were used for this experiment i.e silicon carbide, aluminium oxide and titanium oxide. The mesh size of all abrasives was 400 microns.



Fig 2. Abrasive Particles (Al_2O_3, SiC, TiO_2)

II. RESPONSE CHARACTERISTICS

This experiment studies the effects of parameters (i.e work piece, abrasives, extrusion pressure, number of cycles) on the response (i.e material removal) -

Material removal means the extent of material removed after performing certain number of cycles in the finishing process. Material removal was calculated from the formula

MR= weight of the workpice before machining – weight of the workpiece after machining

III. SCHEME OF EXPERIMENTS

The experiments were modeled to analyze the impact of the parameters being taken on the material removal. An L_9 orthogonal array was used for analysis having 3 levels and 8 degree of freedom. The values of process parameters at three different levels are shown in the table:

Table 1. Values of process parameters at three different levels

Symbol	Process Parameters	Unit	L_1	L_2	L_3
W	Work piece	----------	Al	MS	CI
A	Type of Abrasive	----------	Al_2O_3	SiC	TiO_2
P	Pressure	MPa	15	20	25
N	No. Of Cycle	----------	2	4	6
Abrasive Mesh Size-400, Initial surface roughness -1.12 – 4.95 microns Al- Aluminium, SiC- Silicon Carbide, TiO_2- Titanium Oxide					

Table 2. The L_9 Orthogonal Array for 4 parameters and three levels with their responses

S.N	R.O	Parameters				R			S/N Ratio (db)
		W	A	P	N	R_1	R_2	R_3	
		1	2	3	4				
1	1	1	1	1	1	M_{11}	M_{12}	M_{13}	S/N_1
2	4	1	2	2	2	M_{21}	M_{22}	M_{23}	S/N_2
3	7	1	3	3	3	M_{31}	M_{32}	M_{33}	S/N_3
4	2	2	1	2	3	M_{41}	M_{42}	M_{43}	S/N_4
5	5	2	2	3	1	M_{51}	M_{52}	M_{53}	S/N_5
6	8	2	3	1	2	M_{61}	M_{62}	M_{63}	S/N_6
7	3	3	1	3	2	M_{71}	M_{72}	M_{73}	S/N_7
8	6	3	2	1	3	M_{81}	M_{82}	M_{83}	S/N_8
9	9	3	3	2	1	M_{91}	M_{92}	M_{93}	S/N_9
Total						Σ	Σ	Σ	
R_1, R_2, R_3 shows responses for three repeated trials. M_{ij} are the measured response value									

B. Experimentation

The four process parameters workpiece, type of abrasives, Pressure, No. of Cycle were selected as in table 1. The process parameters are alterd according to the values as shown in table 2. Experiments were performed on the basis of the test condition mentioned in the L_9 OA (Table 2). For every trial condition each experiment was repeated fo three times. Thus operation was performed on the twenty seven work-pieces having their initial surface in range of (1.1 – 2.96 micron). For all trial conditions and their repetitions, the material removal was measured and recorded in Table 3.

Table 3. Experimental results of responses

E. N.	**R. O**	**(MR) (mg)**			**S/N ratio (db)**
		R_1	**R_2**	**R_3**	
1	1	8.93	14.4	7.8	17.7
2	4	7.7	8.8	5.9	15.34
3	7	8.7	4.56	7.6	14.06
4	2	13.6	13.1	17.6	21.4
5	5	2.8	3.2	2.8	7.54
6	8	3.9	3.1	3.4	8.92
7	3	6.74	8.4	10.5	16.46
8	6	3.1	3.8	4.5	9.53
9	9	1.5	2.0	1.9	3.14
Σ		56.9	61.4	61.9	
		TMR =Overall mean of MR = 60.06mg			

IV. ANALYSIS AND DISCUSSION OF RESULT

Taguchi methodology was used to analyze the data. For each parameter average values and S/N ratio of response was calculated for three different levels.

C. Effect on Material Removal

For every parameter at each level, average values of material removal and Signal to noise ratio are calculated and is shown in table 4.

Table 4. Average values of material removal and Signal to noise ratio

Process Parameter	Level	Work piece (W)		Type of Abrasives(A)		Pressure (P)		No. of Cycle	
Type of Data		Raw Data	S/N Ratio	Raw Data	S/N Ratio	Raw Data	S/N Ratio	Raw Data	S/N Ratio
Average Values	L_1	8.26	15.7	11.23	18.52	5.88	12.05	5.03	9.46
	L_2	7.06	12.62	4.73	10.80	8.01	13.29	6.50	13.57
	L_3	4.72	9.71	4.07	8.70	6.15	12.68	8.50	14.99
Main Effects	L_2-L_1	-1.2	-3.08	-6.5	-7.72	2.13	1.24	1.47	4.01
	L_3-L_2	-2.34	-2.91	-0.66	-2.1	-1.86	-0.61	2.00	1.42
Difference $(L_3-L_2)–(L_2-L_1)$		-1.14	0.17	5.84	5.62	-3.99	-1.85	0.53	-2.59

- Figure 3 shows that the material removal at first level (i.e Al) is maximum. The reason of this may be that aluminium is softer in comparison to mild steel and cast iron, that's why maximum material removal is obtained in case of aluminium. Figure also shows that minimum material removal is obtained for cast iron. It is due to the brittleness of the cast iron. Material removal for mild steel lies between the Aluminium and cast iron.

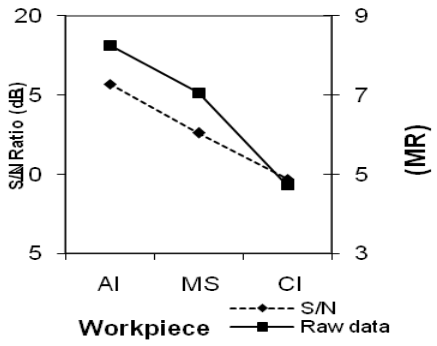


Fig 3. Effect of type of work piece on the S/N Ratio and MR

- Material removal depends on the hardness of abrasive. Among the three (Al_2O_3, SiC, TiO_2) silicon carbide is hardest but it is brittle also. Due to brittleness, it breaks after impact. Aluminium is softer than SiC but more ductile, hence Al_2O_3 causes more material removal. Figure 4 show that Al_2O_3 causes more material removal and minimum occurs in case of TiO_2.

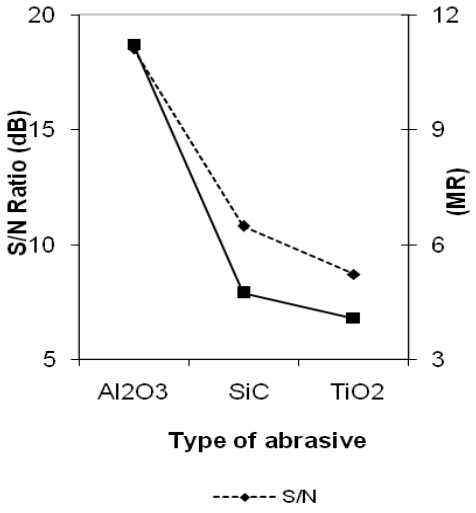


Fig 4. Effect of type of abrasives on the S/N Ratio and MR

- Figure 5 represents that as pressure is increased, material removal also increases due to increase in the impacting force. At certain level indenting capacity of abrasive starts decreasing due to lack of sharpness, hence material removal decreases. Therefore material removal first grows with extrusion pressure and after attaining maximum value, start decreasing with increase in extrusion pressure.

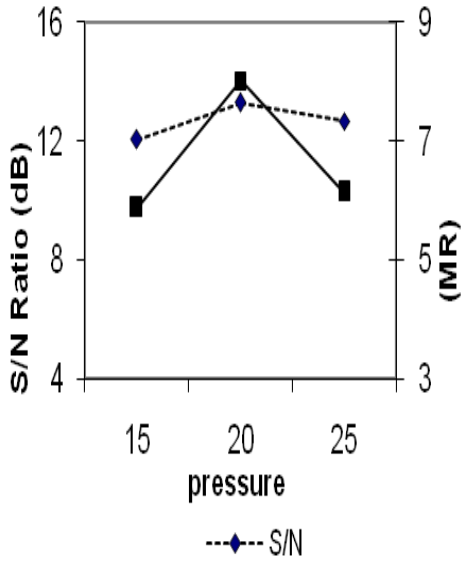


Fig 5. Effect of extrusion pressure on the S/N Ratio and MR

- Figure 6 shows that on increasing the number of cycle , material removal enhances continuously, but after certain level the rate of material removal decreases because as the number of cycle increases, the temperature of media rises and due to that viscosity of media decreases and number of active abrasive particles decreases.

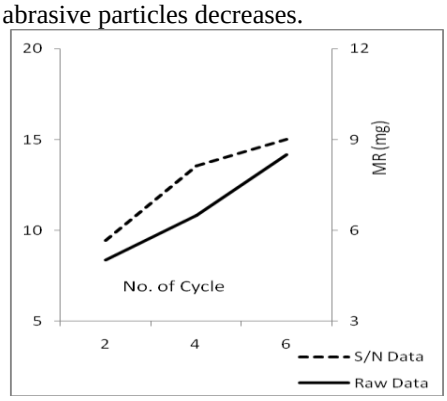


Fig 6. Effect of number of cycle on the S/N Ratio and MR

D. Selection of optimum levels

There may be many parameters which effect the system efficiency but some of them effects significantly to it. This makes important to identify the significant parameters. Here (ANOVA) was used for this purpose. Table 5 & 6 represents the pooled versions of Analysis of variance for both raw and S/N data for the response. These tables represents that the parameters taken above are significant.

Table 5. Pooled ANOVA (S/N) (MR)

Source	SS	DOF	V	F-Ratio	SS'	P%
Work piece	53.79	2	26.89	23.19	51.47	20.21
Type of Abrasive	160.34	2	80.17	69.15	158.2	60.24
Pressure	2.32	*	*	*	0	*
No. of Cycle	49.70	2	24.85	21.43	47.39	18.68
e	2.32	2	1.16	--------	9.28	0.87
Total (T)	266.15	8	-------	-------	266.1	100
*Significant at 95% confidence level, F critical =19(Tabular Value)						

Table6. Pooled ANOVA (Raw Data) (MR)

Source	SS	DOF	V	F-Ratio	SS'	P%
Work piece	58.44	2	29.22	8.78	51.79	12.20
Type of Abrasive	281.67	2	140.83	42.36	275.02	58.81
Pressure	24.27	2	12.13	3.65	17.62	5.06
No. of Cycle	54.70	2	27.35	8.22	48.05	11.42
e	59.84	18	3.32	-----	86.44	12.49
Total (T)	478.93	26	*	-----	478.93	100
*Significant at 95% confidence level, F critical =3.5546(Tabular Value)						

E. Estimation of optimum response characteristics

The results of confirmation experiments are used to validate optimal result. The optimal conditions of the parameters are identified from the prefered response characteristics. Using the confirmation experiments average value of the response is calculated. The results should lie within the 95% confidence level, CI_{CE}. However it is not necessary that the average value of quality characteristic obtained from the confirmation experiments lie within 95% confidence interval, CI_{POP}. After the study it was found that the optimum condition for the maximum material removal is $W_1A_1P_2N_3$ for raw data and S/N data. For the confirmation experiments on the base of data given the optimal settings have been taken as $W_1A_1P_2N_3$ for material material. Based on the optimum selection optimum

response parameters of the material removal have been estimated with the confidence intervals as further.

F. Material Removal (MR)

The mean of the MR at optimum condition is estimated as $\mu = \bar{W}_1 + \bar{A}_1 + \bar{P}_2 + \bar{N}_3 - 3\bar{T}$ …….(Equation No. 1)

= 6.68 mg

$\bar{W}_1$ represents the average material removal for work piece at first level = 8.26 mg

$\underline{A}_1$ represents the average material removal for abrasive at first level= 11.23 mg

$\bar{P}_2$ represents the average material removal for pressure at second level= 8.01 mg

$\underline{N}_3$ represents the average material removal for number of cycle at third level = 8.50 mg

On putting the values, Mean MR = 15.96 mg

(CI_{CE}) and (CI_{POP}) is calculated by using the following equation

$$CI_{CE} = \sqrt{F_a(1, f_e)V_e\left[\frac{1}{n_{eff}} + \frac{1}{R}\right]} \quad \text{……..…….(Equation No. 2)}$$

$$CI_{POP} = \sqrt{\frac{F_a(1, f_e)V_e}{n_{eff}}} \quad \text{……...(Equation No. 3)}$$

Where $F_\alpha(1, f_e)$ = The F- ratio at the confidence level of (1-α) against DOF 1 and error degree of freedom f_e = 3.55 (Tabulated Value)

f_e = error DOF = 18

N = Total no of result =27 (Experiment =9, repeat =3)

R = Sample size = 3

V_e = Error variance = 3.32

$$n_{eff} = \frac{N}{1+[\text{DOF associated in the estimate of mean responce}]}$$

(Equation No.4)

= 3

So, $CI_{CE} = \pm 2.80$

And $CI_{POP} = \pm 1.98$

The 95% confirmation interval of predicted optimal range is:

Mean MR – CI_{CE} <MR< MR + CI_{CE}

13.16< MR <18.76

The 95% confirmation interval of the predicted mean is :

Mean MR – CI_{POP} <MR< MR + CI_{POP}

13.98<MR<17.94

G. Confirmation Experiments

For the maximum material removal, the optimal levels of the process parameter are $W_1A_1P_2N_3$.

W_1= Work piece at first level =Al

A_1 = Abrasive at first level =Al_2O_3

P_2= Extrusion Pressure at Second level =20 MPa

N_3 = No. of Cycle at third level =9

V. CONCLUSION

There is the possibility to improve the material removal by working on the optimal conditions. This improves the system efficiency and also minimizes the loss during the process. The process must be hybridized to get the best results.The result shows that the contribution of the workpiece is 12.2 %, type of abrasive is 58.81 %, pressure is 5.06 %, no. of cycle is 11.42 % on the response (material removal).

References

[1] Brar B.S,Walia R.S, Singh V.P " State of Art Abrasive Flow Machining", National Conference on Advancements and Futuristic Trends in Mechanical and materials engineering (AFTMME' 10), Yadavindra College of Engg., Talwandi Sabo, Distt. Bathinda, Punjab, India, 2010.

[2] Jain R.K, Jain V.K "Abrasive fine finishing process-A review", Int J Manuf Sci Prod vol.2 (1), 1999, pp. 55-68.

[3] Singh S "Studies in metal finishing with magnetically assisted abrasive flow machining", Phd thesis,IIT Roorkee, 2002.

[4] Rhoades L.J, Kohut T.A, Nokovich N.P, Yanda D.W "Unidirectional abrasive flow machining", US patent number 5,367,833, Nov 29th, 1994.

[5] Rhoades L.J, Kohut T.A "Reversible Unidirectional AFM", US patent number 5,070,652, Dec 10th, 1991.

[6] Sankar. M. Ravi, Jain V.K, J.Ramkumar "Abrasive flow machining (AFM): An Overview", INDO-US WORKSHOP on Smart Machine Tools, Intelligent Machining Systems and Multi-scale Manufacturing. December, 2008.

[7] Brar B.S, Walia R.S, Singh V.P "Electro chemical machining in the aid of Abrasive Flow Machining Process", International Journal of Surface Engineering & Materials Technology, Vol. 2 (1), 2012 , pp. 5-9.

[8] Brar B.S, Walia R.S, Singh V.P, Singh Mandeep "Development of a robust Abrasive Flow Machining Process set up ", International Journal of Surface Engineering & Materials Technology, Vol. 1 (1), 2011.

[9] Kohut T "Surface finishing with Abrasive flow machining", SME Technical Paper ,Proc. 4th International Aluminium extraction Technology seminar ,Washington D.C., 1988, pp. 35-42.

Accepted Manuscript



Short communication

Rebuttal to 'Fully-uncoupled independent control of frequency and condition of oscillation: A caution'

Raj Senani, Data Ram Bhaskar, Shanti Swaroop Gupta, Bhopendra Singh

PII: S1434-8411(16)31340-1
DOI: http://dx.doi.org/10.1016/j.aeue.2017.07.010
Reference: AEUE 51967

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

Received Date: 27 November 2016
Accepted Date: 8 July 2017

Please cite this article as: R. Senani, D. Ram Bhaskar, S. Swaroop Gupta, B. Singh, Rebuttal to 'Fully-uncoupled independent control of frequency and condition of oscillation: A caution', *International Journal of Electronics and Communications* (2017), doi: http://dx.doi.org/10.1016/j.aeue.2017.07.010

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

Title page with author details

Rebuttal to 'Fully-uncoupled independent control of frequency and condition of oscillation: A caution'

Raj Senani
Division of Electronics and Communication Engineering,
Netaji Subhas Institute of Technology,
Sector 3, Dwarka. New Delhi 110078, India.
senani@ieee.org

Data Ram Bhaskar
Department of Electronics and Communication Engineering,
Delhi Technological University,
Shahbad Daulatpur,
Main Bawana Road, Delhi 110042, India
drbset@gmail.com

Shanti Swaroop Gupta
Department of Industrial Policy and Promotion,
Ministry of Commerce and Industry,
Government of India,
Udyog Bhawan, New Delhi 110011, India
ssgupta1@gmail.com

Bhopendra Singh
Roorkee College of Engineering,
Roorkee-Haridwar Anal Road,
Uttarakhand 247667, India.
mrbhupendra99@gmail.com

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Rebuttal to 'Fully-uncoupled independent control of frequency and condition of oscillation: A caution'

Raj Senani[a,*], Data Ram Bhaskar[b], Shanti Swaroop Gupta[c] and Bhopendra Singh[d]

Abstract: Based upon the experimental results of the two CFOA-based fully-uncoupled sinusoidal oscillators proposed earlier by Bhaskar-Gupta-Senani-Singh, recently Abuelma'atti and Khalifa issued a caution that the classification of this class of oscillators providing independent control of condition of oscillation (CO) and frequency of oscillation (FO) through separate elements, as *fully-uncoupled*, does not appear to be correct. In this communication, we present the results of a recheck of the experimental data of Abuelma'atti and Khalifa and show that their experimental results are *not repeatable* and in *most of the cases*, the oscillations do not even exist and hence, the measurement data as given by them is misleading. Furthermore, it is shown that the measured oscillation frequencies, as shown by Abuelma'atti and Khalifa in four tables, do not tally with those calculated from their own formula as well as the experimental results taken by us. This communication thus, refutes the claims made in the quoted paper of Abuelma'atti and Khalifa published in this journal earlier.

[a] Division of Electronics and Communication Engineering, Netaji Subhas Institute of Technology, Sector-3, Dwarka. New Delhi 110078, India
[b] Department of Electronics and Communication Engineering, Delhi Technological University, Shahbad Daulatpur, Main Bawana Road, Delhi 110042, India
[c] Department of Industrial Policy and Promotion, Ministry of Commerce and Industry, Udyog Bhawan, New Delhi 110011, India
[d] Roorkee College of Engineering, Roorkee-Haridwar Anal Road, Uttarakhand 247667

* Corresponding author

Indexing terms: Sinusoidal Oscillators, Current feedback op-amps

1. Introduction

Recently, Abuelma'atti and Khalifa [1] issued a *caution* (for another similar caution, see [3]) that the classification of a class of oscillators providing independent control of condition of oscillation (CO) and frequency of oscillation (FO) through separate elements, as *fully-uncoupled*, does not appear to be correct. It may be mentioned that such fully-uncoupled oscillators have been proposed in the past using a variety of active building blocks such as Current Conveyors, Current feedback op-amps (CFOA), operational trans-resistance amplifiers (OTRA) and current differencing buffered amplifiers (CDBA) etc.; see [2]-[13]

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

and the references[1] cited therein. Furthermore, the first author of [1] himself has been the proposer of many such fully-uncoupled oscillators in the past, for instance, see [5]-[7].

Since Abuelma'atti and Khalifa [1] have framed all their arguments and deductions upon the experimental results (given by them in four tables) of the two CFOA-based fully-uncoupled oscillators proposed earlier by the Bhaskar-Gupta-Senani-Singh in [2], we have checked and tried to repeat the *experimental results* given by Abuelma'atti and Khalifa in [1] for *exactly the same* component and parameter values as taken in [1]. To our utter surprise, we found a majority of the experimental results given by Abuelma'atti and Khalifa [1] to be non-repeatable; some of them *cannot* and did not even exist!

In this communication, therefore, we show how most of the experimental data (given in four tables of [1]) on which Abuelma'atti and Khalifa [1] have framed their conclusions are either wrong or non-existing and hence, the caution served by them against fully-uncoupled oscillators of Bhaskar-Gupta- Senani-Singh [2] (and others) is ill-founded.

2. The CFOA-based fully-uncoupled SRCOs of [2] and the arguments of Abuelma'atti and Khalifa as in [1]

Since the caution issued by Abuelma'atti and Khalifa [1] was based upon two fully-uncoupled SRCOs proposed earlier by Bhaskar-Gupta-Senani-Singh in [2], it is worthwhile to reproduce here the quoted oscillator circuits along with their relevant details, for ready reference. The two fully-uncoupled sinusoidal oscillators (from [2]) under question are reproduced here in Fig.1. Assuming the CFOAs to be characterized by $i_y = 0, v_x = v_y, i_z = i_x$ *and* $v_w = v_z$, both the oscillators have the same characteristic equation (CE) which is given by:

[1] For some recent works on sinusoidal oscillators see [14]-[17]

ACCEPTED MANUSCRIPT

$$s^2+\frac{s}{C_1}\left(\frac{1}{R_0}-\frac{1}{R_1}\right)+\frac{1}{R_2R_3C_1C_2}=0 \qquad \text{eqn. (6) of [1]}$$

From the above CE, the CO and FO are given by:

$$\text{CO: } (R_1-R_0)\le 0 \qquad \text{eqn. (7) of [1]}$$

$$\text{FO: } f_0=\frac{1}{2\pi}\left[\frac{1}{R_2R_3C_1C_2}\right]^{1/2} \qquad \text{eqn. (8) of [1]}$$

However, as per Abuelma'atti and Khalifa [1], since the roots of the CE of eqn. (6) are given by:

$$s_{1,2}=\frac{-1/C_1(1/R_0-1/R_1)\pm\sqrt{1/C_1^2((1/R_0)-(1/R_1))^2-4(1/C_1C_2R_2R_3)}}{2} \qquad \text{eqn. (9) of [1]}$$

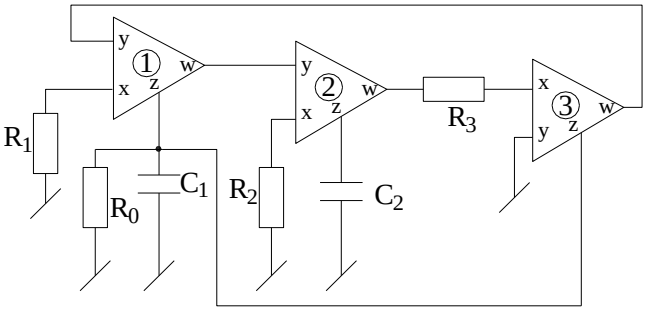
using eqn. (9), the CO is, therefore, given by

$$(R_1-R_0)\le 0 \qquad \text{eqn. (10) of [1]}$$

$$\text{and } \frac{1}{C_1^2}\left(\frac{1}{R_0}-\frac{1}{R_1}\right)^2<\frac{1}{C_1C_2R_2R_3} \qquad \text{eqn. (11) of [1]}$$

while the FO is given by

$$\omega_0^2=\frac{1}{C_1C_2R_2R_3}-\frac{1}{4}\frac{1}{C_1^2}\left(\frac{1}{R_0}-\frac{1}{R_1}\right)^2 \qquad \text{eqn. (12) of [1]}$$



ACCEPTED MANUSCRIPT

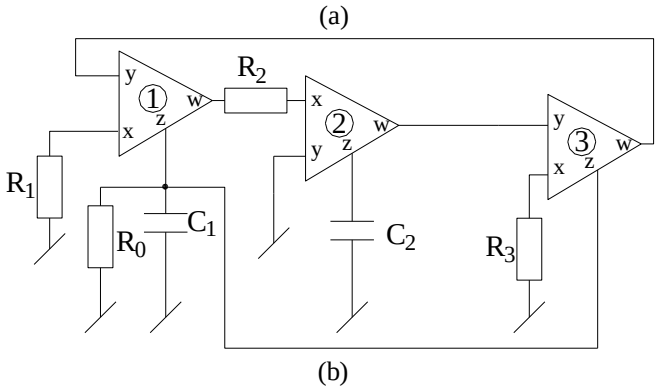


Fig. 1 Fully-uncoupled single resistance controlled oscillators proposed by Bhaskar-Gupta-Senani-Singh [2]

Thereupon, Abuelma'atti and Khalifa in [1], have made the following claim (see the paragraph just after eqn. (12) of [1]):

Claim 1: "Inspection of Eqs. (9)-(11) clearly shows that the two circuits of Fig. 1 do not enjoy the attractive feature of fully uncoupled independent control of frequency and condition of oscillation. Eq. (11) clearly shows that the frequency of oscillation is dependent on resistances R_0 and R_1 that are involved in the condition of oscillation given by Eqs. (10) and (11). Thus, the two oscillator circuits may not be described as enjoying fully uncoupled independent control of the frequency and condition of oscillation."

To substantiate their contention, Abuelma'atti and Khalifa [1] have given the experimental results of both the circuits of Fig. 1 in four Tables where they have claimed and shown that in all the cases FO varies when R_0 is varied, keeping R_1 fixed and vice versa, in both the oscillator circuits. Their second claim is (see the first sentence of the second paragraph of Section 3 of [1]):

Claim 2: "Inspection of Tables 1-4 clearly shows that the frequency of oscillation of the two circuits of Fig. 1 is dependent on both the resistors R_0 and R_1. Thus, contrary to the conclusion made in [2] the two circuits cannot be described as enjoying the attractive feature of fully uncoupled independent control of frequency and condition of oscillation."

ACCEPTED MANUSCRIPT

In the next section we first show that their claim 2 is based upon incorrect (and in many cases, non-existent) data and therefore, is not correct. Subsequently, we will show how their claim 1 is also not justified.

The experimental data, of the two oscillator circuits of Fig.1 under question, was presented by Abuelma'atti and Khalifa [1] in four Tables which were claimed to have been obtained by experimental results (see the first paragraph of section 3 of [1])

In section 3 of [1] Abuelma'atti and Khalifa claimed that:

Claim 3: "In most of the cases the peak-to-peak output voltage is about few volts and the output looks as an excellent sinusoidal wave".

3. The Rebuttal to the experimental results and caution of [1]

We have carefully re-checked the so-called *experimental results* given in Tables 1-4 of [1] with exactly the same component values as taken by Abuelma'atti and Khalifa and which are reproduced here in the second and third columns of the *augmented* Tables1-4. To clarify our observations, we have reconstructed Tables 1-4 which now contain second and third columns exactly same as in [1] whereas the last five columns contain results of our experimental measurements, observations and comments. In each of the four augmented Tables, there is one additional row marked 'X' which corresponds to the case when the concerned circuit generates pure sine wave (these cases were however, not included by Abuelma'atti and Khalifa in their paper [1]).

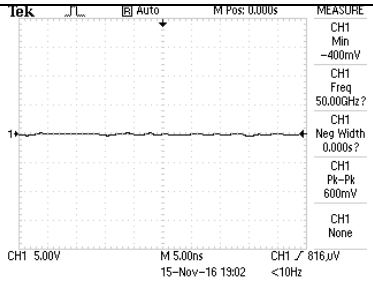
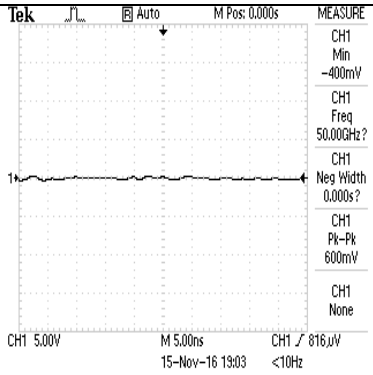
It may be mentioned that the non-ideal characteristic equation (CE) of both the circuits under question was presented in [2] (reproduced here in Appendix 1, for ready reference). The roots of this non-ideal CE have now also been calculated for all the cases depicted in all the augmented Tables 1-4. It may, however, be noted by comparing the entries of columns 4 and

ACCEPTED MANUSCRIPT

5 that the basic nature of the roots does not change *significantly* due to the influence of the various CFOA parasitics[2]. As a consequence, the inference and conclusions drawn from the experimental and other data given in the augmented Tables 1-4 would remain unchanged.

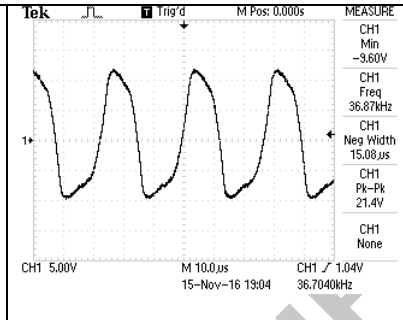
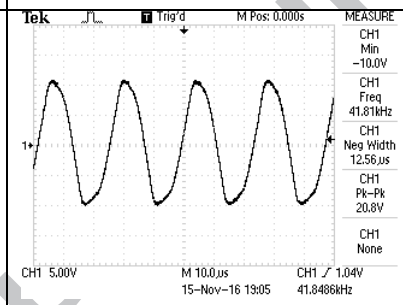
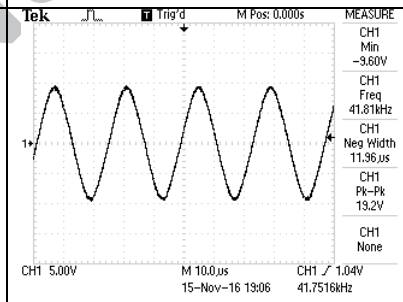
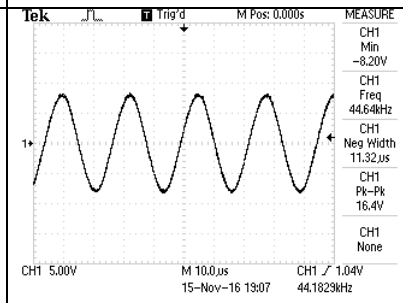
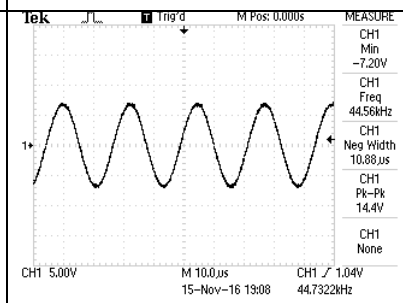
Augmented Table 1

Variation of the frequency of oscillation of the circuit of Fig.1 (a) with R_1 (R_2 = 10kΩ, R_3 = 1 kΩ, C_1 = C_2 =1nF and R_0 = 10 kΩ)

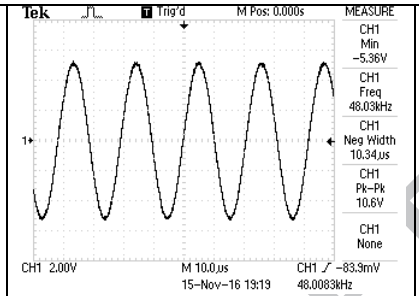
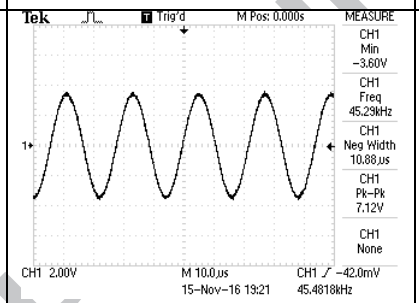
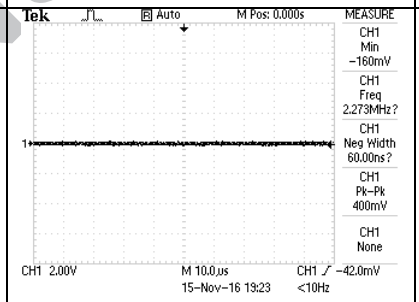
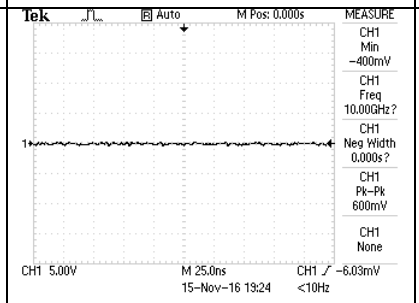
Data reproduced as in Table 1 of Abuelma'atti-Khalifa [1]			Experimental data taken by us				
S. No	R_1 (kΩ)	f_0 (kHz)	Roots of CE (calculated from eqn. 9 of [1])	*Dominant roots calculated from nonideal CE of [2]	f_0 (kHz) calculated from eqn. (12) of [1]	Experimentally obtained output waveform	Remarks
01	1.0	114.0	7.7016×10^5 and $1.2984 \times 10^{5;}$ both the roots are in right half of s-plane and are real!	7.2135×10^5 and 1.2350×10^5; both the roots are in right half of s-plane and are real!	Imaginary !		Circuit *cannot* and really does not oscillate.
02	1.5	98.0	$(2.8335 \pm j1.4040) \times 10^5$,complex conjugate roots deeply in right half of s-plane	$(2.7006 \pm j1.2750) \times 10^5$; complex conjugate roots deeply in right half of s-plane	22.33		No oscillation

[2] In [2] a nonideal expression of the FO incorporating the various parasitics was derived (see eqn. (7) of [2]). Also, the variation of FO w.r.t. resistor R_2 was experimentally determined and plotted in Fig. 3 of [2]. Furthermore, in the last column of Table 1 of [2], the values of FO determined from the nonideal CE were also tabulated. Both these values of FO were found to be quite close but both of them were, on the other hand, smaller than the corresponding ideal values of the FO. Therefore, it is clear that the nonidealities of the circuit elements (CFOAs) do affect the frequency of oscillation and its measured values turn out to be somewhat smaller than the ideal values.

ACCEPTED MANUSCRIPT

03	2.0	88.0	(2.0000±j2.4495)x10^5 ;complex conjugate roots still in right half of s-plane	(1.920217830 ± j2.2885) x10^5; complex conjugate roots still in right half of s-plane	38.94		Circuit oscillates; heavily distorted wave 36.7040 KHz, 21.4V (p-p)
04	3.0	76.0	(1.1670±j2.9391)x10^5 ;complex conjugate roots still in right half of s-plane	(1.126962775 ± j2.7676) x10^5; complex conjugate roots still in right half of s-plane	46.73		Circuit oscillates; distorted wave 41.8486 KHz, 20.8V (p-p)
05	4.3	72.5	(0.6627±j3.0920)x10^5 ;complex conjugate roots still in right half of s-plane	(0.6409 ± j2.9193) x10^5; complex conjugate roots still in right half of s-plane	49.16		Distorted sinusoidal wave; 41.7516 KHz, 19.2V (p-p)
06	5.6	71.5	(0.3794±j3.1394)x10^5 ;complex conjugate roots still in right half of s-plane	(0.3786 ±j2.9651) x10^5; complex conjugate roots still in right half of s-plane	49.88		Nearly sinusoidal wave; 44.1829 KHz, 16.4 V (p-p)
07	6.8	70.0	(0.2353±j3.1535)x10^5 ;complex conjugate roots still in right half of s-plane	(0.2248 ±j 2.9808) x10^5; complex conjugate roots still in right half of s-plane	50.14		Almost sinusoidal wave; 44.7322 KHz, 14.4 V (p-p)

ACCEPTED MANUSCRIPT

X	7.96		(0.1385±j3.1592)x10^5 ;complex conjugate roots slightly in right half of s-plane	(0.1199 ± j2.9870) x10^5; complex conjugate roots slightly in right half of s-plane	50.23		Pure sinusoidal waveform ; 48.0083KHz, 10.6 V (p-p)
08	8.2	68.0	(0.1098±j3.1604)x10^5 ;complex conjugate roots still in right half of s-plane	(0.1019 ± j 2.9877) x10^5; complex conjugate roots still in right half of s-plane	50.25		Somewhat distorted sine wave; 45.4818 KHz, 7.12 V (p-p)
09	10.0	66.0	(0.0000±j3.1623)x10^5 ;roots on ± j ω axis in s-plane	(-0.0057 ± j 2.9896) x10^5; roots slightly in the left half of s-plane	50.28		No oscillation
10	20.0	62.3	(-0.2500±j3.1524) x10^5 i.e., roots are in left half of s-plane!	(-0.2519 ± j 2.9792) x10^5; roots are in left half of s-plane!	50.12		No oscillation

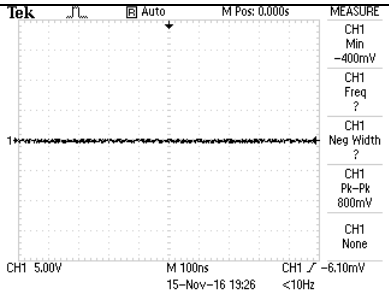
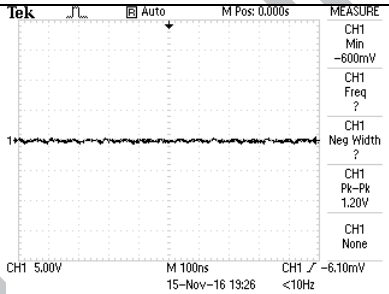
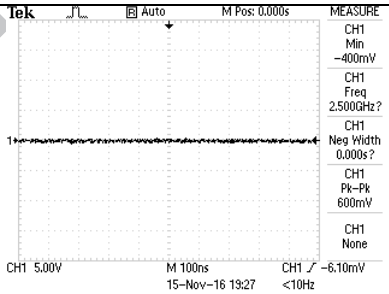
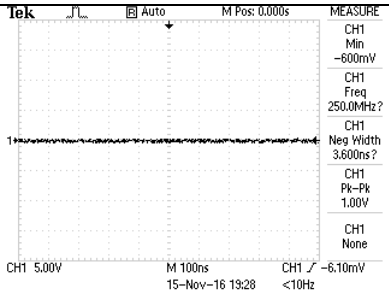
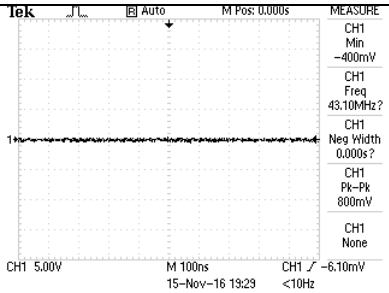
*because of limited space available in the table only the two dominant roots responsible for oscillatory mode have been included here for the sake of easy comparison with the roots calculated in column 4. In all cases, the remaining two roots have been found to be in far away in the left half of the s-plane.

Augmented Table 2

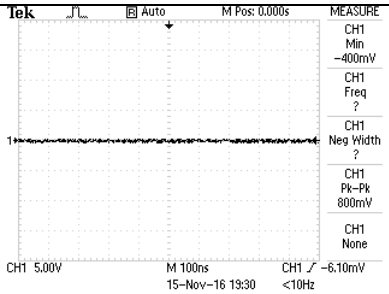
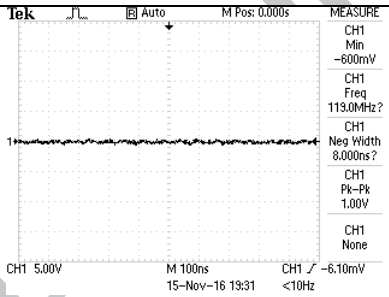
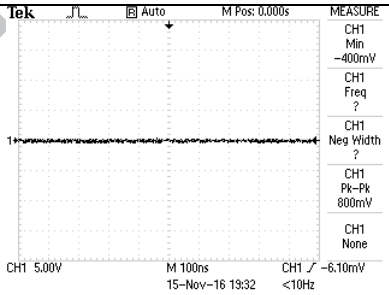
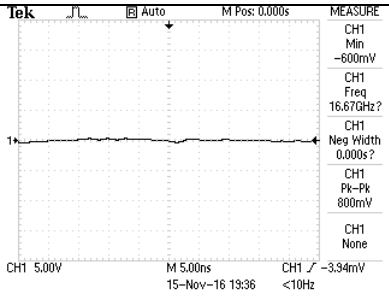
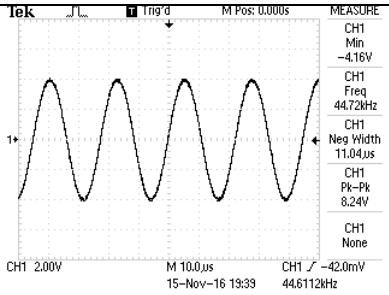
Variation of the frequency of oscillation of the circuit of Fig.1 (a) with R_0 (R_2 = 10kΩ, R_3 = 1 kΩ, C_1 = C_2 =1nF and R_1 = 10 kΩ)

Data reproduced as in Table 2 of Abuelma'atti-Khalifa [1]			Experimental data taken by us				
S. No	R_0 (kΩ)	f_0 (kHz)	Roots of CE (calculated from eqn. 9 of [1])	*Dominant roots calculated from nonideal CE of [2]	f_0 (kHz) calculated from eqn. (12) of [1]	Experimentally obtained output waveform	Remarks

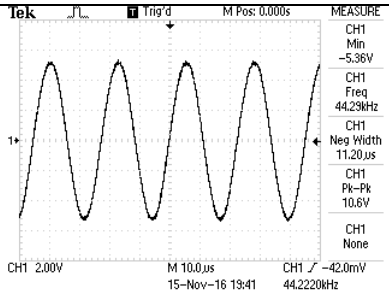
ACCEPTED MANUSCRIPT

01	1.0	152.0	-7.7016×10^5; -1.2984×10^5 i.e., roots are on negative real axis of s-plane	-7.7886×10^5 and -1.1513×10^5; roots are on negative real axis of s-plane	Imaginary!		No question of oscillatory behavior; circuit does not oscillate.
02	1.5	135.0	$-2.8335 \pm j1.4040) \times 10^5$; complex conjugate roots in left half of s-plane	$(-2.8165 \pm j\ 1.0116) \times 10^5$; complex conjugate roots in left half of s-plane	22.33		Circuit does not oscillate.
03	2.0	108.0	$(-2.0000 \pm j2.4449) \times 10^5$;complex conjugate roots in left half of s-plane	$(-1.9898 \pm j\ 2.2341) \times 10^5$; complex conjugate roots in left half of s-plane	38.94		Circuit does not oscillate.
04	3.0	97.3	$(-1.1670 \pm j2.9391) \times 10^5$;complex conjugate roots in left half of s-plane	$(-1.1631 \pm j\ 2.7554) \times 10^5$; complex conjugate roots in left half of s-plane	46.73		Circuit does not oscillate.
05	4.3	88.0	$(-0.6627 \pm j3.0920) \times 10^5$;complex conjugate roots in left half of s-plane	$(-0.6632 \pm j\ 2.9158) \times 10^5$; complex conjugate roots in left half of s-plane	49.16		Circuit does not oscillate.

ACCEPTED MANUSCRIPT

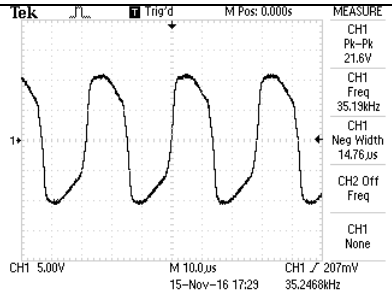
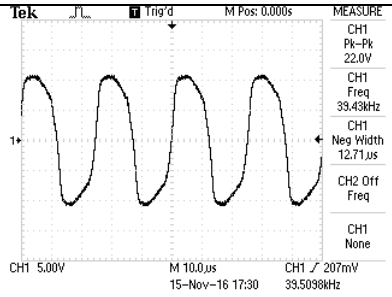
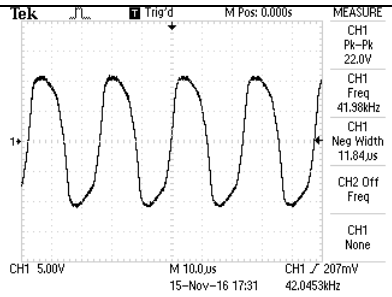
06	5.6	80.0	(-0.3794±j3.1394)x10^5 ;complex conjugate roots in left half of s-plane	(-0.3954 ± j 2.9637) x10^5; complex conjugate roots in left half of s-plane	49.88		Circuit does not oscillate.
07	6.8	75.6	(-0.2353±j3.1535)x10^5 ;complex conjugate roots in left half of s-plane	(-0.2391 ± j 2.9803) x10^5; complex conjugate roots in left half of s-plane	50.14		Circuit does not oscillate.
08	8.2	71.0	(-0.1098±j3.1604)x10^5 ;complex conjugate roots in left half of s-plane	(-0.1146 ± j 2.9875) x10^5; complex conjugate roots in left half of s-plane	50.25		Circuit does not oscillate
09	10.0	66.0	(-0.0000±j3.1623)x10^5 ;Imaginary conjugate roots on ± jω axis in s-plane	(-0.0057 ± j 2.9896) x10^5; roots slightly in the left half of s-plane	50.28		Circuit does not oscillate
X	13.10		(0.1184±j3.1601)x10^5 ;complex conjugate roots slightly in right half of s-plane	(0.1116 ± j 2.9874) x10^5; complex conjugate roots slightly in right half of s-plane	50.24		Generates good quality sine waves; 44.6112 KHz, 8.24 V (p-p)

ACCEPTED MANUSCRIPT

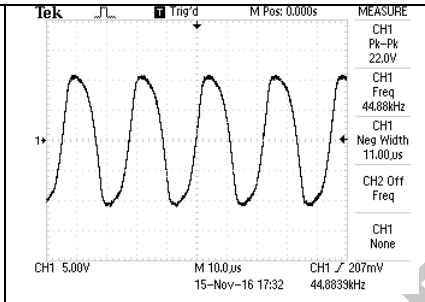
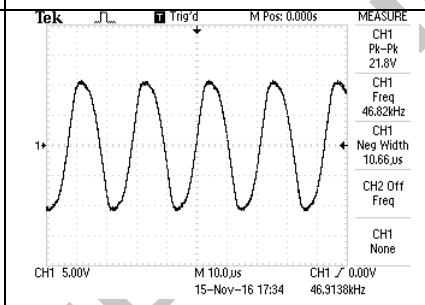
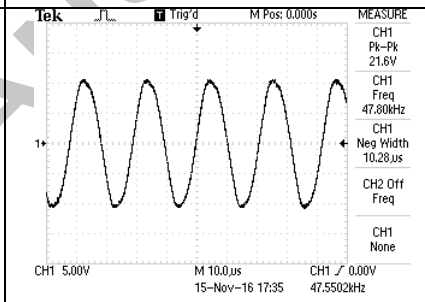
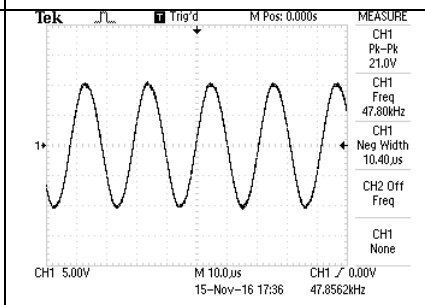
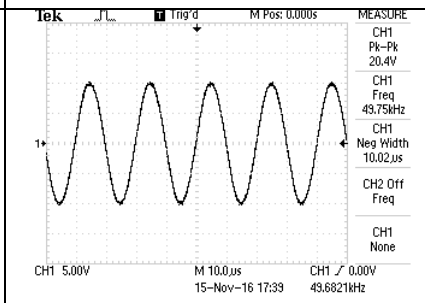
10	20.0	52.0	(0.2500±j3.1524)x10^5 ;complex conjugate roots in right half of the s-plane	(0.2422 ± j 2.9795) x10^5; complex conjugate roots in right half of the s-plane	50.12		Distorted sinewave; 44.2220 KHz 10.6 V (p-p)

Augmented Table 3

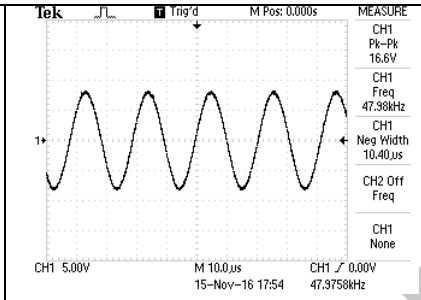
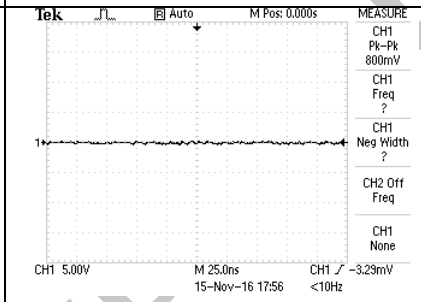
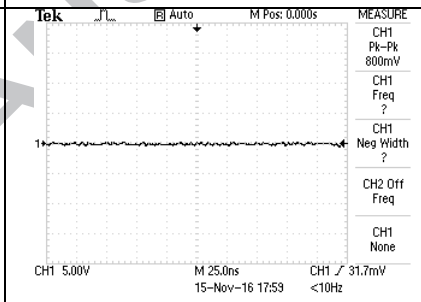
Variation of the frequency of oscillation of the circuit of Fig.1 (b) with R_1 (R_2 = 10kΩ, R_3 = 1 kΩ, C_1 = C_2 =1nF and R_0 = 10 kΩ)

Data reproduced as in Table 3 of Abuelma'atti-Khalifa [1]			Experimental data taken by us				
S. No	R_1 (kΩ)	f_0 (kHz)	Roots of CE (calculated from eqn. 9 of [1])	*Dominant roots calculated from nonideal CE of [2]	f_0 (kHz) calculated from eqn. (12) of [1]	Experimentally obtained output waveform	Remarks
01	1.0	117.4	7.7016x10^5 and 1.2984x10^5; real roots in right half of the s-plane	7.2135 x10^5 and 1.2350 x10^5; real roots in right half of the s-plane	Imaginary !		Circuit oscillates with heavily distorted waveform; 35.2468 KHz with 21.6 V DC
02	1.5	101.4	(2.8335±j1.4040)x10^5 ;complex conjugate roots in right half of s-plane	(2.7006 ± j 1.2750) x10^5; complex conjugate roots in right half of s-plane	22.33		Circuit oscillates with heavily distorted waveform; 39.5098 KHz 22.0 V (p-p)
03	2.0	89.0	(2.0000±j2.4495)x10^5 ;complex conjugate roots in right half of s-plane	(1.9202 ± j 2.2885) x10^5; complex conjugate roots in right half of s-plane	38.94		Circuit oscillates with quite distorted waveform; 42.0453 KHz 22.0 V (p-p)

ACCEPTED MANUSCRIPT

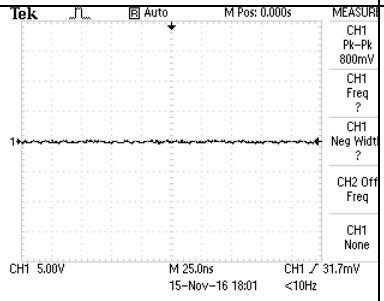
04	3.0	78.5	(1.1670±j2.9391)x10^5 ;complex conjugate roots in right half of s-plane	(1.1269 ± j 2.7676) x10^5; complex conjugate roots in right half of s-plane	46.73		Circuit oscillates with distorted waveform; 44.8839 KHz 22.0 V (p-p)
05	4.3	70.2	(0.6627±j3.0920)x10^5 ;complex conjugate roots in right half of s-plane	(0.6409 ± j 2.9193) x10^5; complex conjugate roots in right half of s-plane	49.16		Circuit oscillates with distorted waveform; 46.9138 KHz 21.8 V (p-p)
06	5.6	67.0	(0.3794±j3.1394)x10^5 ;complex conjugate roots in right half of s-plane	(0.3786 ± j 2.9651) x10^5; complex conjugate roots in right half of s-plane	49.88		Circuit oscillates with less distorted waveform; 47.5502 KHz 21.6 V (p-p)
07	6.8	64.6	(0.2353±j3.1535)x10^5 ;complex conjugate roots in right half of s-plane	(0.2248 ± j 2.9808) x10^5; complex conjugate roots in right half of s-plane	50.14		Circuit oscillates with nearly sinusoidal waveform; 47.8562 KHz 21.0V (p-p)
X	7.73		(0.1468±j3.1589)x10^5 ;complex conjugate roots slightly in right half of s-plane	(0.1382 ± j 2.9862) x10^5; complex conjugate roots slightly in right half of s-plane	50.22		Circuit oscillates with sinusoidal waveform; 49.6821 KHz 20.4 V (p-p)

ACCEPTED MANUSCRIPT

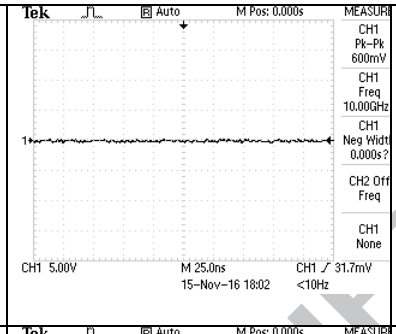
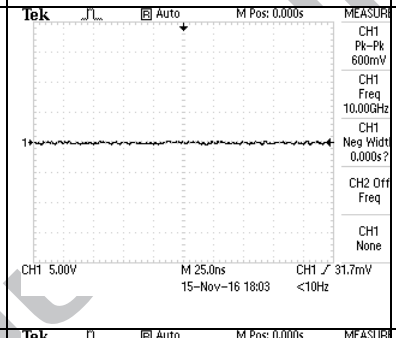
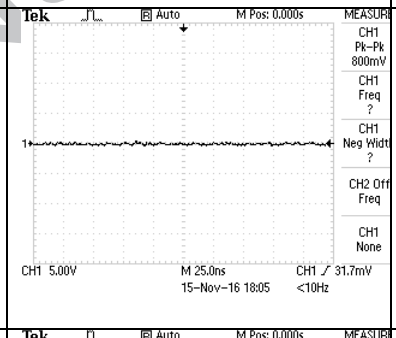
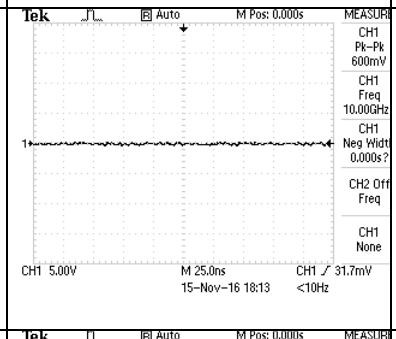
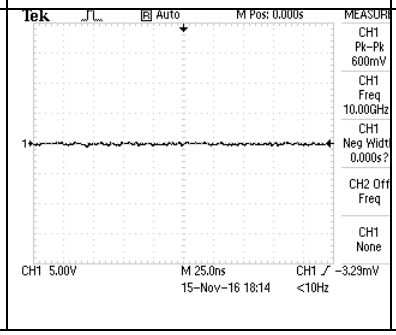
08	8.2	63.0	(0.1098±j3.1604)x10^5 ;complex conjugate roots in right half of s-plane	(0.1019 ± j 2.9877) x10^5; complex conjugate roots in right half of s-plane	50.25		Circuit oscillates with sinusoidal waveform; 47.9758 KHz 16.6 V (p-p)
09	10.0	61.3	(0.0000±j3.1623)x10^5 ;Imaginary conjugate roots on ± j ω axis	(-0.0057 ± j 2.9896) x10^5; roots slightly in the left half of s-plane	50.28		No start up; circuit does not oscillate.
10	20.0	56.0	(-0.2500±j3.1524)x10^5 ;complex conjugate roots in left half of s-plane	(-0.2519 ± j 2.9792) x10^5; complex conjugate roots in left half of s-plane	50.12		Circuit can not and does not oscillate.

Augmented Table 4

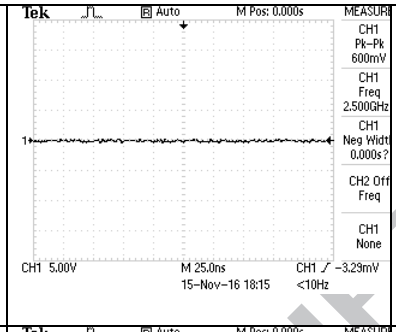
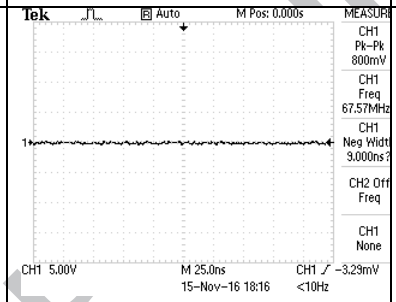
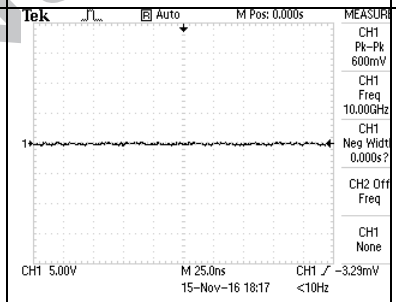
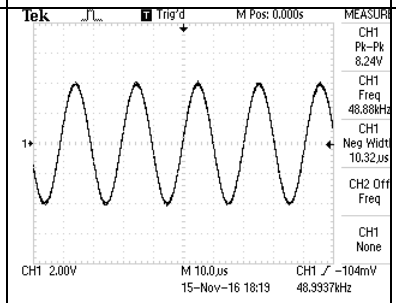
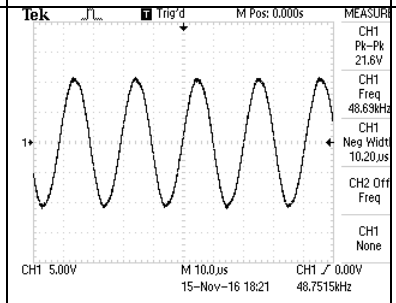
Variation of the frequency of oscillation of the circuit of Fig.1 (b) with R_0 (R_2 = 10kΩ, R_3 = 1 kΩ, C_1 = C_2 =1nF and R_1 = 10 kΩ)

Data reproduced as in Table 4 of Abuelma'atti-Khalifa [1]			Experimental data taken by us				
S. No	R_0 (kΩ)	f_0 (kHz)	Roots of CE (calculated from eqn. 9 of [1])	*Dominant roots calculated from nonideal CE of [2]	f_0 (kHz) calculated from equation (12) of [1]	Experimentally obtained output waveform	Remarks
01	1.0	112.0	-7.7016x10^5 and - 1.2984x10^5 ;roots on negative real axis of s-plane	-7.7886 x10^5 and -1.1513 x10^5; roots on negative real axis of s-plane	imaginary		No question of oscillatory behavior; circuit does not oscillate.

ACCEPTED MANUSCRIPT

02	1.5	106.0	(-2.8335±j1.4040) $x10^5$, ;complex conjugate roots in left half of the s-plane	(-2.8165 ± j 1.0116) $x10^5$; complex conjugate roots in left half of the s-plane	22.33		Circuit does not oscillate.
03	2.0	100.1	(-2.0000±j2.4449) $x10^5$;complex conjugate roots in left half of the s-plane	(-1.9898 ± j 2.2341) $x10^5$; complex conjugate roots in left half of the s-plane	38.94		Circuit does not oscillate.
04	3.0	90.0	(-1.1670±j2.9391) $x10^5$;complex conjugate roots in left half of the s-plane	(-1.1631 ± j 2.7554) $x10^5$; complex conjugate roots in left half of the s-plane	46.73		Circuit does not oscillate.
05	4.3	80.6	(-0.6627±j3.0920) $x10^5$;complex conjugate roots in left half of the s-plane	(-0.6632 ± j 2.9158) $x10^5$; complex conjugate roots in left half of the s-plane	49.16		Circuit does not oscillate.
06	5.6	73.8	(-0.3794±j3.1394) $x10^5$;complex conjugate roots in left half of the s-plane	(-0.3954 ± j 2.9637) $x10^5$; complex conjugate roots in left half of the s-plane	49.88		Circuit does not oscillate.

ACCEPTED MANUSCRIPT

07	6.8	68.8	(-0.2353±j3.1535) $x10^5$;complex conjugate roots in left half of the s-plane	(-0.2391 ± j 2.9803) $x10^5$; complex conjugate roots in left half of the s-plane	50.14		Circuit does not oscillate.
08	8.2	64.4	(-0.1098±j3.1604) $x10^5$;complex conjugate roots in left half of the s-plane	(-0.1146 ± j 2.9875) $x10^5$; complex conjugate roots in left half of the s-plane	50.25		Circuit does not oscillate.
09	10.0	60.5	(-0.0000±j3.1623) $x10^5$;Imaginary conjugate roots on ± j ω axis	(-0.0057 ± j 2.9896) $x10^5$; roots slightly in the left half of s-plane	50.28		No start up; circuit does not oscillate.
X	12.75		(0.1079±j3.1604)$x10^5$;complex conjugate roots slightly in right half of s-plane	(0.1012 ± j 2.9877) $x10^5$; complex conjugate roots slightly in right half of s-plane	50.24		Circuit produces sinusoidal waveform; 48.9937 KHz, 8.24 V(p-p)
10	20.0	47.0	(0.2500±j3.1524)$x10^5$;complex conjugate roots in right half of the s-plane	(0.2422 ± j 2.9795) $x10^5$; complex conjugate roots in right half of the s-plane	50.12		Circuit oscillates with some distortion; 48.7515 KHz, 21.6 V(p-p)

From a quick glance at all the tables, it is found that, contrary to the claims of Abuelma'atti and Khalifa in [1], the experimental results presented by them in Tables1-4 of [1] are non-

ACCEPTED MANUSCRIPT

repeatable and ***most of them do not even exist!*** From a detailed perusal of the data/information given in the augmented Tables 1 to 4, the following is revealed:

(i) Contrary to the statement of Abuelma'atti and Khalifa in [1] (see claim 1 above) that *in most of the cases the peak-to-peak output voltage is about few volts and the output looks as an excellent sinusoidal wave*, in most of the cases of their Tables 1 to 4, when the resistors R_0 and R_1 responsible for CO were varied as per the details given in Tables 1-4, it has been found that the facts are contrary to their claim! In fact, in most of the cases of Augmented Tables 2 and 4 and in some cases of Augmented Tables 1 and 3, the roots of the CE are either on real axis or in the left half of the s-plane. Consequently, there is no question of the circuits oscillating in such cases and they have actually been found *not to oscillate*! In view of this, the *claimed* measured oscillation frequency data of Abuelma'atti and Khalifa [1] as reproduced here in the third column of augmented of Tables 1-4 is extremely surprising because they do not exist!

(ii) It is also apparent that Abuelma'atti and Khalifa [1] have not bothered about the shape of the generated wave while recording the oscillation frequency by varying the resistors R_0 and R_1 involved in the CO. Note that an acceptable sinusoidal waveform should have negligibly small percentage THD (ideally zero) which means that the CO should be adjusted such that the roots of the CE must have their locations in the vicinity of $j\omega$ -axis. It is obvious from the evidence collected by us, as shown in the augmented Tables 1-4 here, that this aspect has been completely ignored by Abuelma'atti and Khalifa [1] while choosing the values of R_0 and R_1 and in measuring the frequency of oscillation since the roots of the CE in majority of the cases are in violation to the above quoted basic requirements.

ACCEPTED MANUSCRIPT

(iii) In all those cases when the roots of the CE are deep inside the right half of the s-plane, the circuits either produce a heavily distorted signal or the outputs simply lie in saturation. In either case, the measurement of frequency is inappropriate and questionable. Hence, the measured frequencies shown by Abuelma'atti and Khalifa [1] for all such cases of their Tables 1-4 (as in the third column in the augmented Tables 1-4 here) are questionable.

(iv) The only acceptable sine wave signals are obtained in those cases when the variable condition-setting resistor either of R_0 and R_1 (as the case may be) has values close to the fixed CO-setting resistor for each of the recorded measurements. It is not surprising therefore that these cases are the ones as reported by Bhaskar-Gupta-Senani-Singh earlier in [2].

(v) In the sixth column of augmented Tables 1-4, the value of the oscillation frequency is determined from formula of eqn. (12) which Abuelma'atti and Khalifa [1] proclaimed as the correct expression for FO. Ironically, in none of the cases the frequencies calculated from this formula match Abuelma'atti-Khalifa's own measured values given in column 2 of their Tables 1-4 of [1] (as reproduced in column 3 of augmented Tables 1-4 here).

(vi) Above all, one fails to understand why one should deliberately try to change the condition setting resistors R_0 and R_1 once these have been adjusted to produce sinusoidal oscillations? Is it just to prove that the oscillation frequency is dependent upon these resistor values also? Note that these resistance values are to be chosen such that the generated oscillations are as close to a sinewave as possible in which case, their contribution in formula of eqn. (12) will be negligibly small and therefore, oscillation frequency calculated either from formula of eqn. (12) or from formula of eqn. (4) both will not differ much. An

ACCEPTED MANUSCRIPT

examination of the data at serial no. X in augmented Tables 1 to 4 can be seen to confirm this contention.

(vii) A question now arises whether the CO and FO given by Eq. (7) and (8) are correct or those in Eq. (10), (11) and (12) are correct? The answer is that with CO taken as an inequality as in Eq. (7) and (10), the expression for ω_0 as in (12) is more accurate. However, if this expression appears to give the impression that since the condition setting resistors R_0 and R_1 also appear in this equation therefore, one has the liberty of changing R_0 and R_1 arbitrarily to any extent and thereby claiming that ω_0 can be varied by changing these resistors is not correct because of the constraint of Eq. (11). In fact, as explained in para (ii), R_0 and R_1 can be changed only to the extent which ensures that the roots of the CE still remain in the vicinity of $\pm j\omega$ axis in which case their contribution in Eq. (12) (in view of the inequality of Eq. (11)) would obviously be very small. Furthermore, by observing the CE of Eq. (6), it is clear that arbitrarily changing R_0 and R_1 as done by Abuelma'atti and Khalifa in [1] may make the roots of the CE either deeply into RHP of s-plane (when $R_0 > R_1$) or the roots may move to LHP of s-plane (when $R_1 > R_0$). Note that in neither of these cases, the circuit can be a sinusoidal oscillator! Since Abuelma'atti and Khalifa [1] have arbitrarily changed R_0 from 1 kΩ - 20 kΩ and R_1 also from 1 kΩ - 20 kΩ and had still claimed that the circuits were producing good quality sine wave *under all cases* that is what looked clearly illogical and hence, was checked numerically and experimentally. The majority of their data shown in the first 3 columns of the augmented Tables 1-4 turned out to be either non-existent or erroneous.

ACCEPTED MANUSCRIPT

4. Concluding remarks

Before making concluding remarks it is useful to reiterate that from equation (6) of [1] it is clear that in both the circuits of Fig. 1 under consideration, the resistors R_0 and R_1 control the CO whereas resistors R_2 and R_3 control the FO. However, Abuelma'atti and Khalifa in [1] conjectured that since R_0 and R_1 also appear in their derived expression (see eqn. (12)) for ω_0, therefore, ω_0 would also change when R_0 and R_1 are changed. In fact, we have refuted this very claim here by showing that most of the experimental data given by Abuelma'atti and Khalifa in [1] is non-existent (since the circuits are ***not*** found to oscillate and generate the ***claimed*** frequencies at most of the values in their four tables of [1]) and hence, their contention regarding variability of ω_0 by R_0 and R_1 is incorrect. Furthermore, for all those combinations of the values of R_0 and R_1 for which the circuits of Fig. 1 do produce sinusoidal oscillations, the contribution of the term involving these resistors in the expression for oscillation frequency given by eqn. (12) of [1] is so small that ω_0 calculated either from eqn. (12) of [1] or from eqn. (6) of [1] will not differ much (as explained in para (vi) of Section 3 above).

Thus, this communication has presented results of a re-check[3] of the experimental results of the two CFOA-based fully-uncoupled oscillator circuits of Fig. 1, as recently presented by Abuelma'atti and Khalifa in [1] and it was shown that

- Firstly, most of the experimental results of [1] are *not repeatable* and in *most of the cases*, the sinusoidal oscillations do not even exist and hence, most of the measurement data as given in [1] is false and misleading.

[3] The experimental results have been carried out at five different times; however, except with the very minor variations the same patterns have been found to be repeated in all the four augmented tables.

ACCEPTED MANUSCRIPT

- Secondly, the measured oscillation frequencies, as shown by Abuelma'atti and Khalifa in [1], do not tally with those calculated from their own formula of equation (12) of [1].

Since the experimental results and the formula of equation (12) as given in [1] have been the main instruments of Abuelma'atti and Khalifa [1] to conclude that fully-uncoupled oscillators may not exist, the counter experimental evidence presented here refutes their contention. It, therefore, follows that the main contention of [1] is ill-conceived and the caution issued by them is ill-founded!

Acknowledgement: The authors wish to thank the anonymous reviewers and Prof. S. Minaei, the Editor-in-chief, for their comments and suggestions which have been very useful in improving the presentation of the manuscript.

References

[1] Abuelma'atti MT, Khalifa ZJ. Fully uncoupled independent control of frequency and condition of oscillation: A caution. Int J Electron Commun (AEU) 2014:1037-1040

[2] Bhaskar DR, Gupta SS, Senani R, Singh AK. New CFOA-based sinusoidal oscillators retaining independent control of oscillation frequency even under the influence of parasitic impedances Analog Integr Circ Sig Process 2012:427-437.

[3] Abuelma'atti MT, Alsuhaibani ES, Obadi ASB, Khalifa ZJ. Independent control of the frequency and condition of oscillation: a caution. Int J Electron 2013: 384-392

[4] Pathak JK, Singh AK, Senani R. Systematic realization of quadrature oscillators using current differencing buffered amplifiers, IET Circuits Devices Syst 2011:5, 203-11

[5] Abuelma'atti MT: Active only sinusoidal oscillator with electronically-tunable fully uncoupled frequency and condition of oscillation, Active Passive Electron Compon 2001: 24: 233-41

[6] Abuelma'atti MT: New Sinusoidal oscillators with fully uncoupled control of oscillation frequency and condition using three CCII+s. Analog Integr Circuits Signal Process 2000:24: 253-61

[7] Abuelma'atti MT, Al-Daghrier H. New Single element controlled sinusoidal oscillators employing CCII+. Microelectron J 1998: 29: 83-6

ACCEPTED MANUSCRIPT

[8] Cam U. Kuntman H. A new CCII-based Sinusoidal oscillator providing fully-independent control of oscillation condition and frequency. Microelectron J 1998: 29: 913-9

[9] Turkoz S. Minaei S. A new current-controlled sinusoidal oscillator using current-controlled conveyor. Frequenz 2000: 54: 132-3

[10] Bhaskar DR. Realization of second-order sinusoidal oscillator/filters with non-interacting controls using CFAs. Frequenz 2003: 57: 12-4

[11] Soliman AM. Current feedback operational amplifier based oscillators. Analog Integr Circ Signal Process 2000: 23: 45-55

[12] Martinez PA. Sabadell J. Aldea C. Celma S. Variable frequency sinusoidal oscillators based on CCII+. IEEE Trans Circ Syst I 46: 1386-1390

[13] Senani R. Bhaskar DR. Singh VK. Sharma RK. Sinusoidal Oscillators and Waveform Generators using Modern Electronic Circuit Building Blocks. Springer International Publishing Switzerland 2016: Ch.4: 206-8

[14] Chen HP. Hwang YS. Ku YT. A systematic realization of third-order quadrature oscillator with controllable amplitude. Int J Electron Commun (AEU) 2017:64-73.

[15] Abaci A. Yuce E. Modified DVCC based quadrature oscillator and lossless grounded inductor simulator using grounded capacitor(s). Int J Electron Commun (AEU) 2017: 86-96.

[16] Rybin YK. Synthesis of electrical oscillators based on principles of symmetry and reflection. Int J Electron Commun (AEU) 2015: 500-506.

[17] Summart S. Thongsopa C. Jaikla W. New current-controlled current-mode sinusoidal quadrature oscillators using CDTAs. Int J Electron Commun (AEU) 2015: 62-68.

[18] B. Metin, K. Pal, O. Cicekoglu, " All-pass filters using DDCC and MOSFET based electronic resistor" Int. J. Circ. Theor. Appl., vol. 39, no. 8, pp. 881-891, 2011.

Appendix 1

The non-ideal CE of both the circuits of Fig. 1, which is incidentally same, was presented in reference [2]. However, the same is being reproduced here for the sake of ready reference. For determining the non-ideal CE, we have considered the finite input resistance R_{xi} at the x-port, i = 1-3, parasitic components R_{yi} in parallel with $1/sC_{yi}$ at the y-port and parasitic components R_{zi} in parallel with $1/sC_{zi}$ at the z-port of all the CFOAs i = 1-3 and also take into account the non-unity current gains between i_x and i_z of the CFOAs as β_1, β_2 and β_3 and then

ACCEPTED MANUSCRIPT

modelling the relation between v_z and v_w of each CFOA by the equation $v_{wi} = v_{zi} + i_{wi}.R_{wi}$, i =1-3 also. By considering all the three CFOAs to be identical (for the sake of simplicity) it has been found that the CE for both the circuits then becomes same which is given by:

$$a_4 s^4 + a_3 s^3 + a_2 s^2 + a_1 s + a_0 = 0$$

where $a_4 = (c_1 + 2c_z)(c_2 + c_z)c_y^{\,2}$

$$a_3 = 2c_y(c_1 + 2c_z)(c_2 + c_z)\left(\frac{1}{R_y} + \frac{1}{R_w}\right) + c_y^{\,2}\left\{\frac{c_1 + 2c_z}{R_z} + (c_2 + c_z)\left(\frac{1}{R_0 \,//\, \frac{R_z}{2}}\right)\right\}$$

$$a_2 = (c_1 + 2c_z)\left(\frac{1}{R_y} + \frac{1}{R_w}\right)\left\{(c_2 + c_z)\left(\frac{1}{R_y} + \frac{1}{R_w}\right) + 2\frac{C_y}{R_z}\right\} + c_y(c_2 + c_z)\left\{2\left(\frac{1}{R_y} + \frac{1}{R_w}\right)\left(\frac{1}{R_0 \,//\, \frac{R_z}{2}}\right) - \left(\frac{\beta_1}{R_1 + R_x}\right)\left(\frac{1}{R_w}\right)\right\} + \frac{c_y^{\,2}}{R_z}\left(\frac{1}{R_0 \,//\, \frac{R_z}{2}}\right)$$

$$a_1 = \left(\frac{1}{R_y} + \frac{1}{R_w}\right)^2\left\{\frac{c_1 + 2c_z}{R_z} + \frac{c_2 + c_z}{R_0 \,//\, \frac{R_z}{2}}\right\} + \frac{2c_y}{R_z}\left(\frac{1}{R_y} + \frac{1}{R_w}\right)\left(\frac{1}{R_0 \,//\, \frac{R_z}{2}}\right) - \frac{(c_2 + c_z)\beta_1}{(R_1 + R_x)R_w}\left(\frac{1}{R_y} + \frac{1}{R_w}\right) - \frac{c_y\beta_1}{R_z R_w}\left(\frac{1}{R_1 + R_x}\right) + \left(\frac{c_y\beta_2\beta_3}{R_w}\right)\left(\frac{1}{R_2 + R_x}\right)\left(\frac{1}{R_3 + R_x + R_w}\right)$$

ACCEPTED MANUSCRIPT

$$a_0 = \left(\frac{1}{R_y}+\frac{1}{R_w}\right)\left\{\begin{array}{l}\left(\frac{1}{R_y}+\frac{1}{R_w}\right)\frac{1}{R_z}\left(\frac{1}{R_0 \,//\, \frac{R_z}{2}}\right) \\ -\left(\frac{\beta_1}{R_z R_w}\right)\left(\frac{1}{R_1+R_x}\right)+\frac{\beta_2\beta_3}{R_w}\left(\frac{1}{R_2+R_x}\right)\left(\frac{1}{R_3+R_x+R_w}\right)\end{array}\right\}$$

The data given in column 5 of augmented Tables 1-4 has been determined from the above equation by taking following parameters and component values: R_3 = 1kΩ, R_0 = 10 kΩ, R_1 = 7.95 kΩ, C_1 = C_2 = 1nF, R_x = 50Ω, R_y = 2MΩ, R_z = 3 MΩ, R_w = 50Ω, C_y = 2pF, C_z = 4pF and $\beta_i = 1\,(i = 1-3)$ with R_2 taken as variable.

It may also be mentioned that a more elaborate analysis including the effect of Rx and Lx as in [18] is also possible but the same has not been considered to be necessary in view of the limited scope of the present manuscript. In view of this, further detailed analysis including Rx and Lx has not been attempted.

ACCEPTED MANUSCRIPT

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 4, Issue 6 , June 2017

Review of some Case Studies of Concentrated Solar Thermal Power Technologies: Status of CSP in India

NirmalPratap Singh, Dr. J P Kesari

Masters in Thermal Engineering, Delhi Technological University, Delhi, India
Associate professor, Delhi Technological University, Delhi, India

ABSTRACT: Solar energy is the most abundant form of energy available freely. It is clean in nature and becomes more attractive considering the current scenario of global warming and climate change worldwide. This paper aims at understanding the technical aspects of the solar thermal technologies like solar reflectors, Heat transfer fluid etc.. The text explores the existing and proven Concentrated Solar Power (CSP) Technologies like Parabolic Trough Collector (PTC) and Linear Fresnel Reflector (LFR), Paraboloid Dish and Solar Power Tower. The current state of CSP Plants in India are explored along with review on some case studies for the existing solar thermal power plants in India and the world.

KEYWORDS: Concentrated Solar Power (CSP), Parabolic Trough, Linear Fresnel, Paraboloid Dish, Rankine Cycle,

I. INTRODUCTION

The problem with fossil fuel is not only that it is limited in nature but also the emission from it which produces Carbon-Di-oxide.CO_2 is one of the most important candidate responsible for Global Warming. It can be estimated that with the increasing concentration of CO_2 the earth can gets warmer by 1-5°C with the advent of the next century [1].The emission emanating from fossil fuels is also dangerous for human health as it causes the air pollution. So the need of hour is that we need to mainstream an energy resource which is not only abundant but also clean in nature. Solar energy has solution for both the concern.

The four commonly used CSP Technologies are parabolic trough collector (PTC), Linear Fresnel Reflector (LFR), Paraboloid Dish and Solar Power Tower (SPT) out of which the first two are belongs to line focus system and last two point focus systems. The PTCs are primarily used in industry since they are the most developed and cheap compared to other CSP Technologies [5]. However, its cost is still more expensive than that of the conventional fossil fuel power plants. Second most installed is SPT technology after PTC, and growing at faster rate among all CSPs technologies [6]. High efficiencies are obtained because the receiver operates at high temperatures and reduced losses. There is a lot of similarity between Linear Fresnel Reflectors and Parabolic Trough Collectors. The difference is the geometry of LFRs which is a long flat or slightly curved mirror placed for concentration of sunlight.

LFR has lower efficiency than other CSP Technologies. The concentration ratio in LFR is significantly lower than that of PTCs. But simplicity of LFR means that they are easy and cheaper to manufacture makes them advantageous over others in certain conditions [7].The Parabolic Dish Reflector (PDR) or Dish Engine is a concave mirror, focuses the incoming sunlight onto a point receiver [2].Paraboloid Dish Systems give the highest efficiency among CSP Technologies.

II. SOLAR THERMAL TECHNOLOGIES

There are four important Solar Thermal Technologies utilizing concentrated for generating energy used by industry.
They are parabolic trough collectors, linear Fresnel reflector, power towers or central receiver systems, and dish/engine.

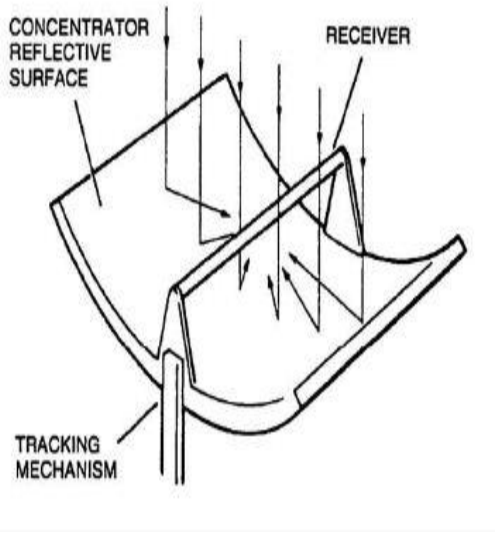
ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 4, Issue 6 , June 2017

A. Parabolic trough collector (PTC)

PTCs is a proven technology used for process heating and other applications commercially around 90 % of the solar thermal power plants are based on parabolic trough systems.[CSP Today, 2013]



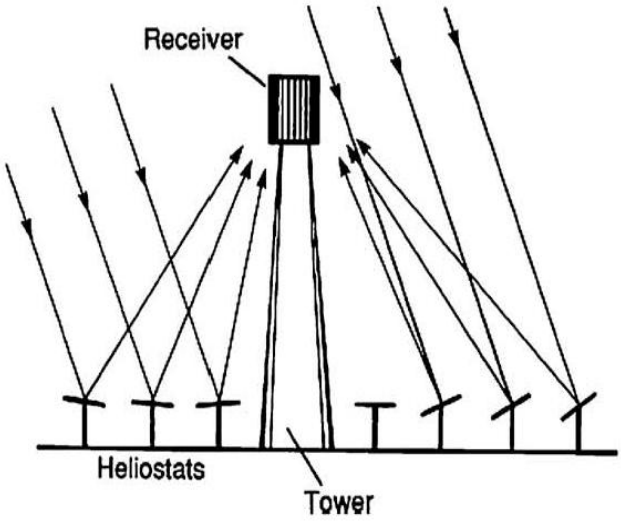


Fig. 1. Parabolic trough system**Fig.2.** Central Receiver Tower System

The main components of PTCs are a parabolic reflecting plate, a transparent cover on reflecting plate and an observer, which passes heat energy to Heat transfer fluid (HTFs). The absorber is fixed permanently at the focus of the parabolic concentrator. The absorber tube is protected by the transparent cover from the heat losses. The Parabolic concentrator is settled on a fixed structure and the single-axis tracking mechanism follows the sun's trajectory. Every cross-section of a PTCs has a focal point. These focal point when joined makes a focal line (Fig.1) . The solar Radiation that passes parallel to optical plane is reflected from the plate in such a way that it passes through the focal line [3].PTC can concentrate direct sunlight to generate working temperatures up to 400°C and achieve concentration ratios in the range of 30-100 [4].

B. Central Receiver Tower

The Solar power tower consists of a field of thousands of mirrors (heliostats) surrounding a tower which holds a heat transfer fluid to concentrate light on a central receiver atop a tower. An array of heliostats concentrate the sunlight onto a receiver settled at the top of a tower. Central Receiver tower is a point focus system (Fig.2). High efficiencies are obtained because the receiver operates at high temperatures and reduced losses. They provide better energy storage capability among CSP technologies. The individual inbuilt tracking mechanism keep the heliostat focused on tower and fluid is heated which runs a turbines. General concentration ratios ranges from 300 to 500 [10].

C. Linear Fresnel Reflector (LFR)

LFRs uses a long flat or nearly flat reflector and a receiver assemble which is fixed to some structure. The assembly may contain single linear receiver tube or multiple linear receiver tubes depending upon the size of the plant. The reflectors are placed at different angles to concentrate the incoming radiation on receiver tubes. Each line of reflectors

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 4, Issue 6 , June 2017

has its own single-axis tracking system to concentrate the radiation correctly on the receiver (Fig.3). Due to its geometric properties the optical efficiency of LFR Field is always lower than that of PTCs Solar field.General concentration ratios is around 30 [11], which is low compared to other CSP technologies.

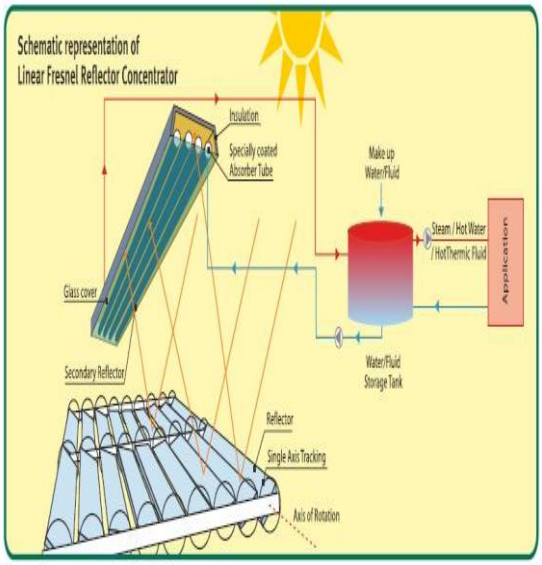


Fig. 3. Schematic representation of Linear Fresnel Reflector Concentrator [8].

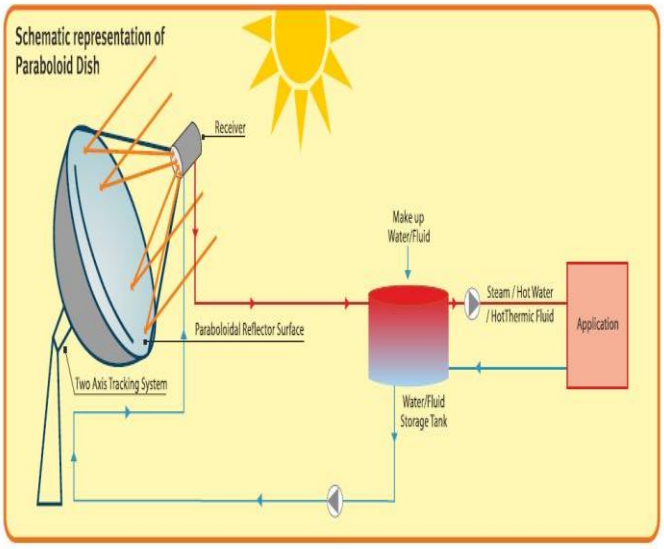


Fig. 4. Schematic representation of Parabolic Dish Reflector [9]

D. The Parabolic Dish Reflector (PDR)

The four main units of PDR is concentrator, receiver, Stirling engine and generator (Refer Fig.).The Stirling engine, receiver, and generator are enclosed in Power Conversion Unit (PCU). The concentrator (which is outside the PCU) focus the solar radiation into the cavity of the receiver. By concentrating the solar energy to actuate the Stirling engine the system produces the electricity.The Parabolic Dish Reflector system is integrated with the dual axis tracking system for tracking the solar radiation [12].The advantage the system lies in the fact that the higher solar to electric efficiencies is achieved. As the ongoing research suggest the Dish Engine has the tremendous potential. The efficiency of 31.25% has been achieved with one PDR, which set the world record for the solar to electrical efficiency [13].

Since the PDR System uses 2-axis tracking mechanism, it capture the highest solar radiation among all CSP Technologies and optical efficiency can be as high as 94%.This lead to higher concentration ratio which is around 500 to 2000 [10].The concern about using this technology is the high cost of mirrors. Some alternative method are being explored to use cheaper mirrors.

The PDR based Solar Industrial process heating system can be installed at automobile industrial plants with high DNI values to make automobile production more green and environment friendly. [25]

III. HEAT TRANSFER FLUIDS (HTFs)

In CSP System the heat is collected from the receiver by a thermal energy carrier called heat transfer fluids (HTFs). The HTF can be taken directly to drive a turbine for production of power or, more often, be used in a heat exchanger to transfer its heat to an another fluid known as the cycle fluid.Desired properties and characteristics of a HTF are: high boiling point, low meltingpoint, and thermal stability, low vapor pressure (<1 atm) at high temperature, low viscosity,

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 4, Issue 6 , June 2017

low corrosion with metal alloys used to contain the HTF, high heat capacity for energy storage, high thermal conductivity and low cost.
The main HTFs based on type of material used are-[14]-(1) water/steam, (2) Organics, (3) molten-salts

A. Water or steam

The use of water or steam as both HTF and working fluid of cycle removes the complexities from the system and leads to enhanced efficiency, with less cost of electricity production. It is used as both HTF and working fluid in the recent and the world's largest CSP plant – the Ivanpah solar power facility which was launched in February 2014. The main issue with the water steam HTF is the scarcity of water in desert regions, and since for better Direct Normal Irradiance (DNI) throughout the year these CSP plants are mostly located in deserts, using water as a HTF poses serious challenges.

B. Organics

Organic materials are also used widely as HTFs in CSP systems. For example Therminol VP1 (Biphenyl/Diphenyloxide) is commonly used in commercial CSP systems. Biphenyl/Diphenyl oxide is a eutectic mixture of two very stable organic compounds; Biphenyl (C12H10) and Diphenyl oxide (C12H10O).
The first solar thermal plant with this organic material as the HTF was commissioned in 2009 at Badajoz, Spain and was named as 'Alvarado 1'.

C. Molten-salts

The most widely used HTF is the Molten Salt. They possess excellent properties like high working temperature (more than 500^{o}C) and heat capacity, low vapor pressure and corrosive property, and good thermal and physical properties at elevated temperatures, which are quite desirable for any HTF [15].Currently the most common molten salt HTF and thermal energy storage media are a mixture of 60% $NaNO_3$ and 40% KNO_3.The first molten-salt as a HTF based power tower systems launched in 1984. They were THEMIS tower (2.5 MW) in France and Molten-salt ElectricExperiment (1 MW) in the United States [16].
The main disadvantage of this salt mixture is the high melting point. The salt can freeze and block the pipeline during winter nights. In order to overcome this problem, auxiliary facilities need to be installed, which could increase the investment and operational costs.

IV. CSP TECHNOLOGIES IN INDIA

Table.1. Status of CSP technologies in India [17].

S. No	Project and name of place	Company/ manufacturer	Capacity	Technology type	Year of commencement
1	Abhijeet Solar Project, Phalodi (Rajasthan)	Corporate Ispat alloys Ltd	50 MW	Parabolic trough	Under construction
2	ACME Solar Tower, Bikaner (Rajasthan)	ACME Group	2.5 MW	Solar tower	Operational 2011
3	Dhursar, Dhursar (Rajasthan)	Reliance Power	125 MW	Linear fresnel reflector	Operational 2014
4	Diwakar, Askandra (Rajasthan)	LancoInfratech	100 MW	Parabolic trough	Under construction
5	Godawari Solar Project, Nokh (Rajasthan)	Godawari green Energy Limited	50 MW	Parabolic trough	Operational 2013
6	Gujarat Solar One, Kutch (Gujarat)	Cargo Solar Power	25 MW	Parabolic trough	Under

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 4, Issue 6 , June 2017

					construction
7	KVK Energy Solar, Project Askandra (Rajasthan)	KVK Energy Ventures Ltd.	100 MW	Parabolic trough	Under construction
8	Megha Solar Plant, Anantapur (Andhra Pradesh)	Megha Engineering and Infrastructure	50 MW	Parabolic trough	Operational 2014
9	National Solar Thermal Power Facility, Gurgaon (Haryana)	IIT, Bombay	1 MW	Parabolic trough	Operational 2012

Currently India has 228.5MW installed capacity of CSP basedpower plants of different capacities and based on the parabolictrough collector, LFR and solar tower technologies. A 275MW CSP power plants capacity based on PTC field technology is under construction.Research is going on in Stirling engine based power plant to test in Ministry of New and Renewable Energy(MNRE) is also taking new initiativesfor the solar thermal electricity by providing new researchfacility for CSP and pilot project on Stirling engine to be tested in Indian climatic condition by MNRE in collaboration with ONGC at National Institute of Solar Energy, Gwalpahari,Gurgaon, India has nine CSP powerplants, seven of them are based on parabolic trough collector, one ison solar tower technology and one is on LFR. According to the collected and communicated data it turns out that parabolic trough collector and LFR are performing well as compared to solar powertower technology [17].

V. EXAMPLES: A REVIEW

A. Godawari Solar Project, Nokh

Godawari CSP Plant of 50 MW is the first ever CSP plant in India located near the Nokh Village in Jaisalmer district of Rajasthan (India).The project started in June 26, 2011 and started giving power supply in June 5, 2013 coinciding with the world environment day.Technical Specification of the project is written in the table.2.
It employs Parabolic Trough CSP Technology with state of-the-art SKAL-ET 150 trough structure. The total aperture area of the trough is 3, 92,400m² with 480 Solar Collector Assemblies (SCAs) and total occupied land is 150 hectares.
The plant runs on Rankine cycle. Dowtherm A is used as a HTF which gather heat from parabolic trough and transfer it to cycle fluid i.e. water to generate steam. The steam generated flows through the pipe and strike the turbine blades to generate power. This Project is capable of lighting 2.5 million homes in India. System Advisory Model (SAM) for plant performance analysis. Is used which observes the real power output of the plant. [18]
Godawari power is the first successful CSP plant in India and one of the finest example of concentrated solar thermal power in the country.

Table.2. Overview and technical characteristics of Godawari Solar Project, Nokh [19]

Project	Godawari Solar Project
Location and country	Nokh, Rajasthan (India)
Lat/Long Location	27°36′ 5.0″ North, 72°13′ 26.0″ East
Land Area	150 hectares
Technology	Parabolic Trough
Turbine Capacity	50 MW
Heat Transfer Fluid Type	Dowtherm A
Cooling Method	Wet cooling
Solar Field Inlet Temperature	293 ^{O}C
Solar Field Outlet Temperature	390°C
Storage Type	None
Solar-Field Aperture Area	392,400 m²
# of Solar Collector Assemblies (SCAs)	480
# of Loops	120
# of SCAs per Loops	4

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 4, Issue 6 , June 2017

SCAs Aperture Area	817 m²
SCA Length	144 m

B. National Solar Thermal Power Facility

CSP plant National Solar Thermal Power Facility of 1MW is commissioned at Gurgaon near New Delhi, India. The plant has the unique feature of integration of two different solar fields- parabolic trough collectors and linear Fresnel reflectors- without having the conventional fossil fuel backup. [21] Overview and technical characteristics of the plant is given in the table.3.The parabolic trough collector (PTC) field having 8175 m^2 of total aperture area heat-up the synthetic oil (as HTF) which enters the PTC field at about 232°C and exits at 393°C.
The linear Fresnel reflector (LFR) having total aperture area of field 7020 m^2 with direct steam generation (DSG) produces saturated steam at 44 bar and 256°C and together with the hot oil from PTC field superheated steam is formed at 350°C, 42 bar. This saturated steam runs a turbine to generate electricity. The PTC field contributes around 60 % and LFR field around 40 % of the heat. The plant was developed as a research-cum-demonstration facility. [22]

Table.3. Overview and technical characteristics of National Solar Thermal Power Facility, Gurgaon [19]

Project	National Solar Thermal Power Facility
Location and country	Gurgaon, Haryana (India)
Lat/Long Location	28°25′ 39.0″ North, 77°9′ 33.0″ East
Technology	Parabolic Trough, Linear Fresnel Reflectors
Turbine Capacity	1 MW
Heat Transfer Fluid Type	Therminol VP-1
Solar Field Inlet Temperature	293 ^{O}C
Solar Field Outlet Temperature	393°C
Storage Type	None
Solar-Field Aperture Area	8000 m²
# of Solar Collector Assemblies (SCAs)	12
# of Loops	3
# of SCAs per Loops	4
SCAs Aperture Area	667 m²
SCA Length	120 m

C. ACME Solar Tower

ACME CSP Power Plant of 2.5 MW is based on power tower technology and located in Bikaner, Rajasthan. The construction of the project started in 2009 and became operational in year 2011. Technical Characteristics of the plant are mentioned in table.4.The 46m high solar tower gets solar radiation from 14,280 heliostats with each heliostat aperture area of 1.136 m² which makes approximately 16,222 m² (14,280 X1.136)of solar field aperture area.
This solar thermal plant is considered as the first demonstration project India had after the launch of Jawaharlal Nehru National Solar Mission (phase-I bidding) on January 11, 2010. [24]
The plant works on the steam Rankine cycle with wet cooling (cooling tower is used) and maximum cycle pressure reaches to 60 bar and temperature to 440°C

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 4, Issue 6 , June 2017

Table.4. Overview and technical characteristics of ACME Solar Tower Project, Bikaner [19]

Project	ACME Solar Tower
Location and country	Bikaner, Rajasthan (India)
Lat/Long Location	28°11′ 2.0″ North, 73°14′ 26.0″ East
Land Area	12 hectares
Technology	Power tower
Turbine Capacity	2.5 MW
Heat Transfer Fluid Type	Water/Steam
Cooling Method	Wet cooling (Cooling Tower)
Receiver Inlet Temperature	218 ^{O}C
Receiver Outlet Temperature	440°C
Power Cycle Pressure	60 bar
Storage Type	None
Heliostat Solar-Field Aperture Area	16,222 m²
# of Heliostats	14,280
Heliostat Aperture Area	1.136 m²
Tower Height	46 m

D. Julich Power Tower

CSP plant Julich Power Tower of 1.5 MW is situated in Julich, Germany and is the country's only plant based on solar power tower technology. Overview and technical characteristics of the plant is given in the table.5.Construction of the Julich Power Tower was begun in July 2007 and installed in December 2008. Heliostat field is composed of total 2153 heliostats having each one of 8.2 m^2 aperture area which makes around total aperture area of 17650 m^2.On the top of the 60 m tower the power block steam turbine coupled with generator produces power in the form of electrical energy. The air as a HTF is heated to 680 ◦C and is used to heat water in the power cycle up to 480 ◦C. The efficiency of the plant is not much high since the heat transfer coefficient of air is low. The plant is equipped with 1.5 hour heat storage capacity by a packing of ceramic elements which are heated by hot air flowing between them. [20]

Table.5. Overview and technical characteristics of *Jülich* Solar Tower, *Jülich* [19]

Project	Saguaro Power Plant
Location and country	*Jülich*, Rhineland (Germany)
Lat/Long Location	50°54′ 54.0″ North, 6°23′ 16.0″ East
Land Area	17 hectares
Technology	Power Tower
Turbine Capacity	1.5 MW
Heat Transfer Fluid Type	Air
Cooling Method	Dry cooling
Receiver Outlet Temperature	680°C
Storage Type	Ceramic heat sink (Capacity-1.5hour)
Heliostat Solar-Field Aperture Area	17,650 m²
# of Heliostats	2,153
Heliostat Aperture Area	8.2 m²
Tower Height	60 m

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 4, Issue 6 , June 2017

E. *Ivanpah*Solar Power Plant

The world's largest solar thermal Ivanpah CSP plant based on solar power tower of 377 MW is located at Mojave Desert in Primm, Nevada.The construction work started in October 2010, and started producing power in February 2014. The facility is comprised of three separate plants built in phases between 2010 and 2013 [20]. Overview and technical characteristics of the plant is given in the table.6.The 392 MW gross capacity plant in 3,500 acres deploys 173,500 heliostats with total solar field aperture area of 2,600,000 m^2 focuses incident solar radiation on the boiler of three centralized tower of 140 m height. The produced steam runs turbine and in turn power is generated [19].Ivanpah solar plant uses dry cooling process to condense the steam back to liquid water and feed into boiler, thus around 90 % of the water is conserved as compared to other solar thermal technologies using conventional wet cooling systems. Dry cooling system comprises air cooled condenser having large forced draught fans to circulate air into the heat exchanger (condenser) to cool and condense the steam [23].

Table.6. Overview and technical characteristics of *Ivanpah* Solar Power Plant,*Primm* [19]

Project	*Ivanpah* Solar Electric Generating System
Location and country	*Primm*, Nevada (USA)
Lat/Long Location	35°33′ 8.5″ North, 115°27′ 30.97″ West
Land Area	3,500 Acres
Technology	Power Tower
Turbine Capacity	377 MW
Heat Transfer Fluid Type	Water
Cooling Method	Dry cooling
Receiver Inlet Temperature	249
Receiver Outlet Temperature	566°C
Power Cycle Pressure	160 bar
Storage Type	None
Heliostat Solar-Field Aperture Area	2,600,000 m^2 m^2
# of Heliostats	173,500
Heliostat Aperture Area	15.0 m^2
Tower Height	140

VI. CONCLUSION

Concentrated solar power plants have great potential for supplying and fulfilling a significant energy demand in near future. The excellent thing about CSP is the clean and carbon emission free energy. The portfolio of CSP technologies comprises of mature parabolic trough technology, Linear Fresnel System and Central receiver Tower systems. The Paraboloid dish system have high solar to electricity conversion efficiencies but not tested for large industrial scale. Heat Transfer fluid like water and organics are used in specific purpose. Molten Salt is most widely used HTF.

The development of CSP based power plants in India is not very inspiring and needs proper research focusing and adapting Indian conditions. Finally some case studies proved that the successful installation and operation of solar thermal power plant is possible and can be seen as a demonstration plant.

REFERENCES

[1]Wuebbles DJ, Jain AK. Concerns about climate change and the role of fossil fuel use. Fuel Processing Technology 2001; 71:99-119.

[2]Kesari JP, Gupta M, Jain A, Ojha AK. Review of the concentrated solar thermal technologies: Challenges and opportunities in India. Volume II, Issue I, January 2015, ISSN 2321 - 2705

[3]Günther M, Joemann M, Csambor S. Parabolic Trough Technology. Advanced CSP Teaching Materials (Chapter 5), Institute for Electrical Engineering, Rational Energy Conversion, University of Kassel, WilhelmshöherAllee 73, 34121 Kassel.

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 4, Issue 6 , June 2017

[4]Duffie JA, Beckman WA. Solar engineering of thermal processes. 3rd Ed. NewYork: John Wiley & Sons; 2006.
[5] Price H, Lüpfert E, Kearney D, Zarza E, Cohen G, Gee R, et al. Advances inParabolic trough solar power technology. Journal of Solar Energy Engineering 2002; 124(2):109–125.
[6] Avila-Marin AL, Fernandez-Reche J, Tellez FM. Evaluationof the potential of central receiver solar power plants: configuration, optimization and trends. Applied Energy 2013; 112:274–288.
[7]Buie D, Dey CJ, Mills DR. Optical considerations in line focus Fresnel concentrators. In: 11th Proceedings of the 11th SolarPACES International Symposium on Solar Thermal Concentrating Technologies, 4- 6th September 2002, Zurich Switzerland.
[8] UNDP-GEF:Solar Energy-Sustainable, clean and safe-"Linear Fresnel Reflector Concentrator for medium and high temperature Applications"-UNDP-GEF project on CSH, Ministry of new and Renewable energy Government of India.
[9] UNDP-GEF: Solar Energy-Sustainable, clean and safe-"Paraboloid Dish for medium and high temperature Applications"-UNDP-GEF project on CSH, Ministry of New and Renewable Energy, Government of India.
[10]Kalogirou SA. Solar thermal collectors and applications. Progress in Energyand Combustion Science 2004; 30(3):231-295.
[11] Mills DR, Lievre PL, Morrison GL. First results from compact linear Fresnel reflector installation. In: Proceedings of ANZSES Solar2004. 2004.
[12]Kaltschmitt M, Streicher W, Wiese A. Renewable energy: technology, economics, and environment. Berlin; New York: Springer; 2007.
[13] Taggart S. CSP: dish projects inch forward. Renewable Energy Focus2008; 9(4):52-54.
[14]Maria IRS. Concentrating Solar Thermal Technologies: Analysis and Optimisation by CFD Modelling. Springer, Green Energy Technology, ISBN: 978-3-319-45882-3
[15]Peng Q, Ding J, Wei X, Yang J, Yang X. The preparation and properties of multi component molten salts. Applied Energy 2010; 87:2812–2817.
[16] Rebecca Dunn. A Global Review of Concentrated Solar Power Storage. Australian National University, Department of Engineering, Building 31 North Rd, ACT 0200, Australia.
[17]MNRE.Ministry of New and Renewable Energy--Achievements 2012.
[18]Sun Focus Magazine, Vol.4 issue-2 (Oct-Dec 2016)
[19] NREL Website: https://www.nrel.gov/csp/solarpaces/
[20] Wikipedia: https://en.wikipedia.org/wiki/List_of_solar_thermal_power_stations
[21] IIT Bombay Website: http://www.ese.iitb.ac.in/~NSTPP/?q=content/power-plant
[22] Nayak JK, Kedare SB, Banerjee R, Bandyopadhyay S, Desai NB, Paul S, Kapila AA . 1 MW national solar thermal research cum demonstration facility at Gwalpahari, Haryana, India. Current Science 2015; 109(8):1445-1457.
[23]http://www.brightsourceenergy.com/stuff/contentmgr/files/0/5a6bf289d056bfb79dca9bd1136ef23f/folder/dry_cooling_faq___july_2012.pdf
[24]The state of concentrated solar power in .A roadmap to developing solar thermal technologies in India.
[25]Uppal A, Kesari J P, ZunaidM.Designing of Solar Process Heating System for Indian Automobile Industry.International Journal of Renewable Energy Research.Vol.6, No.4, 2016

AUTHOR'S BIOGRAPHY

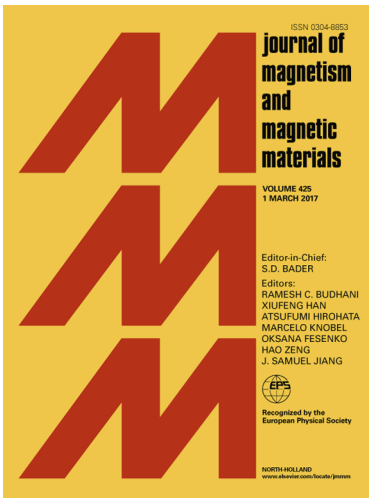


Dr. J.P. Kesariis an Associate Professor at Dept. of Mechanical Engg. At Delhi Technological University, Delhi. He is a B.Tech. From I.I.T. Kharagpur, M.Engg.from Asian Institute of Technology, Bangkok, Ph.D.in Solar Energy from Toulouse, France and a Post-Doctoral Fellow from CNRS, Paris. Presently he is Faculty Advisor of Solaris Team at DTU.



NirmalPratap Singh is a B.Techfrom IIT(ISM) Dhanbad in Mechanical Engineering, currentlypursuing M.Tech in Thermal Engineering from Delhi Technological University, Delhi. His area of interest includes Solar Thermal, Alternate Energy Resources, and Solar Industrial Process Heat

Accepted Manuscript



Role of Magnetic Exchange Interaction due to Magnetic Anisotropy on Inverse Spin Hall Voltage at $FeSi_{3\%}$/Pt Thin Film Bilayer Interface

Jyoti Shah, Saood Ahmad, Rishu Chaujar, Nitin K. Puri, P.S. Negi, R.K. Kotnala

PII: S0304-8853(17)30806-5
DOI: http://dx.doi.org/10.1016/j.jmmm.2017.07.060
Reference: MAGMA 62991

To appear in: *Journal of Magnetism and Magnetic Materials*

Received Date: 6 March 2017
Revised Date: 10 July 2017
Accepted Date: 19 July 2017

Please cite this article as: J. Shah, S. Ahmad, R. Chaujar, N.K. Puri, P.S. Negi, R.K. Kotnala, Role of Magnetic Exchange Interaction due to Magnetic Anisotropy on Inverse Spin Hall Voltage at $FeSi_{3\%}$/Pt Thin Film Bilayer Interface, *Journal of Magnetism and Magnetic Materials* (2017), doi: http://dx.doi.org/10.1016/j.jmmm.2017.07.060

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

Role of Magnetic Exchange Interaction due to Magnetic Anisotropy on Inverse Spin Hall Voltage at $FeSi_{3\%}$/Pt Thin Film Bilayer Interface

Jyoti Shah[1], Saood Ahmad[1*], Rishu Chaujar[2], Nitin K. Puri[2], P. S. Negi[1] and R. K. Kotnala[1]

[1] *CSIR-National Physical Laboratory, New Delhi-110012, India.*
[2] *Department of Applied Physics, Delhi Technological University, Delhi-110042, India.*

*Corresponding Author

Abstract

In our recent studies inverse spin Hall voltage (ISHE) was investigated by ferromagnetic resonance (FMR) using bilayer $FeSi_{3\%}$/Pt thin film prepared by pulsed laser deposition (PLD) technique. In ISHE measurement microwave signal was applied on $FeSi_{3\%}$ film along with DC magnetic field. Higher magnetization value along the film-plane was measured by magnetic hysteresis (M-H) loop. Presence of magnetic anisotropy has been obtained by M-H loop which showed easy direction of magnetization when applied magnetic field is parallel to the film plane. The main result of this study is that FMR induced inverse spin Hall voltage 12.6 μV at 1.0 GHz was obtained across Pt layer. Magnetic exchange field at bilayer interface responsible for field torque was measured 6 x 10^{14} $\Omega^{-1}m^{-2}$ by spin Hall magnetoresistance. The damping torque and spin Hall angle have been evaluated as 0.084 and 0.071 respectively. Presence of Si atom in $FeSi_{3\%}$ inhomogenize the magnetic exchange field among accumulated spins at bilayer interface and feebly influenced by spin torque of $FeSi_{3\%}$ layer. Weak field torque suppresses the spin pumping to Pt layer thus low value of inverse spin Hall voltage is obtained. This study provides an excellent opportunity to investigate spin transfer torque effect, thus motivating a more intensive experimental effort for its utilization at maximum potential. The improvement in spin transfer torque may be useful in spin valve, spin battery and spin transistor application.

Keywords: "Inverse Spin-Hall Effect", " Ferromagnetic resonance", "Magnetization", "damping torque", Spin Hall angle", "Spin Hall Magnetoresistance".

Introduction

Spintronics deals with manipulation and transport of spin current in single or multilayer paramagnetic / ferromagnetic thin films. Spin current can be generated by spin Hall Effect and detected by the inverse spin Hall Effect (ISHE). In bilayer thin film spin current flows towards the interfaces, where it can be absorbed as a spin-transfer torque (STT) on the ferromagnet. This STT affects the magnetization damping or even switches the magnetization [1]. Spin current when crosses the bilayer interface, Fermi energy splitting of spin up and spin down takes place due to difference in chemical potential [2]. Diffusion of spins across interface undergoes absorption and reflection which is controlled by the magnetization direction of ferromagnetic film [3]. Consequently, change in resistance has been observed by sweeping magnetic field direction known as magnetoresistance [4-5]. The spin current transmission across the bilayer interface is controlled by spin mixing conductance $G_{\uparrow\downarrow}$, which is quantified as $G_{\uparrow\downarrow}=G_r+iG_i$ where G_r is a real part, signifies torque at interface, and G_i is known as exchange magnetic field by which accumulated spin precess around the magnetization of ferromagnetic film [6]. Conversion of spin current into charge current is enhanced by adding ferromagnetic layer in which spin wave is generated by ferromagnetic resonance (FMR) excitation [7] and get induced in to Pt layer by spin pumping at interface. The magnitude of spin pumping at interface depends mainly on magnetic homogeneity at interface, bilayer resistance, magnetic anisotropy, spin-orbit scattering. Generated spin current also varies with orientation of charge current and magnetic anisotropy. The magnetization dynamics in ferromagnetic layer is described by a generalized Landau-Lifshitz-Gilbert Eq. (1).

$$\frac{\partial M}{\partial t} = -\gamma M \times H_{eff} + \alpha \left[M \times \frac{\partial n}{\partial t} \right] \qquad (1)$$

ACCEPTED MANUSCRIPT

Where, n is the unit vector in the direction of magnetization M, γ is the absolute value of the gyromagnetic ratio, and α is the magnetic damping constant. The first term represents precession torque in field H_{eff}, which is sum of H_{ext} and demagnetization field, while second term represents Gilbert damping torque [8]. Dynamics of magnetic moment not only depends on above mentioned microscopic parameter but influenced by macroscopic effects like resistance, magnetic anisotropy, and magnetic inhomogeneity. Effect of current in plane and current perpendicular to plane resulted into giant magnetoresistance in FM/NM multilayer geometry. In GMR, magnetic anisotropy of ferromagnetic layer has a significant role in spin current conduction. Similarly in microwave induced ISHE, magnetic anisotropy has a substantial function in spin pumping mechanism. Here we report our investigation on, ISHE of $FeSi_{3\%}$/Pt bilayer induced by ferromagnetic resonance. ISHE has been measured by sweeping microwave frequency from 1 GHz to 10 GHz at fixed applied DC magnetic field by measuring induced voltage on Pt layer. ISHE measurement of bilayer thin film has been analysed by determining spin mixing conductance at bilayer interface by performing magnetoresistance measurements [9].

Experimental

Bilayer $FeSi_{3\%}$/Pt film has been deposited by very sophisticated technique pulse laser deposition (PLD). It's very rare to deposit $FeSi_{3\%}$/Pt by PLD technique. It took several trials to optimize bilayer $FeSi_{3\%}$/Pt film deposition. In the current work inverse spin Hall voltage is observed in bilayer $FeSi_{3\%}$/Pt film. Further the conduction behavior of spins with magnetic field direction has been carried out by spin Hall magnetoresistance (SMR). The role of applied magnetic field on the interfacial accumulated spins which changes the conduction in Pt layer has been analyzed. Inverse spin Hall voltage (ISHE) was investigated by FMR using indigenously designed and developed non-resonant microwave device jig [10] on $FeSi_{3\%}$/Pt bilayer thin film. The low field torque suppresses the spin pumping to Pt layer thus low value of ISHE voltage is obtained. The preparation of $FeSi_{3\%}$/Pt bilayer thin film was carried out on silicon substrate by pulsed laser deposition technique. Silicon (110) substrate of dimension 9mm x 4mm has been taken for deposition using KrF laser of wavelength 298 nm in pulsed mode. The substrate was first heated at an elevated temperature to absorb the contaminations. After this, thin films were deposited in vacuum of order $5x10^{-6}$ mTorr pressure at 5-6 J/cm^2 laser energy density with 1Hz frequency followed by annealing at 400^{o}C in high vacuum for 2 hours. The substrate temperature was kept at 200^{o}C during film deposition. The layer thickness of FeSi and Pt were controlled during this process as 40 nm and 30 nm respectively. Crystal structure of $FeSi_{3\%}$ film was investigated by X-ray diffraction (XRD) analysis using Bruker AXS D8 Advanced powder X-ray diffractometer at room temperature with Cu Kα radiation (40 kV, 35 mA), scanned for 2θ from 10° to 80º with a step size of 0.001º per sec. The surface roughness of thin film was analysed by atomic force microscopy (AFM) Nanoscope-5. Magnetoresistance of $FeSi_{3\%}$/Pt interface has been measured by four probe method. The magnetic anisotropy of thin film was determined from hysteresis loops in film-plane and perpendicular to applied magnetic field upto 5kOe using Vibrating Sample Magnetometer (VSM).

Subsequently, thin film was mounted on a microstrip based sample holder having electrodes in order to measure output dc voltage on the sample. The designed pattern and high bandwidth of the sample-holder enables us to carry out FMR measurements with microwave frequency excitation from 1 GHz to 10 GHz [10]. Ferromagnetic resonance describes absorption of microwave radiation in a magnetic material containing strongly exchange coupled electrons; absorption is measured as a function of an applied magnetic field. By this method one measures magnetic properties by detecting the precessional motion of the magnetization in a ferromagnetic sample. The applied static magnetic field (H_{ext}) causes the total magnetic moment to precess around the direction of the precession torque field (H_{eff}) before relaxation processes damp this precession and the magnetization aligns with H_{eff}. If a thin film sample is irradiated with a transverse RF field and if the RF frequency coincides with the precessional frequency, the resonance condition is fulfilled and the maximum microwave power signal is absorbed by the sample. The motion of the magnetization is described by the Eq.(1). In FMR measurement setup, a signal from the microwave source was applied using a bias tee through a coaxial cable into the sample holder, which is further shorted to RF ground [3,10]. The precision measurements were carried out using microwave source, precision nanovoltmeter at 10 mW power level from 0.1 GHz to 10 GHz along with the sweeping dc magnetic field upto 0.2 T. The sample was mounted on a specially designed sample holder, which was simulated for its S parameters for impedance matching when subjected to the Electric field. Microwave power was applied in the presence of a magnetic field and measured the FMR voltage spectra across the film. In plane static magnetic field is applied and the dc voltage due to ISHE was measured on Pt film under FMR condition. FMR voltage spectra were measured across the film using microwave

ACCEPTED MANUSCRIPT

signal of 10 mW at the desired frequencies when the external magnetic field (H_{ext}) was applied upto 0.2T along the bilayer film plane.

Results and discussion

In spin Hall Effect, an RF current flowing through the nonmagnetic layer with strong spin-orbit coupling generates a spin current perpendicular to the interface. When this spin current reaches the interface, there is a spin transfer torque on the magnetization of the magnetic layer. To understand the influence of the bilayer interface on magnetization dynamics, many researchers investigated FMR induced spin pumping from ferromagnet to non-magnetic, detection with enhanced damping or dc voltage due to ISHE. The spin-current transmission across the bilayer interface is controlled by spin-mixing conductance. The magnitude of the damping torque is parameterized by the spin-Hall angle. It has been demonstrated that the spin-Hall-induced damping torque can switch the magnetization in a reliable and efficient way, facilitating the development of magnetic memory and non-volatile spin logic technologies.

X-ray diffraction pattern of $FeSi_{3\%}$ film is shown in **Fig.1(a).** Peak correspond to ~ 44 degree is assigned to $FeSi_{3\%}$ thin film. XRD measurement has shown that iron thin film has oriented crystalline structure. The AFM image of $FeSi_{3\%}$ is shown in **Fig.1(b)**.The surface roughness of $FeSi_{3\%}$ film was 17nm confirms good quality of film. Magnetization measurements of $FeSi_{3\%}$/Pt film has been carried out along and perpendicular direction of film plane as shown in **Fig.2**. The MH curve shows highest magnetic moment of 121emu/cc along the film-plane and 43emu/cc in perpendicular direction respectively. $FeSi_{3\%}$/Pt film shows coercivity of 205 Oe in parallel to the field direction. This may be due to pinning of domain walls by Si atoms which is more pronounced in parallel direction of thin film. Domain wall pinning by Si atoms also impedes the low field saturation of $FeSi_{3\%}$ film. Also, Pt thin film over $FeSi_{3\%}$ film anchored the surface atoms of magnetic domains resulting into increase in the coercivity. The film-plane magnetic moment is higher compared to moment recorded in perpendicular direction. It confirms an in-plane magnetic anisotropy growth in thin film. Due to higher film-plane magnetization in $FeSi_{3\%}$/Pt film FMR measurement taken in film-plane direction. The schematic illustrations of FMR experimental setup for electrical measurements and thin film representing torques acting on magnetization are shown in **Fig.3(a,b)**. Microwave signal was applied in x-z plane and external magnetic field along the y axis of a thin film sample. The dc voltage is measured along the x axis on Pt film. The resulting magnetization vector (M) is in y-z plane at an angle θ from the y axis. In FMR measurement, microwave signal and external magnetic field were transversely applied parallel to film-plane. Microwave RF induced magnetic field (H_{rf}) in FMR measurement was calculated from the dimension of sample. The injected charge current in thin film induces precession of spins around the magnetization direction. Precession of spins induces transfer of angular momentum which generates spin wave (magnon) [11]. The generated spin wave transfer at interface excited by FMR is known as spin pumping. Spin current induced by spin pumping in paramagnetic Pt layer generates charge current due to spin-orbit scattering which ultimately gives rise to an electrical potential difference. The magnitude of ISHE voltage generated in thin film was 12.6μV at 1.0 GHz, 10.4μV at 2.7 GHz and 5.5μV at 2.9 GHz for an incident microwave power of 10 mW corresponding to sweeping dc magnetic field upto 0.2T. Using FMR, voltage spectra of $FeSi_{3\%}$/Pt film has been plotted as shown in **Fig.4**, which is field derivative of absorbed microwave power at 1GHz frequency. The magnitude of FMR resonance field and shape of FMR spectra depend upon the magnetic field direction, magnetic anisotropy, magnetic inhomogeneity and quality of the bilayer thin film.

Ferromagnetic resonance is the absorption of microwave power by a magnetic film, measured as a function of the applied sweeping dc magnetic field. It is one of the methods to characterize the ferromagnetic layer. FMR technique has been used to determine the magnetic damping constant (α) of thin film from resonance magnetic field measure of voltage spectra. The α has been determined as 0.056 from FMR linewidth (ΔH) using the relation, $\alpha = (\gamma/2\pi f)\Delta H$, by linear fits to the linewidth data. The demagnetization field has been evaluated as 0.152T for $FeSi_{3\%}$/Pt bilayer thin film by one parameter fit to resonant frequencies. Broadband FMR measurements have been used to determine Gilbert damping factor (G) important for achieving high speed magnetization switching [12]. It is calculated using $G = \alpha\gamma M_s$, where α is damping coefficient, γ is gyromagnetic ratio and its value is 1.76 x 10^{11} rad.s^{-1}.T^{-1} and M_s=121kA/m is the saturation magnetization of thin film. The G has been calculated as 1.19 x 10^8 s^{-1}. The calculated values of damping constant and Gilbert damping factor for $FeSi_{3\%}$ film are in good agreement with the reported values of α = 0.028 and G as 1.26 x 10^8 s^{-1} for Fe thin films [8-12]. The FMR linewidth enhancement with frequency indicates that attenuation in magnetic thin film is governed by Gilbert damping, which is determined from the slope of ΔH(f). The microwave power linear dependence of induced Hall voltage at 1GHz confirms that

ACCEPTED MANUSCRIPT

spin pumping increases with increase in microwave power as shown in **Fig.5(a).** Further, frequency dependence of induced voltage with respect to in plane dc magnetic field has been shown in **Fig.5(b)**. Recently, there have been significant disagreements in the values of spin Hall angle (SHA) in Pt reported by different groups, ranging from 0.012 to 0.12. One of the effective techniques used to evaluate SHA is spin torque ferromagnetic resonance measurement. Spin wave injected into adjacent Pt layer has been quantified by measuring SHA. The relationship between charge current, j_c and transverse spin current, j_s is given by Eq.(2).

$$j_c = \theta_{SH}\left(\frac{2e}{\hbar}\right) j_s \times \sigma \qquad (2)$$

Where, θ_{SH}, e, ħ and σ are SHA, elementary charge, reduced Plank constant and spin polarization vector of the spin current respectively. Using Eq.(2) demagnetization field and magnetic saturation of thin film, SHA in Pt was evaluated as 0.071. SHA for Pt in $FeSi_{3\%}$/Pt thin film is in good agreement with reported values by Xin Fan *et al* & Ando *et al* [13-14].

When the spin current interacts with the magnetic moment in a ferromagnetic metal, the angular momentum and energy conservation give rise to the spin transfer torque. Depending on the direction, spin current either stabilizes or destabilizes the magnetization. Spin-current injection into Pt layer due to FMR depends upon spin-mixing conductance $G_{\uparrow\downarrow}$ that regulates SHA in Pt [15]. Using real part of spin mixing conductance and FMR linewidth evaluation, damping torque is achievable [16-19], whereas the change in resonance field yields a direct measure of the field torque [20-22]. For a Pt layer much thicker than its spin diffusion length, damping torque is proportional to the real part of the spin-mixing conductance, SHA and conductivity of the Pt layer [23-24]. Damping torque ($D\tau$) is calculated using Eq.(3).

$$D\tau = \frac{2\,\mathrm{Re}|G_{\uparrow\downarrow}|}{\sigma_{\mathrm{Pt}}/\lambda_{Pt}}\theta_{\mathrm{SH}} \qquad (3)$$

Where, θ_{SH} is the spin-Hall angle of Pt, λ_{Pt} is the spin diffusion length of Pt and σ_{Pt} is the conductivity of Pt layer. The damping torque has been calculated as 0.084. This value is similar to recently reported value of damping torque 0.087 by Kasai *et. al* [25]. A tabulated comparison of the reported work with our measurement results is shown in **table.1**.Though damping torque originates from spin-current generation within Pt layer, the magnitude of the torque depends on the transmission of spin current across the interface. The damping torque was measured to understand the influence of interface on magnetization. In addition to spin current injection, interfacial scattering effects [26-27], may contribute to both stronger damping and lower spin injection in Fe/Pt. Thus spin pumping is reduced and low value of inverse spin Hall voltage is obtained. These results of damping torque and spin mixing conductance are comparable with recently reported results [28-30], indicating that damping torque may be controlled by manipulating the bilayer interface.

Magnetoresistance (MR) of $FeSi_{3\%}$/Pt bilayer has been measured by applying dc magnetic field parallel to applied current as shown in **Fig.6**. MR of a material arises due to change in the value of its electrical resistance with applied DC magnetic field. The dependence of the resistance on the angle between current and magnetization in metallic ferromagnet is called anisotropic magnetoresistance (AMR). The physical origin of AMR is spin-orbit interaction, in contrast to giant magnetoresistance, which reflects the change in resistance. Effect of magnetization on ferromagnetic layer by spin reaching at interface due to spin orbit scattering has been schematically shown in inset of **Fig.6**. At interface spin current interacts with magnetic moment of ferromagnetic film. Spin electrons are absorbed when magnetic moment m of film is perpendicular to the spin polarization σ and spins reflected when m is parallel to the σ [31-32]. Magnetoresistance of $FeSi_{3\%}$/Pt bilayer interface has been measured by applying both current and magnetic field parallel to the film plane and longitudinal voltage was measured on Pt layer. The magnetic moment parallel to spin polarization direction shows increase in resistance by absorbing accumulated spins at interface, otherwise it exhibits decrease in resistance by reflecting the accumulated spins responsible for additional charge current. At applied field H lower than coercive value (H_c); the film shows maximum fall in magnetoresistance revealing contribution due to magnetization of $FeSi_{3\%}$ thin film. Below coercive field magnetic domains are not aligned with the applied direction. As a result magnetization is not parallel to the spin polarization direction and most of the spins accumulated at interface reflected into Pt layer generating the additional charge current thus resistance decreased. After H_c, magnetoresistance linearly increases due to magnetic moment of $FeSi_{3\%}$

ACCEPTED MANUSCRIPT

begins to align with applied magnetic field. Magnetic exchange interaction in ferromagnetic $FeSi_{3\%}$ film splits the majority and minority spins [33-36]. Spins accumulated at interface may get adsorbed or reflect depending on direction of magnetic layer spin. When the applied magnetic field direction is parallel to the charge current spin Hall magnetoresistance decreased exhibiting polarized spins accumulated at interface get reflected. While for reverse field, a larger change in magnetoresistance is observed. It may be due to spin accumulation at interface which may further get pinned by Si atoms on applying magnetic field.

Exchange magnetic field G_i responsible for accumulated spins to precess around magnetic moment m of ferromagnetic layer has been estimated by extrapolating high field Hall resistance to zero magnetic field [30]. Exchange field G_i [35-36], responsible for field torque of accumulated spins at interface has been experimentally measured by magnetoresistance $6x10^{14}$ $\Omega^{-1}m^{-2}$. It signifies accumulated spins at interface possesses discontinuous exchange field due to presence of Si atoms. Also interfacial spins precess about resultant magnetization due to magnetization in both directions in $FeSi_{3\%}$ film. It ultimately suppresses the spin transfer torque to Pt layer. Magnetization and quality of bilayer interface has the key role in transforming the spin current produced in $FeSi_{3\%}$ layer that is finally measured as DC voltage in Pt layer.

Conclusions

We have experimentally demonstrated that in $FeSi_{3\%}$/Pt thin film, magnetic exchange field induces a low field torque which has been obtained by magnetoresistance. Weak field torque due to spin-orbit interaction may enhance magnetic damping and reduces the spin transfer torque. FMR induced low ISHE voltage has been observed near the resonance field in $FeSi_{3\%}$/Pt thin film. The strength of spin pumping from ferromagnetic $FeSi_{3\%}$ layer to Pt layer depends on spin transfer torque at bilayer interface. The experimental result shows that the spin current interacts with FeSi magnetic moment and exerts anti-damping torque on it. Hence, magnetic moment can be manipulated by spin torque. In-plane magnetic anisotropy has been observed for ferromagnetic $FeSi_{3\%}$ film. Presence of Si atoms in $FeSi_{3\%}$ film impedes the magnetic exchange field at interface suppressing field torque. Weak field torque play an important role in magnetic exchange interaction at the interface, which reduces spin pumping to Pt layer thus low inverse spin Hall voltage has been obtained. The FMR based technique used here allows precise quantification of spin-orbit torques directly applicable to spin-current-driven devices.

Acknowledgements

Authors thank to the Director, CSIR-National Physical Laboratory India for his kind support and motivation to carry out this work.

References

1. Vignale, G.; *J. Supercond. Novel Magn.,* **2010**, 23: 3-10.
DOI: 10.1007/s10948-009-0547-9

2. Felser, C.; Fecher, G. H.; and Balke, B; *Angew. Chem., Int. Ed.*, **2007**, *46(5), 668–699*.
DOI:10.1002/anie.200601815

3. Ahmad, S; Shah, J; Puri, N, K.; Negi, P, S.; Kotnala, R, K.; *Appl. Phys. Lett.***, 2013,** *103, 032405*.
DOI:10.1063/1.4813833

4. Chen, Y.-T.; Takahashi, S.; Nakayama, H.; Althammer, M.; Goennenwein, S. T. B.; Saitoh, E.; and Bauer, G. E. W.; *Phys. Rev. B 87*, **2013**, 144411.
DOI:10.1103/PhysRevB.87.144411

5. Nakayama, H.; Althammer, M.; Chen, Y.-T.; Uchida, K.; Kajiwara, Y.; Kikuchi, D.; Ohtani, T.; Geprägs, S.; Opel, M.; Takahashi, S.; Gross, R.; Bauer, G. E. W.; Goennenwein, S. T. B.; and Saitoh, E.; *Phys. Rev. Lett*., **2013**, *110, 206601*.
DOI:10.1103/PhysRevLett.110.206601

6. Meyer, S.; Schlitz, R.; Geprägs, S.; Opel, M.; Huebl, H.; Gross, R.; and Goennenwein, S. T. B. ; *Appl. Phys. Lett. 106*, **2015**, *132402*.
DOI:10.1063/1.4916342

7. Liu, L.; Moriyama, T.; Ralph, D. C.; and Buhrman, R. A.; *Phys. Rev. Lett. 106*, **2011**, *036601*.
DOI:10.1103/*PhysRevLett*.106.036601

ACCEPTED MANUSCRIPT

8. Lee, S. J.; Tsai, C. C.; Cho, H.; Seo, M.; Eom, T.; Nam, W.; Lee, Y. P.; and Ketterson, J. B.; *J. Appl. Phys. 106*, **2009**, *063922*.
DOI:10.1063/1.3204648

9. Nan, T.; Emori, S.; Boone, C. T.; Wang X.; Oxholm T. M.; Jones, J. G.; Howe, B. M.; Brown, G. J.; and Sun, N. X.; *Phys. Rev. B*, **2015**, *91, 214416*.
DOI:10.1103/*PhysRevB*.91.214416

10. Ahmad, S.; Shah, J.; Katiyar A.; Chaujar R.; Puri, N. K.; Negi, P.S; and Kotnala, R.K.; *Indian J. Pure Appl. Phys.*, **2016**, *Vol. 54, pp 60-65*.

11. Brataas, A.; Kent, A. D.; and Ohno, H.; *Nat. Mater.*, **2012**, *3311*.
DOI:10.1038/*nmat3311*

12. Urban, R.; Wolterdorf, G.; and Heinrich, B.; *Phys. Rev. Lett.* 87, **2014**, *217204*.
DOI:10.1038/ncomms4768

13. Ando, K.; and Saitoh, E.; *Nat. Commun. 3*, **2012**, 1640.
DOI:10.1038/*ncomms1640*

14. Wang, H. L.; Du, C.H.; Pu, Y.; Adur, R.; Hammel, P.C.; and Yang, F.Y.; *Phys. Rev. Lett. 112*, **2014**, 197201.
DOI:10.1103/*PhysRevLett*.112.197201

15. Fan, X .; Wu, J.; Chen, Y.; Jerry, M. J. ; Zhang, H.; & Xiao J. Q. ; *Nat. Commun. 4*, **2013**, *1799*.
DOI:10.138/ncomms2709

16. Ando, K.; Takahashi, S.; Harii, K.; Sasage, K.; Ieda, J.; Maekawa, S.; and Saitoh, E.; *Phys. Rev. Lett. 101*, **2008**, *03660.*
DOI:10.1103/PhysRevLett.101.036601

17. Rojas-Sánchez, J.-C.; Reyren, N.; Laczkowski, P.; Savero, W.; Attané, J.-P.; Deranlot, C.; Jamet, M.; George, J.-M.; Vila, L.; and Jaffrès, H.; *Phys. Rev. Lett. 112*, **2014**, *106602*.
DOI:10.1103/PhysRevLett.112.106602

18. Nilsen, G. J.; Okamoto, Y.; Masuda, T.; Rodriguez-Carvajal, J.; Mutka, H.; Hansen T.; and Hiroi Z.; *Phys. Rev. B 91*, **2015**, *174435.*
DOI:10.1103/PhysRevB.91.174435

19. Oogane, M.; Wakitani, T.; Yakata, S.; Yilgin, R.; Ando, Y.; Sakuma, A.; and Miyazaki, T.; *Jpn. J. Appl. Phys*., **2006**, *45, 5A.*
DOI:10.1143/JJAP.45.3889

20. Pelzl, J.; Meckenstock, R.; Spoddig, D.; Schreiber, F.; Pflaum, J.; and Frait, Z.; *J. Phys.: Condens. Matter 15*, **2003**, *S451*.
DOI:10.1088/0953-8984/**15**/5/003

21. Shu-juan, Y.; Kun, XU.; Li-ming, YU.; Chao, J.; Shi-xun, C.; and Jin-cang, Z.*; J. Shanghai Univ.*, **2007**, *11(3).*
DOI:10.1007/s11741-007-0316-z

22. Czeschka, F.D.; Dreher, L.; Brandt, M. S.; Weiler, M.; Althammer, M.; Imort, I.-M.; Reiss, G.; Thomas, A.; Schoch, W.; Limmer, W.; Huebl, H.; Gross, R.; and Goennenwein, S. T. B.; *Phys. Rev. Lett. 107*, **2011**, 046601.
DOI:10.1103/PhysRevLett.107.046601

23. Weiler, M.; Shaw, J. M.; Nembach, H. T.; and Silva, T. J.; *IEEE Magn. Lett.* **2014**, *5, 1*.
DOI:10.1109/LMAG.2015.2438773

24. Wang, Y.; Deorani, P.; Qiu, X.; Kwon, J. H.; and Yang, H.; *Appl. Phys. Lett. 105*, **2014**, 152412.
DOI:10.1063/1.4898593

25. Kasai, S.; Kondou, K.; Sukegawa, H.; Mitani, S.; Tsukagoshi, K.; and Otani, Y.; *Appl. Phys. Lett.104*, **2014**, *092408*.
DOI:10.1063/1.4865425

26. Ando, K.; Takahashi, S.; Harii, K.; Sasage, K.; Ieda, J.; Maekawa, S.; and Saitoh, E.; *Phys. Rev. Lett. 101*, **2008**, *036601*.
DOI:10.1103/PhysRevLett.101.036601

27. Petit, S.; Baraduc, C.; Thirion, C.; Ebels, U.; Liu, Y.; Li, M.; Wang, P.; and Dieny, B.; *Phys. Rev. Lett. 98*, **2007**, 077203.
DOI:10.1103/PhysRevLett.98.077203

28. Haney, P. M.; Lee, H.-W.; Lee, K.-J.; Manchon, A.; and Stiles, M. D.; *Phys. Rev. B 87*, **2013**, *174411*.
DOI:10.1103/PhysRevB.87.174411

29. Pai, C.-F.; Nguyen, M.-H.; Belvin, C.; Vilela-Leão, L. H.; Ralph, D. C.; and Buhrman, R. A.; *Appl. Phys. Lett. 104*, **2014**, *082407*.
DOI:10.1063/1.4866965

ACCEPTED MANUSCRIPT

30. Niimi, Y.; Kawanishi, Y.; Wei, D.H.; Deranlot, C.; Yang, H.X.; Chshiev, M..; Valet, T.; Fert, A.; and Otani, Y.; *Phys. Rev. Lett. 109*, **2012**, *156602*.
DOI:10.1103/PhysRevLett.109.156602

31. Vlietstra, N.; Shan, J.; Castel, V.; van Wees, B. J.; and Youssef, J. B.; *Phys. Rev. B 87*, **2013**, *184421*.
DOI:10.1103/PhysRevB.87.184421

32. Althammer, M.; Meyer, S.; Nakayama, H.; Schreier, M.; Altmannshofer, S.; Weiler, M.; Huebl H.; Geprägs, S.; Opel, M.; Gross R.; Meier, D.; Klewe, C.; Kuschel, T.; Schmalhorst, J.-M.; Reiss, G.; Shen, L.; Gupta, A.; Chen, Y.-T.; Bauer, G. E. W.; Saitoh, E.; and Goennenwein, S. T. B.; *Phys. Rev. B 87*, **2013**, *224401*.
DOI:10.1103/PhysRevB.87.224401

33. Obata, K; Tatara, G; *Phys Rev. B 77*, 2008, 214429.
DOI:10.1103/*PhysRevB*.77.214429

34. Manchon, A; Zhang, S; *Phys. Rev. B 79*, 2009, 094422.
DOI:10.1103/*PhysRevB*.79.094422

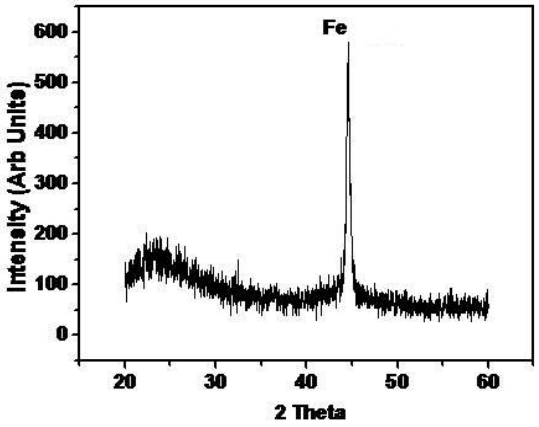
35. Ralpha, D.C.; Stilesb, M.D.; *J. Magn. Magn. Mater. 320*, **2008,** *1190–1216*.
DOI:10.1016/j.jmmm.2007.12.019

36. Haney, P. M.; Lee H.-W.; Lee, K.-J.; Manchon, A.; and Stiles, M. D.; *Phys. Rev. B 87*, **2013**, *174411*.
DOI:10.1103/PhysRevB.87.174411

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

SUPPORTING INFORMATION



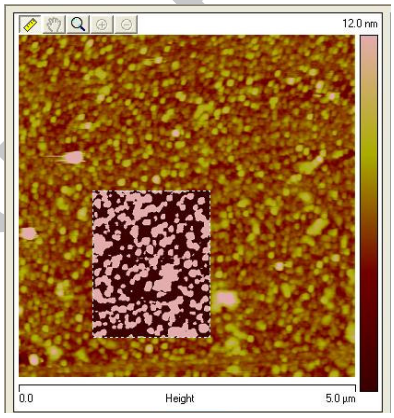


Fig.1a: XRD pattern of FeSi3% thin film , **Fig.1b:** AFM image of FeSi3% thin film

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

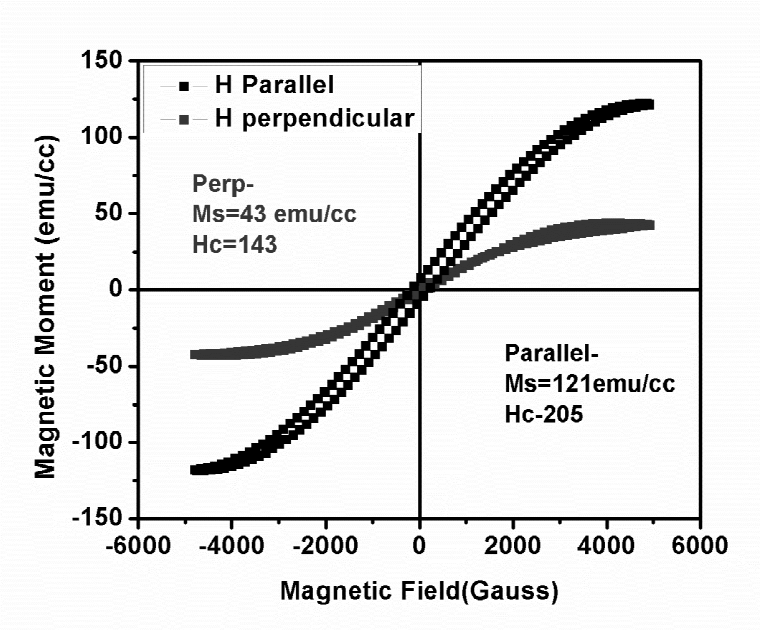


Fig.2. MH curve plotted parallel and perpendicular to the applied field for $FeSi_{3\%}$/Pt bilayer thin film

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

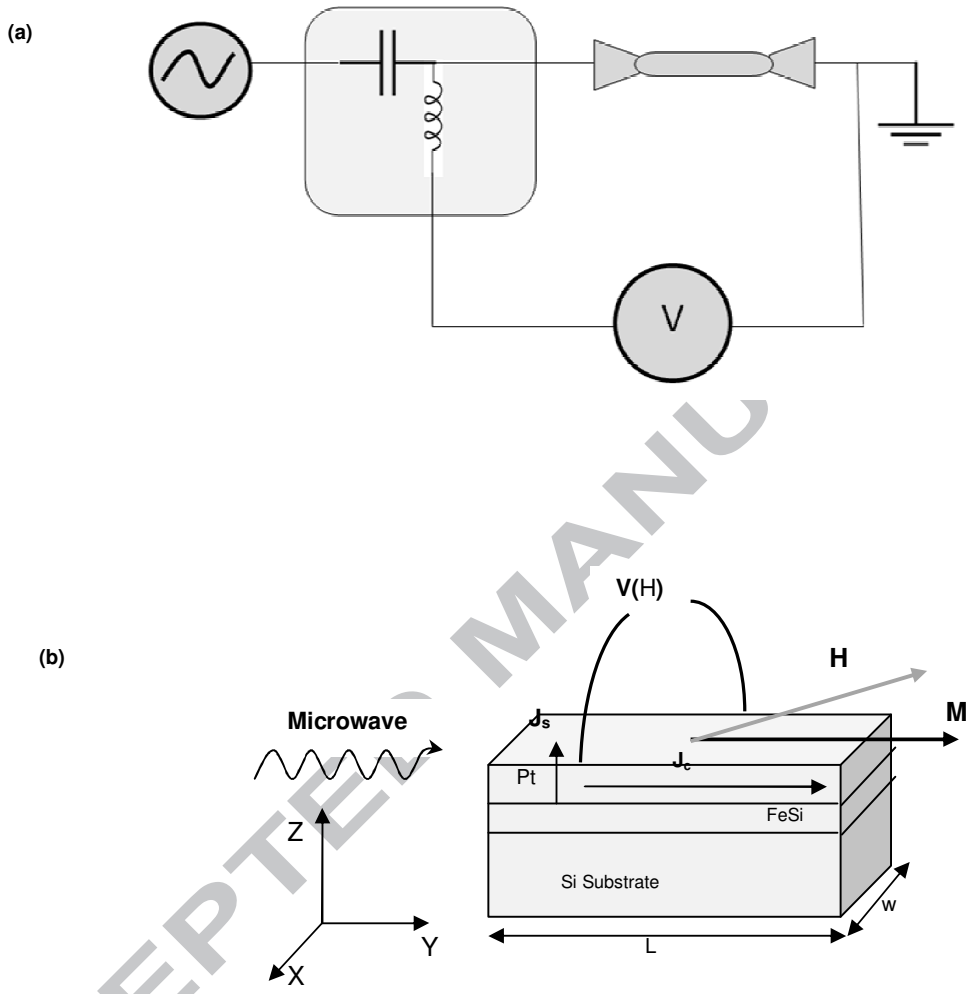


Fig.3. (a) Block diagram of FMR experimental setup, (b) Schematic illustration of bilayer thin film model

ACCEPTED MANUSCRIPT

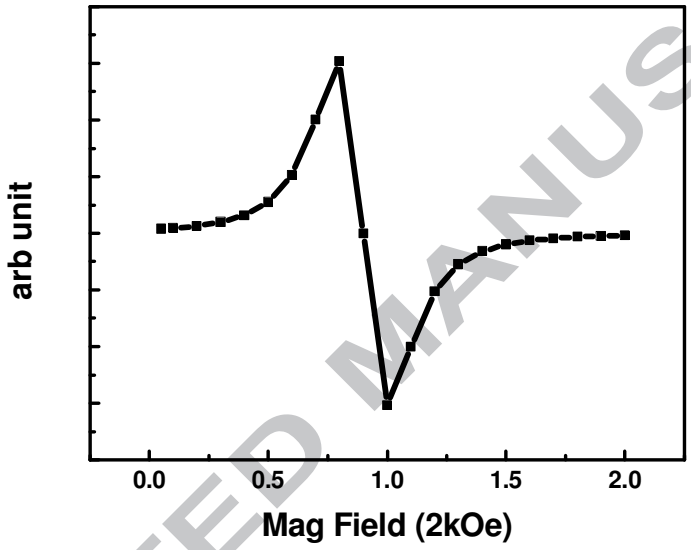


Fig.4. FMR voltage spectra of FeSi/Pt at 1GHz

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

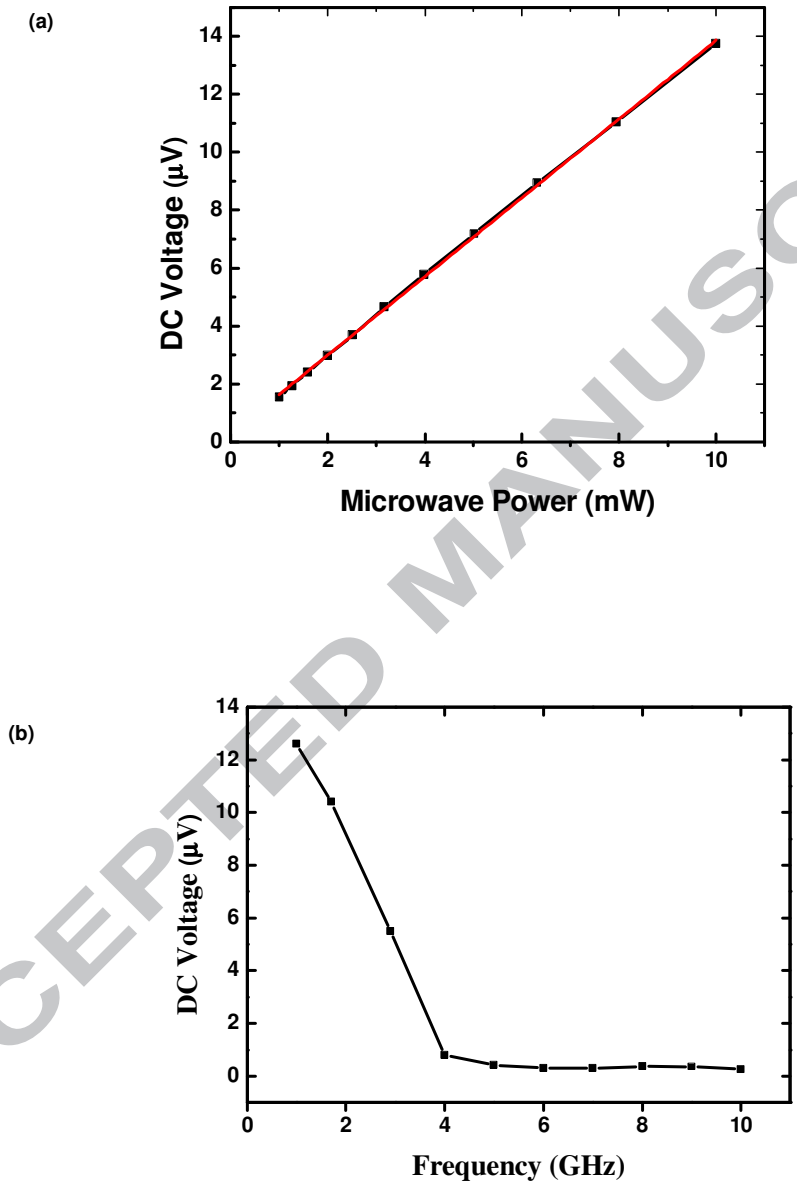


Fig.5 (a) Microwave power linear dependence of induced spin-Hall voltage at 1 GHz, (b) Frequency dependence of output Hall voltage wrt in plane dc magnetic field

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

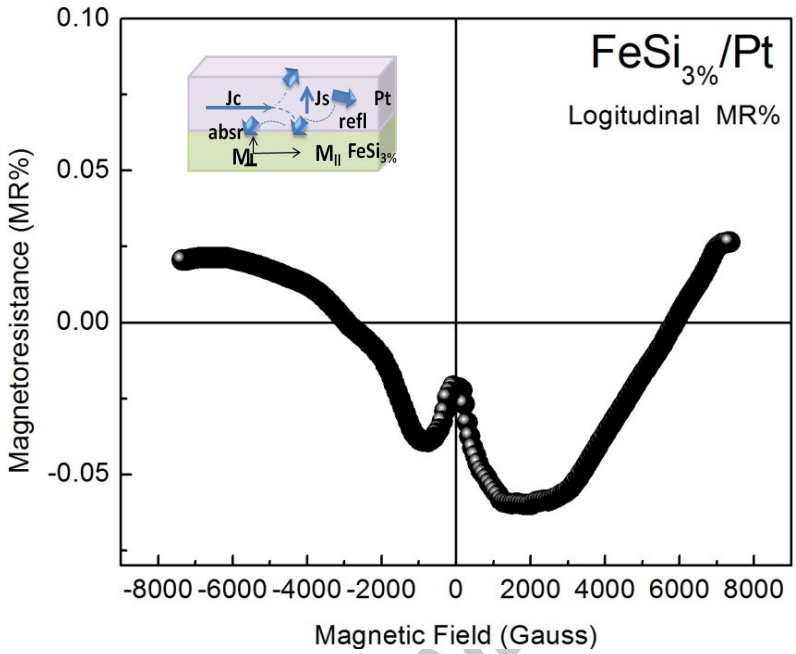


Fig.6. Magnetoresistance (MR%) vs. applied magnetic field for $FeSi_{3\%}$/Pt bilayer interface taken parallel to the applied field; inset shows schematic of magnetoresistance at bilayer with applied current and magnetic field parallel to each other.

ACCEPTED MANUSCRIPT

Table 1.Parameters related to FMR Induced ISHE

Parameters	Measurement Results	Reported Values	References
Damping Torque ($D\tau$)	**0.084**	**0.087**	**[25]**
Real Part of Spin mixing conductance (Gr) ~ ($10^{14}\ \Omega^{-1}m^{-2}$)	**6.0**	**8.1**	**[15]**
Magnetic damping (α)	**0.056**	**0.028**	**[8-9]**
Gilbert damping Factor (G) ~ ($10^{8}\ s^{-1}$)	**1.19**	**1.26**	**[8-9]**
Spin Hall Angle (θ_{SH})	**0.071**	**0.08**	**[13]**

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

HIGHLIGHTS

- Study of Inverse Spin Hall voltage in $FeSi_{3\%}$/Pt bilayer thin film
- Presence of Magnetic anisotropy has been obtained by Magnetic –Hysteresis loop.
- Ferromagnetic resonance induced inverse spin Hall voltage 12.6 μV at 1.0 GHz was obtained across Pt layer.
- The damping torque and spin Hall angle have been evaluated as 0.084 and 0.071 respectively.
- Weak field torque suppresses the spin pumping to Pt layer thus low value of inverse spin Hall voltage is obtained.

ACCEPTED MANUSCRIPT

International Journal of Electronics, Electrical and Computational System
IJEECS
ISSN 2348-117X
Volume 6, Issue 6
June 2017

Simulation for Position Control of DC Motor using Fuzzy Logic

Pankaj Kumari Meena
PG Scholar
Department of Electrical Engg., Delhi Technological University
New Delhi, India

Dr. Bharat Bhushan
Professor
Department of Electrical Engg.
Delhi Technological University
New Delhi,India

Abstract-The purpose of this project is to control the position of DC Motor by using Fuzzy Logic Controller (FLC) with MATLAB application. The scope includes the simulation and modeling of DC motor, fuzzy controller and conventional PID controller as benchmark to the performance of fuzzy system. The position control is an adaptation of Closed Circuit Television (CCTV) system.[1] Fuzzy Logic control can play important role because knowledge based design rules can be easily implemented in the system with unknown structure and it is going to be popular since the control design strategy is simple and practical. This make FLC an alternative method to the conventional PID control method used in nonlinear industrial system. The results obtained from FLC are compared with PID control for the dynamic response of the closed loop system. Overall performance show that FLC perform better than PID controller.

Keywords: Fuzzy Logic Controller, Position Control, Fuzzy Logic Toolbox, DC Servo Motor, MATLAB

I. INTRODUCTION

In recent years DC motors are used in high performance drive system because of its advantages. A DC servo motor which is usually a DC motor of low power rating is used as an actuator to drive a load[2]. It is having high ratio of starting torque to inertia and faster dynamic response. Because of their high reliability, flexibility and low cost, DC servo motors are widely used in industrial applications, robot manipulators and home appliances where speed and position control of motor are required. It is a common actuator in control system, it is directly provides the rotary and transitional motion and the field excitation is provided by permanent magnet and remains constant under all operating conditions. DC servo drive systems normally use the full four quadrant operations which allow bidirectional speed control with regenerative braking capabilities. The conventional methods have the following difficulties, it depends on the accuracy of the mathematical model of the system and the expected performance is not met due to the load disturbances. The classical linear control shows good performance only at one operating speeds. Fuzzy logic is a technique to embody human like thinking into a control system[4]. Fuzzy control has been primarily applied to the control of processes through fuzzy linguistic performances. The fuzzy controller gives the better dynamic performance as well as error reduction[5][12].

II. MATHEMATICAL MODEL OF THE DC SERVOMOTOR

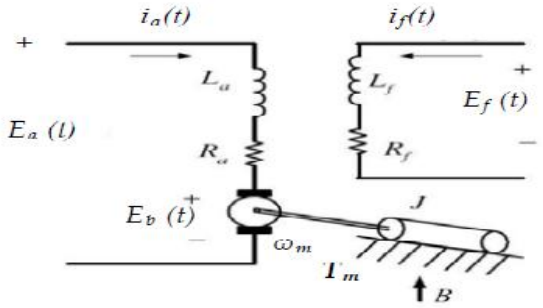


Fig.1. Schematic Representation of the DC servomotor

Fig.1. represented the servo motor model[1][2][3][6].

Let's consider:-

$E_a(t)$ =Input voltage

International Journal of Electronics, Electrical and Computational System
IJEECS
ISSN 2348-117X
Volume 6, Issue 6
June 2017

$i_a(t)$ =Armature current

R_a = Armature resistance

L_a = Armature inductance

$E_b(t)$ =Back e.m.f

T_m = Developed Torque

ω_m =Motor angular velocity

J=Motor moment of inertia *B*=Viscous friction coefficient K_b=Back e.m.f constant

K_T=Torque constant

Here, the differential equation of armature circuit is-

$$E_a(t) = R_a.i_a(t) + L_a.di_a(t)/dt + E_b(t) \quad (1)$$

The Torque equation is-

$$T_m(t) = J.d\omega_m(t)/dt + B.\omega_m(t) \quad (2)$$

The torque developed by motor is proportional to the product of the armature current and field current

i.e. $T_m(t) = K_f.i_f.i_a$ (3)

Where, K_f is constant. In armature – controlled D.C. motor the field current (i_f) is kept constant

i.e. $T_m = K_T.i_a$ (4)

Where, $K_T = K_f.i_f$ is torque constant. The back e.m.f. of motor is proportional to the speed

i.e. $E_b(t) = K_b.\omega_m$ (5)

Where, K_b is back e.m.f. constant. In order to create the block diagram of system initial conditions are zero and Laplace transform is implemented to the equations. i.e.

$$E_a(s) = R_a.I_a(s) + sL_a.I_a(s) + E_b(s)$$

$$I_a(s) = Ea(s) - E_b(s)/sL_a + R_a \quad (6)$$

$$T_m(s) = sJ.\omega_m(s) + B.\omega_m(s)$$

$$\omega_m(s) = T_m(s)/sJ + B \quad (7)$$

$$T_m(s) = K_T.I_a(s) \quad (8)$$

$$E_b(s) = K_b.\omega_m(s) \quad (9)$$

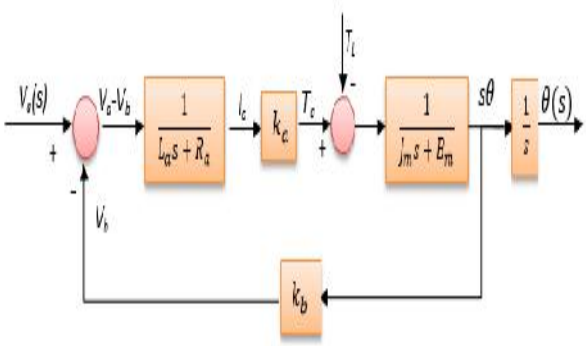


Fig.2.Block Diagram of DC motor

III. Fuzzy Logic Controller

The Fuzzy Logic Controller is designed to have two fuzzy state Variables and one control variable for achieving position control of the DC Servo Motor[5][10][12]. These two input variable are the error and change in error. The fuzzy logic controller initially converts the crisp error and change in error variables into fuzzy variables and then are mapped into linguistic labels. Membership functions of input 'velocity', 'error' and output 'control' are shown in fig. which consists of two inputs and one output. Different membership functions are used for the fuzzy set of the input and output vectors. There are seven linguistic variables for the input velocity, error, and output control. All the MFs are symmetrical for positive and negative values of the variables. The linguistic labels are divided into seven groups.

The seven linguistic variables are NB- negative big, NM - negative medium, NS - negative small, Z-zero, PS- positive small, PM – positive medium, PB - positive big.

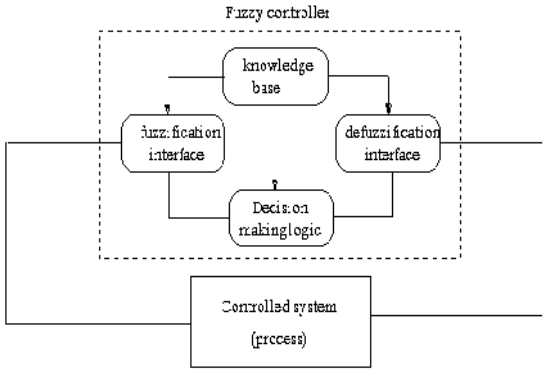


Fig.3. Basic structure of FLC

International Journal of Electronics, Electrical and Computational System
IJEECS
ISSN 2348-117X
Volume 6, Issue 6
June 2017

The fuzzy controller block for a DC motor is shown in figure 5. The three principal elements to a fuzzy logic controller are 1. Fuzzification module (Fuzzifier) 2. Rule base and Inference engine 3. Defuzzification module (Defuzzifier).

The fuzzification module assigns suitable linguistic values for the input data, this may be viewed as labels of fuzzy sets. The knowledge-base, consisting of input and output membership functions, along with the rule-base provides information for the appropriate fuzzification operations. In order to obtain the desired fuzzy control action, defuzzification is performed on the fuzzy input variables with the help of rulebase and inference mechanism. The rules commonly used are the IF, THEN rules. The inference mechanism is the kernel of fuzzy logic controller. The control rules are evaluated by an inference mechanism. Individual rule based inference (firing) is used here. The Mamdani algorithm is used for the implementation of the inference mechanism. Defuzzification is a technique of converting the final combined fuzzy conclusion into a crisp one. The defuzzified output is in turn applied to the plant.

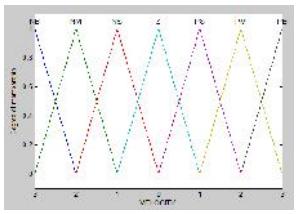


Fig.4(a)

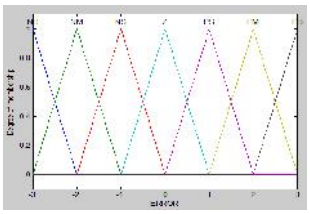


Fig.4(b)

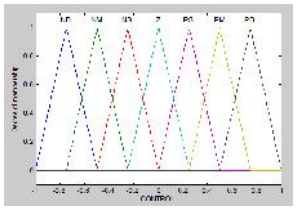


Fig.4(c)

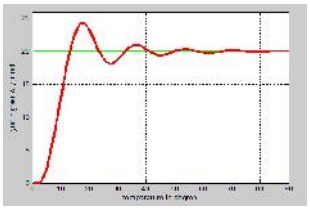


Fig.4(d)

Fig.4.Triangular Membership Functions of ‘e’, ‘ce’ and ‘u’ and output trajectory

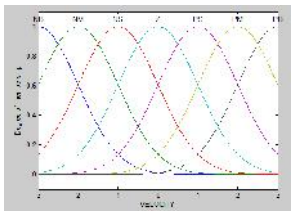


Fig.5(a)

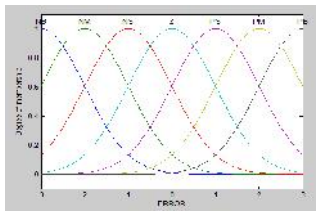


Fig.5(b)

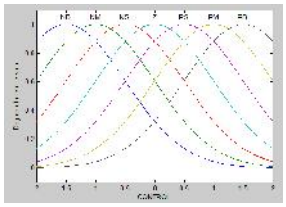


Fig.5(c)

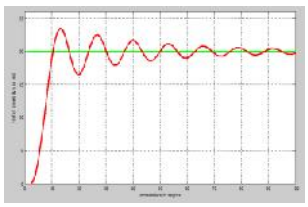


Fig.5(d)

Fig.5.Gaussian Membership Functions of ‘e’, ‘ce’ and ‘u and output trajectory

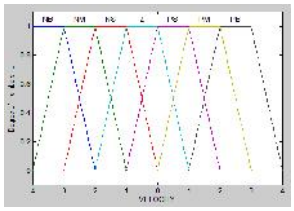


Fig.6(a)

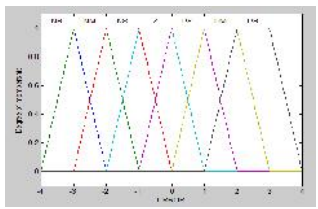


Fig.6(b)

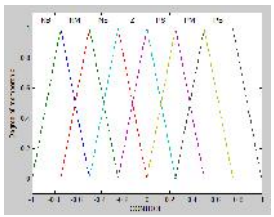


Fig.6(c)

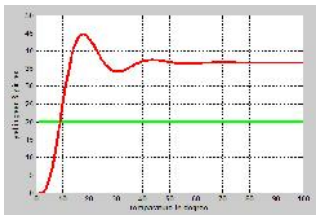


Fig.6(d)

Fig.6.Trapezoidal Membership Functions of ‘e’, ‘ce’ and ‘u’ and output trajectory

The fuzzy-rule-based matrix used for the proposed DC servo motor specific FLC algorithm is shown in table I [9]. The top row and left column of the matrix indicate the fuzzy set of the variables ‘velocity’ and ‘error’ respectively, and the MFs of the output variable ‘control’ are shown in the body of the

International Journal of Electronics, Electrical and Computational System
IJEECS
ISSN 2348-117X
Volume 6, Issue 6
June 2017

matrix. There are 7*7=49 possible rules in the matrix.

ce / e	NB	NM	NS	ZE	PS	PM	PB
NB	PB	PB	PB	PS	PS	PS	ZE
NM	PB	PM	PM	PS	PS	ZE	NS
NS	PB	PM	PS	ZE	ZE	NS	NS
ZE	PM	PS	ZE	ZE	ZE	NS	NS
PS	PM	PS	ZE	ZE	NS	NS	NM
PM	PS	PS	NS	NS	NS	NS	NM
PB	ZE	NS	NS	NM	NB	NB	NB

Table I: The control rules of Fuzzy Logic Controller

After designing the rule, we can get the surface viewer in Fig. that represent the rule of FLC. Where as the x-axis is error 'e' and y-axis change is error 'ce' and z-axis is output 'u'.

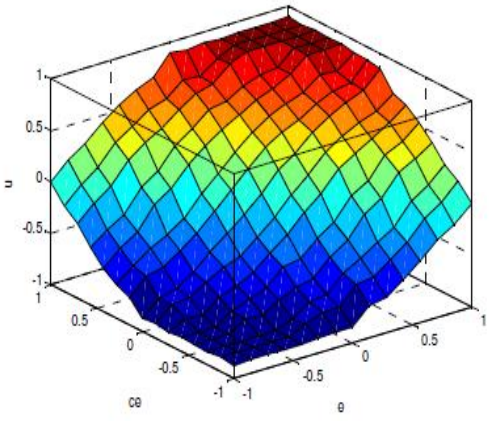


Fig.7. Surface Viewer of FLC

IV. CONCLUSION

A fuzzy logic based controller for a DC servomotor has been studied. The results have been compared to the conventional controller. The design of the fuzzy logic controller has been explained and the performance was evaluated by simulation. The simulation results indicate that FLC provides the best performance in comparison with PI controller, and the shape of the FLC surface is smoother than that of PI controller[7][8][13][14] . In Fuzzy logic control it is not necessary to change the control parameters at any conditions. It does not happen in conventional PI. It is clear that the fuzzy PI controller is more advantageous than conventional PI because the settling time is very low compared to the conventional PI Controller. It gives better dynamic response. This proposed scheme is very suitable for applications of industrial position control drives.

REFERENCES

[1] M. GOPAL, "Digital Control and State Variable Methods." Fourth edition
[2] I.J. Nagrath & M. Gopal, "Control Systems Engineering," New Age publication, Fifth edition.
[3] Math Works, Fuzzy Logic Toolbox User's Guide
[4] S.F. Abd-Alkarim "Application of Fuzzy Logic in Servo Motor",Al-Khwarizmi Engineering Journal 2007, Vol.3, No.2, pp. 8 -16
[5] H. Aloui, A. Ben Jridi, N. Chaker and R. Neji, "Fuzzy Control for Speed Tracking of an Electrical Vehicle enclosing an Axial-Flux Synchronous Motor", IEEE 14th International Conference on Sciences and Techniques of Automatic Control & Computer Engineering, 2013, pp. 36-42.
[6] K.Premkumar, B.V. Manikandan "Adaptive Fuzzy Logic Speed Controller for Brushless DC motor", IEEE International Conference on Power, Energy and Control, 2013, pp. 290.
[7] Modern control engineering 5th edition, K.Ogata prentice hall, 2012
[8] Zakharov Alexei, Halasz Sandor,"Robust Speed Fuzzy Logic Controller for DC Drive", 0-7803-3627-5/97/$5.00 IEEE, pp. 385-389, 1997
[9] Lee C. "Fuzzy Logic in Control Systems", Fuzzy logic controller, part II .IEEE Trans. On Systems, Man. And cybernetics 1990.
[10] Nasser T. and others "Design and Applications of Fuzzy Logic for The Speed Control of a Direct-Drive DC Motor", 4th International Conference, Cattaee, Amman-Jordan 2002.
[11] Zhao ZY, Tomizuka M.Isaka S."Fuzzy Gain Scheduling of PID Controllers" Proceedings of the IEEE.Conference on Control Applications. Dayton,Ohio 1992.
[12] Dr. Shereen F. Abd-Alkarim, "Application of Fuzzy Logic in Servo Motor", Al-Khwarizmi Engineering Journal,Vol.3, No.2, pp8 -16 (2007).
[13] G.-R. Yu, R.-C. Hwang, and C.-P. Lin, "Optimal Fuzzy Control of the Spindle Motor in a CD-ROM Drive Using Genetic Algorithms," presented at Asian Control Conference, 2004.
[14] "Control Tutorial for Matlab: DC Motor Speed Modeling,"
[15] I. K. Bousserhanel, A. Hazzabl, M. Rahli, M. Kamli, and B. Mazari, "Adaptive PI Controller using Fuzzy System Optimized by Genetic Algorithm for Induction Motor Control," presented at CIEP-IEEE, Puebla, Mexico, 2006.

4th International Conference on Emerging Trends in Engineering, Technology, Science and Management
Institution of Electronics and Telecommunication Engineers, IETE Janakpuri, New Delhi, India (ICETETSM-17)
16th July 2017, www.conferenceworld.in ISBN: 978-93-86171-54-2

Software Defined Internet of Things (IoT) in 5G Millimeter Wave (mmWave) Communication System

S. K. Agrawal[1], Dr. Kapil Sharma[2]

[1]*Samsung Research Institute Noida, (India)*

[2]*Computer Engineering Department, Delhi Technological University, (India)*

ABSTRACT

This research paper discusses software defined secure Internet of Things (IoT) data transmission in 5G millimeter wave (mmWave) communication system. The proposed software defined system dynamically interleaves the data before transmission in in 5G mmWave communication system. Software defined secure 5G mmWave communication system randomly transmits the dynamically interleaved data and the interleaved parameters to the desired IoT device. The IoT receiver performs dynamic de-interleaving of the received data using the received interleaved parameters.

Keywords - Millimeter Wave (mmWave); software defined (SD); Software Defined Random Skip Count (SDRSC)

I. INTRODUCTION

5G mmWave Communication is an emerging technique that helps to increase data capacity across the globe by having mmWave communication link. 5G mmWave communication will facilitate various advance IoT applications like IoT, Virtual Reality (VR) , Augmented Reality (AR) and etc. Further IoT devices are also involved in various transactions, data exchange, payments and so on. Further, 5G mmWave IoT Communication enabled devices can be used to view real time information and to purchase items. IoT devices are equipped with transceivers are capable of sending an encrypted data to the other IoT devices [1-5].

5G mmWave IoT communication system requires security for data. The 5G IoT data can be captured and decoded which will lead to loss of sensitive data. The data transmitted using 5G mmWave IoT communication system may be demolished using a jammer. The 5G mmWave IoT communication data can be prone to data modification attacks if not transmitted securely. 5G mmWave IoT communication can suffer man-in-the-middle attack and important transaction may take the incorrect results. Various kind of attacks can critically affect the 5G mmWave IoT communication. Interleaving has been used for avoiding the noise and it permits arrangement of data in specific row and column formats. In 5G mmWave IoT communication; there is a demand and need for a secure method for ensuring data security in 5G mmWave IoT communication system [6-9].

5G Millimeter Wave (mmWave) communication system growth also inviting attention of IoT along with security. 5G IoT will provide pervasiveness and ubiquity with mmWave communication system. IoT for a common user will increase interaction with user common devices and will interact in virtual or augmented mode having IoT connectivity. The method proposed here is reliable and highly precise secured to prevent IoT data loss and IoT data leakage [10-15].

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

This paper is organized as follows: Section II describe the dynamic interleaving in 5G mmWave IoT communication system. Sections III describe the dynamic interleaving steps in 5G mmWave IoT communication system. Further section IV describe the system architecture for dynamically data interleaving in 5G mmWave IoT communication system. This research object is to enable a secured 5G mmWave IoT communication system among IoT devices. It improves security in 5G mmWave IoT communication system by implementing dynamic interleaving using randomization.

II. DYNAMIC INTERLEAVING IN 5G MMWAVE IOT COMMUNICATION SYSTEM

Accordingly the research provides a system for facilitating a secure 5G mmWave IoT communication between a first IoT device and a second IoT device. In this system, the first IoT device transmits secure interleaved data dynamically to the second IoT device, using randomization technique. Further, the first device transmits the dynamically interleaved data to the second device. The first device also transmits dynamically interleaved parameters to the second device. Further, upon receiving the dynamically interleaved data and parameters from first device, the second device de-interleaves received data.

The research system enables secured IoT data communication by having randomizer based dynamic interleaving tool. In this system, data to be transmitted using 5G mmWave IoT communication system is interleaved in a specific format and interleaved data is sent to the other IoT device. The IoT transmitter also sends dynamic interleaving parameters which are used for interleaving the data, to the destination IoT device through secured 5G channel or by having a certificate authority (CA). The dynamically interleaved IoT data and dynamic interleaving parameters can be sent to the destination IoT device using same 5G communication channel. The dynamically interleaved IoT data and dynamic interleaving parameters can be sent to the destination IoT device through different channels and by using separate 5G communication channels for IoT data transmission increases security in the data transmission.

5G mmWave IoT communication system dynamic interleaving can be performed in 2-dimensional matrix. In a 2-dimensional model, the IoT data to be transmitted using 5G can be arranged in a 2 dimensional (2D) matrix by using SDR based architecture. 2D matrix has form of Rows and Columns. In this security method, the 5G mmWave IoT communication system has to calculate parameters such as quantity of rows (m), quantity of columns (n), data filling / releasing format. System further has capability for Random Row (Dm) / column Selection for filling the matrix and Random Column (Dn) / Row Selection for releasing the IoT data from matrix and so on for arranging the IoT data in matrix in the random fashion using SDR.

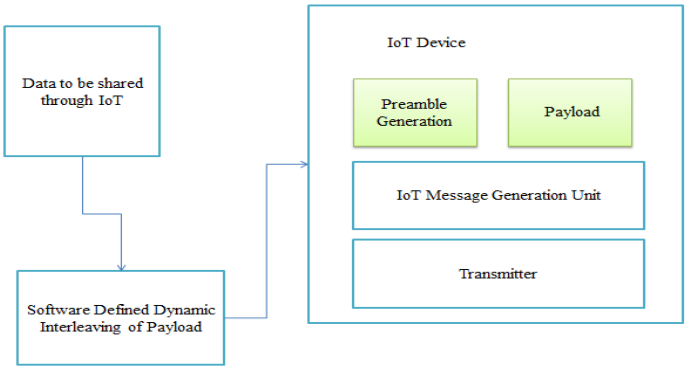


Figure 1: Dynamically Data Interleaving Before Transmission for 5G mmWave IoT Communication System

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

FIG. 1 illustrates a block diagram which shows various components of a IoT device. The IoT device comprises transmitter antenna; software defined dynamic interleaving of payload, IoT message generation unit, preamble generation and etc. The IoT transmitter antenna can be used to send and receive IoT and IoT message generation unit facilitates transfer of messages from the IoT device. The IoT device can send a new message and/or view a received message using the secure module. The IoT deice facilitates 5G mmWave IoT communication system between IoT devices.The proposed system enables secured data communication by incorporating a randomization based dynamic interleaving scheme. Using the dynamic interleaving scheme, the 5G mmWave IoT communication system devices interleaves the IoT message to be transmitted, in a specific format and transmit to the destination IoT device. The 5G mmWave IoT communication system device improves IoT data security in transmission.

III. DYNAMIC INTERLEAVING STEPS IN 5G MMWAVE IOT COMMUNICATION SYSTEM

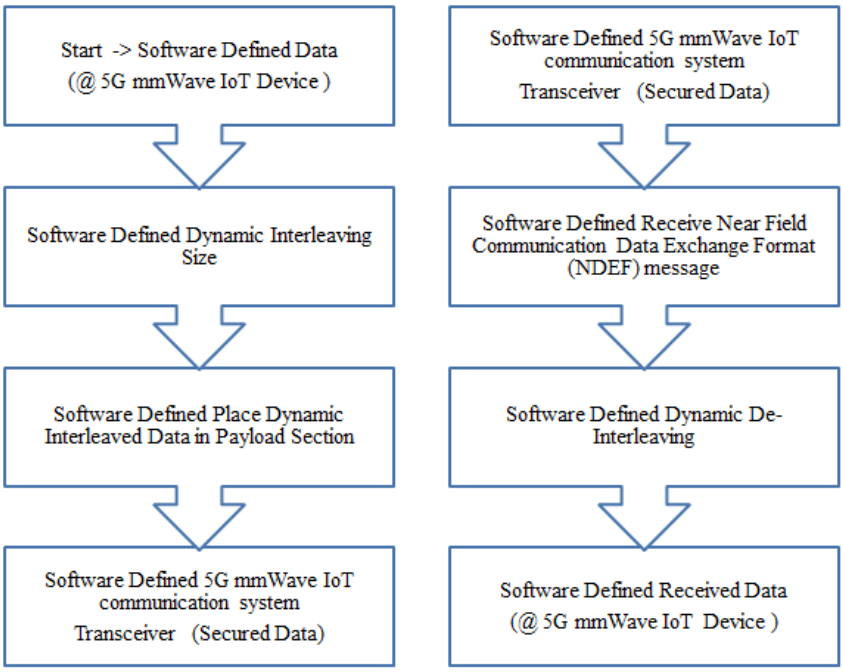


Figure 2: Dynamically Data Interleaving Between 5G mmWave IoT Communication Devices

FIG. 2 illustrates a various stages and steps of software defined secure Internet of Things (IoT) in 5G mmWave communication system. The IoT device comprises preamble generation element, dynamic interleaved payload generation element, IoT message generation element , transceiver, IoT message detection element , preamble detection element and dynamic de-interleaved payload generation element. The preamble generation element 301 generates preamble of the IOT message to be transmitted. The preamble part of the message can be used to obtain synchronization between the transmitted and received messages. An example of the preamble of a message can be header of the message. Further, length of the preamble affects packet overhead of the message and can accordingly affect transmission time of the message. The IoT dynamic interleaved payload generation element generates dynamically interleaved payload IoT data. The dynamic interleaved payload generation element implements a dynamic interleaving algorithm in order to ensure high level security in the IoT data

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

transfer. The IoT dynamic interleaved payload generation element interleaves the data to be transmitted into a specific format using randomization technique. IoT dynamic interleaved payload generation element can safeguard the data at the time of any attack or communication failure or hacking.

IoT message generation element generates the IoT message to be transmitted to a destination IoT device. The IoT message can be generated by arranging in order the preamble and the dynamically interleaved payload in a particular sequence as specified by the 5G communication protocol used for data transmission between the IoT devices. The IoT transmitter transmits the IoT message generated by the IOT message generation element and also transmits interleaved parameters to the destination IOT device. IoT transmitter can transmit the IOT message and the interleaved parameters using same 5G communication channel at different time. IoT transmitter can transmit the interleaved parameters through different secured 5G channels. IoT data transmission can also take place through an authorized certification authority (CA). IoT receiver element receives the dynamically interleaved data and/or the interleaved parameters transmitted by the IoT transmitter. The IoT preamble detection element processes the received interleaved information and detect the preamble in the received IOT message.

The IOT message is passed onto the dynamic de-interleaved payload generation element. The dynamic de-interleaved payload generation element performs de-interleaving of the received IOT message to identify the payload data. The dynamic de-interleaved payload generation element can use the interleaved parameters received from the IoT transmitter of the sending IoT device. Using the interleaved parameters , the dynamic de-interleaved payload generation element, the dynamic de-interleaved payload generation element perform de-interleaving of the received IOT message. IOT message detection element processes the received information and constructs the original IOT message.

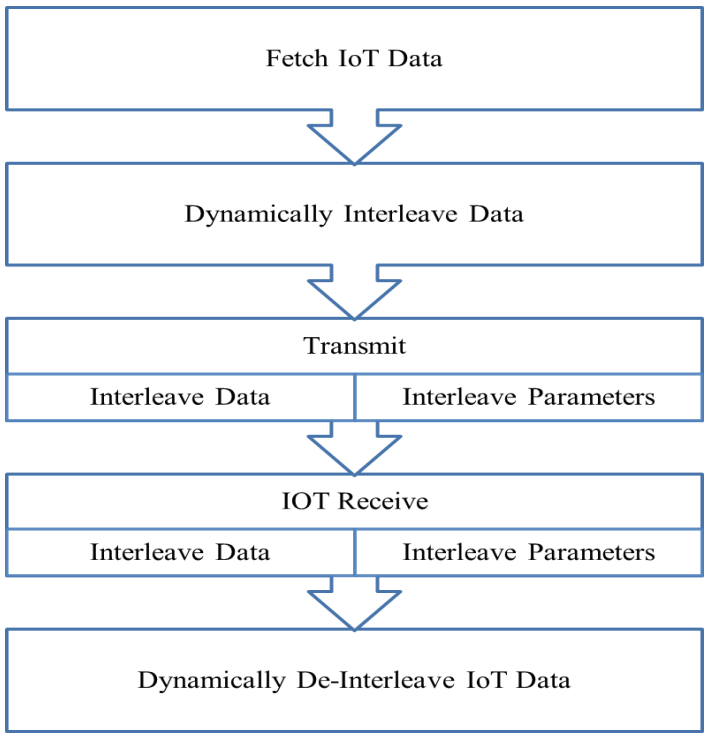


Figure 3: Dynamically Data Interleaving Steps

FIG. 3 illustrates research flow diagram which depicts various steps involved in the process of IoT dynamic interleaving and dynamic de-interleaving of data. IoT data is fetched by the dynamically interleaved payload

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

generation element and dynamically interleaved payload generation element performs dynamic interleaving of the data. IoT dynamic interleaving of data is performed by using randomization of data in interleaving fashion. The dynamic interleaved data is transmitted to the destination IoT device using a transmitter in the sender IoT device.

IoT data can be dynamically interleaved in the form of a 2-dimensional matrix. Software defined secure IoT system can use dedicated randomizers for the purpose of 2D interleaving and dynamic interleaving can also be done after generating the message or packet. The proposed dynamic interleaving mechanism does not require addition of any extra additional bits for security of the IoT data. This helps to maintain same IoT data size even if the IoT data is dynamically interleaved. Software defined secure IoT system parameters are not always shared with other IoT devices. If both IoT devices are paired or agreed to share the data based on certain parameters , the certain parameters can be used for the data sharing and no need to send the parameters every time with the data. In case of any attack or loss of data, the parameters are shared with the IoT device.

The interleaved data is received by the IoT receiver element in the receiver IoT device. The SDR receiver shall be aware that the dynamic interleaving is done after packet generation or before the packet generation. Further, the IoT receiver checks if the received data is new or if the data transmission is affected by hacking or data loss or any such issues. If the received data is new or if the data transmission is affected by hacking or data loss or any such issues, the receiver receives new IoT dynamic interleaved parameters from the data transmitter. If not, the receiver use (706) already received dynamic interleaved parameters to dynamically de-interleave the received de-interleaved data. In an , the proposed mechanism does not require the IOT receiver to receive dynamic interleaved parameters each time the data is transmitted. The IoT dynamic de-interleaved payload generation element present in the receiver IoT device performs dynamic de-interleaving of the received data using the dynamic interleaving parameters received from the transmitter IoT device.

II. SYSTEM ARCHITECTURE FOR DYNAMICALLY DATA INTERLEAVING IN 5G MMWAVE IOT COMMUNICATION SYSTEM

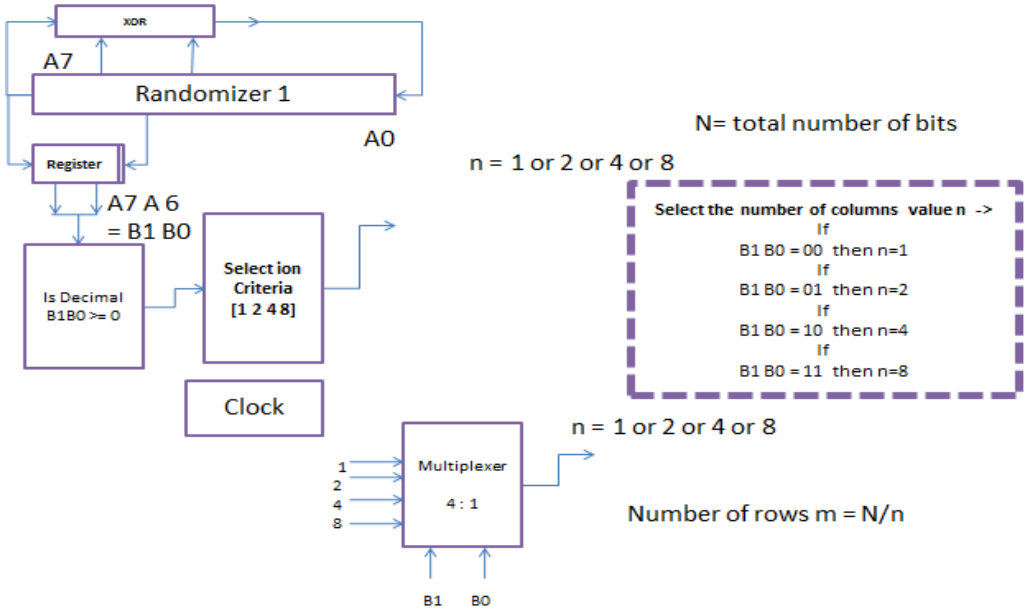


Figure 4: System Architecture for Dynamically Data Interleaving

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

Figure 4 illustrates software defined secure IoT system architecture for measuring number of rows and number of columns using randomizers. For a 2-dimensional matrix, if number of rows (m) is calculated using randomizer, then the number of columns (n) is calculated using a suitable equation or else if the number of columns (n) is calculated using randomizers, then the number of rows (m) is calculated using a suitable equation. In the proposed software defined secure IoT system architecture for calculating number of rows or columns, the randomizer is initiated by inputting a bit sequence. A random IoT clock shift is performed to the inputted IoT bit sequence and two bits from pre-defined location in the bit sequence are fetched for IoT. Fetched bits are compared with a IoT look up table and table can comprise number of rows (m) to be selected corresponding to the fetched bit sequences for IoT. Based on the value in table the system selects/decides the number of rows (m) for the 2D matrix. Software defined secure IoT system architecture for calculating data filling/releasing format for a 2D matrix, the randomizer is initialized by inputting a bit sequence of suitable length to the randomizer. The selected IoT data filling/releasing format is used for data filling/releasing to and from the 2D matrix. The bits can be picked from any place in the sequence, as set by the software defined secure IoT system and bits can be tapped from any set location.

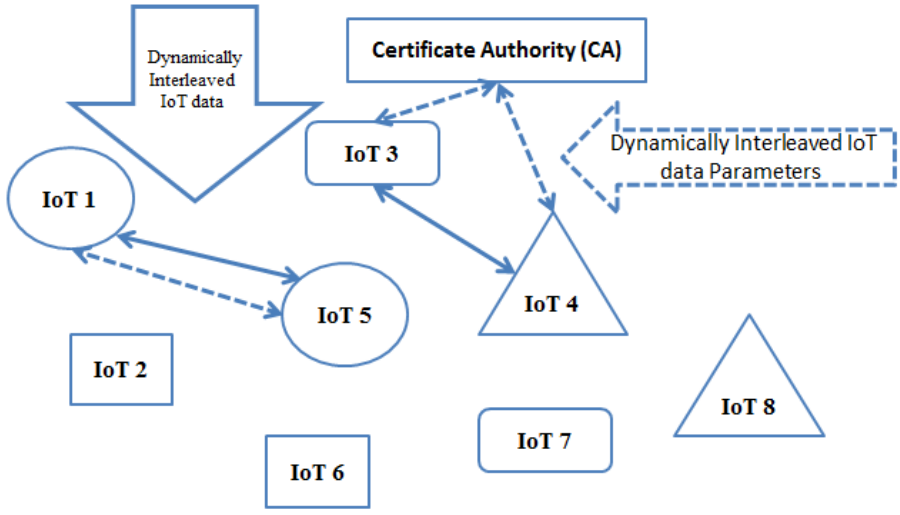


Figure 5: System Architecture for Sharing Dynamically Data Interleaving Data and Parameters

Figure 5 illustrates SDR system architecture for sharing dynamically interleaved IoT data and parameters between sender IoT and receiver IoT. The data to be transmitted from the IoT device to the destination IoT device is dynamically interleaved in the IoT devices like as shown IoT 1 to IoT 8. The dynamically interleaved data and the dynamic interleaving parameters are shared with the destination IoT device in one to one manner or through central server or CA. The dynamic interleaved IoT data and the dynamic interleaved IoT parameters can be shared through different 5G channels for better security. The dynamic interleaved IoT data and the dynamic interleaved IoT parameters can be shared through same channels having different time slots for better security. The dynamic interleaved IoT parameters corresponding to a dynamically interleaved data stream is shared only once between the IoT devices. The IoT parameters can be retransmitted for a IoT data only if any hacking takes place. The IoT parameters can be transmitted before the dynamically interleaved IoT data so that the receiver can perform real time de-interleaving of the IoT data. The channel used for sharing dynamically interleaved IoT parameters shall be secured channel and authorized certificate authority (CA) in 5G software defined secure IoT system.

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

V. CONCLUSION

This paper discussed about facilitating a secure Internet of Things (IoT) data transmission in 5G millimeter wave (mmWave) communication system. In this paper, the first IoT device dynamically interleave the information and transmits to the second IoT device by having randomized interleaving . The IoT device transmits the dynamically interleaved data to the other device and also transmits dynamically interleaved parameters to the second IoT device. After receiving the dynamically interleaved data and parameters from first IoT device, the second device de-interleaves received data. Accordingly the paper discusses a secure Internet of Things (IoT) data transmission in 5G millimeter wave (mmWave) communication system. In Future work; to improve security level in IoT data transmission, the system can use software defined random skip count (SDRSC) values to generate new random skip count value based parameters to dynamically interleave the IoT data .

VI. ACKNOWLEDGMENT

This paper is specifically meant for educational purpose for completing Ph.D. degree with Delhi Technological University (DTU), Delhi India. This paper is prepared for submission to under Ph.D. progress. The authors are very much thankful to Samsung Research Institute Noida, India and Delhi Technological University (DTU) earlier recognized as Delhi-College of Engineering (DCE), New Delhi, India.

REFERENCES

[1] Zhouye Pi et al., “An Introduction to Millimeter- Wave Mobile Broadband Systems, ” IEEE Communications Magazine, vol.49, no.6, pp.101- 107, Jun. 2011.

[2] Andrews J.G. et al., “What Will 5G Be?, ” IEEE JSAC Special Issue On 5G Wireless Communication Systems, May 2014.

[3] Theodore S. Rappaport et al., “Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!, ” IEEE Access, vol.1, pp.335-349, May 2013.

[4] Wonil Roh et al., “Millimeter-wave Beamforming as an Enabling Technology for 5G Cellular Communications: Theoretical Feasibility and Prototype Results, ” IEEE Communications Magazine, vol. 52, no. 2, pp. 106-113, Feb. 2014.

[5] Samsung, “5G Vision, ” DMC R&D Center, Samsung Electronics Co., Ltd., Feb. 2015.

[6] Wei L. et. al., “Key Elements to Enable Millimeter Wave Communications for 5G Wireless Systems, ” IEEE Wireless Communications, vol.1, pp.136-143, Dec. 2014.

[7] Palattella, Maria Rita, et al. "Internet of things in the 5G era: Enablers, architecture, and business models." IEEE Journal on Selected Areas in Communications 34.3 (2016): 510-527.

[8] Roman, Rodrigo, Jianying Zhou, and Javier Lopez. "On the features and challenges of security and privacy in distributed internet of things." Computer Networks 57.10 (2013): 2266-2279.

[9] Babar, Sachin, et al. "Proposed embedded security framework for internet of things (iot)." Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology (Wireless VITAE), 2011 2nd International Conference on. IEEE, 2011.

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

[10] S. K. Agrawal and K. Sharma, "Software Defined Millimeter Wave 5th Generation Communications System", Application and Theory of Computer Technology, [S.l.], v. 2, n. 1, p. 46-56, Jan. 2017. ISSN 2514-1694

[11] S. K. Agrawal and P. Garg, "Calculation of Channel Capacity Considering the Effect of Different Seasons for Higher Altitude Platform System", Wireless Personal Communications, Springer, Wireless Personal Communications: vol. pp-719-729, Issue 4 2010.

[12] S. K. Agrawal and K. Sharma, "5G Millimeter Wave (mmWave) Communication System with Software Defined Radio (SDR)", International Conference on Recent Trends in Engineering & Science (ICRTES-16), 29th-30th September 2016, ISBN : 978-93-86171-06-1

[13] S. K. Agrawal and K. Sharma, "5G Millimeter Wave (mmWave) Communication System with Software Defined Radio (SDR)", International Journal for Innovative Research in Science & Technology, Volume 2, Issue 9, September 2016

[14] J. Mitola. III, "Technical Challenges in the Globalization of Software Radio,'' IEEE Commun. Mag., Feb. 1999, pp. 84–89.

[15] Yang L-L, Hanzo L. Software-defined-radio-assisted adaptive broadband frequency hopping multicarrier DSCDMA. IEEE Commun Mag 2002;40(3):174–83.

cogent engineering

Stabilization of nonlinear inverted pendulum system using MOGA and APSO tuned nonlinear PID controller

Sudarshan. K.Valluru and Madhusudan Singh

Accepted Manuscript Version

This is the unedited version of the article as it appeared upon acceptance by the journal. A final edited version of the article in the journal format will be made available soon.

As a service to authors and researchers we publish this version of the accepted manuscript (AM) as soon as possible after acceptance. Copyediting, typesetting, and review of the resulting proof will be undertaken on this manuscript before final publication of the Version of Record (VoR). Please note that during production and pre-press, errors may be discovered which could affect the content.

© 2017 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

Publisher: Cogent OA
Journal: *Cogent Engineering*
DOI: http://doi.org/10.1080/23311916.2017.1357314

ACCEPTED MANUSCRIPT

Check for updates

Stabilization of nonlinear inverted pendulum system using MOGA and APSO tuned nonlinear PID controller

Sudarshan. K.Valluru and Madhusudan Singh

Department of Electrical Engineering, Delhi Technological University, Delhi-110042, India

E-mail: sudarshan_valluru@dce.ac.in, madhusudan@dce.ac.in

Abstract An inverted pendulum system (IPS) is a highly nonlinear dynamical open loop unstable system, typically used as a benchmark to verify the performance of controllers. The IPS emulates the behaviour of an altitude control of a space booster or rocket on take-off. The problem is to develop suitable controllers to maintain the stabilization and swing up of an inverted pendulum on a cart. This paper presents the evolutionary tuning methods of nonlinear PID (NL-PID) controller for IPS with the multi-objective genetic algorithm (MOGA) and adaptive particle swarm optimization (APSO) algorithm. The function of NL-PID controllers is to keep the pendulum in an upright position by maintaining the pendulum at same state and angle at zero degrees. The comparison of responses and performance of MOGA tuned NL-PID and APSO tuned NL-PID controllers for an IPS are described. The mathematical modeling and simulation analysis of the IPS is presented in detail to test the effectiveness of controller tuning algorithm. The APSO based tuning of the NL-PID controller has lesser chattering, noise and fast settling time than MOGA based tuning of the controller.

Keywords *Inverted pendulum system; NL- PID controller; Tuning; Multi-objective genetic algorithm; Adaptive particle swarm optimization algorithm*

ACCEPTED MANUSCRIPT

1. Introduction

Most of the physical systems are nonlinear dynamic systems which exhibit multiple equilibrium points. The stability analysis of the nonlinear system in the presence of unknown disturbances, uncertainties in system model and system parameter variations are more complex when compared to the linear systems (Sastry.S, 1999; Vidyasagar.M., 1993). The inverted pendulum system is one of the established benchmark problems in control literature as well as one of the most complicated dynamical systems according to report of the International Federation of Automatic Control (IFAC) theory committee. It is a standard benchmark problem (E.J.Davison, 1990; Jonathan P. How, 2015) to test controllers and their tuning algorithms due to the presence of higher order nonlinearities and instability problem (Aracil, J. Gordillo, 2004; Geva & Sitte, 1993). It has a resemblance to many complex practical engineering systems such as helicopter control, ballistic missile or space flight launching control and humanoid mobile robots, etc.

The inverted pendulum system is an upright pendulum with its fulcrum fixed at the top of the cart that can move horizontally and controlled such that the pendulum is always balanced and upright. The cart is free to move in the horizontal plane in response to an externally applied force. Consider a situation that the pendulum is in an upright position on the cart without control action, now a slight position change of cart would destabilize the pendulum causing instability in the system. The objective of the controller is to regulate control input i.e. the torque on the cart and keep the pendulum in an upright position by maintaining pendulum angle at same state and zero degrees. Extensive research work on control of dynamical systems have been reported and available in three subgroups such as classical, adaptive and artificial intelligence based controllers. PI, PID, and LQR controllers are classical controllers and suggested for linear systems control. The classification of sliding mode and model predictive controllers belongs to adaptive controllers for nonlinear regimes. The artificial intelligence based controllers such as fuzzy controllers, neuro-controllers, hybrid neuro-fuzzy controllers and metaheuristic algorithm tuned PID controllers (Astrom & Kumar, 2014) are more suitable for nonlinear dynamical systems. Despite the advent of many controllers, PID controller is still one of the most simple and easy to implement in the stabilization of inverted pendulum system. It offers simple

ACCEPTED MANUSCRIPT

structure, practicability, robustness, cost-effective solution with good performance and simple tuning methods. Most of the conventional tuning methods require tedious numerical computations to get appropriate PID parameters. Tuning of PID parameters are very critical, and it is anticipated to find optimal parameters to give satisfactory response for an inverted pendulum. In Recent past, besides the conventional tuning methods, several researchers have attempted for optimisation of PID gain parameters exclusively based on metaheuristic algorithms. The genetic algorithm (GA) (Daley & Liu, 1999; Shen, 2002) and particle swarm optimisation algorithm (Akihiro Oi, Nakazawa, & Matsui, 2008; Banks, Vincent, & Anyakoha, 2007, 2008; Zhao, Li, & Qian, 2005) tuning methods yields optimal results than conventional tuning techniques. These metaheuristic algorithms are highly robust since the efficiency does not depend upon the characteristics of the inverted pendulum system. The application of GA and standard PSO for tuning of linear PID controller in the linear dynamical system has emerged significantly in last few decades.

Chinichian and Kashani presented design and analysis of state variable feedback controller for a state space linear model of inverted pendulum system mounted on a motor driven cart (Chinichian, M., Kashani, 1989). Murata et al. presented the dynamical modeling and nonlinear observer based controller with linear feedback gains to stabilize the inverted pendulum mounted on cart system (Murata, N. Matsuno, H. Moran, A. Hayase, 1995). Tuning of PID controller is difficult due to parameters variations Doris et al. and Omatu et al. have applied genetic algorithm tuned neuro PID controller to stabilize the pendulum mounted cart system (Deris, S. Omatu, S. Kitagawa, 1995; Omatu & Yoshioka, 1998). Medrano illustrates the design of a robust digital controller with a blend of state space and frequency methods to balance the inverted pendulum system on moving cart (Medrano-Cersa, 1999). An energy based control strategy is described by Åström and Furuta to control and swing up of an inverted pendulum system on the cart (Åström & Furuta, 2000). Patnaik and Behera have reported the design and implementation of a multi-objective genetic algorithm based classic quadratic regulator, standard H_∞ and mixed H_2 / H_∞ controllers for the inverted pendulum (Patnaik & Behera, 2008). The response of PID plus LQR has been compared to that of simple PID on a pendulum cart system by Prasad et al. (Prasad, L.B., Tyagi, B., Gupta, 2011). Aouaouda et al. consider the tracking problem of uncertain Takagi-Sugeno fuzzy continuous systems subject to unknown inputs (S

ACCEPTED MANUSCRIPT

Aouaouda, Chadli, Khadir, & Bouarar, 2012). Sofiane Bououden et al. introduced a new approach for designing an adaptive fuzzy model predictive control using standard particle swarm optimization algorithm (Sofiane Bououden, Mohammed Chadli, Fouad Allouani, 2013). Durand et al. proposed energy control and event-based state feedback control strategies for unstable inverted pendulum system on cart system and demonstrated the computational cost efficiency on real-time test bed (Durand, Guerrero-castellanos, & Marchand, 2013). Mohammed Chadli et al. presented a fault-tolerant tracking controller for a verticle take-off and landing aircraft flight in uncertain conditions (Chadli, Aouaouda, Karimi, & Shi, 2013). Aouaouda et al. described the design of a robust fault estimation and fault tolerant control for the vehicle lateral dynamics subject to external disturbance and unknown faults (Sabrina Aouaouda, Chadli, & Karimi, 2014). Sofiane Bououden et al. proposed ant colony optimization for designing an adaptive fuzzy model predictive control (Bououden, Chadli, & Karimi, 2015). Aouaouda et al. have designed a multi-objective H_- / H_∞ sensor fault detection observer for a discrete-time nonlinear system subject to parameter uncertainties(S Aouaouda, Chadli, Shi, & Karimi, 2015). An output feedback controller is proposed by Lee et al. to stabilize and trajectory control for inverted pendulum on cart system (Lee, Mukherjee, & Khalil, 2015). Korkmaz et al. disfavor the classical tuning of PID besides this, they proposed the design of nonlinear PID (NL-PID) and compared with GA tuned PID controller (Korkmaz.Mehmet, Ömer Aydoğdu, 2012). Different PSO algorithms are stated by Menhas et al. to design optimal PI and PID controllers for a MIMO system(Menhas, Wang, Fei, & Pan, 2012). Valluru et al. have made a comparative study of untuned PID, NARMA L-2, and standard PSO tuned linear PID for a nonlinear dynamical system (Sudarshan.K.Valluru, MadhusudanSingh, & Bhushan, 2015).

This paper presents a performance evaluation of the inverted pendulum system with MOGA tuned NL-PID, and APSO tuned, NL-PID controllers instead of routine GA and standard PSO tuning of linear PID controller. The present study focuses on detailed investigations, and performance evaluation of the NL-PID controlled inverted pendulum system for optimal gain tuning by new variants of metaheuristics i.e. MOGA and APSO based optimization of controller parameters. The simulation results validate the system performance index regarding chattering, noise and settling time.

ACCEPTED MANUSCRIPT

The paper is organized in five sections commencing with the introduction followed by Section 2, which analyzes mathematical modeling of inverted pendulum system and its control strategies. In section 3 the design of nonlinear PID controller for an inverted pendulum system is presented. Section 4 discusses pseudo codes of the MOGA, APSO algorithms and pendulum angle response of IPS is investigated with NL-PID controller. Section 5 presents the conclusion.

2. Mathematical modeling of inverted pendulum system

The inverted pendulum system (Chatterjee, Patra, & Joglekar, 2002) involves a cart which is capable of moving backward and forward with a pendulum fulcrum at the bottom of its limited length such that the pendulum can move along with cart in the same plane. The pendulum attached to the cart is free to fall along the cart axis of motion. The pendulum movement is constrained in a vertical plane while the cart can move only in the horizontal plane. The control input force F is used to control the motion and hence the displacement x of the cart in the horizontal plane while indirectly controlling the angular position of the pendulum angle θ_p. The value of the input force depends upon control voltage (u) applied to DC motor, which drives the cart. The cart position and pendulum angle are monitored and used to change control signal so that the pendulum remains in stable region and upright at zero angles and does not affect even by small disturbance on the cart. A schematic representation combined with free body diagram is shown in the Fig.1

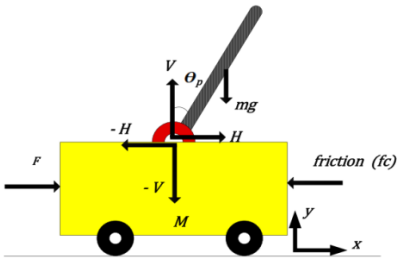


Fig.1 Schematic of Inverted Pendulum

The modeling of the inverted pendulum system is challenging due to its inherent nonlinearities and instability. The mathematical modeling is based on the motion patterns through physical and mechanical means to set up the input-output relationship in the system.

ACCEPTED MANUSCRIPT

2.1 Inverted pendulum system model equations

The nomenclature of the parameters and variables used in modeling of inverted pendulum system are given in Table. 1

Table 1 Parameters and variables of inverted pendulum system

F	Force applied to the cart by motor
M	Mass of the cart
f_c	Frictional force of the cart wheels on the track
m	mass of the pendulum or pole
$2l$	Length of the pendulum or pole
I	Moment of inertia of the pendulum or pole at the center of gravity
x	Cart position coordinate
y	Pendulum or pole position coordinate
θ_p	Pendulum or pole angle measured from vertical axis
H	Horizontal component of the reaction force at the fulcrum
V	Vertical component of the reaction force at the fulcrum

The dynamic equations (Guida, Nilvetti, & Pappalardo, 2010; Teixeira & Zak, 1999) of the system using Newton's second law described as:

$$\ddot{x} = -mal\ddot{\theta}_p \cos\theta_p + mal\,\dot{\theta}^2 \sin\theta_p - af_c + aF \quad (1)$$

$$\ddot{\theta}_p = \frac{mgl\sin\theta_p + m^2l^2a\dot{\theta}_p^2\sin\theta_p\cos\theta_p}{I - m^2l^2a\cos^2\theta_p + ml^2} \quad (2)$$

Where $a = \frac{1}{(m+M)}$

The equations (1) and (2) denotes the nonlinear dynamic model of the inverted pendulum system.

2.2 Control Strategy and Stability Analysis

The nonlinear dynamic equations for $\ddot{x}$ and $\ddot{\theta}_p$ may be represented into a state space form by considering the four state variables as:

$x_1 = \theta_p$= pendulum angle measured from vertical

$x_2 = \dot{\theta}_p =$ Pendulum angular velocity

$x_3 = x =$ Cart position

ACCEPTED MANUSCRIPT

$x_4 = \dot{x} =$ Cart velocity

Therefore the nonlinear dynamic equations in the state space form rewritten as follows

$$\dot{x_1} = x_2 \tag{3}$$

$$\dot{x_2} = \frac{g\sin x_1 - mlax_2^2 sinx_1 \cos x_1}{\frac{4l}{3} - mla\cos^2 x_1} - \frac{mla\cos x_1 (F - f_c)}{\frac{4ml^2}{3} - m^2 l^2 a\cos^2 x_1} \tag{4}$$

$$\dot{x_3} = x_4 \tag{5}$$

$$\dot{x_4} = \frac{\frac{-mag\sin(2x_1)}{2} + \frac{4ml}{3} ax_2^2 sinx_1 + \frac{4a(F - f_c)}{3}}{\frac{4}{3} - macos^2 x_1} \tag{6}$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} x_2 \\ \frac{g\sin x_1 - mlax_2^2 sinx_1 \cos x_1}{\frac{4l}{3} - mla\cos^2 x_1} \\ x_4 \\ \frac{\frac{-mag\sin(2x_1)}{2} + \frac{4ml}{3} ax_2^2 sinx_1}{\frac{4}{3} - macos^2 x_1} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{a\cos x_1}{\frac{4l}{3} - mla\cos^2 x_1} \\ 0 \\ \frac{\frac{4a}{3}}{\frac{4}{3} - macos^2 x_1} \end{bmatrix} (F - f_c) \tag{7}$$

Applying the first order Taylor series approximations, nonlinear dynamics of the inverted pendulum system shall be linearized around (Prasad, L.B., Tyagi, B., Gupta, 2011) the equilibrium point $x = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix}^T = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}^T$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{3g}{4l - 3mla} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{-3mag}{4 - 3ma} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{3a}{3mla - 4l} \\ 0 \\ \frac{4a}{4 - 3ma} \end{bmatrix} F \tag{8}$$

The parameters of the system chosen as $l = 0.3m$, $M = 0.5Kg$, $m = 0.2Kg$, $a = 1.428\ Kg^{-1}$, $I = 0.006\ \ Kg.m^2$. The system matrices are

$$[A] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 31.18 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -2.67 & 0 & 0 & 0 \end{bmatrix}; [B] = \begin{bmatrix} 0 \\ -4.547 \\ 0 \\ 1.8172 \end{bmatrix}; [C] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}; [D] = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

According to direct Lyapunov method(Chadli.M., 2013; Khalil, 2002), the controller state can define as $\theta_p(x_1, x_2)$ which is a function of the control is input

ACCEPTED MANUSCRIPT

$(F - f_c)$. Now by assuming the system equilibrium at origin i.e., $\theta_p(0,0)$ and choosing positive definite matrix P as $\begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}$, the energy equation is

$$V(x) = \frac{1}{2} X^T P X \tag{9}$$

Where $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$$\dot{V} = [x_1 x_2] \begin{bmatrix} x_2 \\ \frac{g \sin x_1 - mlax_2^2 sinx_1 \cos x_1}{\frac{4l}{3} - mla\cos^2 x_1} - \frac{mla\cos x_1 (F - f_c)}{\frac{4ml^2}{3} - m^2 l^2 a\cos^2 x_1} \end{bmatrix} \tag{10}$$

Let $V{:}\mathcal{D} \rightarrow \mathbb{R}$ be a continuously differential function such that $\dot{V} < 0$ to verify the asymptotic stability. To prove this there exists a scalar $\beta > 0$ (Khalil, 2002), then the asymptotic stability of the system satisfies the control input.

$$\theta_p(x_1, x_2) > \left[\frac{\frac{4l}{3} - mla\cos^2 x_1}{mla\, cosx_1} \right] \left[\frac{\beta}{x_2} + x_1 - \left(\frac{g \sin x_1 - \frac{mla}{2} x_2^2 sin2x_1}{\frac{4l}{3} - mla\cos^2 x_1} \right) \right] \tag{11}$$

To ensure asymptotic stability, the control input of inverted pendulum system must be as follows

$$\theta_p(x_1, x_2) > \left[\frac{0.4 - 0.08568\cos^2 x_1}{0.08568\, cosx_1} \right] \left[\frac{\beta}{x_2} + x_1 - \left(\frac{9.81 \sin x_1 - 0.04284 x_2^2 sin2x_1}{0.4 - 0.08568\cos^2 x_1} \right) \right] \tag{12}$$

3. Design of nonlinear PID controllers for inverted pendulum system

The PID controllers are widely used to stabilize the unstable systems. The PID controllers for unstable systems are having the capability of self-tuning by either conventional or intelligent technique to attain the stability in the neighborhood of operating envelope. However, the conventional tuning requires tedious numerical computation to achieve best possible controller gains, during the change of parameter variations. Therefore, the PID controller may perform poorly if the unstable system operates at different operating points, which need that the controller to be repeatedly re-tuned. The system may lose the stability at that operating point if the re-tuning made through complex and slow tuning methods. In these situations, an alternate approach to eliminate the problem is to use

ACCEPTED MANUSCRIPT

nonlinear PID with intelligent tuning techniques such as bio-inspired meta-heuristics, which are suitable to achieve faster and precise computation with best controller gains. However, when the system structure is complex with parameter variations, the linear PID do not exhibit good efficacy compared to that of the nonlinear PID. The nonlinear PID controller changes its parameters quickly over time according to system error and can provide an additional degree of freedom to achieve improved performance. Currently, different variants of GA and PSO algorithms tuned PID controllers have shown renewed interest among the researchers to automate the tuning process in the design of the controller.

3.1 NL-PID controller design

Mostly the PID controllers are employed to control unstable systems. While linear PID controllers are often sufficient for stabilizing unstable systems during small linear operating envelope, however for more high precision control, the competency and robustness of linear PID remained insufficient. The major failure of linear PID controller comes due to the constraints in the mathematical model i.e. contradiction between increased gain and the faster response of the system. If a suitable law is formulated for the system, which removes this limitation leading to a better control action for the system (J.Q.Han, 1994).A general nonlinear PID module is introduced (Chen.W.-H., D. J. Balance, P. J.Gawthrop, J. J. Gribble, 1999; Su, Y.X, Dong Sun, 2005)as in equation (13)

$$f(e,\alpha,\delta) = \begin{cases} |e|^{\alpha} sign(e) & |\gamma| \geq \delta \\ \frac{e}{\delta^{1-\alpha}} & |\gamma| < \delta \end{cases} \quad (13)$$

Where $sign(e) = \begin{cases} 1, & e \geq 0 \\ -1, & e < 0 \end{cases}$

e is the error signal, δ describes the linear range of the function f, here function f can accommodate a greater range of nonlinear characteristics which is determined by α. Hence now control signal being generated by the NL-PID is given in equation (14)

$$u(t) = K_p e f_p(e_p,\alpha_p,\delta_p) + K_i f_i(e_i,\alpha_i,\delta_i) \int e dt + K_d f_d(e_d,\alpha_d,\delta_d) \frac{de}{dt} \quad (14)$$

ACCEPTED MANUSCRIPT

The function $f(e,\alpha,\delta)$ denotes the rate of error feedback, to compensate the nonlinearity of the systems considered the value of α_p is taken in the range of $\alpha_p \epsilon$ [0,1] since it needs lower gain when the error is high and vice versa. The integral saturation problem of the integral term can be rectified by using α_i in the range of $\alpha_i \epsilon[-1,0]$. The value of differential term is chosen as $\alpha_d > 1$ so that when the steady state is reached the effect of the differential term is reduced. The systematic structure design of NL-PID is shown in Fig.2

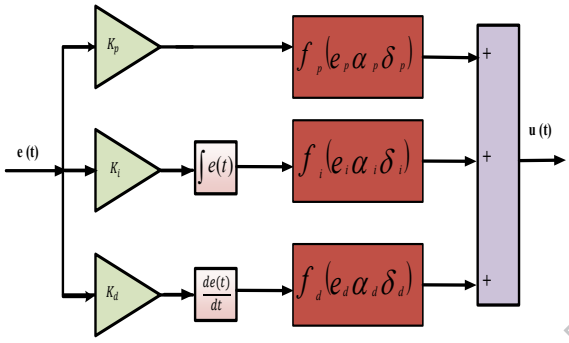


Fig.2 Structure design of NL-PID

4. Metaheuristics tuned NL-PID controllers for inverted pendulum system

Bio-inspired metaheuristic algorithms are dissimilar as compared to the traditional algorithms regarding a multi solid solution while later have a single point solution; metaheuristic algorithms can all together evaluate some points in the search space with lesser convergence time and converge to global optima rather than local optima. Further, the metaheuristic algorithms reproduce offspring to progress the search space in addition to solution. Many metaheuristic algorithms such as Genetic Algorithm, Ant colony algorithm, particle swarm optimizer algorithm, firefly algorithm, and artificial bee colony algorithm(Xin-She Yang, 2010) and their advanced variants have been developed in recent past. These algorithms differ from each other by variables to be tuned, offspring generation, and replacement mechanism. Out of all these algorithms, the most familiar is a genetic algorithm and particle swarm optimization algorithm for PID tuning of several systems. The variants of Multi-Objective Genetic Algorithm (MOGA) and Adaptive Particle Swarm Optimization Algorithm (APSO) have been considered here for the tuning of NL-PID controllers for inverted pendulum system. The advantage of these algorithms is a shortfall of system matrix eigenvalues.

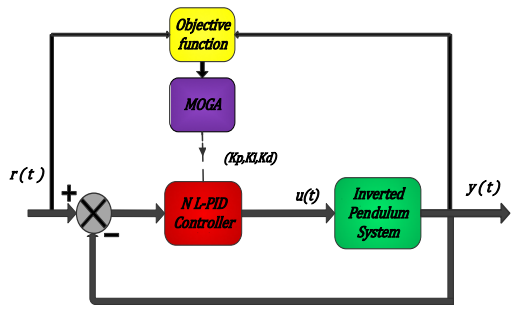
ACCEPTED MANUSCRIPT

4.1 MOGA tuned NL-PID Controllers

The multi-objective genetic algorithms were one type of combinatory optimized search algorithm, which require discretized search space to solve real decision variables of system optimization problems for both single objective and multi-objective functions. The MOGA selects individuals based on random selection from the population and consists of genetic transformations such as crossover, mutation to create a new set population. Hence initially there is no clue, regarding the best answer for an optimization problem. Over the successive generations, the best among the lot are selecting, operators are worked upon on them and evaluated again until they evolve to the global optimal or suboptimal solution. The objective of using Multi-Objective Genetic Algorithm (Sudarshan K Valluru, 2011) here is to determine the optimal value of NL-PID parameters (K_p,K_i and K_d) such that the objective functions $J(x)$ is minimized. The overall result of the finding of the optimal parameters would lead to fine tuning of the PID controller and simultaneously improvement in the transient as well as steady state response of the system under consideration. The objective function $J(x)$becomes an instrument to evaluate the performance of NL-PID controller with the determined value of gain parameters, resulting in an optimized controller with the best individual is given in equation (15)

$$J(x) = \ (r(n) - y(n))^2 + \ (max[r] \ - max[y])^2 \qquad (15)$$

Where $r(n)$ is n[th] input of the system and $y(n)$ is the n[th] output of the system. A general block diagram of MOGA tuned NL-PID is shown in Fig.3 which uses the torque as reference signal to the inverted pendulum, the error between actual output and the reference signal along with proportional(K_p), integral(K_i), derivative (K_d) gains generate the control signal.



ACCEPTED MANUSCRIPT

Fig.3 MOGA tuned NL-PID controller for inverted pendulum

The pseudo code of MOGA employing to tune NL-PID controllers for inverted pendulum system can be summarized as follows:

1. Set the range of PID controller parameters as objective function $J(x)$, $x = (K_p, K_i, K_d)^T$.
2. Encode the solution into chromosomes as binary strings.
3. Define the fitness $F \propto J(x)$ for maximization.
4. Generate the initial population of size N for the NL-PID controller parameters.
5. Initial probabilities of crossover (p_c) and mutation(p_m).
6. Evaluate the initial population.
7. **While** (t< Max. number of generations)

   Generate new solution by crossover and mutation

   If $p_c >$ rand, Crossover; **end if**

   If $p_m >$ rand, Mutation; **end if**

   The new solutions are accepted if fitness increases.

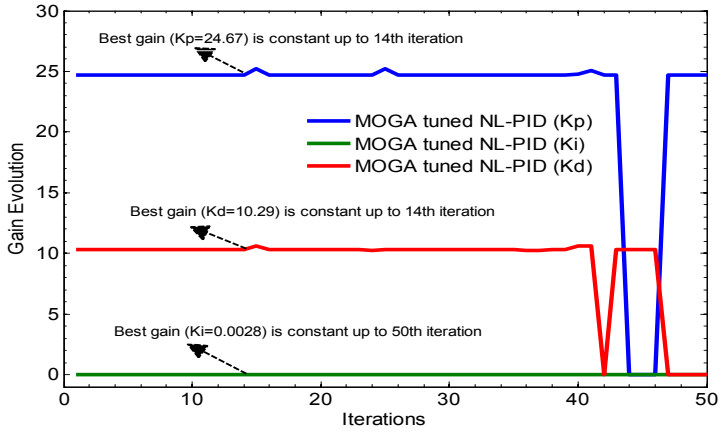
   Select the current best for new generation

   end while.
8. Decode the binary strings and visualize the new values of K_p, K_i, K_d.

The size of the population, dimension, mutation rate, selection rate, maximum iterations used in the MOGA process to find best optimizing tuning parameters of NL-PID controllers for inverted pendulum system taken as 10,3,0.2,0.3,100 respectively. The estimated NL-PID controller gain parameters by multi-objective genetic algorithm, which is tabulated in Table.2 and shown in Fig.4

Table 2. MOGA tuned NL- PID controller best gain parameters

Controller	Time Elapsed for MOGA (sec)	K_p	K_i	K_d
NL-PID	295.7	24.67	0.0028	10.29



ACCEPTED MANUSCRIPT

Fig.4 Best gain evolution by MOGA tuned NL-PID controlled inverted pendulum

The control signal $u(t)$ for MOGA tuned NL-PID as shown in Fig.5. The pendulum angle response of inverted pendulum system with MOGA tuned NL-PID as shown in Fig.6.

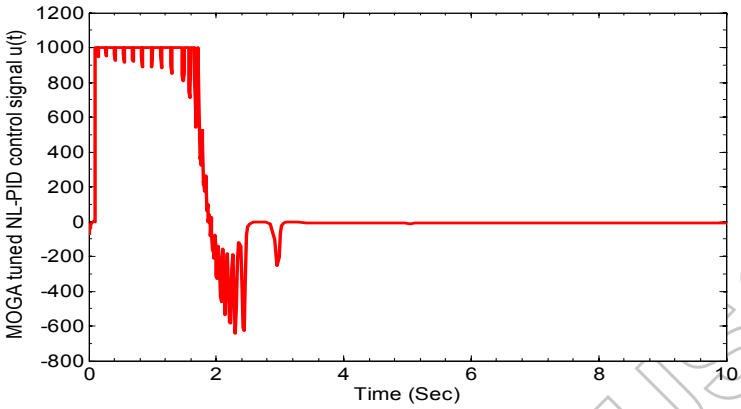


Fig. 5 Control signal of MOGA tuned NL-PID controller

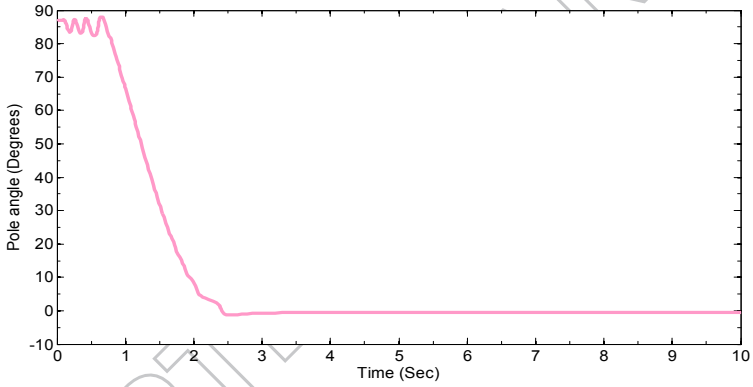


Fig.6 Pendulum angle response of inverted pendulum system with MOGA tuned NL-PID controller

4.2 APSO tuned NL-PID Controllers

APSO algorithm is effective for optimizing a broad range of objective functions, inspired from the swarming phenomena of fish schooling and bird flocking. The multiple agents known as particles, swarm around the search space starting from an initial guess. APSO has some similarities with MOGA, but APSO is much simpler because it uses real number randomness and global communications instead of mutation/crossover. The APSO algorithm starts with the initialization of group of arbitrary particles in the search space, and each particle represents a potential solution for the optimization of an objective function.

ACCEPTED MANUSCRIPT

During each iteration, all the particles in the search space discover a feasible solution. After, the particle updates its position according to the present velocity vector which contains its previous value and decides to take into account the past local and global best solutions. By keeping the best solutions, the particle moves towards local best solution attained by its fellow particles but also the global best. Hence if a particle has discovered a new best solution, then all other particles would try to move toward it. The four important terms in APSO (Zhan, Zhang, Li, & Chung, 2009)(Li & Zhi-hui Zhan, YingLin, 2015)(Chen, Hu, Wang, Gao, & Li, 2016) for the particles in the swarm are p_i (position), v_i (current velocity), Pb_i (local best position), Gb_j (global best position).each particle is updated according to the above four features in each iteration, assuming an cost function J to be minimized. The innovative velocity of a particle is determined by the equation (16)

$$v_{i,j}(n+1) = C \times \left\{ w v_{i,j}(n) + c_1 r_1 \left[Pb_{i,j}(n) - p_{i,j}(n) \right] + c_2 r_2 \left[Gb_j(\mathrm{n}) - p_{i,j}(n) \right] \right\} \quad (16)$$

Where n is the number of iterations, w is inertia weight, c_1 and c_2 are the acceleration co-efficients (called cognitive and social component respectively) and $r_1$and r_2 are two quasi uniform random numbers between (0, 1) . $v_{i,j}$ is the velocity of the j^{th} dimension of the i^{th} particle. The new position of the particle is updated by using equation (17)

$$p_{i,j}(n+1) = p_{i,j}(n) + v_{i,j}(n+1) \quad (17)$$

The local best of every particle is updated by the equation (18)

$$Pb(n+1) = \begin{cases} Pb(n), & J(p(n+1)) \geq J(Pb(n)) \\ Pb(n+1), & otherwise \end{cases} \quad (18)$$

Here the population size is denoted by s; then the optimized global best can be found by the equation (19)

$$Gb(n+1) = \min_{Pb} J\left(Pb_i(n+1)\right), \ 1 \leq i \leq s \quad (19)$$

Tuning of NL-PID can be done efficiently by using particle swarm optimizer to find the values of optimal gain parameters (K_p,K_i and K_d) which will minimize the objective function (J). The block diagram of APSO tuned NL-PID for cart pendulum system is shown in Fig.7 which uses the step reference signal to the

ACCEPTED MANUSCRIPT

inverted pendulum system, the error between actual output and the reference signal along with proportional(K_p), integral (K_i), derative (K_d) gains generated the control signal $u(t)$.

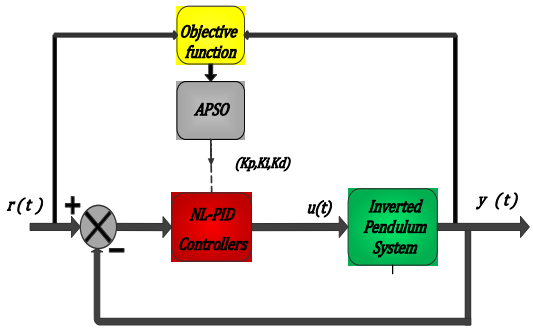


Fig.7 APSO tuned NL-PID controller for inverted pendulum system

The APSO pseudo code to tune the L-PID and NL-PID controllers for inverted pendulum system is summarized as follows:

1. Set the range of PID controller parameters $J(x), x = (K_p, K_i, K_d)^T$
2. Initialize the locations of p_i and velocity v_i for swarm size n, acceleration factors c_1 and c_2 to the controller parameters K_p, K_i, K_d each representing dimensions of particle position vector in real parameter space.
3. **for** each particle, ***i* do**, by testing each particle for fitness by evaluating the objective function $J(x)$.
4. Set the initial local best (Pb_i), and global best (Gb_j)
5. **While** stopping criteria is not reached **do** **for** every particle ***i* do**
   Generate innovative velocity $v_{i,j}(n+1)$ according to equation (16)
   Generate new positions $p_{i,j}(n+1)$ according to equation (17)
   Evaluate objective functions at new positions
   Renew the current best of each particle
   end for
   Find the current global best
   end while
6. Output the final results of optimized global best

The size of the swarm, dimension, maximum iterations, cognitive factor, social factor, constriction factor used in the adaptive particle swarm optimizer to find the best-optimized parameters for NL-PID controller for the inverted pendulum system are 10, 3, 50, 2.05, 2.05, and 2 respectively. The best gain parameters of APSO tuned NL-PID controllers are given in Table.3 and gained evolution is shown in Fig. 8

Table 3 APSO tuned NL-PID controller best gain parameters

ACCEPTED MANUSCRIPT

Controller	Time Elapsed for APSO Algorithm(Sec)	K_p	K_i	K_d
NL-PID	46.85	12.551	1.078	7.248

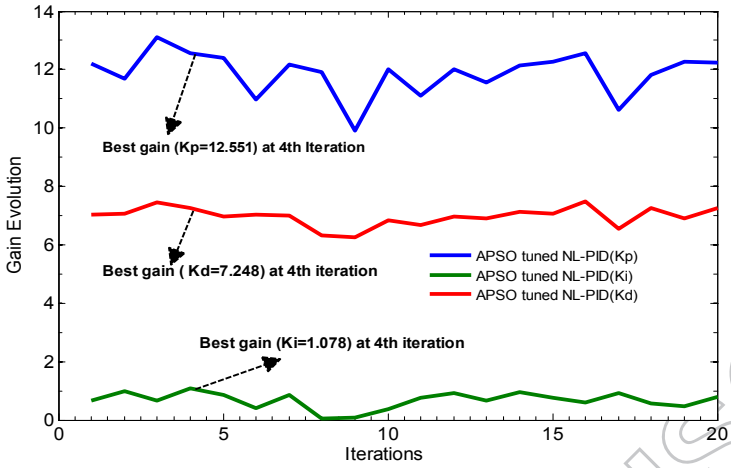


Fig.8 Best gain evolution by APSO tuned NL-PID controlled inverted pendulum system

The control signal $u(t)$ for APSO tuned NL-PID as shown in Fig. 9. The pendulum angle response of inverted pendulum system with APSO tuned NL-PID as shown in Fig.10.

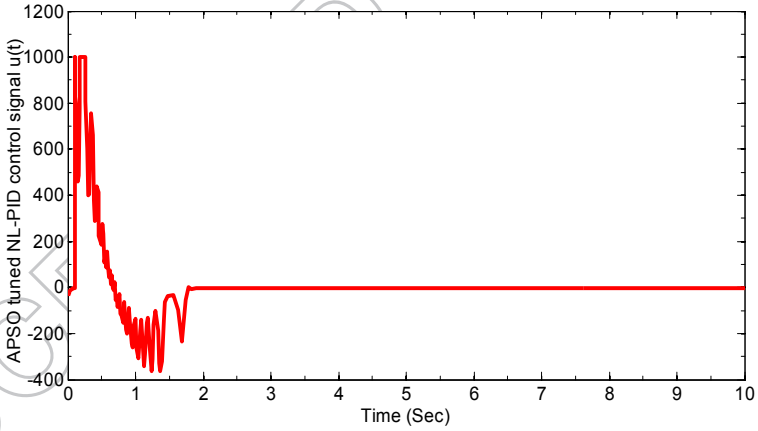


Fig.9 Control signal of APSO tuned NL-PID controller

ACCEPTED MANUSCRIPT

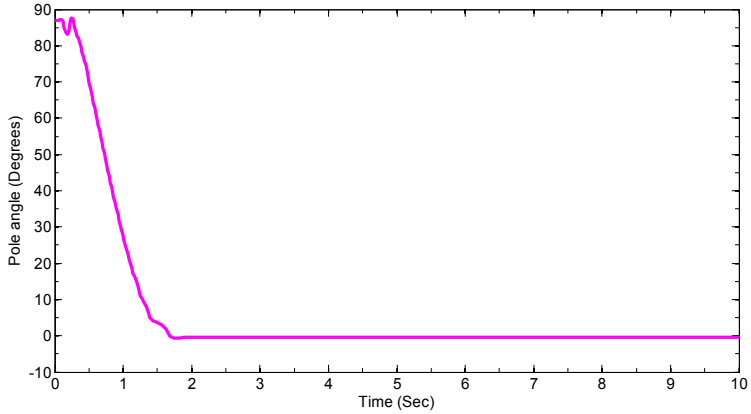


Fig.10 Response of inverted pendulum system with APSO tuned NL-PID controller

4.3 Performance comparison of MOGA and APSO tuned NL-PID controller for inverted pendulum system

A comparison performance of the IPS for pendulum angle response using multi-objective genetic algorithm and adaptive particle swarm optimization algorithm tuned NL-PID controllers are carried out through simulation studies in MATLAB and summarized in Table.4.

Table.4 Performance indices of L-PID and NL-PID

Performance Index	MOGA tuned NL-PID controller	APSO tuned NL-PID controller
Chattering time (Sec)	0.752	0.328
Settling time(Sec)	2.5	1.7

5. Conclusion

This investigation presents comprehensive tuning strategy and performance evolution of NL-PID controller in conjunction with MOGA and APSO to stabilize the pendulum angle of IPS. APSO based tuning of NL-PID gives lesser settling time, and lesser chattering in response than MOGA tuned NL-PID controller. In addition to this, it is observed that the utilization of an NL-PID controller improves the dynamic performance of IPS. The APSO tuned NL-PID controller provides the optimum performance of the IPS and stabilize the system in operating state quickly with less chattering in the initial response.In the present scope of work, efforts are made to validate the performance and stabilization of IPS with MOGA and APSO based tuning of NL-PID controllers. However, the authors are motivated to implement these control algorithms on real-time prototype IPS in the laboratory.

ACCEPTED MANUSCRIPT

References

Akihiro Oi, Nakazawa, C., & Matsui, T. (2008). Development of PSO-based PID Tuning Method. In *International Conference on Control, Automation Systems* (Vol. 2, pp. 1917–1920). Seoul, Korea.

Aouaouda, S., Chadli, M., & Karimi, H. (2014). Robust static output-feedback controller design. *IET Control Theory and Applications*, *8*(9), 728–737. http://doi.org/10.1049/iet-cta.2013.0709

Aouaouda, S., Chadli, M., Khadir, M. T., & Bouarar, T. (2012). Robust fault tolerant tracking controller design for unknown inputs T – S models with unmeasurable premise variables. *Journal of Process Control*, *22*(5), 861–872. http://doi.org/10.1016/j.jprocont.2012.02.016

Aouaouda, S., Chadli, M., Shi, P., & Karimi, H. R. (2015). Discrete-time H_ /H∞ sensor fault detection observer design for nonlinear systems with parameter uncertainty. *International Journal Of Robust And Nonlinear Control*, *25*, 339–361. http://doi.org/10.1002/rnc.3089.2014

Aracil, J. Gordillo, F. (2004). The inverted pendulum: a benchmark in nonlinear control. In *A Proceedings of World Automation Congress*. (Vol. 16, pp. 468–482).

Åström, K. J., & Furuta, K. (2000). Swinging up a pendulum by energy control. *Automatica*, *36*(2), 287–295.

Astrom, K. J., & Kumar, P. R. (2014). Control: A perspective. *Automatica*, *50*(1), 3–43. http://doi.org/10.1016/j.automatica.2013.10.012

Banks, A., Vincent, J., & Anyakoha, C. (2007). A review of particle swarm optimization. Part I: background and development. *Natural Computing*, *6*(4), 467–484.

Banks, A., Vincent, J., & Anyakoha, C. (2008). A review of particle swarm optimization. Part II: Background and development. *Natural Computing*, *7*(1), 109–124.

Bououden, S., Chadli, M., & Karimi, H. R. (2015). An ant colony optimization based fuzzy predictive control approach for nonlinear processes. *Information Sciences*, *299*, 143–158. http://doi.org/10.1016/j.ins.2014.11.050

Chadli.M., P. B. (2013). *Multiple Models Approach in Automation : Takagi-Sugeno Fuzzy Systems*. (Bernard Dubuisson, Ed.) (1st Edition). New Jersey: John Wiley & Sons.

ACCEPTED MANUSCRIPT

Chadli, M., Aouaouda, S., Karimi, H. R., & Shi, P. (2013). Robust fault tolerant tracking controller design for a VTOL aircraft. *Journal of the Franklin Institute*, *350*(9), 2627–2645. http://doi.org/10.1016/j.jfranklin.2012.09.010

Chatterjee, D., Patra, A., & Joglekar, H. K. (2002). Swing-up and stabilization of a cart – pendulum system under restricted cart track length. *Systems &Control Letters*, *47*, 355–364.

Chen.W.-H., D. J. Balance, P. J.Gawthrop, J. J. Gribble, J. O. R. (1999). Nonlinear PID predictive controller. *IEE Control Theory Applications*, *146*(6), 603–611.

Chen, C., Hu, L., Wang, L., Gao, S., & Li, C. (2016).Adaptive Particle Swarm Algorithm for ParametersTuning of Fractional Order PID Controller. *Telkomnika*,*14*(2),478-488.

Chinichian, M., Kashani, R. (1989). State space controller design for a spatial inverted cart/pendulum system. In *IEEE Proceedings of the 32nd Midwest Symposium on Circuits and Systems,* (pp. 633–636).

Daley, S., & Liu, G. (1999). Optimal PID tuning using direct search algorithms. *Computing & Control Engineering Journal*, *10*(2), 51–56.

Deris, S. Omatu, S. Kitagawa, K. (1995). Stabilization of inverted pendulum by the genetic algorithm. In *Proceedings of IEEE International Conference on Intelligent Systems for the 21st Century Man and Cybernetics.* (pp. 4372–4377).

Durand, S., Guerrero-castellanos, J. F., & Marchand, N. (2013). Event-Based Control of the Inverted Pendulum : Swing up and Stabilization. *Control Engineering and Applied Informatics*, *15*(3), 96–104.

E.J.Davison. (1990). *Benchmark problems for the control system design: report of the IFAC Theory Committee*.

Geva, S., & Sitte, J. (1993). Cartpole experiment benchmark for trainable controllers. *IEEE Control Systems Magazine*, *13*, 40–51.

Guida, D., Nilvetti, F., & Pappalardo, C. M. (2010). Dry friction of bearings on dynamics and control of an inverted pendulum. *Journal of Achievements in Materials and Manufacturing Engineering*, *38*(1), 80–94.

J.Q.Han. (1994). Nonlinear PID Controller. *Acta Automatica Sincia*, *20*(4), 487–490.

Jonathan P. How. (2015). Benchmarks. *IEEE Control Systems Magazine*, 6–7.

Khalil, H. (2002). *Nonlinear Systems* (3rd ed.). Upper Saddle River, NewJersey:

ACCEPTED MANUSCRIPT

Prentice-Hall Inc.

Korkmaz.Mehmet, Ömer Aydoğdu, H. D. (2012). Design and Performance Comparison of Variable Parameter Nonlinear PID Controller and Genetic Algorithm Based PID Controller. In *International Symposium on Innovations in Intelligent Systems and Applications* (pp. 1–5). Trabzon, Turkey.

Lee, J., Mukherjee, R., & Khalil, H. K. (2015). Output feedback stabilization of inverted pendulum on a cart in the presence of uncertainties. *Automatica*, *54*, 146–157.

Li, Y., & Zhi-hui Zhan, YingLin, J. (2015). Comparisons Study of APSO OLPSO and CLPSO on CEC2005 and CEC2014 Test Suits. In *IEEE Congress on Evolutionary Computation* (pp. 3179–3185).

Medrano-Cersa, G. A. (1999). Robust computer control of an inverted pendulum. *IEEE Control Systems*, *19*(3), 58–67.

Menhas, M. I., Wang, L., Fei, M., & Pan, H. (2012). Comparative performance analysis of various binary coded PSO algorithms in multivariable PID controller design. *Expert Systems with Applications*, *39*(4), 4390–4401.

Murata, N. Matsuno, H. Moran, A. Hayase, M. (1995). Control of inverted pendulum-wheeled cart systems using nonlinear observers. In *Proceedings of the 34th SICE Annual Conference* (Vol. 5, pp. 1133–1136).

Omatu, S., & Yoshioka, M. (1998). Stability of inverted pendulum by neuro-PID control with genetic algorithm. In *Proceedings of IEEE World Congress on Neural Networks and Computational Intelligence.* (Vol. 3, pp. 2142–2145).

Patnaik, A., & Behera, L. (2008). Evolutionary Multiobjective Optimization Based Control Strategies For An Inverted Pendulum On A Cart. In *Proceedings of IEEE World Congress on Computational Intelligence* (pp. 3141–3147).

Prasad, L.B., Tyagi, B., Gupta, H. O. (2011). Optimal Control of Nonlinear Inverted Pendulum Dynamical System with Disturbance Input using. In *Proceedings of IEEE International Conference on Control System, Computing and Engineering* (pp. 540–545).

Sastry.S. (1999). *Nonlinear systems analysis stability and control*. NewYork, USA: Springer-Verlag Inc.

Shen, J.-C. (2002). New tuning method for PID controller. *ISA Transactions*, *41*(2002), 473–484.

Sofiane Bououden, Mohammed Chadli, Fouad Allouani, S. F. (2013). A New

ACCEPTED MANUSCRIPT

Approach For Fuzzy Predictive Adaptive Controller Design Using Particle Swarm Optimization Algorithm. *International Journal of Innovative Computing, Information, and Control*, *9*(9), 3741–3758.

Su,Y.X, Dong Sun, B. Y. D. (2005). Design of an enhanced nonlinear PID controller. *Mechatronics*, *15*(2005), 1005–1024.

Sudarshan.K.Valluru, MadhusudanSingh, & Bhushan. (2015). Comparative Analysis of PID, NARMA L-2 and PSO Tuned PID Controllers for Nonlinear Dynamical System. *J. Automation & Systems Engineering*, *9*(2), 94–108.

Sudarshan K Valluru. (2011). *Introduction to Neural Networks, Fuzzy Logic &Genetic Algorithms* (2nd Reprint). Mumbai: Jaico Publishing House.

Teixeira, M. C. M., & Zak, S. H. (1999). Stabilizing Controller Design for Uncertain Nonlinear Systems Using Fuzzy Models. *IEEE Transactions on Fuzzy Systems*, *7*(2), 133–142.

Vidyasagar.M. (1993). *Nonlinear system analysis*. Englewood Cliffs N.J., U.S.A.: Prentice-Hall Inc.

Xin-She Yang. (2010). *Engineering Optimization* (1st edition). New Jersey: John Wiley & Sons, Inc.

Zhan, Z. H., Zhang, J., Li, Y., & Chung, H. S. H. (2009). Adaptive Particle Swarm Optimization. *IEEE Transactions on Systems, Man Cybernetics*, *39*(6), 1362–1381. http://doi.org/10.1109/TSMCB.2009.2015956

Zhao, J., Li, T., & Qian, J. (2005). Application of Particle Swarm Optimization Algorithm on Robust PID Controller Tuning. In K. C. and Y. S. O. L.wang (Ed.), *Advances in Natural Computation* (pp. 948–957). Springer-Verlag Berlin Heidelberg.

ACCEPTED MANUSCRIPT

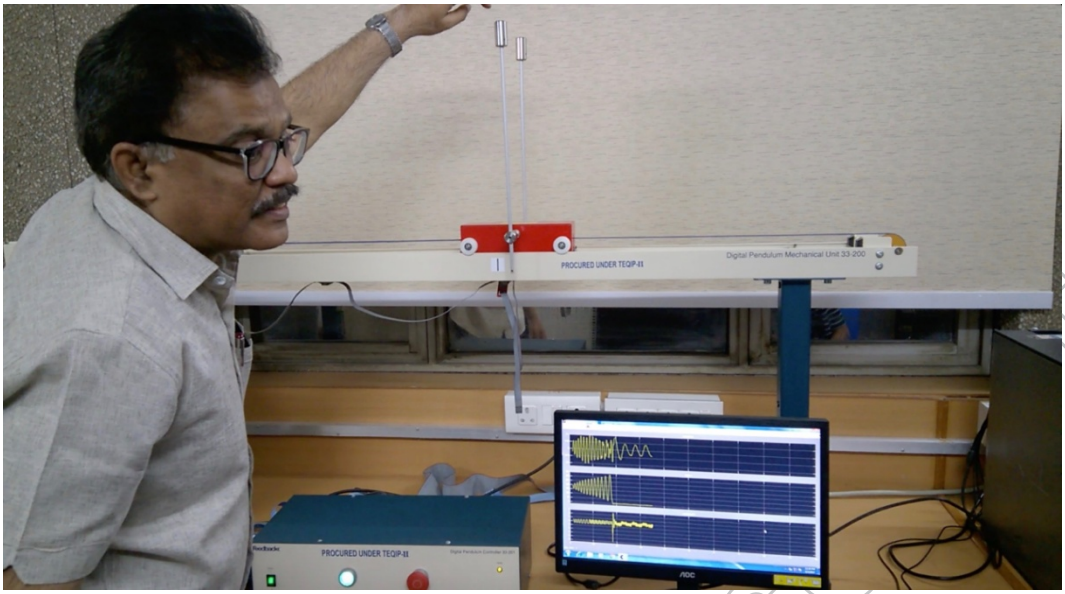
Sudarshan K.Valluru has been an Associate Professor of Electrical Engineering since 2009 at Delhi Technological University, New Delhi, India. His area of interests is Meta-Algorithms Applications to Control of Dynamical Systems, Intelligent, and Soft Computing Techniques. He is Fellow Member of Institution of Electronics and Telecommunication Engineers, Life Member of Institution of Engineers (I), and Life Member of Indian Society for Technical Education and also a member of IEEE(USA).

Madhusudan Singh has been working as Professor of Electrical Engineering since 2007 at Delhi Technological University, New Delhi, India. His research interests are in the area of modeling and analysis of Electrical Machines and Dynamical Systems, Voltage control aspects of Self-Excited Asynchronous Generator, Power Electronics, and Drives. He is a Fellow Member of Institution of Electronics and Telecommunication Engineers, Fellow Member of Institution of Engineers (I), and Life Member of Indian Society for Technical Education and also a Senior member of IEEE(USA).

ACCEPTED MANUSCRIPT

It is interesting to watch the rocket launching at the space center, many of us watching on television the launching process of the rocket/satellites and their establishment in the orbit of the earth. Sometimes our mind stuck on certain aspects while watching the scenario of rocket flight. The aspects include what force is controlling the launching rocket which is carrying the satellite, how the rocket of such huge weight is projected upright, how the satellite is placed into the space orbit. All these aspects for a budding control engineer are hard to understand initially, but once we start embracing the results obtained from control engineering, to learn and design dynamic controllers are becoming inevitable nowadays. The nonlinear inverted pendulum system(IPS) is a challenging benchmark control problem to design and test such new dynamic controllers. It resembles the many real-world dynamical systems, such as self-erecting rocket, missile launchers, Segway robotic platforms.

ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT

Theoretical modeling to study the impact of different oxidizers (etchants) on the plasma-assisted catalytic carbon nanofiber growth

Ravi Gupta and Suresh C. Sharma

Citation: Physics of Plasmas **24**, 073504 (2017); doi: 10.1063/1.4990556
View online: http://dx.doi.org/10.1063/1.4990556
View Table of Contents: http://aip.scitation.org/toc/php/24/7
Published by the American Institute of Physics



Theoretical modeling to study the impact of different oxidizers (etchants) on the plasma-assisted catalytic carbon nanofiber growth

Ravi Gupta and Suresh C. Sharma
Department of Applied Physics, Delhi Technological University (DTU), Shahbad Daulatpur, Bawana Road, Delhi 110042, India

(Received 9 May 2017; accepted 14 June 2017; published online 29 June 2017)

An analytical model based on the various surface deposition processes and plasma sheath kinetics of the plasma species (electrons, positively charged ions, radicals, and neutrals) has been developed to investigate the effects of different plasmas (different etchants) on the catalyzed plasma aided growth of carbon nanofibers (CNFs). In particular, the model accounts the poisoning of the catalyst nanoparticle, i.e., the formation of the amorphous carbon layer on the catalyst active surface due to the continuous dissociation of incoming hydrocarbon species from the plasma. It is observed that oxidizers (H_2O and O_2) in the typical hydrocarbon/hydrogen (C_2H_2 + H_2) plasma act as the dominant etchants and remove the amorphous carbon layer from the catalyst surface and, thus, preserve and enhance the catalyst activity. However, the growth rate of CNFs is much higher when O_2 is added as an etchant in the reactive plasma as compared to H_2O. This is due to the dual role played by the oxygen, i.e., (i) removal of amorphous carbon from the catalyst active surface, (ii) removal of hydrogen radicals that interact with the carbon species generated on the catalyst surface and suppress their diffusion through the catalyst nanoparticles. The CNF grows much longer in the presence of O_2, therefore, etching of CNF tip and deformation of catalyst nanoparticle is the maximum, and hence, the CNF tip diameter is least. Moreover, in the present investigation, we also found that the relative concentrations of H_2O or O_2 species in the reactive plasma have significant effects on the CNF growth. Our theoretical results are in good agreement with the experimental observations. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4990556]

I. INTRODUCTION

Rolled graphene sheets in the form of one-dimensional carbon nanostructures, i.e., carbon nanofibers (CNFs) and carbon nanotubes (CNTs), have attracted great interest for applications in field emission devices inferable from their exceptional electrical properties.[1–3] Catalyst-assisted plasma enhanced chemical vapor deposition (PECVD) process is thought to be the most promising growth technique to grow vertically aligned CNFs/CNTs.[4–6] In PECVD growth process, a thin layer of metal catalyst is deposited over the substrate surface, and on plasma processing in the presence of etching gas, granulated into the metal catalyst nanoparticles.[7] Each catalyst nanoparticle serves as the nucleation seed for the CNFs/CNTs growth by decomposing the carbon precursors over its active surface coming from the reactive plasma that comprises a carbon source gas and an etching gas.[8,9] Catalyst activity during the growth, i.e., catalyst lifetime influences the CNFs/CNTs growth rate and catalyst lifetime decreases due to the formation of the amorphous carbon layer on the catalyst surface. Removal of the amorphous carbon is one of the major concerns to enhance the CNFs/CNTs growth rate.

The etching gas etches the amorphous carbon layer formed on the active surface and retains the activity of the catalyst. Usually, hydrogen and nitrogen are utilized as the etching gases.[10–12] In the catalytic chemical vapor deposition growth process, it is observed that introducing a small amount of oxidizer such as H_2O and O_2 can colossally expand the CNFs/CNTs growth rates. The presence of water vapors in the growth atmosphere amplifies the catalyst lifetime and preserves the catalyst action because of which CNFs/CNTs growth time and their length increases.[13–15] Besides, water vapors were also found to change their structure from multiwall to more slender nanotubes.[16] A few reviews have shown the possibility to grow ultra long CNFs/CNTs via plasma enhanced CVD process within sight of oxidizers.[13,17–21] It is accounted for that high concentration of hydrogen radicals in PECVD process act as the sink for carbon species, and thus, suppresses the number of carbon species accessible for precipitation in the form of graphene layers around the catalyst nanoparticle. Oxygen in the PECVD process keeps up the deterministic control over carbon to hydrogen ratio by scavenging these hydrogen radicals and also removes the amorphous carbon layer formed on the catalyst nanoparticle and in this manner supports the considerably higher CNF/CNT growth.[20] Kim *et al.*[21] investigated the role of water and oxygen on the growth of vertically aligned CNTs and found that addition of oxygen and water in the hydrocarbon/hydrogen plasma increased the CNTs growth rate up to thrice and twice, respectively, than that of CNT grown in typical hydrocarbon/hydrogen plasma. For the future industrial fabrication, it is vital to see how the presence of oxidizers and their relative level in the plasma influence the CNFs/CNTs growth.

In this study, we attempt to present the kinetic model to compare the effects of utilizing oxygen and water as the

etching gases, on the plasma assisted catalyzed growth of vertically aligned CNFs accounting the plasma sheath effects and various surface processes including the formation of the amorphous carbon layer on the catalyst surface and poisoning of the catalyst nanoparticle. The first order kinetic differential equations given in Sec. II have been simultaneously assessed, and their results in the view of time evaluation analysis are discussed in Sec. III. Finally, the conclusion is given in Sec. IV.

II. MODEL

In this work, we have adopted the following scenario to enlighten the role of different plasmas on the growth of carbon nanofibers (CNFs). (i) We have accounted for the CNF growth in hydrocarbon + hydrogen (C_2H_2+H_2) plasma, hydrocarbon + hydrogen + water (C_2H_2+H_2+H_2O) plasma, and hydrocarbon + hydrogen + oxygen (C_2H_2+H_2+O_2) plasma denoted as type I, type II, and type III plasmas, respectively. These plasma compositions contain electrons, singly ionized positively charged ions, radicals, and neutrals that are listed in Table I. (ii) The nucleation of CNFs is considered on the nickel metal catalyst nanoparticles placed over the negatively biased substrate surface. (iii) The temperature of the metal catalyst nanoparticle and the substrate is assumed to be the same. (iv) The sheath electric field due to space charge separation is assumed along the direction perpendicular to the substrate surface. (v) The Maxwellian distribution of ions and electron is assumed. (vi) It is assumed that the ion number density is greater than the electron number density in the plasma sheath region. (vii) To investigate the structure of plasma sheath and the dynamics of plasma particles (electrons, ions, radicals, and neutrals), the sheath kinetics given in Refs. 22–24 are considered in the present paper. Using the assumptions mentioned earlier, the main equations used in the model are as follows:

A. Charging rate equation of CNFs

$$\partial_t(q) = \sum_X (I_{\text{tip-}iX} + I_{\text{cur-}iX}) + \sum_Y (I_{\text{tip-}iY} + I_{\text{cur-}iY}) + \sum_{Z=1}^{2} (I_{\text{tip-}iZ} + I_{\text{cur-}iZ}) - \gamma_e (I_{\text{tip-e}} + I_{\text{cur-e}}), \quad (1)$$

where q is the charge developed on CNFs surface, $I_{\text{tip_}ij}$ and $I_{\text{cur-}ij}$ are the ion collection currents at the tip and the curved surface of CNFs, respectively,[24] j refers to X, Y, and Z type of positively charged ions, $I_{\text{tip_e}}$ and $I_{\text{cur_e}}$ are the electron collection currents at the tip and the curved surface of CNFs,[24] respectively, and γ_e is the electron sticking coefficient at the CNF surface. The positively charged ions of type X and Y are taken into account when plasma of type I is considered. However, additional species of type Z are taken into account when plasma of type II and III are considered. From Table I, it can be seen that species of type Z are different for plasma of type II and III, respectively. In Eq. (1), the first three terms denote the charge developed on the CNFs surface due to the accumulation of positively charged ions, and the last term indicates the reduction in charge on account of the accumulation of electrons from the plasma at the CNFs surface. It is assumed that all the ions, incident on the CNF surface, transfer their charge and get neutralized.

TABLE I. Plasma species considered in the present model. Species of type A and B are taken into account when CNF growth is investigated in type I plasma. However, additional species of type C are taken into account when CNF growth is examined in type II and type III plasmas.

			C	
Type	A	B	II plasma	III plasma
Ions	$C_2H_2^+$, $C_4H_2^+$, C_2H^+, CH^+	H_2^+, H^+	H_2O^+, OH^+, O^+	OH^+, O^+
Neutrals	C_2H_2, C_4H_2	H_2	H_2O	O_2
Radicals	CH, C_2H	H	OH, O	OH, O

B. Kinetic equation of electron density

$$\partial_t(n_e) = \sum_{X=1} \beta_X n_X + \sum_Y \beta_Y n_Y + \sum_Z \beta_Z n_Z - \sum_{X=1} \alpha_X n_e n_{iX} - \sum_{Y=1} \alpha_Y n_e n_{iY} - \sum_{Z=1} \alpha_Z n_e n_{iZ} - \gamma_e \frac{n_{cnf}}{\lambda_d} (I_{\text{tip-e}} + I_{\text{cur-e}}), \quad (2)$$

where β_j is the coefficient of ionization of neutral atoms,[25] α_j is the coefficient of recombination of electrons and positively charged ions,[26] n_e is the electron number density, n_{ij} is the ion number density (i.e., number of ions per unit volume), n_j is the neutral atom density, n_{cnf} is the number density (i.e., number of CNFs per unit area) of CNFs, and $\lambda_d (= 3 \times 10^{-3}$ cm) is the plasma sheath width.

Equation (2) indicates the growth of electron number density in the reactive plasma. The first three terms of Eq. (2) denotes the gain in electron number density per unit time due to ionization of neutral atoms. The last four terms of Eq. (2) indicates the electron number density decay rate due to electron-ion recombination and electron collection current at the CNFs surface, respectively.

C. Kinetic equations of positively charged ion

$$\partial_t(n_{iX}) = \beta_X n_X - \alpha_X n_e n_{iX} - \frac{n_{cnf}}{\lambda_d} (I_{\text{tip-}iX} + I_{cur\text{-}iX}) + \sum_{\ell XY} k_\ell n_X n_{iY} - \sum_{QYX} k_Q n_Y n_{iX}, \quad (3)$$

$$\partial_t(n_{iY}) = \beta_Y n_Y - \alpha_Y n_e n_{iY} - \frac{n_{cnf}}{\lambda_d} (I_{\text{tip-}iY} + I_{\text{cur-}iY}) + \sum_{QYX} k_Q n_Y n_{iX} - \sum_{\ell XY} k_\ell n_X n_{iY}, \quad (4)$$

$$\partial_t(n_{iZ}) = \beta_Z n_Z - \alpha_Z n_e n_{iZ} - \frac{n_{cnf}}{\lambda_d} (I_{\text{tip-}iZ} + I_{\text{cur-}iZ}), \quad (5)$$

where k_j is the ion-neutral reaction rate coefficient. Equations (3)–(5) refer to the growth rate of positively charged ions of type X, Y, and Z, respectively, in the plasma on account of gain in the ion number density per unit time

due to ionization of neutral atoms (first term), decay rate of positively charged ion number density due to electron-ion recombination (second term) and ion collection current at the CNF surface (third term), increase/decrease in ion number density due to reaction between positively charged ions and neutrals in the plasma, these reactions have different reaction rate coefficients (fourth and fifth terms).

D. Kinetic equations of neutral atoms

$$\begin{aligned}\partial_t(n_X) = {} & \alpha_X n_e n_{iX} - \beta_X n_X + \frac{n_{cnf}}{\lambda_d}(1-\gamma_{iX})(I_{\text{tip-}iX} + I_{cur\text{-}iX}) \\ & - \frac{n_{cnf}}{\lambda_d}\gamma_X(I_{\text{tip-}X} + I_{cur\text{-}X}) + IF_X - OF_X - J_{\text{ai-x}} \\ & + J_{\text{desp-x}} - \sum_{\ell XY} k_\ell n_X n_{iY} + \sum_{QYX} k_Q n_Y n_{iX}, \end{aligned} \tag{6}$$

$$\begin{aligned}\partial_t(n_Y) = {} & \alpha_Y n_e n_{iY} - \beta_Y n_Y + \frac{n_{cnf}}{\lambda_d}(1-\gamma_{iY})(I_{\text{tip-}iY} + I_{cur\text{-}iY}) \\ & - \frac{n_{cnf}}{\lambda_d}\gamma_Y(I_{\text{tip-}Y} + I_{cur\text{-}Y}) + IF_Y - OF_Y - J_{\text{ai-Y}} \\ & + J_{\text{desp-Y}} + J_{\text{thd}} - \sum_{QYX} k_Q n_Y n_{iX} + \sum_{\ell XY} k_\ell n_X n_{iY}, \end{aligned} \tag{7}$$

$$\begin{aligned}\partial_t(n_Z) = {} & \alpha_Z n_e n_{iZ} - \beta_Z n_Z + \frac{n_{cnf}}{\lambda_d}(1-\gamma_{iZ})(I_{\text{tip-}iZ} + I_{cur\text{-}iZ}) \\ & - \frac{n_{cnf}}{\lambda_d}\gamma_Z(I_{\text{tip-}Z} + I_{cur\text{-}Z}) + IF_Z - OF_Z - J_{\text{ai-Z}} \\ & + J_{\text{desp-Z}}, \end{aligned} \tag{8}$$

where γ_{ij} is the ion sticking coefficient, $J_{\text{ai-j}}$ is the adsorption flux on the catalyst substrate surface, $J_{\text{desp-j}}$ is the desorption flux from the catalyst substrate surface, J_{thd} is the hydrogen ion flux due to thermal dehydrogenation, $I_{\text{tip-j}}$ and $I_{cur\text{-j}}$ are the neutral atom collection currents at the tip and curved surface of the CNF,[24] $IF_j(=\frac{4.4\times10^{17}\times f_j}{V})$ and $OF_j(=\frac{R_p\times n_j}{V})$ are the jth species inflow rate and outflow rate into/from the plasma chamber, respectively. f_j is the inlet flow of corresponding gas, $V(\approx 2\times10^4\,\text{cm}^3)$ is the volume of the chamber, and R_p is the pumping rate.[27] Equations (6)–(8) refer to the growth of neutral atoms number density of type X, Y, and Z in the plasma on account of gain in neutral atom number density per unit time due to electron-ion recombination (first term), decay rate of neutral atom number density due to ionization (second term), gain in neutral density per unit time due to neutralization of ions collected at the CNF surface (third term), decay in neutral atom number density due to accretion of neutrals at the CNF surface (fourth term), inflow and outflow rate of neutral species from plasma chamber (fifth and sixth term), adsorption of ions to the catalyst surface (seventh term) and desorption of ions from the catalyst surface (eighth term). The Ninth term of Eq. (7) corresponds to increase in hydrogen ion number density in the plasma due to thermal dehydrogenation.[28,29] The remaining terms correspond to the increase and decrease in neutral atoms number density due to the reaction between ions and neutrals in the plasma as mentioned above.

E. Growth rate equations of carbon and hydrogen species generated on the catalyst surface

$$\begin{aligned}\partial_t(n_C) = {} & \Bigg[\sum_X J_X(1-\theta_t) + \sum_X n_{sX}\upsilon\exp\left(\frac{-\delta E_{td}}{k_B T_S}\right) \\ & + \sum_{iX}\left(\sum_X \frac{n_{sX}y_d}{\upsilon_0}\right)J_{iX} + \sum_{iX} J_{iX} + J_C \\ & + \sum_{iY}\left(\sum_{iX}\frac{J_{iX}\sigma_{ads}}{\upsilon}\right)J_{iY} - \sum_Y\left(\sum_X n_{sX}\sigma_{ads}\right)J_Y \\ & - n_C\upsilon\exp\left(\frac{-\delta E_{ev}}{k_B T_S}\right) - \sum_X n_X\upsilon\exp\left(\frac{-\delta E_{dhc}}{k_B T_S}\right) \\ & - n_C\sigma_{ads}\upsilon n_H - \sum_\gamma n_C\sigma_{ads}J_\gamma \\ & - \frac{1}{(\pi D)}\sum_{goc} k_g n_O n_C\Bigg] \times \frac{d\Omega_C}{dt}, \end{aligned} \tag{9}$$

$$\begin{aligned}\partial_t(n_H) = {} & \sum_Y J_Y(1-\theta_t) + \sum_{iX}\left(\sum_X\frac{n_{sX}y_d}{\upsilon_0}\right)J_{iX} \\ & + \sum_Y n_{sY}\exp\left(\frac{-\delta E_{td}}{k_B T_S}\right) + \sum_{iY} J_{iY} \\ & - \sum_Y n_{sY}\upsilon\exp\left(\frac{-\delta E_{dH}}{k_B T_S}\right) - n_C\sigma_{ads}\upsilon n_H \\ & - \sum_Y n_{sY}\sigma_{ads}J_Y - \sum_Y n_{sY}\sigma_{ads}J_{iY}, \end{aligned} \tag{10}$$

where $J_j\left(=\frac{n_j v_{th}}{4}\right)$ is the flux of neutral impinging species on the catalyst nanoparticle surface from the plasma, $n_{sj}(=\theta_j\upsilon_0)$ is the surface concentration of neutral atoms,[28,30] θ_j is the surface coverage by the respective species, υ_0 is the adsorption sites per unit area, $J_{ij}\left(=n_{ij}\sqrt{\frac{k_B T_{ij}}{m_{ij}}}\right)$ is the ions flux on the catalyst nanoparticle surface from the plasma,[28] where j denotes the species of type X, Y, Z, and carbon species. υ is the thermal vibration frequency, n_C and n_H are the number of carbon species and hydrogen radicals per unit area, respectively, generated on the catalyst nanoparticle surface, n_O is the flux of water/oxygen species that impinge the catalyst nanoparticle surface, $\sigma_{ads}(=6.8\times10^{-16}\text{cm}^2)$ is the cross section for the interaction among the various species,[28] $\delta E_{td}(=2.1\,\text{eV})$ is the hydrocarbons dissociation energy, $\delta E_{\text{ev}}(=1.8\,\text{eV})$ carbon evaporation energy, $\delta E_{dhc}(=1.8\,\text{eV})$ is the hydrocarbon desorption energy, $\delta E_{th}(=1.87\,\text{eV})$ is the energy for thermal dehydrogenation of hydrocarbons, and $\delta E_{dH}(=1.8\,\text{eV})$ is the hydrogen atom desorption energy.[31]

The nucleation of the CNF is a fundamental stage of synthesizing the CNFs and the catalyst nanoparticles placed over the substrate surface act as the nucleation sites for the growth of CNFs. The hydrocarbons and other species from the plasma dissociate on the catalyst nanoparticle surface to generate carbon species. These carbon species diffuse through the catalyst particle surface and precipitate at the rear of the catalyst nanoparticle, and hence, nucleate the CNF.[8,9] Thus, the number of carbon species generated on the catalyst nanoparticle surface is the one of the most remarkable growth determining factors. Due to continuous

dissociation of hydrocarbons and generation of carbon species, a thin carbon film known as amorphous carbon layer is formed on the catalyst nanoparticle surface, which prevents the further dissociation of hydrocarbons and diffusion of carbon species through the catalyst, and thus leads to the poisoning of the catalyst nanoparticle and stops the CNFs growth.[32] The Eq. (9) represents the number of carbon species per unit area (n_C) generating on the catalyst nanoparticle on accounting various surface processes i.e., hydrocarbons adsorption on the catalyst nanoparticle surface (first term), thermal dissociation of hydrocarbons (second term), ion induced dissociation of hydrocarbons (third term), ion dissociation (fourth term), direct carbon flux towards catalyst surface (fifth term),[28] and generation of carbon species due to hydrocarbon ions and hydrogen ions interaction, i.e., ion-ion interaction (sixth term). Moreover, many processes lead to the loss of carbon species, such as, loss of hydrocarbons neutrals due to interaction with hydrogen (seventh term), carbon evaporation (eight term), and desorption of hydrocarbons neutrals from catalyst nanoparticle surface.[31] The last three terms denote the loss of carbon species due to interaction with the atomic hydrogen generated on the catalyst nanoparticle surface, interaction with the species of type Z, and due to reaction with water/oxygen molecules, respectively. These interactions act as the sink for carbon species that may otherwise form an amorphous carbon layer on the catalyst nanoparticle surface. Moreover, the term $\frac{d\Omega_c}{dt}(=1-n_c\pi D^2)$ is the blocking term that denotes the blocking of hydrocarbons dissociation on the catalyst nanoparticle surface when the number of carbon species (n_c) generated on the catalyst nanoparticle surface increases rapidly and forms the amorphous carbon layer.

In the plasma-assisted catalytic growth process, the relative number density of carbon species (n_c) and hydrogen radicals (n_H) generated on the catalyst nanoparticle surface have significant implications for the CNF growth. Therefore, the model accounts the various processes responsible for the generation of hydrogen species or radicals on the catalyst nanoparticle surface discussed in Eq. (10), such as hydrogen adsorption on catalyst nanoparticle surface[30] (first term), hydrogen produced due to ion-induced dissociation and thermal dissociation of hydrocarbons (second and third term),[31] and direction hydrogen ion flux (fourth term). However, there are many other surface processes, which are responsible for the loss of hydrogen species, such as desorption of hydrogen radicals from catalyst nanoparticle surface (fifth term), interaction of hydrogen radicals with carbon species (sixth term), and interaction of adsorbed hydrogen with incoming hydrogen neutrals and ions from the plasma (seventh and eight term).[31]

F. Growth rate equation of carbon nanofiber

$$\partial_t(\pi d^2 h) = [D_s + D_b + k\nu D] \times \frac{P(4\pi D)n_C}{\nu\rho_{cnf}I_{ix}} + \left[\frac{D_m}{\pi D^2} + n_H(\pi D^2)\sum_y I_{iy} + \sum_z(\gamma_z I_z + \gamma_{iz} I_{iz})\right] \times \frac{m_{cat}}{\rho_{cat}}, \quad (11)$$

where D is the catalyst nanoparticle diameter, d is the CNF tip diameter, h is the CNF height, $D_S[=D_{S0}\exp(\frac{-\varepsilon_S}{k_BT_S})]$ is the surface diffusion coefficient, $D_{S0}=a_0^2\nu$, $a_0(=0.34\,\text{nm})$ is the carbon atoms inter-atomic distance, $\varepsilon_S(=0.3\,\text{eV})$ is the energy barrier for carbon species diffusion through the catalyst nanoparticle surface,[30] $D_b[=D_{b0}\exp(\frac{-\varepsilon_b}{k_BT_S})]$ is the bulk diffusion coefficient, $D_{b0}=\frac{\nu D}{2\pi}$, ε_b $(=1.6\,\text{eV})$ is the energy barrier for carbon species bulk diffusion through catalyst nanoparticle,[30] P $(\approx 20\,\text{GPa})$ is the pressure exerted by graphene layers on catalyst nanoparticle,[33] $k=A_k\exp(\frac{-\delta\varepsilon_{inc}}{k_BT_S})$ is the carbon species incorporation speed into the graphene layers,[34] $A_k=a_0\nu$, $\delta\varepsilon_{inc}(=0.4\,\text{eV})$ is the carbon species energy barrier to diffuse along the nanofiber-catalyst interface, ρ_{cnf} is the nanofiber density, m_{cat} is the mass of the metal catalyst particle, and ρ_{cat} is the density of metal catalyst nanoparticle. $D_m=D_{m0}\exp(\frac{-\varepsilon_{SD}}{k_BT_S})$ is the metal atoms diffusion coefficient,[35] D_{m0} is the constant, and ε_{SD} is the metal atoms self-diffusion activation energy.

The catalyzed CNF growth through the PECVD process is dynamical and complex process during which the carbon species generated on the catalyst nanoparticle surface [as discussed in Eq. (9)] diffuse through the catalyst and ultimately precipitate to form graphene layers around the rear surface of the catalyst nanoparticle and forms the CNF. Therefore, the terms in the first square bracket of Eq. (11) represent the surface as well as bulk diffusion of carbon species through catalyst nanoparticle and incorporation of carbon species to form graphene layers, respectively. Due to continuous dissociation of hydrocarbons and diffusion of carbon species, carbon precipitation grows in the vertical as well as in the planar direction. During the growth, the catalyst nanoparticle undergoes shape transformation due to the compressive stress exerted by the continuously adding graphene layers. Therefore, the pressure term is also taken into account outside the first bracket. Moreover, self-diffusion of the metal catalyst nanoparticle atoms is also an important factor for catalyst nanoparticle shape transformation (first term in the second square bracket). Continuous supply carbon species results in the continuous vertical as well as planar growth. However, to stop the planar growth, hydrogen plays an important role. Highly reactive atomic hydrogen and ions act as the etchants and etch the terminal carbons. However, the carbons under the shadow of the catalyst nanoparticle remain protected. Thus, new graphene layers form in between the catalyst particle and existing graphene layers and support the vertical growth. The second term in the second square bracket denotes the etching of CNF by the hydrogen species. The last term in the second square bracket represents the sticking of atoms and ions of type "z" that also act as good etchants and etch the CNF tip as well as side walls.

III. RESULTS AND DISCUSSION

Using the analytical equations presented in Sec. II, we will investigate and analyze how the different plasma compositions affect the CNF growth rate and structure. To understand the better insight of such effects, we varied the types of neutrals and ions for various plasma compositions and their

initial concentration in the reactive plasma and simultaneously solved the first order differential Eqs. (1)–(11) using Mathematica Software for typical glow discharge parameters and appropriate boundary conditions at t = 0, $n_e = 1.5 \times 10^{12}\,\mathrm{cm}^{-3}$, $n_{iX} = 0.12\,n_e$, $0.1\,n_e$, and $0.08\,n_e$ for $C_2H_2^+$, C_2H^+, and $C_4H_2^+$, respectively, $n_{iY} = 0.35\,n_e$ for H_2^+ and H^+. The initial concentration of H_2O/O_2 species in the reactive plasma is considered to be less than 1% of the hydrocarbon species,[13,20] i.e., $n_{iZ} = 0.001\,n_e$, for H_2O^+, OH^+, and O^+, $n_X = 1.4 \times 10^{13}\,\mathrm{cm}^{-3}$, $1.2 \times 10^{13}\,\mathrm{cm}^{-3}$, $1.0 \times 10^{13}\,\mathrm{cm}^{-3}$ for C_2H_2, C_2H, and C_4H_2, respectively, $n_Y = 5.6 \times 10^{14}\,\mathrm{cm}^{-3}$ for H_2 and H, respectively, $n_Z \approx 5 \times 10^{10}\,\mathrm{cm}^{-3}$ for H_2O, O_2, OH, and O respectively, $T_e = 2.5\,\mathrm{eV}$, $T_i = 2500\,\mathrm{K}$, $T_n = 2000\,\mathrm{K}$, $m_{iX} = m_X = 26$ amu, 25 amu, 50 amu for, C_2H_2, C_2H, and C_4H_2 respectively, $m_{iY} = m_Y = 2$ amu, 1 amu for H_2 and H, respectively, $m_{iZ} = m_Z = 18$ amu, 32 amu, 17 amu, 16 amu for H_2O, O_2, OH, and O, respectively, atomic mass of nickel $m_{cat} = 58.69$ amu, density of catalyst nickel $\rho_{ct}(=8.96\,\mathrm{g/cm}^3)$, coefficient of recombination of electrons and ions $\alpha_X \approx \alpha_Y \approx \alpha_Z(=1.12 \times 10^{-7}\,\mathrm{cm}^3/\mathrm{s})$, $\kappa = -1.2$, substrate temperature $T_S = 800\,\mathrm{K}$, sticking coefficient of electrons $\gamma_e(=1)$, sticking coefficient of ions $\gamma_{iX} = \gamma_{iY} = \gamma_{iZ}(=1)$, sticking coefficient of neutral atoms acetylene $\gamma_X = \gamma_Y = \gamma_Z(=1)$, acetylene gas ($C_2H_2$) inflow rate = 50 sccm, and hydrogen gas (H_2) inflow rate = 100 sccm.

In Fig. 1, the variation of the CNF height with growth time for different plasmas is shown. From Fig. 1, one can see that the CNF grown in $C_2H_2+H_2+O_2$ plasma attains maximum height followed by the CNF grown in $C_2H_2+H_2+H_2O$ and $C_2H_2+H_2$ plasma. To understand why CNF growth rate is the maximum in $C_2H_2+H_2+O_2$ plasma and minimum in $C_2H_2+H_2$ plasma, we have plotted the curves of the number of carbon species per unit area generated on the catalyst nanoparticle surface as a function of growth time for different plasmas (see Fig. 2). From Fig. 2, one can see that the rate at which the carbon species generated on the catalyst nanoparticle surface is the maximum for $C_2H_2+H_2+O_2$ plasma followed by $C_2H_2+H_2+H_2O$ and $C_2H_2+H_2$ plasma. In $C_2H_2+H_2$ plasma, CNF growth continues due to continuous dissociation of hydrocarbons on the front surface of catalyst nanoparticle to generate carbon species and due to continuous diffusion of these carbon species through the

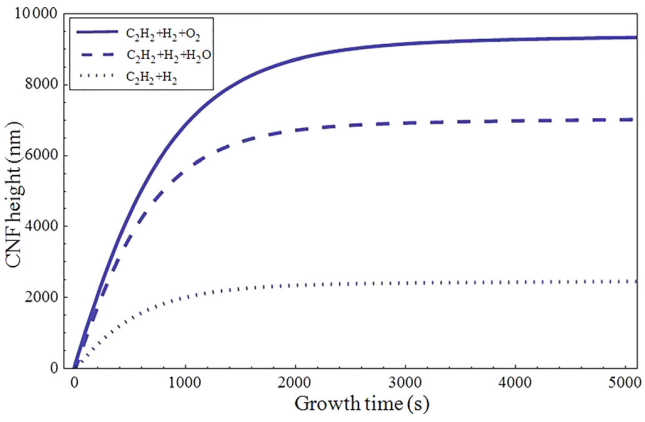


FIG. 1. Time evaluation of height of the CNF grown in different plasma mixtures.

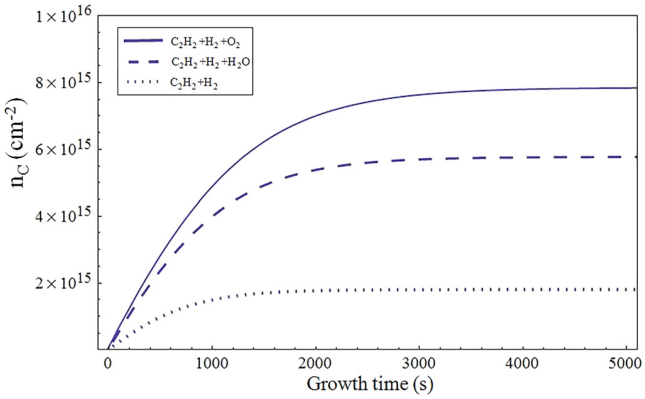


FIG. 2. Time evaluation of generation of carbon species per unit area (n_C) on the catalyst nanoparticle surface as a function of different plasma mixtures.

catalyst. Atomic hydrogen etches the amorphous carbon from the catalyst surface and maintains the catalyst activity for further dissociation of hydrocarbons coming from the plasma and contributes to the continuous CNF growth. During the growth, the number of carbon species generated on the catalyst nanoparticle surface increases continuously due to which etching of amorphous carbon by hydrogen becomes less effective, and this leads to covering the catalyst nanoparticle surface with amorphous carbon layer and block further dissociation of incoming hydrocarbons from the plasma. Thus, CNF growth terminates due to the poisoning of catalyst nanoparticle. However, in $C_2H_2+H_2+H_2O$ plasma, water acts as the catalyst enhancer, i.e., water molecules enhance and preserve the catalyst activity during the CNF growth. Water molecules etch away the amorphous carbon layer covering the catalyst nanoparticle by means of $C + H_2O \rightarrow CO + H_2$ (Ref. 11) with large reaction rate coefficient, and thus, maintain catalyst activity that increases the hydrocarbons dissociation on the catalyst surface. Therefore, the number of carbon species generated on the catalyst surface increases (cf. Fig. 2) and results in the much-enhanced CNF growth as compared to $C_2H_2+H_2$ plasma. Furthermore, the addition of water vapors in the plasma not only removes the carbonaceous impurities (amorphous carbon) from the catalyst surface but also increases the number of hydrogen radicals on the catalyst surface $(C + H_2O \rightarrow CO + H_2)$. These highly reactive hydrogen radicals interact with the carbon species on the catalyst surface and act as a sink for carbon species that may otherwise be diffused through the catalyst to contribute in the CNF growth. However, $C_2H_2+H_2+O_2$ plasma affords highly productive CNF growth; this is because of the two important roles played by the oxygen; (i) removal of amorphous carbon from the catalyst surface by converting carbon species into CO species by the means of various reactions,[20,36] and thus, prevents catalyst poisoning, (ii) removal of the highly reactive hydrogen radicals, i.e., $(H + O_2 \rightarrow OH + O)$ with large reaction rate $k_g = 10^{17}\mathrm{cm}^3/\mathrm{mol\,s}$ (Refs. 20 and 36) that prevents the sinking of the carbon species due to carbon-hydrogen interaction. These highly effective contributions of oxygen continuously increase the hydrocarbons dissociation on the catalyst surface to generate the maximum number of carbon

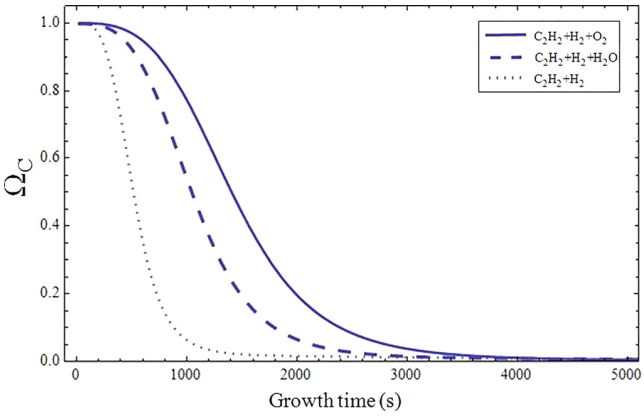


FIG. 3. Temporal variation of blocking function (Ω_C) for different gas mixtures.

species on the catalyst surface (cf. Fig. 2) and results in much higher CNF growth compared to $C_2H_2+H_2+H_2O$ and $C_2H_2+H_2$ plasma. During the growth, the concentration of oxygen species in the plasma decreases (as their concentration is about 1% of the hydrocarbon species) and formation of amorphous carbon on the catalyst surface increases. Consequently, dissociation of the hydrocarbon species on the catalyst surface decreases, which ultimately drives the CNF growth in the saturation region.

To make it clear on how different plasmas maintain the catalyst activity during the growth, we have plotted the curves of blocking function Ω_C as a function of growth time for various plasma mixtures in Fig. 3, which shows that blocking function attains the saturation (i.e., the situation when catalyst particle is completely covered with the amorphous carbon film) quickly in the case of $C_2H_2+H_2$ plasma and blocks the further dissociation of the incoming hydrocarbons from the plasma on the catalyst surface. Thus, catalyst particle loses its activity, and therefore, more hydrocarbon species remains undissociated in the plasma as can be seen from Fig. 4. However, in $C_2H_2+H_2+H_2O$ plasma, the blocking function slowly attains the saturation, and thus, the catalyst particle maintains its activity for a much longer time. This leads to more dissociation of hydrocarbons on the catalyst surface, and therefore, fewer numbers of undissociated hydrocarbons are left in the plasma (cf. Fig. 4). In $C_2H_2+H_2+O_2$ plasma, the blocking function attains the saturation at the slowest rate and thus maintains the high catalyst activity due to which much more hydrocarbons dissociate on the catalyst surface, and therefore, least number of hydrocarbons remained undissociated in the plasma as can be seen in Fig. 4. From the results obtained in Fig. 1, the calculated CNF growth rates (i.e., saturated height per unit time taken to achieve saturation) are 1.52 nm/s, 3.315 nm/s, and 4.34 nm/s for $C_2H_2+H_2$, $C_2H_2+H_2+H_2O$, and $C_2H_2+H_2+O_2$, respectively. This shows that on the addition of water and oxygen in the reactive plasma, CNFs grow with much higher rates (almost double).

Figure 5 shows the time evolution of the CNF tip diameter for the different plasmas. From Fig. 5, it can be seen that the CNF tip diameter decreases with the growth time. This can be understood from the catalyst particle shape and size transformation, which remained on the tip of the CNF during the growth. Due to compressive stress applied by the continuously formed graphene layers around the catalyst particle, the shape of the catalyst particle changes continuously to teardrop like shape from the initial spherical shape that causes the remarkable reduction in the rear size of the catalyst particle. Moreover, hydrogen acts as the etchant and etches the CNF side walls as well as the tip, which results in the reduction in the CNF tip diameter. From Fig. 5, one can also see that the CNF tip diameter is least for the $C_2H_2+H_2+O_2$ plasma followed by $C_2H_2+H_2+H_2O$ and $C_2H_2+H_2$ plasma. The significant reduction in the CNF tip size during the growth in $C_2H_2+H_2+H_2O$ plasma in contrast with the $C_2H_2+H_2$ plasma is attributed to the fact that the

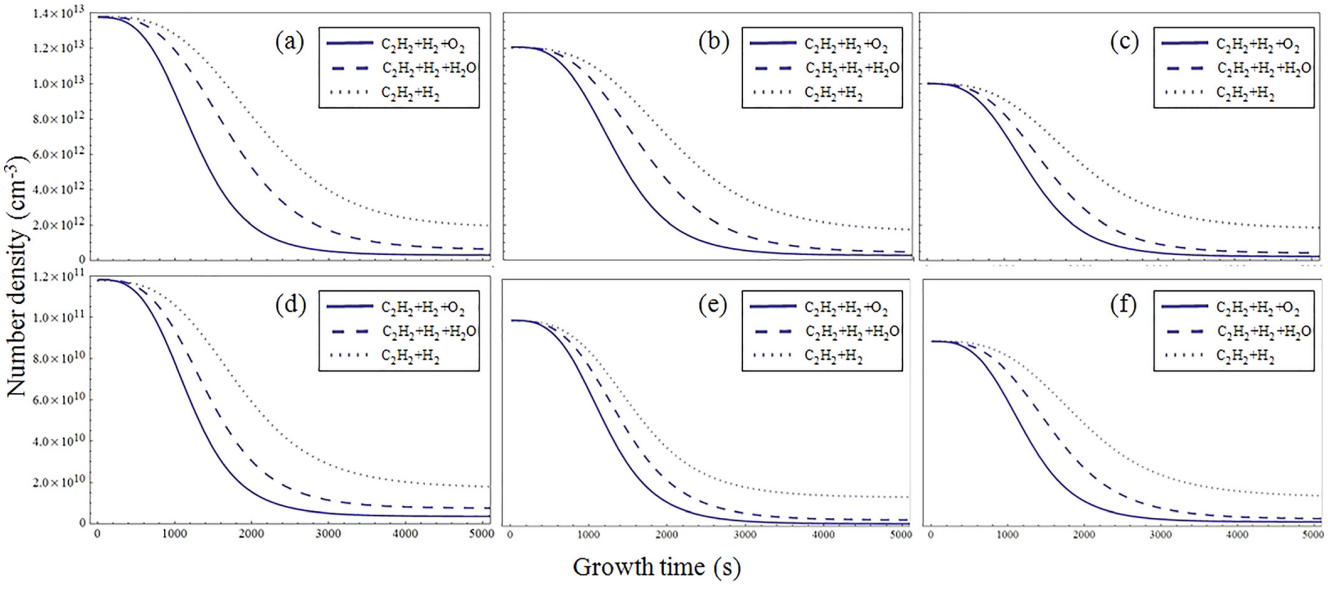


FIG. 4. Temporal variation of plasma species number density in different plasma mixtures for (a) C_2H_2, (b) C_2H, (c) C_4H_2, (d) $C_2H_2^+$, (e) C_2H^+, and (f) $C_4H_2^+$.

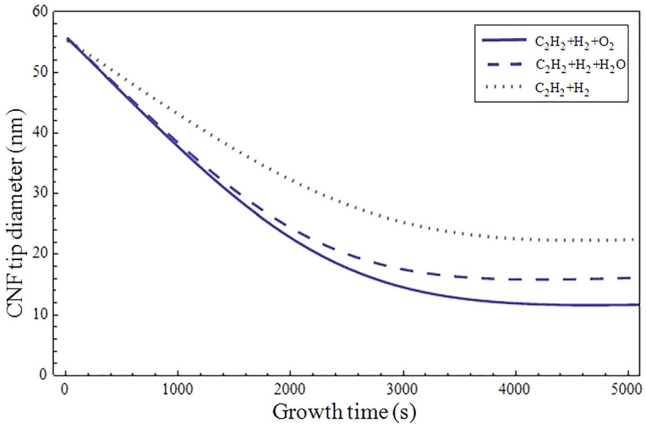


FIG. 5. Time evolution of diameter of CNF grown in three different plasma mixtures.

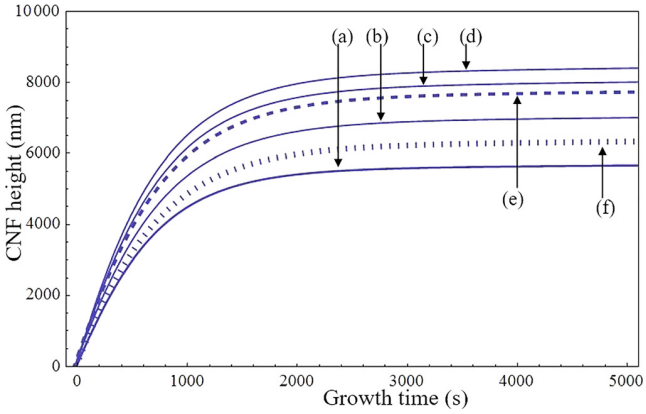


FIG. 6. Temporal variation of height of CNF grown in $C_2H_2+H_2+H_2O$ plasma as function of different water concentration (number density) in the plasma, i.e., (a) $4.7 \times 10^{10}\,cm^{-3}$, (b) $5.5 \times 10^{10}\,cm^{-3}$, (c) $6.0 \times 10^{10}\,cm^{-3}$, (d) $6.5 \times 10^{10}\,cm^{-3}$, (e) $7.5 \times 10^{10}\,cm^{-3}$, and (f) $8.5 \times 10^{10}\,cm^{-3}$.

number of hydrogen radicals augments on the catalyst nanoparticle surface by means of $C + H_2O \rightarrow CO + H_2$, thus relative etching of the CNF increases, which results in the reduction in the CNF tip diameter. Moreover, OH species formed in the $C_2H_2+H_2+H_2O$ plasma also act as the

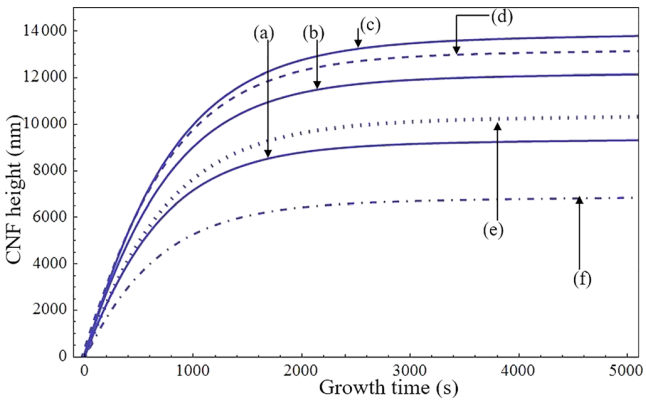


FIG. 7. Temporal variation of height of CNF grown in $C_2H_2 + H_2 + O_2$ plasma as function of different oxygen concentration (number density) in the plasma, i.e., (a) $4.5 \times 10^{10}\,cm^{-3}$, (b) $6.0 \times 10^{10}\,cm^{-3}$, (c) $7.3 \times 10^{10}\,cm^{-3}$, (d) $8.0 \times 10^{10}\,cm^{-3}$, (e) $8.5 \times 10^{10}\,cm^{-3}$, and (f) $9.5 \times 10^{10}\,cm^{-3}$.

etchants and etch the CNF and reduce the CNF tip diameter. However, the CNF tip diameter is least in the $C_2H_2+H_2+O_2$ plasma among all three plasma mixtures. In this way, in $C_2H_2+H_2+O_2$ plasma, the CNF grows for a much longer time in contrast to other plasmas; relative etching of the CNF by etchants (hydrogen radicals and OH species) is higher, which brings about the lesser tip diameter.

A family of time evolution curves discussed above reveals that CNF can be grown with much higher rate and with lesser tip diameter when H_2O/O_2 is added to the hydrocarbon + hydrogen plasma. The kinetics of water/oxygen assisted catalytic CNF growth must be accurately known to synthesize the highly efficient CNFs for their practical applications in field emission devices. The CNF growth can be

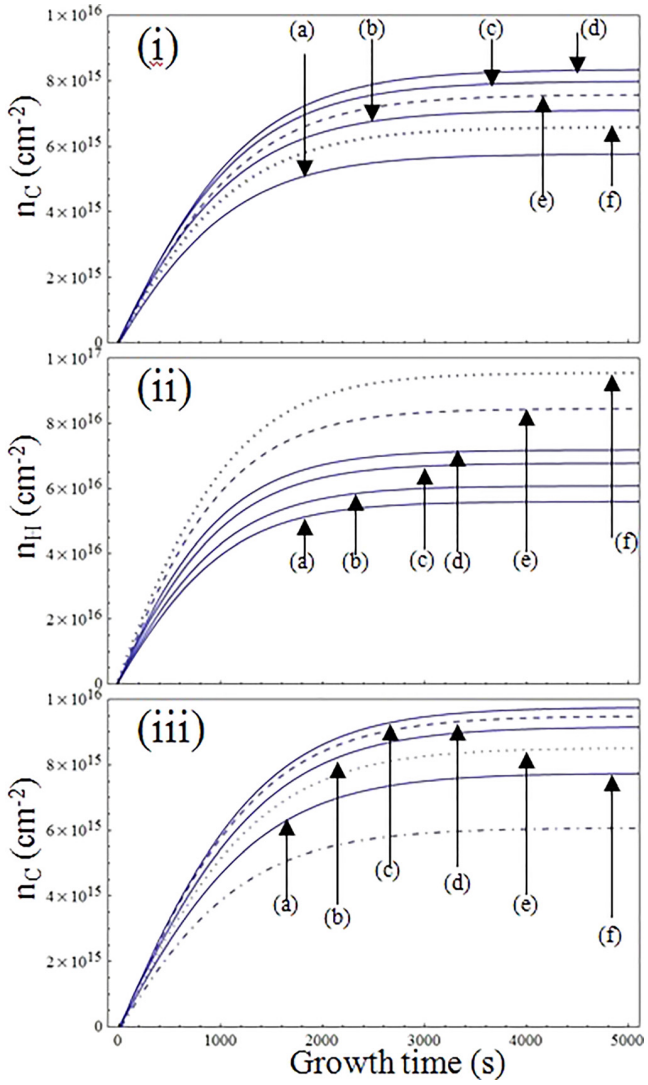


FIG. 8. (i) and (ii) Shows the temporal variation of carbon species number density (n_C) and hydrogen species number density (n_H), respectively, generated on the catalyst nanoparticle surface in $C_2H_2 + H_2 + H_2O$ plasma as function of different water concentration (number density) in the plasma, i.e., (a) $4.7 \times 10^{10}\,cm^{-3}$, (b) $5.5 \times 10^{10}\,cm^{-3}$,(c) $6.0 \times 10^{10}\,cm^{-3}$, (d) $6.5 \times 10^{10}\,cm^{-3}$, (e) $7.5 \times 10^{10}\,cm^{-3}$, and (f) $8.5 \times 10^{10}\,cm^{-3}$. (iii) Temporal variation of carbon species number density generated on the catalyst nanoparticle surface in $C_2H_2 + H_2 + O_2$ plasma as function of different oxygen concentration number density) in the plasma, i.e., (a) $4.5 \times 10^{10}\,cm^{-3}$, (b) 6.0×10^{10} (c) $7.3 \times 10^{10}\,cm^{-3}$, (d) $8.0 \times 10^{10}\,cm^{-3}$, (e) $8.5 \times 10^{10}\,cm^{-3}$, and (f) $9.5 \times 10^{10}\,cm^{-3}$.

deterministically controlled by suitably varying the growth parameters. The water/oxygen level in the plasma is one of the most significant growth parameters. In Fig. 6, the time evolution of the CNF growth has been carried out at different water concentrations, i.e., different water molecules number density in the reactive plasma. The curves plotted in Fig. 6 exhibit the complex behavior: as water number density increases up to $6.51 \times 10^{10}\,\mathrm{cm}^{-3}$ (which is about 1% of the hydrocarbon species number density in the plasma), the CNF height, i.e., the CNF growth rate increases, but with further increase in water concentration CNF height decreases significantly. The increase in the CNF growth rate till $6.51 \times 10^{10}\,\mathrm{cm}^{-3}$ exemplifies the effect of water in maintaining the catalyst activity. Beyond this water level, the CNF growth dropped steadily, which is due to the rapid increase in the hydrogen species on the catalyst surface [cf. Fig. 8(i)]. These hydrogen species ruin the carbon diffusion and precipitation, and thus retards the growth process. Moreover, with the high water level in the plasma, the oxidation of carbons from the catalyst surface increases by means of $C + H_2O \rightarrow CO + H_2$; therefore, the number of carbon species available for growth decreases significantly [cf. Fig. 8(ii)]. Thus, the last three terms of Eq. (9) become dominants over the all other terms responsible for carbon generation due to which water loses its primary characteristic. Likewise, the effect of the oxygen level in the plasma on the CNF growth has been seen in Fig. 9, which indicates that CNF height increases up to optimum oxygen concentration, i.e., $7.3 \times 10^{10}\,\mathrm{cm}^{-3}$. This is due to the fact that adequate measure of oxygen expels the amorphous carbon from the catalyst surface and accordingly preserves and enhances the catalyst activity which is its essential trademark. But beyond this ideal level, oxygen species oxidizes the carbon from catalyst surface [last two terms of Eq. (9) becomes dominants] and act as the sink for carbon species [cf. Fig. 8(iii)] that otherwise may form graphene layers and may contribute to the growth. Hence, the oxygen setbacks its essential characteristic and poisons the growth process. Thus, one can interpret the results of Figs. 7–9 that H_2O/O_2 enhances the CNF growth with the large extent when their relative concentration is almost 1% of the hydrocarbon species in the reactive plasma.

A. Comparison with available experimental literature

In our analytical model, we have investigated the effects of different plasmas (i.e., $C_2H_2+H_2$, $C_2H_2+H_2+H_2O$, and $C_2H_2+H_2+O_2$) on the growth and tip diameter of the vertically aligned CNFs. The theoretical results obtained from our model are in good agreement with the available experimental observation, as follows. Hata *et al.*,[13] Yoshihara *et al.*,[14] Yun *et al.*,[15] Hussain *et al.*[16] observed the increase in CNTs growth rate (CNF height per unit deposition time) on the addition of water in the CVD growth process. He *et al.*[37] observed the much enhanced CNF growth and more elongation in catalyst nanoparticle in the presence of water in acetylene/hydrogen plasma mixture. Zhang *et al.*[20] revealed that the CNF growth rate increases when oxygen is added to the hydrocarbon/hydrogen reactive plasma. Kim *et al.*[21] compared the growth of CNTs in the above three plasma mixtures and found that growth rate is maximum in $C_2H_2+H_2+O_2$ plasma followed by $C_2H_2+H_2+H_2O$, and $C_2H_2+H_2$ plasma. Futaba *et al.*[38] observed the increase in

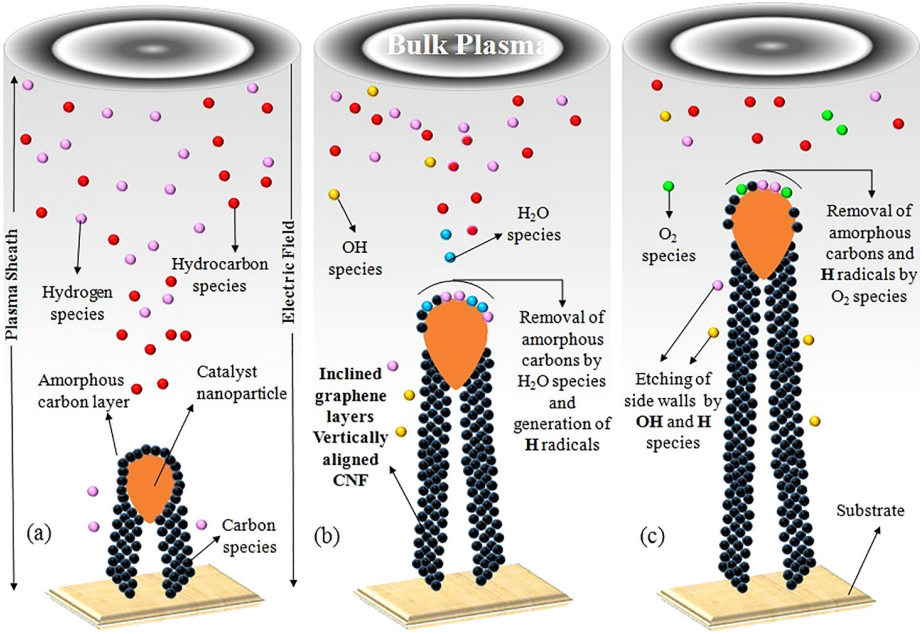


FIG. 9. Bottom to top approach, i.e., atom to atom operation of CNF growth in three different plasma gas mixture. (a) CNF growth in typical hydrocarbon (C_2H_2)/hydrogen (H_2) plasma, in which continuous dissociation of hydrocarbon species (neutrals and ions) on the catalyst nanoparticle surface generate the excess carbon species and the ultimately amorphous carbon layer is formed on the catalyst nanoparticle surface and terminates the CNF growth. (b) CNF growth in $C_2H_2 + H_2 + H_2O$ plasma, in which water species etch the amorphous carbon from the catalyst surface by means of many chemical reactions and preserve the catalyst activity and enhance the CNF growth. (c) CNF growth in $C_2H_2 + H_2 + O_2$ plasma, in which oxygen species not only remove the amorphous carbon from the catalyst surface but also remove the excess hydrogen radicals (H) from the catalyst surface which act as the sink for carbon species and thus CNF grows with much higher rate as compared to growth in (a) and (b).

CNT growth rate up to some optimum water level in growth atmosphere and found the decrease in growth beyond that level.

IV. CONCLUSION

An analytical model to study the impact of different plasma mixtures on the growth kinetics of the CNFs has been developed. The time evolution analysis of the various complex equations presented in Sec. II have been summarized in Fig. 9, which indicates the CNF growth in three different plasmas. Figure 9(a) shows that CNF growth in normal C_2H_2+H_2 plasma terminates due to the formation of the amorphous carbon layer on the catalyst active surface. However, on the addition of water in the reactive plasma, i.e., in C_2H_2+H_2+H_2O plasma, water molecules etch the amorphous carbons by means of $C + H_2O \rightarrow CO + H_2$ from the catalyst surface, and thus, preserve the catalyst activity and enhance the CNF growth as shown in Fig. 9(b). On the other hand, Fig. 9(c) shows that the CNF growth rate is the maximum in C_2H_2+H_2+O_2 plasma. This is because of the dual role played by the oxygen during the CNF growth, i.e., etching of amorphous carbon layer from the catalyst surface and removal of hydrogen species from the catalyst surface, which act as the sink for carbon species. Also, in the present study, we found that the tip diameter is the minimum for CNF grown in C_2H_2+H_2+O_2 plasma followed by C_2H_2+H_2+H_2O and C_2H_2+H_2 plasmas. Moreover, the relative concentrations of oxidizers in the reactive plasma significantly influence the CNF growth, i.e., CNF growth rate increases when the concentration of oxidizers increases up to some optimum level after that CNF growth rate rapidly decreases. This is because of the oxidation of the carbon species generated on the catalyst surface that might ultimately be diffused and precipitated in the form of the graphene layers around the catalyst nanoparticle.

The present work can be used for the better understating of the plasma aided catalyst-assisted growth of vertically aligned CNFs and to fabricate the highly desirable field emitter devices based on one-dimensional carbon nanostructures (CNFs/CNTs). More importantly, the conclusions obtained here are not only restricted to the growth of CNTs and CNFs and may extend to the catalytic growth of other carbon nanostructures in the low-temperature reactive plasma.

[1]W. A. Heer, A. Chatelain, and D. Ugarte, Science **270**(5239), 1179 (1995).
[2]S. Fan, M. G. Chapline, N. R. Franklin, T. W. Tombler, A. M. Cassell, and H. Dai, Science **283**(5401), 512 (1999).
[3]T. Y. Tsai, C. Y. Lee, N. H. Tai, and W. H. Tuan, Appl. Phys. Lett. **95**(1), 013107 (2009).
[4]V. I. Merkulov, A. V. Melechko, M. A. Guillorn, M. L. Simpson, D. H. Lowndes, J. H. Whealton, and R. J. Raridon, Appl. Phys. Lett. **80**, 4816 (2002).
[5]H. W. Wei, K. C. Leou, M. T. Wei, Y. Y. Lin, and C. H. Tsai, J. Appl. Phys. **98**, 044313 (2005).
[6]M. Meyyappan, L. Delzeit, A. Cassell, and D. Hash, Plasma Sources Sci. Technol. **12**, 205 (2003).
[7]H. Cui, X. Yang, M. L. Simpson, D. H. Lowndes, and M. Varela, Appl. Phys. Lett. **84**, 4077 (2004).
[8]C. S. Cojocaru, A. Senger, and F. L. Normand, J. Nanosci. Nanotechnol. **6**, 1331 (2006).
[9]N. M. Rodriguez, J. Mater. Res. **8**, 3233 (1993).
[10]D. H. Seo, S. Kumar, A. E. Rider, Z. Han, and K. Ostrikov, Opt. Mat. Express **2**, 700 (2012).
[11]T. M. Minea, S. Point, A. Granier, and M. Touzeau, Appl. Phys. Lett. **85**, 1244 (2004).
[12]L. Valentini, J. M. Kenny, L. Lozzi, and S. Santucci, J. Appl. Phys. **92**, 6188 (2002).
[13]K. Hata, D. N. Futaba, K. Mizuno, T. Namai, M. Yumura, and S. Iijima, Science **306**(5700), 1362 (2004).
[14]N. Yoshihara, H. Ago, and M. Tsuji, J. Phys. Chem. C **111**, 11577 (2007).
[15]Y. Yun, V. Shanov, Y. Tu, S. Subramaniam, and M. J. Schulz, J. Phys. Chem. B **110**, 23920 (2006).
[16]S. Hussain, R. Amade, and E. Bertran, Mater. Chem. Phys. **148**, 914 (2014).
[17]L. Song, G. Toth, R. Vajtai, M. Endo, and P. M. Ajayan, Carbon **50**, 5521 (2012).
[18]L. Kurzepa, A. L. Raus, J. Patmore, and K. Koziol, Adv. Funct. Mater. **24**, 619 (2014).
[19]C. Zhang, Y. Fu, Q. Chen, and Y. Zhang, Front. Mater. Sci. China **2**(1), 37 (2008).
[20]G. Zhang, D. Mann, L. Zhang, A. Javey, Y. Li, E. Yenilmez, Q. Wang, J. P. McVittie, J. Gibbons, and H. Dai, Proc. Natl. Acad. Sci. U.S.A **102**(45), 16141 (2005).
[21]Y. Kim, W. Song, S. Y. Lee, S. Shrestha, C. Jeon, W. C. Choi, M. Kim, and C. Y. Park, Jpn. J. Appl. Phys., Part 1 **49**, 085101 (2010).
[22]H. Mehdipour, K. Ostrikov, and A. E. Rider, Nanotechnology **21**, 455605 (2010).
[23]M. A. Lieberman and A. J. Lichtenberg, *Principles of Plasma Discharges and Material Processing* (Wiley-Interscience, New York, 1994).
[24]R. Gupta, S. C. Sharma, and R. Sharma, Plasma Sources Sci. Technol. **26**, 024006 (2017).
[25]M. S. Sodha, S. Misra, and S. K. Mishra, Phys. Plasmas **17**, 113705 (2010).
[26]M. S. Sodha, S. Misra, S. K. Misra, and S. Srivastava, J. Appl. Phys. **107**, 103307 (2010).
[27]I. B. Denysenko, S. Xu, J. D. Long, P. P. Rutkevych, N. A. Azarenkov, and K. Ostrikov, J. Appl. Phys. **95**, 2713 (2004).
[28]I. Denysenko and K. Ostrikov, J. Phys. D: Appl. Phys. **42**, 015208 (2009).
[29]L. Vattuone, Y. Y. Yeo, R. Kose, and D. A. King, Surf. Sci. **447**, 1 (2000).
[30]I. Denysenko and N. A. Azarenkov, J. Phys. D: Appl. Phys. **44**, 174031 (2011).
[31]Z. Marvi, S. Xu, G. Foroutan, and K. Ostrikov, Phys. Plasmas **22**, 013504 (2015).
[32]V. I. Merkulov, D. K. Hensley, A. V. Melechko, M. A. Guillorn, D. H. Lowndes, and M. L. Simpson, J. Phys. Chem. B **106**(41), 10570 (2002).
[33]M. Yudasaka, R. Kikuchi, T. Matsui, Y. Ohki, S. Yoshimura, and E. Ota, Appl. Phys. Lett. **67**, 2477 (1995).
[34]O. A. Louchev, T. Laude, Y. Sato, and H. Kanda, J. Chem. Phys. **118**, 7622 (2003).
[35]See http://matter.org.uk/matscicdrom/manual/df.html for metal atom diffusion coefficient.
[36]C. Benndorf, P. Joeris, and R. Kroger, Pure Appl. Chem. **66**, 1195 (1994).
[37]Z. B. He, J. L. Maurice, C. S. Lee, A. Gohier, D. Pribat, P. Legagneux, and C. S. Cojocaru, Carbon **49**, 435 (2011).
[38]D. N. Futaba, K. Hata, T. Yamada, K. Mizuno, M. Yumura, and S. Iijima, Phys. Rev. Lett. **95**, 056104 (2005).

American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-6, Issue-7, pp-52-57
www.ajer.org
Research Paper Open Access

Thermodynamic (Energy-Exergy) Analysis Of Nine MW Coal Based Thermal Power Plant Using Entropy Generation Principle.

Kaushalendra Kumar Dubey[#*1], R. S. Mishra[*2]
[#]*D/o Mechanical Engineering, Sharda University, UP-201306, India*
[1]*dubey.kaushalendra@gmail.com*
[*]*D/o Mechanical Engineering, Delhi Technological University, Delhi-110046, India*
[2]*professor_rsmishra@yahoo.co.in*

ABSTRACT: *The global concern about energy and fuel cost,environment safety, availability of conventional sources for power generation and fulfillment of energy demand will enhance the adoption of coal &gas based power plants. The rising attention of efficient and resourceful use of energy will support the adopted power generation techniques. The exergy modeling of thermal systems provide the information about,cause of losses, to scale the process, and rectifying the components.In this paper, a thermodynamic analysis of coal based nine MW thermal power plant is carried out. The energy-exergy analysis have been carried out for the boiler, turbine, condenser and pump majorly. The result of thermodynamic analysis is computed in terms of 24.12% plant energy efficiency, 35% of plant exergy efficiency and 11% of boiler energy efficiency respectively. The various exergy destruction in the components have been presented in this paper.*
Keywords: *Energy-Exergy Analysis, ,Irreversibility Analysis,. Rankine thermal power plant.*

I. INTRODUCTION

The large amount of heat is transferred between all utilities of thermal power plant and its greatly affect the plant performance. The First Law of Thermodynamic (FLT) is quantative evaluation of energy, it treat the work and heat interaction with equivalent forms of energy between system and surrounding. The FLT does not provide the information about internal losses, thermal inefficiency due to high grade energy of plant dump in to atmosphere and sustainability. The real performance of thermal system and quality of energy transfer process estimated by the approach of Second Law of Thermodynamic (SLT) called concept of exergy. The term 'Exergy'is maximum possible work obtained from a system at a given state when interacting with a environment, its useful quantity that extract from the SLT. It clearly indicates the inefficiencies of a process by locating the degradation of energy. In short exergy efficiency is true or real efficiency and energy efficiency is approximation of real efficiency. The energy-exergy analysis of 9 MW small size of coal based thermal power unit refers live operational data (Average monthly data of February 2015) of Adani Power Gujrat State of India in present analysis. Major units of plant (boiler,turbines,condensar and pump) have been investigated in proposed thermodynamic model. The main objective of energy-exergy modeling to identifying the magnitude of process, cause of losses and rectifying the components. The numerous researches have been conducted on energy-exergy analysis of various thermal systems, power plant components, and renewable energy conversion systems. Dicer and Rosen [2] stated that, why exergy is important for thermal analysis?.They stated that exergy flow rate of a flowing commodity is the maximum rate of work which is obtained from it and it passes reversibly to the environmental state, exchanging heat and materials only with the surroundings. exergy analysis is the theoretical limitations imposed upon a system, clearly pointing out that no real system can conserve exergy and that only a portion of the input exergy can be recovered. Also, exergy analysis quantitatively specifies practical limitations by providing losses in a form in which they are a direct measure of lost exergy.A.Bejan [3] introduced EGM (entropy generation minimization) method, and concluded whenever entropy production is minimized, useful energy is maximized. Khaliq and Kaushik [5] concluded the combustion chamber exergy destruction is 50% of overall exergy destruction of cycle, and effect of intermediate pressure-ratio and effect two stage of reheating on combined GT-ST power cycle performance by using 2nd law thermodynamic approach. Yang et al [6] compared conventional and advanced approach of exergy analysis of supercritical coal thermal power plant. Conventional approach identified exergy destruction of all components, whereas advanced exergyanalysis conclude thermodynamic interactions among thermal utilities for energy saving potential. Ahmadi and Dincer [7] analysed efficiency of dual pressure combined cycle

power plant (CCPP),1st law and 2nd law thermodynamic efficiencies are reduced with the addition of duct burner to HRSG, but power output of CCPP is improved. S.C.Kaushik et al and Aljundi [8,9] review begins with concept of exergy, entropy generation, irreversibility of thermodynamic system and exergy destroyed, and and identify the components having major energy-exergy losses in steam power plant. Gulhane and Thakur [10] investigated amount and source of irreversibilities in 35 TPH capacity of boiler. And also found 76% of exergy destruction at peak load on boiler but 1st and 2nd law efficiency are enhanced. Osueke et al [11] computed maximum exergy destruction 87.3% in boiler, and plant exergy efficiency 11.03% with the effect of water flow in boiler on steam generation and plant output of 75MW of steam power plant,which is located in Sapele-Nigeria. R kumar [12] critically summarized in review of 4-E (energy,exergy,exergonomic and economic) concept for different-different power generation techniques. He addressed the boiler and condenser component have major exergy destruction in case of rankine power cycle, whereas combustion chamber is larger part for exergy destruction in case of gas-turbine power cycle.

II. BRIEF OF STEAM POWER PLANT

The working of coal fired thermal power plant based on Rankine thermal power cycle. The plant size have medium level consisted by major parts 01 boiler unit,03 turbine unit,02 dearators,01 condenser unit, 02 HP heaters,01 pump,etc. Another thermal utilities also provide, the detail of plant layout, dimensions and flow rate values of units are not available due to company privacy policy. The proposed energy-exergy analysis is modeled for major parts of Rankine cycle,boiler,turbines, condenser and pump as per operational data of plant. The plant data is collected in month of February 2015,which is shown in table-1 (a) &(b).

Table-1 (a) Operation data of 9MW steam power plant

Time	CW FLOW	Temperature		Ambient Conditions			Condenser		
		CW - IN	CW - OUT	DBT	RH%	WBT	Vacuum	Temp	Flow
	(m^3/hr)	(ºC)	(ºC)	(ºC)	(%)	(ºC)	(Kg/cm^2 g)	(ºC)	(TPH)
10:30 AM	2350	29	40	26	57	19.9	-0.93	44.85	15.1
11:00 AM	2350	29	40	28	50	19.9	-0.93	44.96	15.3
11:30 AM	2371	29	40	28	49	19.8	-0.94	43.79	13.9
12:00 PM	2382	29	40	29	47	20.4	-0.94	44.11	14.2
12:30 PM	2383	28.5	40	29	46	20.3	-0.94	43.43	14.3
1:00 PM	2389	28.5	40	28	54	20.8	-0.94	42.13	10.5
1:30 PM	2392	28.5	40	28	54	21.1	-0.94	42.62	11.9
2:00 PM	2382	28.5	40	29	47	20.6	-0.94	43.56	15.1
2:30 PM	2387	28	39	28	46	19.9	-0.94	42.43	12.5

Table-1 (b) Operation data of 9MW steam power plant

Time	Inlet Steam			Extraction 1 & 2			Extraction 3		
	Pressure	Temp	Flow	Pressure	Temp	Flow	Pressure	Temp	Flow
	(Kg/cm^2 g)	(ºC)	TPH	(Kg/cm^2 g)	(ºC)	TPH	(Kg/cm^2 g)	(ºC)	TPH
10:30 AM	61.81	472	68.02	14.21	322.21	15.69	2.53	181.58	39.87
11:00 AM	61.64	473.6	68.90	14.16	324.0	15.46	2.52	183.65	39.94
11:30 AM	62.06	474.4	68.05	14.05	324.0	15.66	2.52	183.60	39.85
12:00 PM	62.31	473.9	67.87	14.03	321.91	14.87	2.52	182.44	39.22
12:30 PM	62.37	472.92	68.54	14.09	321.8	15.33	2.52	182.26	39.72
1:00 PM	61.57	472.96	70.74	14.27	321.8	18.46	2.52	179.84	43.57

1:30 PM	61.47	477.16	68.74	14.12	328.96	15.53	2.52	186.29	39.72
2:00 PM	61.81	474.09	68.30	14.05	319.08	14.63	2.52	180.03	38.55
2:30 PM	61.77	474.05	70.08	14.28	324.91	16.20	2.52	183.80	41.05

III. ENERGY AND EXERGY ANALYSIS OF NINE MW COAL BASED THERMAL POWER PLANT

All equations of analysis is adopted from fundamental approach of 1st law and 2nd law of thermodynamics, and properties (enthalpy,entropy,specific heat,specific volume,etc) of steam formation referred from steam property table(cited from www.engineeringtoolbox.com/www.spiraxsarco.com) as per steam condition (superheated and saturated) at different pressure-temperature level. PK Nag [1] and Shappiro [4] given energy-exergy equations for power-plant analysis-

$\eta_{PLANT} = W_{NET} / Q_{BOILER}$ ……..……(1)

$W_{NE}=W_{TURBINETOTAL}-W_{PUMP}$……..…..…(2)

$W_{TURBINETOTAL}=W_{TURBINE_1}+W_{TURBINE_2}+W_{TURBINE_3}$……..…..…(3)

$W_{TURBINE_1}$=m_dot_turbine1_steamflow(h_temp inlet_turbine1-h_temp exit_turbine1)……(4)

$W_{TURBINE_2}$=m_dot_turbine2_steam flow(h_tempinlet_turbine2-h_temp exit_turbine2)……(5)

$W_{TURBINE_3}$ = m_dot_turbine3_steamflow (h_temp inlet_turbine3- h_{2s})……..……..…(6)

From enthalpy and entropy balance equation-

$h_{2s}=h_{f2}+x_2h_{fg2}$…..(X2 is steam dryness fraction at condenser)……..……..……(7)

s1=s2s=sf2+x2 sfg2……...……..……(8)

h1=h_temp inlet_turbine3 = 2827.18 KJ/Kg….temp at 180 ^{0}C (inlet temp of turbine3)

s1=s2 = 7.353 kJ/kg-K from steam properties at condenser line pressure of 0.098 bar (absolute)

h3=hf3=hf2=184 kJ/kg from steam properties

s4=s3=sf2=0.629 kJ/kg-K

hfg2= 2397.95 kJ/kg

sfg2= 7.591 kJ/kg-K

Sp Volume of saturated steam at condenser line v=0.001009 m3/kg

Put all values of entropy and get x2=0.91 and estimate h2s from eq. no7

h2s=2376KJ/kg

Now total turbine work (mass flow rate of steam for different turbine provide in plant data table-1 (a)&(b),providing all values of enthalpy as per inlet temperatures Values of turbines)

$W_{TURBINE\ TOTAL}$ = 960 kJ/kg or 8.5 MW (flow rate of steam varying daily with time)

$W_{PUMP}=v\times dP=v\times(P_{BOILERLINE}-P_{CONDENSERLINE})$=m_dot_pump(h4shf3)……..,,……(9)

h4s = 184.06 kJ/kg

W_{PUMP}=184 kJ/kg or 437KW at 8.5TPH of water flow through pump

W_{NET} = 8.2-0.43 = 7.77 OR 7.8MW

Heat Generated by boiler

Q_{BOILER}=m_dot_boilersteamflow (h_boiler temp - h4s)……..……..……(10)

(Steam flow through boiler is 65TPH from data table)

Q_{BOILER} = 3196 kJ/Kg or 58627 KW

For $Q_{CONDENSER}$ …. $W_{TURBINE\ TOTAL} - W_{PUMP} = Q_{BOILER} - Q_{CONDENSER}$……..……..(11)

η_{PLANT}= W_{NET} / Q_{BOILER} = 24.12%

η_{Boiler}= Q_{BOILER} / Q_{FUEL}

Q_{FUEL} = (m_air + m_fuel)×fuel calorific value…....(4620 kcal/kg as per data)……..(12)

η_{Boiler} = 11%

(13 TPH of coal feeding with 97 TPH of air)

ExergyBalance Equation-

Exergy flow rate of inlet steam-

ϵ_{f_1n} = m_dot_steam × cp_boiler steam × To $[\frac{Tboiler}{To} - 1 - ln\frac{Tboiler}{TO}]$……..(13)

(Steam generation inlet temperature is 485±5^{0}C, 75TPH and specific heat of steam 2.21 kJ/kg-k, atmospheric temp To is 298K)

ϵ_{f_1n} = 387 kJ/Kg or 7715KW

Exergy flow rate of exhaust steam-

ϵ_{f_out} = m_dot_steam × cp_boiler steam × To $[\frac{Tboiler}{To} - 1 - ln\frac{Tboiler}{TO}]$……(14)

(Steam temperatureexhaust is 320 ^{0}C, 75TPH and specific heat of steam 2.21 kJ/kg-k, atmospheric temp To is 298K)

$€_{f_out}$= 260 KJ/Kg or 5240KW

Rate of exergy decrease- $€_{f_1n}$ - $€_{f_out}$= 127 KJ/Kg or 2475 KW……(15)

Rate of exergy loss during steam generation= $\frac{€f_out}{€f_in}$ = 0.47……..……(16)

Rate of exergy increase of steam = Exergyutilization rate= $€_{f_Usefull}$

$€_{f_Usefull}$ = m_dot_steam [h1-h4s-To (s1-s4s)]…….……(17)

$€_{f_Usefull}$= 1172.10 KJ/Kg or 23547.5KW

Exergy Destruction in boiler=$€_{Des_boiler}$ = Rate of exergy increase of steam - Rate of exergy decrease

$€_{Des_boiler}$=1045.1 KJ/Kg or 21072.5 KW

Exergy flow rate of wet steam to condenser

$€_{wetsteam\ to\ condenser}$ = $m_{_\ wet\ steam\ to\ condenser}$ [h2s-hf3-To (s_2-sf_3)]…….……...…….(18)

(26THP of wet steam passing through condenser)

$€_{wet\ steam\ to\ condenser}$ = 188.25 KJ/Kg or 1350.44 KW

Irreversibility estimation of each components-

I_{RR}= ToΣΔS = To (ΔS_{boiler}+ $\Delta S_{turbine}$+ $\Delta S_{condenser}$+ ΔS_{pump})……...……..……..(19)

And ΔS = cpln($\frac{Tout}{Tin}$) from 2nd law of thermodynamics……...……..……(20)

(cp will vary with temperature and pressure condition of working fluid, putting all value of cp and Tout&Tin of components)

I_{RR_Boiler}= To$(\Delta S)_{_boiler}$ = 1086 KJ/Kg

$I_{RR_turbine}$= To$(\Delta S)_{_turbine}$ = 419 KJ/Kg

$I_{RR_condenser}$= To$(\Delta S)_{_cond}$ = 124 KJ/Kg

I_{RR_pump}= To$(\Delta S)_{_pump}$ = 61.17 KJ/Kg

Now EDR ratio = $\frac{ToΣΔS}{Qboiler}$……..……..…..(21)

EDR ratio = 0.65

Exergetic efficiency = η_{EX} = 1- EDR_{ratio}……..……..……..…..(22)

□$_{EX}$ = **1- EDR_{ratio}= 0.35**

Exergy Destruction Rate = $\frac{ToΣΔS}{Wnet}$……(23)

Exergy Destruction Rate = 2.73

IV. RESULT AND DISCUSSION

The property of steam of parts of plant is taken from steam property chart with variation of pressure-temperature and saturated & superheated state of steam condition as per available data. All values of enthalpy, entropy and specific heat shown in table -2 to Table-5 respectively.

Table-2 Boiler Observation

Pressure (Kg/cm^2 gauge)	Temp (ºC)	Flow (Kg/s)	Enthalpy (kJ/Kg)	Entropy (kJ/Kg-K)	Sp Heat (kJ/Kg-K)
61.81	472	18.89	3353.1	6.7778	2.418
61.64	473.6	19.10	3357.32	6.7848	2.415
62.06	474.4	18.9	3358.73	6.7838	2.416
62.31	473.9	18.85	3337.13	6.7796	2.418
62.37	472.92	19.05	3354.68	6.7759	2.42
61.57	472.96	19.65	3355.9	6.7836	2.415
61.47	477.16	19.09	3366.17	6.7981	2.409
61.81	474.09	18.97	3358.23	6.7841	2.415
61.77	474.05	19.46	3358.7	6.7851	2.415
61.87	473.90	19.11	3357.69	6.7833	2.416

Table-3 Condenser observation

Vacuum Pr	Temp	Flow	Enthalpy (KJ/Kg)			Entropy (KJ/Kg-k)			Sp Heat
(Kg/cm^2 gauge)	(°C)	(kg/s)	h_f	h_{gf}	h_g	S_f	S_{fg}	S_g	(kJ/Kg-K}
-0.934	44.85	4.2	164.9	1866.9	2031.8	0.554	5.772	6.326	1.505
-0.933	44.96	4.25	166.6	1885.9	2052.6	0.559	5.831	6.391	1.52
-0.937	43.79	3.35	159.9	1809.7	1969.6	0.553	5.595	6.133	1.459
-0.936	44.11	3.94	161.6	1828.8	1990.4	0.542	5.654	6.197	1.474
-0.939	43.43	3.98	148.1	1771.1	1928.2	0.526	5.478	6.004	1.351
-0.944	42.13	2.92	153.1	1676.1	1824.5	0.497	5.183	5.68	1.397
-0.941	42.62	3.3	159.1	1733.5	1886.7	0.514	5.36	5.874	1.549
-0.937	43.56	4.2	151.5	1809.5	1969.6	0.537	5.595	6.133	1.382
-0.942	42.43	3.48	154.8	1714.5	1866	0.509	5.301	5.18	1.412
-0.94	43.54	3.74	154.8	1752.6	1907.4	0.52	5.419	5.939	1.551

Table-4 Turbine 1& 2 Observation

Pressure	Temp	Flow	Enthalpy	Entropy	Sp Heat
(Kg/cm^2 gauge)	(°C)	(Kg/s)	(kJ/Kg)	(kJ/Kg-K)	(kJ/Kg-K)
14.21	**322.21**	**4.35**	**3086.33**	**7.0028**	**2.202**
14.16	**324.0**	**4.29**	**3090.39**	**7.0111**	**2.2**
14.05	**324.0**	**4.35**	**3080.65**	**7.0148**	**2.116**
14.03	**321.91**	**4.15**	**3086.14**	**7.007**	**2.2**
14.09	**321.8**	**4.25**	**3085.7**	**7.005**	**2.201**
14.27	**321.8**	**5.12**	**3085.28**	**6.999**	**2.204**
14.12	**328.96**	**4.31**	**3101.38**	**7.03**	**2.195**
14.05	**319.08**	**4.06**	**3079.82**	**6.996**	**2.203**
14.28	**324.91**	**4.50**	**3092.11**	**7.01**	**2.201**
14.14	**323.19**	**4.38**	**3088.65**	**7.008**	**2.2**

Table-5 Turbine-3 Observation

Pressure	Temp	Mass Flow Rate	Enthalpy	Entropy	Sp Heat
(Kg/cm^2 gauge)	(°C)	Kg/s	kJ/Kg	kJ/Kg-K	kJ/Kg-K
2.53	**181.58**	**11.07**	**2824.18**	**7.153**	**2.097**
2.52	**183.65**	**11.09**	**2828.67**	**7.164**	**2.093**
2.52	**183.60**	**11.06**	**2828.57**	**7.164**	**2.093**
2.52	**182.44**	**11.03**	**2826.14**	**7.159**	**2.095**
2.52	**182.26**	**11.03**	**2825.76**	**7.158**	**2.095**
2.52	**179.84**	**12.1**	**2820.69**	**7.147**	**2.099**
2.52	**186.29**	**11.03**	**2834**	**7.178**	**2.089**
2.52	**180.03**	**10.7**	**2821.09**	**7.147**	**2.099**
2.52	**183.80**	**11.4**	**2822.7**	**7.151**	**2.098**
2.52	**182.61**	**11.17**	**2826.5**	**7.159**	**2.095**

The numerical computation of nine MW coal based thermal power plant is carried out and performance parameters in terms of plant efficiency boiler efficiency, exergetic efficiency of plant and irreversibilities in terms of exergy destruction in various components, exergy destruction ratio of plant is shown in table-respectively.

Table-6 Parametric Result

	Thermodynamics Parameters	Computed Results
	Net Out Put (Wnet)	7.8MW
	Total Turbine Out Put (Wturbine)	8.5 MW
	Boiler Heat Generation (Qboiler)	3196KJ/KG
	Plant Efficiency (η_{PLANT})	24.12&
	Boiler Efficiency (η_{BOILER})	11%
	Plant Exergy Destruction Rate	2.17
	Plant Exergetic Efficiency (η_{EXERGY})	35%

V. CONCLUSIONS & RECOMMENDATION

The exergy analysis plays deterministic role in identifying the magnitude of thermodynamic processes , cause of losses and rectifying the exergy destruction rate in the components due to energy saving in components (i.e.boilers, condensers in steam power plant)

1. The maximum exergy destruction(irreversibility) as 34%.
2. Exergy Destruction is found to be 2.17 (which is the based on output analysis) and o.65 based on input.
3. The plant efficiency is 24.12%
4. The second law efficiency is 35%

ACKNOWLWDGEMENT

We express warm thanks to MrBhawin, EnegryAuditor,DarshanInstitue,Badodra,Gujrat,India, who provide all technical facilities related to plant analysis.

REFERENCES

[1]. P.K Nag,"Power Plant Engineering"Tata McGraw Hill,New Delhi,2007.

[2]. M.A.Rosen and I. Dincer"Exergoeconomic analysis of power plants operating on various fuels"Applied Thermal Engineering,23,pp 643-658,2003.

[3]. A.Bejan "Fundamentals of ExergyAnalysis,EntropyGeneration,Minimization and Generation of Flow Architecture" International Journal of Energy Research,Vol,26,No.7,2002,pp545-565.

[4]. MichealJ.Moran and Howard N.Shapiro "Fundamentals of Engineering Thermodynamics" 6thedition,SI Version, John Willey & Sons,Corporation,2010.

[5]. Khaliq and S.C Kaushik "Second-law based thermodynamic analysis of brayton/rankine combined power cycle with reheat" Applied Energy, Vol 78,No.2,pp 179-197,2004.

[6]. Y.Yang,L.Wang,C.Dong,G.Xu,T.Morousk and G.Tsatsaronis "Comprehensive exergy based evaluation and parametric study of a coal-fired ultra supercritical powerplant".Applied Energy,2013.

[7]. P.AhmadiandI.Dincer"Thermodynamic analysis and Thermoeconomic optimization of a dual pressure combined cycle power plant with supplementary firing unit"Energy Conversion Management,52,pp 2296-308,2011

[8]. S.C Kaushik,V.S.Reddy and S.K Tyagi "Energy and Exergy analysis of thermal power plants-A Review"Renewable and Sustainable Energy Reviews,15,pp 1857-1872,201.

[9]. I.H. Aljundi"Energy and exergy analysis of a steam power plant in Jordan" Applied Thermal Engineering,29,pp 324-328,2009.

[10]. Sarang J Gulhane,Amitkr Thakur "Exergy analysis of boiler in cogeneration thermal power plant"Americal journal of engineering research,e-issn-2320-0847,p-issn-2320-0936,vol-02,issue-10,pp-385-392,2013.

[11]. Christian O.Osueke, Anthony O. Onokwai and AdeyinkaO.Adeoye "Energy and exergy analysis of 75 MW steam power plant in Sapele (Nigeria) "International journal of innovative research in advanced engineering.ISSN-2349-2163,vol-2,issue-6,pp-169-179,2015

[12]. Ravindra kumar "A critical review on energy,exergy,exergonomics and economic (4-E) analysis of thermal power plants" Engineering Science and Technology, an International Journal,Elsevier,issn-2215-0986,pp 283-292, Feb 2017.

Author's Information.

Prof.(Dr.) R.S. Mishra Ph.D-SOLAR ENERGY IIT DELHI-1986., HOD Mechanical Engg in Delhi



Technological University,Delhi,India . He has published more than 250 research papers in repute journals & conferences and also written 55 books on ASTROLOGY, VADIC Sciences, FENG SHUAI. Professor Mishra is dynamic personality of area ofsolar thermal system design-development and advanced refrigeration techniques.

Kaushalendra K Dubey is associated with is the Dept of Mechanical Engineering, Sharda University, Greater



Noida, India as assistant professor. He is a research fellow of Delhi Technological University, Delhi, India. Author qualification is Master of technology in Thermal Engineering with honor's from JamiaMilliaIslamia, Delhi, India..He has faculty research fellowship from Indian Institute of Technology,Delhi in 2016 in Energy and Heat Recovery Research. Author has published more than 25 research papers in repute journals and conferences and active volunteer for work on energy research and technology through industrial and academic project

International Journal of Research in Engineering and Innovation (IJREI)
journal home page: http://www.ijrei.com
ISSN (Online): 2456-6934

Thermodynamic study of R134a in Vapour Compression Refrigeration System in Summer Climate

Kaushalendra Kumar Dubey[1] R.S Mishra[2]

Department of Mechanical Engineering, Delhi Technological University, Delhi-110042
Email id.: professor_rsmishra@yahoo.co.in, dubey.kaushalendra@gmail.com

Abstract

Thermodynamic analysis of an ideal vapour compression refrigeration system is done using refrigerants R134a.This Energy-Exergy analysis is carried out by developed mathematical model. Various parameters are numerically computed are first law efficiency in terms of coefficient of performance (COP) of the system both ideal case and actual experimental and exergetic efficiency and exergy destruction ratio have been computed . The eco-friendly R134a refrigerant gives lower COP than R22. Comparing with R22, R134a takes more compressor power and second law efficiency (exergetic efficiency) of VCRS is also lower than using R22. Due to higher GWP and ODP of R22, R134a can be used for domestic / commercial and industrial applications.
© 2017 ijrei.com. All rights reserved

Keywords: Hydrofluorocarbon refrigerants, ozone-friendly, performance characteristics, ODP, GWP,

1. Introduction

Refrigerant is a substance which is used as a working substance in refrigeration system to move heat from lower temperature to higher temperature to get cooling effect. Midgley discovered halogenated hydrocarbon refrigerants having favourable thermodynamic properties. But these halogenated hydrocarbon refrigerants attack the atmospheric stratosphere and destroy ozone layer. With the discovery of ozone hole in stratosphere as stated by Molina and Rowland et al [1] and the Montreal protocols (Paul, *et al* [2]), the CFC and HCFC refrigerants are to be phased out due to their higher global worming potential (GWP) and higher ozone depletion potential (ODP). For these reasons, recently ODP and GWP plays a vital role in the development of new environment friendly alternative refrigerants other than CFC and HCFC refrigerants for their higher ODP and GWP. So, develop countries stopped production of CFC and HCFC refrigerants and looking for alternative environment friendly refrigerants. So, HC and HFC refrigerants with zero ODP and low GWP are considered for long term purpose. Although the ODP of some HFCs is zero, their GWP related to the greenhouse investigated on domestic refrigerator using propane and iso-butane as a refrigerant mixture and found that mass effect is large. On the other hand, HC refrigerants have a flammability issue, which restricts the usage in existing systems. However this flammability issue can be avoided by blending HC refrigerants with HFC refrigerants. It is also found that an HC/HFC mixture makes a very good solution with mineral oil and contribution to global warming of HC/HFC mixture is very low due to very lowGWP (about one third of HFC) and hence R134a is found to be most suitable alternative refrigerant for R12. Refrigerant R134a has very similar thermodynamic properties such as molecular weight, critical temperature, boiling point as R12, with zero ODP and less GWP as compared to CFC12 are shown in table 1 (Calm and Hourahan, et al [3]). This gave confidence to the researchers to consider R134a as a suitable replacement to R12 for short term basis. Hammad and Alsaad et al [4] investigated on domestic refrigerator using LPG (24.4% propane, 56.4% butane and 17.2 % iso-butane) as refrigerant. They concluded that this is the environment friendly refrigerant mixture and that can replace R12 in domestic refrigerator. Jung et al. (Jung, *et al*, 1996)
fraction range of 0.2 to 0.6 of propane increase the COP up to 2.3% compared to R12. Mani and Selladurai (Mani and

Corresponding author: Kaushalendra Kumar Dubey[1]
Email Address: dubey.kaushalendra@gmail.com

Selladurai, 2008) experimented on vapour compression refrigeration system using new refrigerant mixture of propane and iso-butane for substitution of R12 and R134a. They found that this mixture had a refrigeration capacity 19.9% to 50.1 % higher than R12 and 28.6% to 87.2% higher than R134a and COP improved by 3.9% to 25.1% than R12 at lower evaporator temperature and 11.8% to 17.6% at higher temperature. Chen and Prasad (Chen and Prasad, 1999) theoretically analysed vapour compression refrigeration system using R134a and R12 as refrigerants and reported that the COP for R134a is slightly (3%) lower than the R12.Richardson and Butterworth (Richardson and Butterworth, 1995) experimentally investigated the performance of hydrocarbon refrigerants in a hermetic vapour compression refrigeration system and came to the conclusion that propane and propane/isobutene mixtures can be used in an unmodified R12 system and can achieved better COPs than R12 under the same operating conditions. They found that around 50% propane and 50% isobutene mixture not only just give the similar saturation characteristics but also gives a better COP when proportion of propane is increased. They also mentioned that hydrocarbon refrigerants have a flammability issue, they can be safely used in hermetic vapour compression refrigeration

The aim of this study is to investigate the performance of vapour compression refrigeration systems using R12, R125, R134a, R143a and R152a as refrigerants based on energy and exergy concept. Various parameters like COP, refrigerating capacity, compressor work, cycle efficiency, exergy efficiency are computed and compared in this work.

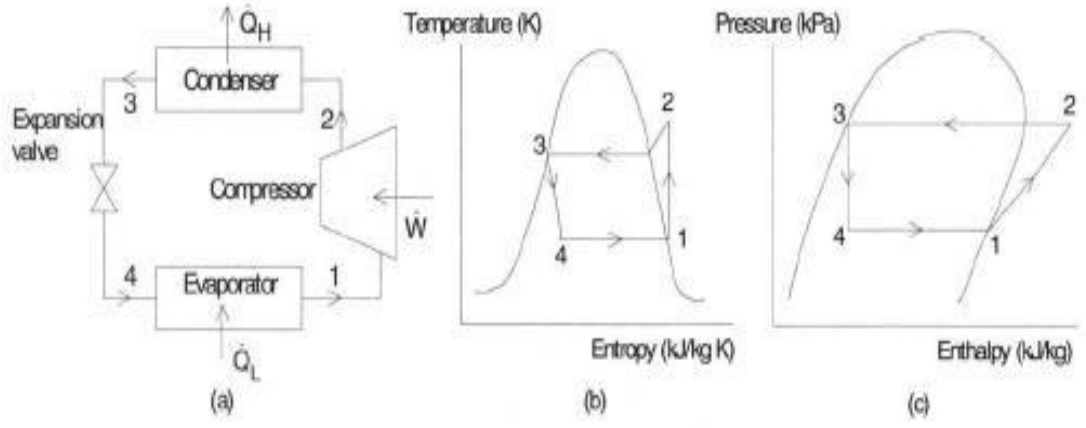


Figure 1 (a) A basic Vapor Compression refrigeration Cycle. (b) T-S diagram (c) its P-H diagram. [1]

Vapor compression refrigeration cycle on p-h diagram. The p-h diagram is frequently used in the analysis of vapor compression refrigeration cycle and usually consists of the four processes.

Process 1-2 is the compression.
Process 2-3 is the Condensation.
Process 3-4 is the expansion.
Process 4-1 is the evaporation.

T_1 and T_2 are the suction and discharge temperatures.
P_e and P_c are the evaporating and condensing pressures.
The different enthalpies involved in the cycle are,
h_1: enthalpy of suction pipe or end of evaporation.
h_2: end of suction pipe or end of compressor.
h_3: inlet of throttling valve or end on heat exchanger.
h_4: inlet of evaporation or end of expansion.

1.1 Thermal Analysis of Vapour Compression Refrigeration System

Figure 1 shows the p-h diagram of a complete ideal vapour compression refrigeration system. Various calculations are done based on this system using different refrigerants. COP of vapour compression refrigeration system is a very important criterion for performance analysis. It represents the refrigeration effect per unit compressor work.

COP_{VCRS} = Refrigeration Effect/ Compressor Work= R_E/W

Where R_E is the refrigeration effect and W is the compressor work.

The refrigerating effect, RE can be expressed as

$R_E = \dot{m} (h_1 - h_4)$

The compressor work, W can be expressed as

$W = \dot{m} (h_2 - h_1)$

Where $(h_2 - h_1)$ is the difference of enthalpy which includes the effect of compressor efficiency and $\dot{m}$is the mass flow rate of the refrigerants which can be expressed as

$\dot{m} = RE/ (h_1 – h_4)$

Where $(h_1–h_4)$ is the difference of enthalpy in the cycle. The exergetic efficiency, η_{II} is computed in terms of second law efficiency , which is a the ratio of COP of VCRS to the COP of Carnot cycle, then

$$\eta_{II} = COP_{VCRS}/ COP_{Carnot} = \frac{\frac{R_E}{W}}{\frac{T_e}{T_c - T_e}}$$

and $COP_{Carnot} = \frac{T_e}{T_c - T_e}$

Where T_e is evaporator temperature and Tc is condenser temperature

(The COP of Carnot refrigeration cycle is a function of evaporator and condenser temperatures only and is independent of the nature of the working substance. The Carnot COP sets an upper limit for refrigeration systems operating between two constant temperature thermal reservoirs (heat source and sink). From Carnot's theorems, for the same heat source and sink temperatures, no irreversible cycle can have COP higher than that of Carnot COP.)

The cycle efficiency (also called as second law efficiency) is a good indication of the deviation of the standard VCRS cycle from Carnot cycle. Unlike Carnot COP, the cycle efficiency depends very much on the shape of T s diagram, which in turn depends on the nature of the working fluid.

Exergy or availability of a system is the maximum obtainable work output from the system. So exergy loss is the very important criterion to evaluate the thermodynamic performance of the vapour compression refrigeration system. System performance improved if exergy loss is less. For that reason it is our aim to minimize the exergy loss to improve the system thermodynamically. The exergy can be expressed following Chen and Prasad (Chen and Prasad, 1999) as

$e = (h-h_0) –(s-s_0)$

where kinetic and potential energy are excluded.

The exergy loss can be calculated by calculating exergy loss in each component of the system. Exergy loss in compressor can be expressed as

$\Delta e_w=(h_1-h_2)+T_0(s_2-s_1)+W$

Exergy loss in condenser can be expressed as

$\Delta e_c= (h_2-h_3)+T_0(s_3-s_2)$

Exergy loss in expansion valve can be expressed as

$\Delta e_v= (h_3-h_4)+T_0(s_4-s_3)$

Exergy loss in evaporator can be expressed as

$\Delta e_w= (h_4-h_1)(T_0/T_r)+T_0(s_1-s_4)$

And the total exergy loss can be expressed as

$\Delta E = \dot{m}(\Delta e_w +\Delta e_c+\Delta e_v+\Delta e_w)$

Exergetic efficiency can be expressed as

$$\eta_x = \frac{RE\left|\left(1-\frac{To}{Tr}\right)\right|}{W}$$

1.2 Properties and Environmental Impacts of Selected Alternative Refrigerants

The vapour compression refrigeration system operates on refrigerant R12 which is most important CFC refrigerant identified for phase out in the country by HFC refrigerant. The institute has been one of the few institutes selected for funding for research work under World Bank project TEQIP. A research work was undertaken for finding the most suitable ozone-friendly Hydrofluorocarbon refrigerant for replacing ozone depleting refrigerant R12. Then most suitable refrigerant has been selected for replacing harmful refrigerant. Four non-ozone depleting HFC refrigerants (R125, R134a, R143a and R152a) were selected from methane and ethane derivatives and their performances in vapour compression refrigeration system were investigated. Properties and environmental impacts of selected zero ozone-depleting refrigerants used as substitutes in the refrigeration systems are listed in the Table (1a) to Table-1(c) [3-6] respectively.

Table-1(a) : Some properties and environmental impacts of selected alternative refrigerants.[3][5][6]

Refrigerants	**Chemical formula**	**Molecular mass**	**Freezing point (°C)**
R12	CF_2CL2	121	-157.5
R125	$C_2\ HF_5$	120	-102.99
R134a	$C_2H_2F_4$	102	-96.67
R143a	$C_2H_3F_4$	84	-111

Table-1(b) : Some properties and environmental impacts of selected alternative refrigerants.[3][5][6]

Refrigerants	**Chemical formula**	**Boiling point (°C)**	**Critical point (°C)**
R12	CF_2CL2	-29.8	112
R125	$C_2\ HF_5$	-48.1	66.18
R134a	$C_2H_2F_4$	-26.1	101.06
R143a	$C_2H_3F_4$	-47.2	72.89
R152a	$C_2H_4F_2$	-24.0	113.26

Table-1(c): Some environmental properties of selected alternative refrigerants [3][5][6.

Refrigerants	Ozone Depletion Potential (ODP)	Global Warming Potential (GWP)
R12	1	8100
R125	0	2800
R134a	0	1300
R143a	0	3800
R152a	0	140

2. Results and Discussions

Experiment was conducted in the laboratory for obtaining thermal performance using eco-friendly R134a, refrigerant. The following data have been recorded.
Atmospheric pressure: 1 bar
Atmospheric temperature: T_0 =30°C =303 K
Condenser pressure: 135psi
Evaporator pressure: 8psi
Compressor delivery temperature: 61.9°C
Compressor suction temperature: -12.7°C
Ref. leaving condenser temperature: 29.6°C
Evaporator inlet temperature:-16.7°C
Mass flow rate: 5 kg/min
Evaporator volts: V_e/V = 196
Evaporator AMPS: I_e/n =2.24
Motor volts: V_m/V =198
Motor amps I_m/A =2.892
Motor compressor speed: 2800rpm
The Enthalpy and Entropy have been computed on above state temperature data as shown in the table-2.

Table-2 Enthalpy and Entropy of various components of VCRS

States	***Temperature in ^0C***	***Enthalpy in KJ/Kg***	***Entropy in KJ/Kg-K***
Outlet of evaporator-1	-12.7	391	1.7355
Compressor outlet-2	61.9	427	1.7011
Condenser outlet-3	29.6	240	1.1437
Evaporator Inlet-4	-16.7	178.6	0.9179

- Coefficient of performance
  $COP_{ideal\ cycle} = (h_1-h_4)/((h_2-h_1)/0.8) = 4.719$
  (Assuming 80% volumetric efficiency of compressor)
  $COP_{Actual} = Q_e/Power = V\times I/Power = V\times I/(Motor\ volts \times Motor\ amps \times Cos\varnothing) = 0.983$
  $\eta_{JItheoritical} = (COP_{Actual} / COP_{ideal\ cycle}) = 0.2083$
- Exergy loss in each component of the system
  Exergy loss in compressor
  $\Delta e_w = (h_1-h_2) + T_0(s_2-s_1) + W$
  Δe_w=133.5768 KJ/KG

  Exergy loss in condenser
  $\Delta e_c = (h_2-h_3)+T_0(s_3-s_2)$
  Δe_c= 18.1078 KJ/KG

  Exergy loss in expansion valve
  $\Delta e_v = (h_3-h_4)+T_0(s_4-s_3)$
  Δe_v=-7.0274KJ/KG

  Exergy loss in evaporator
  $\Delta e_w = (h_4-h_1)(T_0/T_2)+T_0(s_1-s_4)$
  Δe_w=-3.3565KJ/KG

  And the total exergy loss
  $\Delta E = \dot{m}(\Delta e_w + \Delta e_c + \Delta e_v + \Delta e_w)$
  ΔE = 706.5035 KJ/KG
- Exergectic Efficiency (η_x)

$$\eta_x = \frac{RE\left|\left(1-\frac{To}{T2}\right)\right|}{W}$$

$\eta_{_II_exp}$=0.1791= 17.91%

3. Conclusion

The experimental work presented in this paper on vapour compression refrigeration system, the following conclusions were made.

1. The eco-friendly R134a refrigerant gives lower COP than R22.
2. Comparing with R22, R134a takes more compressor power.
3. Second law efficiency (Exergetic efficiency of VCRS is also lower than using R22.
4. Due to higher GWP and ODP of R22, R134a can be used for domestic / commercial and industrial applications.

Acknowledgement

Authors) thanks to my respected teachera Prof (Dr.) Vijay Gupta, Vice-Chancellor, Sharda University, Greater Noida and Prof. N K Bansal, Ex-VC of MVDU katra for guidance.

References

[1] M.J.Molina, FS.Rowland, (1974) Stratospheric sink for chlorofluoromethane: chlorine atom catalyzed destruction of ozone, Nature, 249,810–812.

[2] S.Paul, A.Sarkar, B.K.Mandal, (2013) Environmental impacts of halogenated refrigerants and their alternatives: recent developments, International Journal of Emerging Technology and Advanced Engineering, 3: Special Issue 3: ICERTSD 2013:400-409.

[3] J.M.Calm, G.C.Hourahan, (2001) Refrigerant data summery, Engineering Systems, 18,74–88.

[4] M.A.Hammad, M.A.Alsaad, (1999) The use of hydrocarbon mixtures as refrigerants in domestic refrigerators, Applied Thermal Engineering, 19,1181–1189.

[5] Bhatti, M.S. (1999). A historical look of chlorofluorocarbonrefrigerants. ASHRAE Transactions, part 1, 1999.pp.1186- 1206

[6] Cavallini A, Working fluids for mechanical refrigeration. International Journal of Refrigeration 19 (1996) 485-496.

[7] McMullan JT, Refrigeration and the environment issues and strategies for the future, International Journal of Refrigeration 25 (2002) 89-99.

JOURNAL OF Applied Polymer SCIENCE

Understanding the role of internal microstructure in capsule-based healing of polymeric composites

Manorama Tripathi,[1,2] **Devendra Kumar**,[2] **Prasun Kumar Roy**[1]
[1]Centre for Fire, Explosive and Environment Safety, DRDO, Delhi 110054, India
[2]Department of Applied Chemistry and Polymer Technology, Delhi Technological University, Delhi 110042, India
Correspondence to: M. Tripathi (E-mail: tripathim07@gmail.com or manorama@cfees.drdo.in)

ABSTRACT: The efficacy of microcapsules possessing varied internal microstructure toward introducing extrinsic self-healing functionality in epoxy has been compared. Unsaturated polyester was encapsulated in microcapsules by adopting different methodologies. Microcapsules formed using dispersion polymerization exhibited "reservoir" microstructure while solvent evaporation led to the formation of "monolithic" microcapsules. Theoretical model was developed to predict the amount of healant released in the event of microcapsule rupture, which clearly highlighted the benefits associated with reservoir type microcapsules, especially at lower core contents. At larger core contents (≥50% vol/vol), all the micro-droplets within the monolithic structure coalesced to form a healant reservoir. Self-healing composites were prepared by introducing both types of microcapsules in an epoxy matrix and the healing efficiency was quantified. In line with the theoretical predictions, reservoir type microcapsules led to much higher healing efficiencies in comparison to monolithic microcapsules. © 2017 Wiley Periodicals, Inc. J. Appl. Polym. Sci. **2017**, *134*, 45471.

KEYWORDS: composites; theory and modeling; thermosets

Received 12 April 2017; accepted 25 June 2017
DOI: 10.1002/app.45471

INTRODUCTION

Self-healing polymers belong to the category of smart responsive materials, which possess an inbuilt ability to recover their properties. These materials are finding increasing usage as products in wide class of industries like coatings, electronics, energy, and transportation.

What renders this approach particularly attractive is envisioning the repair process in the absence of the crack detection step. In this context, one of the approaches of extrinsic self-healing is particularly interesting, where the healing agent is dispersed within the polymeric matrix in different containment structures i.e., hollow glass tubes,[1] micro-vascular networks,[2] and microcapsules,[3–6] with the latter being most extensively studied. It is to be noted, however, that at increased loadings of microcapsules, the composite may exhibit inferior mechanical properties, which can actually reduce the lifetimes. The process of repair is initiated by the rupture of containment structure resulting from crack propagation. The mending monomer is released into the crack plane, and subsequently undergoes polymerization to re-establish its structural integrity.

Considering a microcapsule-based healing system, the method of encapsulation defines its internal microstructure, which in turn affects the self-healing efficiency. Conventionally, healant encapsulated microcapsules are prepared using dispersion polymerization route wherein the hydrophobic healant remains encapsulated within a spherical fragile shell. During the dispersion polymerization process, the shell wall constituents undergo condensation reaction within the aqueous phase to form a low-molecular-weight pre-polymer, which deposit preferentially over the healant on the oil–water interface i.e., urea or melamine with formaldehyde. Eventual crosslinking within the chains forms a spherical solid wall around a liquid core forming a "reservoir" of healant.

Another relatively less complex methodology, which almost routinely employed for the encapsulation of drugs for pharmaceutical applications, is "solvent evaporation."[7] Interestingly, the use of this process for healant encapsulation is not so popular, but has garnered a lot of attention lately.[8–12] It is to be noted that for the purpose of healant encapsulation, the essential criteria reduces to the compatibility of the fragile microcapsule shell wall with the base resin, and other features like biocompatibility and degradability can be dispensed, which are an absolute necessity for pharmaceutical applications. Solvent evaporation leads to the formation of a monolithic structure, where the

Additional Supporting Information may be found in the online version of this article.
© 2017 Wiley Periodicals, Inc.

microdroplets of healants are dispersed within the shell homogenously.

For healing purposes, only the resin available on the cross-sectional ruptured plane is allowed to flow out. It is clear that the extent of release of healant is primarily defined by the internal morphology of the microcapsule, surprisingly, a systematic study in this context is lacking. Theoretical models for healant delivery exist for rupture of a reservoir-type microcapsule,[13] however, to the best of our knowledge, such models for a monolithic type geometry have not yet been developed, which instigated us to take up this study. In the present article, we discuss the role of the microcapsule internal structure toward healant release for imparting self-healing efficiency.

For providing an experimental validation of the model, microcapsules encapsulating unsaturated polyester with varied internal microstructures were prepared and dispersed within an epoxy resin to prepare healable formulations. The healing efficiency was quantified as the ratio of impact strength before and after healing and our results clearly highlighted the advantages associated with "reservoir" type microcapsules, especially when the core contents are low.

EXPERIMENTAL

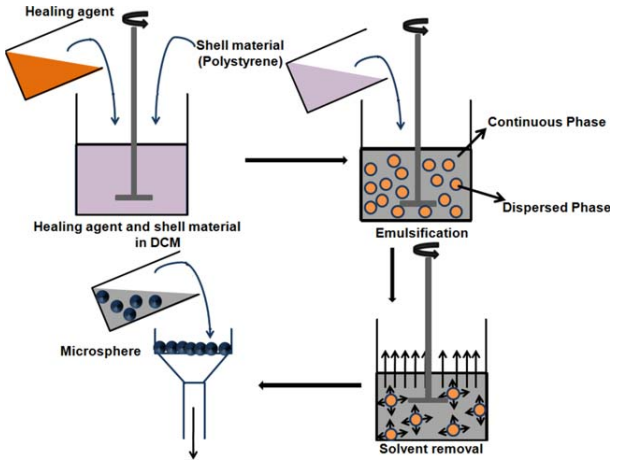
Materials

Aliphatic epoxy resin [Ciba Geigy, (Basel, Switzerland) Araldite CY 230; epoxy equivalent 200 eq g^{-1}] and triethylenetetramine (TETA)-based hardener (HY 951; amine content 24 eq kg^{-1}) was used as received. Sodium lauryl sulfate, styrene, potassium persulfate, and sodium metabisulfite were of "AR" grade, E. Merck (Hohenbrunn, Germany). Urea, formalin (37% formaldehyde in water), sodium hydroxide, resorcinol, ammonium chloride, dichloromethane, and poly(vinyl alcohol) (PVA) (M_w ~14,000) was obtained from CDH, Mumbai, India. EMA (Sigma Aldrich, St. Louis, USA) and ethyl acetate (E. Merck) was used without any further purification. Isophthalic acid-based unsaturated polyester resin (USP RPL211) with a styrene content of 32% was purchased from Revex (Delhi, India). Double distilled water was used throughout the course of this work.

Microencapsulation of USP

USP was encapsulated in polystyrene (PS) microcapsules by solvent evaporation technique using the procedure reported in the literature,[9] a general schematic of which is presented in Scheme 1 PS, the shell wall material, was prepared using emulsion polymerization as per the procedure discussed previously.[14] The procedure adopted along with the detailed characterization results are presented in Supporting Information Figure S1. Predetermined amount of USP and PS was dissolved in dichloromethane (30 mL), which was added to 50 mL of aqueous PVA solution (2.5% wt/wt) under agitation at room temperature to produce an oil/water emulsion. The resulting mixture was then poured into 150 mL aqueous PVA solution with continuous agitation. After evaporation of the solvent, the reaction mixture was cooled, and the USP encapsulated PS microcapsules were filtered and washed repeatedly with water followed by drying under vacuum.

The procedure employed for encapsulation of USP in poly(urea-formaldehyde) shells has been detailed previously.[15] In brief, an



Scheme 1. A pictorial representation of the solvent evaporation process for microcapsule preparation. [Color figure can be viewed at wileyonlinelibrary.com]

emulsifier solution was prepared by mixing EMA solution (2.5% wt/vol) with water at 70–80 °C (~450 rpm) using a magnetic stirrer. The solution was diluted to 0.5% wt/vol following which urea, ammonium chloride, and resorcinol was added, which underwent condensation polymerization to form the shell wall. After complete dissolution, the pH was adjusted to 3.2 ± 0.2. Subsequently, a diluted solution of healing agent solution (USP:ethyl acetate:: 1:1 wt/wt) was slowly introduced while maintaining a constant stirring speed. Thereafter, requisite amount of formalin was added and the temperature was increased to 50–55 °C, without altering the rate of agitation. The reaction was allowed to proceed for extended periods, followed by filtration and air-drying of the microcapsules, which resulted in the formation of free-flowing powder.

Preparation of Self-Healing Composites

Epoxy composites containing varying amounts of microcapsules (5%–25% wt/wt) were prepared by room temperature curing of epoxy with TETA. The microcapsules along with AIBN [2,2′—azobis(2—methylpropionitrile)] (1% wt/wt microcapsule) were dispersed in epoxy resin and ultrasonicated to obtain homogeneous slurry. This was followed by the addition of stoichiometric amount of TETA hardener (13% wt/wt). After degassing under vacuum, the suspension was transferred to greased silicone molds and allowed to cure at 30 °C for 24 h. For the purpose of comparison, neat epoxy specimens were also prepared.

CHARACTERIZATION

The surface morphology of USP encapsulated microcapsules was studied using a scanning electron microscope (Zeiss EVO MA15) under an acceleration voltage of 1 kV. Samples were mounted on aluminum stubs and sputter-coated with gold and palladium (10 nm) using a sputter coater (Quorum SC7620) operating at 10–12 mA for 120 s. Average microcapsule size and standard deviations were determined by measuring the diameters of about 100 microcapsules using ImageJ software. Optical images of suspended micro-droplets were captured using optical microscope (MOTIC B3–223PL). Structural characterization of

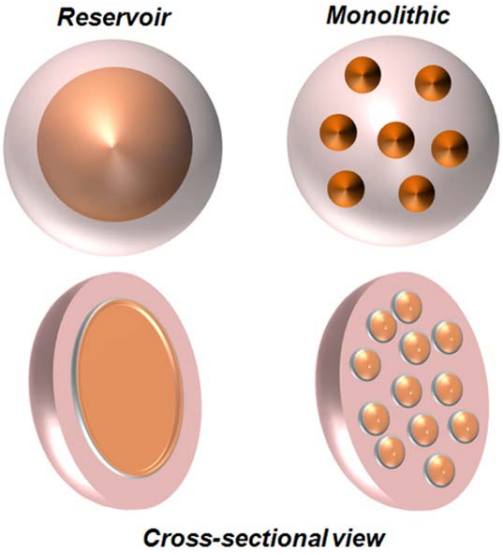


Figure 1. Pictorial representation of reservoir type and monolith type cross-section. [Color figure can be viewed at wileyonlinelibrary.com]

microcapsules was performed by recording Fourier transform infrared (FTIR) spectra in the wavelength range 600–4000 cm^{-1} using a Thermo scientific FTIR (NICOLET 8700) analyzer with an attenuated total reflectance (ATR) crystal accessory. Perkin Elmer Diamond simultaneous thermo gravimetry-differential thermal analyzer (STG-DTA) was used to investigate the thermal degradation behavior of the microcapsules under N_2 atmosphere in the temperature range 50–800 °C at a heating rate of 10 °C/min. A sample mass of 5.0 ± 0.5 mg was used for each experiment. Calorimetric studies were performed on a differential scanning calorimeter (TA instruments Q 20, Hertfordshire, UK). For dynamic differential scanning calorimetry (DSC) scans, samples (10 ± 2 mg) were sealed in aluminum pans, and heated from 50 to 200 °C at 10 °C/min. All experiments were performed under flowing N_2 (50 mL/min) to minimize oxidative degradation of the sample. Prior to the experiments, the instrument was calibrated for temperature and enthalpy using standard indium and zinc. Thermal equilibrium was regained within 1 min of sample insertion, and the exothermic reaction was considered to be complete when the recorder signal leveled off to the baseline. The total area under the exothermic curve was determined to quantify the heat of curing.

Self-healing efficiency was determined as the ratio of impact strength of healed and virgin sample as per the reported procedure.[16]The Izod impact strength of the as-prepared specimens (63.5 mm × 12.5 mm × 3.0 mm) was determined as per ASTM D 256 using an impact strength testing machine (International Equipments, India). Post-fracture, both the faces were aligned and gently pressed by hand for a few seconds and clamped to ensure close contact for a period of 6 h at 110 °C. Photographs of the type of clamped samples have been shown in our previous article.[15] The healed specimens were re-subjected to impact tests after 24 h and the self-healing efficiency was quantified. A control set of experiments was performed with an aim to quantify the intrinsic healing ability of neat epoxy in the absence of microcapsules. For this purpose, fractured epoxy samples were exposed to conditions mentioned previously, and the impact strength both before and after fracture was compared. For each composition, at least five identical specimens were tested and the average results along with the standard deviation have been reported.

RESULTS AND DISCUSSION

Internal Morphology of Microcapsule

It is clear that the distribution of the healant delineates the internal morphology of the microcapsule, which in turn is influenced by the methodology adopted toward its encapsulation. Interfacial methods, like urea-formaldehyde (UF) encapsulation leads to the formation of "reservoir-type" morphology, where the entire amount of healant forms the core, surrounded by a polymeric shell.

On the other hand, there is a distinct possibility of formation of "monolithic" internal structure in microcapsules obtained by solvent evaporation. The cross-sectional view of both types of microcapsules is presented in Figure 1. It is clear that in the event of rupture of microcapsules with "reservoir" type morphology, the entire amount of healant is released into the crack place. On the other hand, only the micro-droplets placed within the cross-sectional plane are available for healing purposes for microcapsules with a monolithic structure.

Quantification of Healing Agent Delivered

Rule *et al.* proposed an analytical model for predicting the amount of healant delivered in the event of microcapsule rupture,[13] where it is assumed that the entire contents encapsulated within the microcapsules are released in the crack plane, which restricts the predictions to that of microcapsule with "reservoir" type microstructure. According to this model, the total amount of healing agent delivered per unit crack area (m) is proportional to both the microcapsule mass fraction as well as microcapsule diameter (Supporting Information Figure S2, Table S1).

$$m = \rho_s \cdot \phi \cdot D \quad (1)$$

where

ρ_s = density of the sample,

ϕ = mass fraction of microcapsules and

D = microcapsule diameter

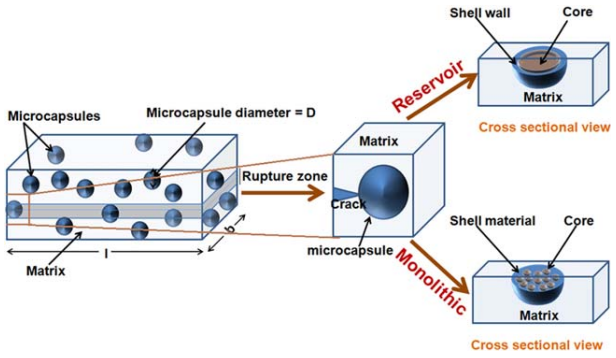


Figure 2. Pictorial representation of a cuboidal sample containing microcapsules. Microcapsules lying within the shaded region are expected to rupture leading to the release of healant. [Color figure can be viewed at wileyonlinelibrary.com]

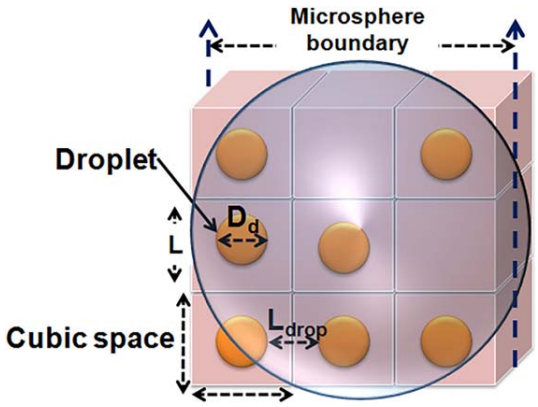


Figure 3. Model for estimating distance between droplets within a microcapsule. [Color figure can be viewed at wileyonlinelibrary.com]

However, for monolithic microcapsules, the scenario is different, as only a fraction of the total encapsulated healant is available for healing purposes (Figure 2).

As has already been mentioned, in a monolithic microcapsule, only the healant present in the exposed droplets (formed as a result of crack propagation) flow into the crack plane. For the purpose of quantification, a single microcapsule is assumed to be composed of cubic spaces each of length (L) containing a single droplet (Figure 3).[7]

The droplets are assumed to be perfectly spherical (diameter, D_d) with the distance between each droplet being L_{drop}. For a fixed core content ($core_{mic}$%, vol/vol), the relationships can be expressed as

$$core_{mic} = \frac{\left(\frac{\pi}{6}\right) D_d{}^3}{L^3} \tag{2}$$

The distance between two droplets (L–D_d) can further be expressed as a function of $core_{mic}$ and D_d as follows:

$$L_{drop} = D_d \left[\left(\frac{\pi}{6 \times core_{mic}} \right)^{1/3} - 1 \right]. \tag{3}$$

The effect of increasing microdroplet dimensions (D_d) on the distance between drops (L_{drop}) is presented in Figure 4. It can be seen that with increasing droplet size, the distance between the adjacent drops increases. The effect of increasing core content on the distance between two adjacent droplets is also presented in the figure. It is interesting to note that microcapsules with core content ~50%, irrespective of the dimensions of the dispersed microdroplets, essentially exhibit "reservoir" type morphology, as the distance between two droplets is essentially zero.

For the purpose of determining the amount of healant released in the event of rupture of monolithic type microcapsule, a cross-sectional area is considered (Supporting Information Figure S3). It can be visualized that on a 2D plane, the area of a square of side = L, containing a single microcapsule of diameter D_d, is $(L_{drop} + D_d)^2$. For such a scenario, the total number of microcapsules (N) can be estimated in terms of the ratio of the area of the cross-sectional plane to that of a single square. Assuming that all the microcapsules are ruptured, the total volume of the healant can be calculated as $N \times \frac{4}{3}\pi r^3$, where r refers to the radius of a single microdroplet.

A comparison in terms of the amount of healant delivered as a function of microcapsule loading for both "reservoir" and "monolith" type configuration for two representative microcapsule diameters (150 and 200 μm) is presented in Figure 5. At core contents 50% and higher, the predictions of monolithic and reservoir are essentially same.

However, it can be seen that the difference between reservoir and monolithic morphology is pronounced at relatively smaller dimensions and lower core contents (<50%). Plots corresponding to different microdroplet diameter (D = 20–80 μm) for a fixed core content ($core_{mic}$ = 40%) are presented in the figure

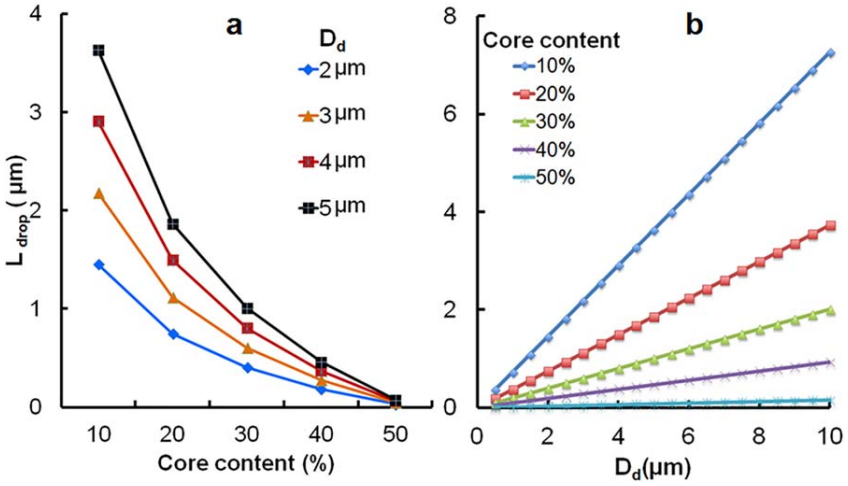


Figure 4. Variation of distance between microdroplets (a) as a function of increasing core content at different micro-droplet diameters, (b) as a function of microdroplet diameter at different core content for microcapsules with monolithic internal structure. [Color figure can be viewed at wileyonlinelibrary.com]

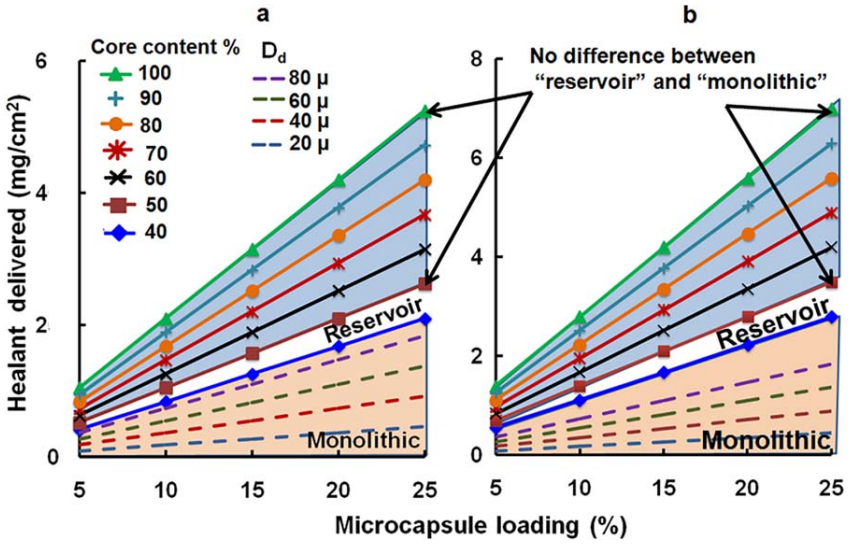


Figure 5. Healant delivered as a function of volume fraction and microdroplet diameter due to rupture of microcapsules of $D =$ (a) 150 μm and (b) 200 μm with different internal microstructure. [Color figure can be viewed at wileyonlinelibrary.com]

and the detailed calculations are presented in the Supporting Information Table S2. It is apparent that as the microdroplet dimensions increase, the monolith-type configuration essentially converges to that of a reservoir.

It is to be noted that this difference between the internal microstructure i.e., monolith and reservoir is not important in the context of pharmaceutical applications, where the encapsulating shell undergoes *in vivo* degradation or dissolution thereby releasing the encapsulated drug in a controlled fashion, which in turn is decided by the rate of dissolution. However for applications, like self-healing, where the release of entire content is a desirable feature, "reservoir" type microstructure is definitely more beneficial, especially for smaller microdroplets.

Experimental Validation of Model Using Unsaturated Polyester as Healing Agent

The potential of USP as a healant for self-healing applications has been demonstrated in our previous article,[15] where microcapsules with "reservoir" type microstructure were prepared by dispersion polymerization route. For preparation of USP encapsulated monolithic microcapsules, solvent evaporation route was adopted, the details of which are presented in Supporting Information. The effect of increasing stirring speed on the microcapsule dimensions is presented in Figure 6. The microcapsules yield, as determined by gravimetric analysis, was found to be 72% ± 5%. Spherical microcapsules with practically no inter-capsule bonding can be observed in the scanning electron microscopy (SEM) images. In addition, microcapsules appear to exhibit a smooth texture with some sub-micron debris on their shell.

The average particle size distribution of the microcapsules was found to shift to lower dimensions with increasing mixing speed, in line with previous observations[12,17,18] (Figure 7). This shift can be attributed to the exertion of higher hydrodynamic stresses on the droplets during the microcapsule formation process. The average particle size decreased from 135 ± 15 μm to 95 ± 10 μm as the agitation speed was increased from 350 to 550 rpm. In comparison, for "reservoir" type microcapsules, prepared using dispersion polymerization route, the particle size was found to decrease from220 ± 34 to 72 ± 26 μm as the agitation speed was increased from 250 to 450 rpm

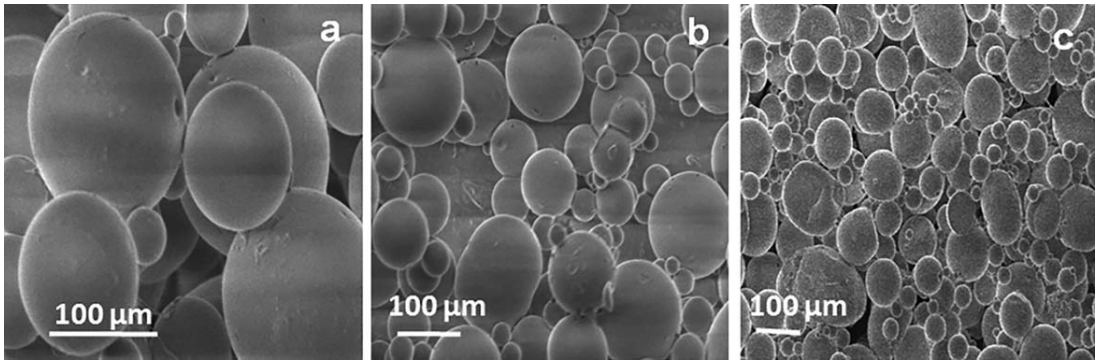


Figure 6. SEM image of microcapsules prepared under different stirring speeds (a) 350, (b) 450, (c) 550 rpm.

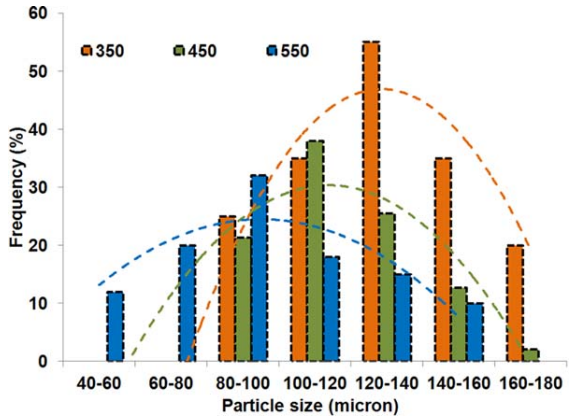


Figure 7. Particle size distribution of microcapsules prepared using solvent evaporation route. [Color figure can be viewed at wileyonlinelibrary.com]

Chemical Reactivity of Encapsulated Resin

The chemical reactivity of the encapsulated USP was assessed through calorimetric studies. The microcapsules were crushed and subjected to a controlled heating program from room temperature to 200 °C. Studies were also performed on crushed microcapsules in the presence of a free radical initiator, AIBN (1% wt/wt of microcapsule) and the obtained traces are also included in Figure 8 The reason behind choosing AIBN as a preferred initiator for USP curing has been discussed previously.[15] It is clear from the trace that the released USP undergoes crosslinking as indicated in the curing exotherm, confirming its chemical activity after rupture of the microcapsule, irrespective of the shell wall used for encapsulation. However, in the absence of the initiator, curing could not commence, under the temperature range (<200 °C) studied as indicated by the absence of the exotherm. For comparison purposes, the DSC trace of neat USP resin, both in the presence and absence of AIBN is also presented, the results of which are included in the Supporting Information Figure S4.

The released USP was found to undergo crosslinking in the presence of AIBN as indicated in the curing exotherm, confirming its chemical activity. AIBN was capable of initiating the curing reactions at temperatures as low as 80 °C, with T_{peak} at 106 °C. As expected, the released USP did not undergo curing in the absence of the initiator. It is to be noted that these microcapsules do not alter the curing kinetics of epoxy resins as has been established by non-isothermal DSC studies previously.[19]

Thermal Properties

Thermogravimetry (TG) traces of USP encapsulated microcapsules are presented in Figure 9. For comparison purposes, the TG trace of the liquid resin (USP) and cured USP along with the shell wall materials i.e., poly(urea-formaldehyde) and PS are also included. Three distinct steps are evident in the decomposition profile associated with neat USP. The first mass loss at $T < 200$ °C is due to the loss of the volatile diluent i.e., styrene, present within the polyester resin. Dynamic heating leads to the polymerization of the polyester with styrene, and the subsequent mass loss at 300–400 °C is attributed to the degradation of the short PS segments formed. The concluding mass loss at 400–550 °C occurs due to the decomposition of the crosslinked structure of USP. The TG trace of USP-encapsulated PS microcapsules exhibits two-step profile. The initial mass loss (300–400 °C) is due to degradation of shell wall material i.e., PS and subsequent mass loss (400–550 °C) occurs due to the decomposition of core material i.e., USP.

USP-encapsulated poly(urea-formaldehyde) microcapsules exhibit a two-step profile. It is apparent that release of the core polyester occurs at slightly elevated temperatures in comparison to the unencapsulated liquid. The encapsulating shell wall of poly(urea formaldehyde) ruptures at 240 °C, leading to the release of the core components, which further undergo similar degradation profile as exhibited by cured polyester. Poly(urea-formaldehyde) exhibit a ~5% mass loss at $T \leq 150$ °C, presumably due to the loss of adsorbed water followed by a precipitous decomposition at about 242 °C (T_{max} ~270 °C).

Self-Healing Efficiency

Calorimetric studies reveal that the curing process is completed at temperature above 100 °C. Further studies also indicated that a period of 6 h was sufficient for the samples to regain their mechanical integrity; hence quantification of healing efficiency was performed under these conditions. For the purpose of quantification, the impact tested samples were joined and maintained at 110 °C for extended periods.

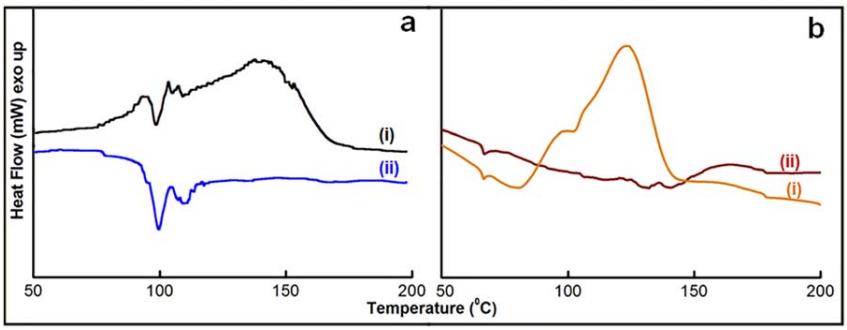


Figure 8. DSC trace associated with the curing of USP in the presence of AIBN (a) monolithic microcapsules of USP encapsulated in PS and (b) USP "reservoir" encapsulated in urea-formaldehyde. The trace obtained in the absence of AIBN is also included as (i). [Color figure can be viewed at wileyonlinelibrary.com]

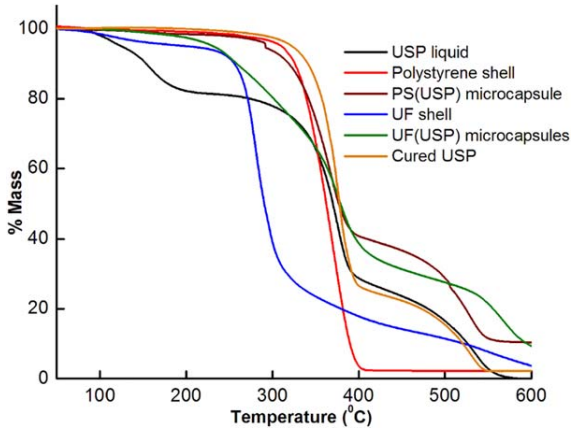


Figure 9. TG traces of polyester resin both before and after microencapsulation. [Color figure can be viewed at wileyonlinelibrary.com]

Healing efficiency was quantified in terms of the ratio of impact strength both before and after self-healing. The effect of introducing microcapsules on the impact strength of epoxy is presented in Figure 10. As reported previously,[20] introduction of microcapsules lead to a decrease in the impact strength. It was however interesting to note that USP-encapsulated PS microcapsules with monolithic microstructure led to lesser deterioration in the properties as compared to USP-encapsulated poly(urea-formaldehyde) microcapsules, which can be attributed to the fragile nature of the latter. It is imperative to mention here that PS microspheres have been reported to act as thermoplastic toughener,[17] which leads to an improvement in the impact strength of the neat resin.

The variation of healing efficiency with increasing loading for either type of microcapsule is presented in Figure 10. Healing efficiency is expected to increase with increasing microcapsule loading and the same is evident in both the cases. However, the healing efficiency is much higher in samples prepared using reservoir-type microcapsules. Figure 10 shows that at 10% loading, the healing efficiency for samples containing reservoir-type microcapsule is ~67% while that for samples containing monolithic microcapsules is less than 5%. Healing efficiency reached a maximum (100% ± 2%) at 20% loading for samples prepared using reservoir-type structure whereas the healing efficiency was markedly lower, (~10% for similar loadings) in case of monolithic-type structure.

Fractured epoxy samples, prepared in the absence of microcapsules, did not exhibit any healing, thereby indicating absence of intrinsic healing in the resin. It is to be noted that introduction

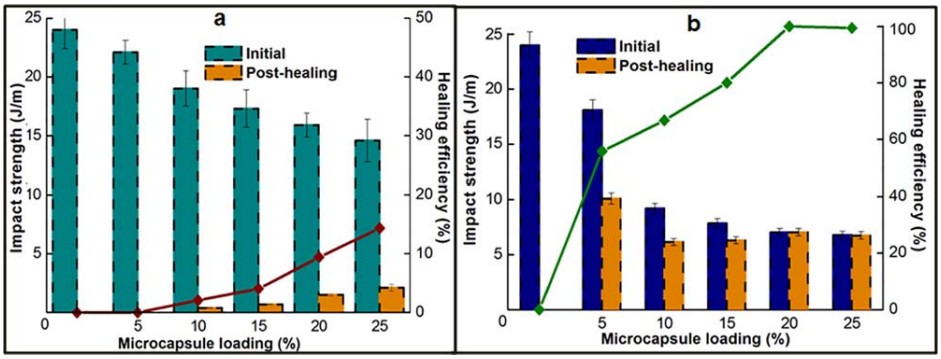


Figure 10. Variation of impact strength (primary axis) and healing efficiency (secondary axis) as a function of microcapsule loading (a) monolithic morphology, (b) reservoir morphology. [Color figure can be viewed at wileyonlinelibrary.com]

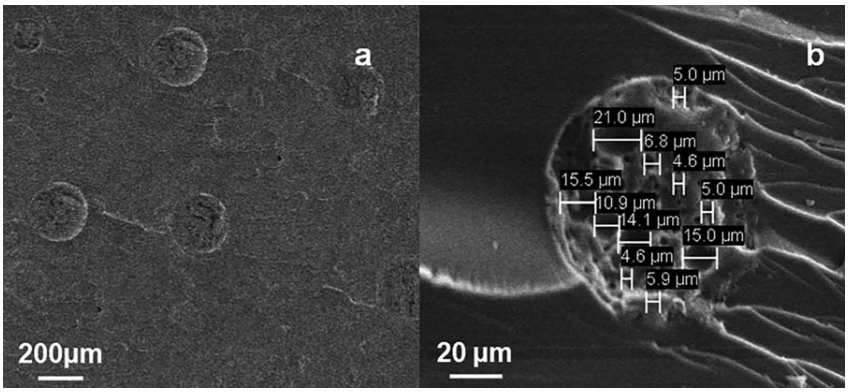


Figure 11. SEM image of fractured plane containing monolithic microcapsule (20% loading) showing (a) dispersion of microcapsules, (b) microdroplets within the monolithic microcapsules.

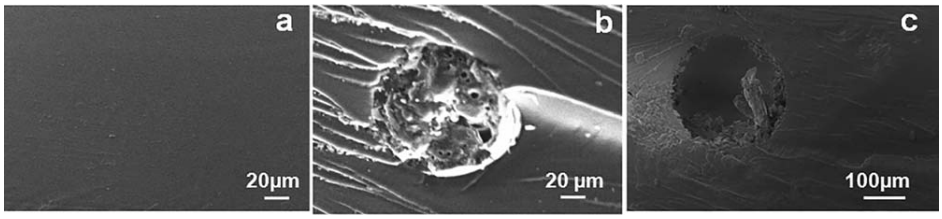


Figure 12. SEM images of fracture surface (a) neat epoxy, (b) epoxy containing microcapsules with monolithic internal structure, (c) containing microcapsules with reservoir type structure.

of microcapsules result in a reduction of impact strength, however the extent of decrease in impact strength was much higher in the case of reservoir-type microcapsules.

The morphology of the fractured plane of a sample containing 20% (wt/wt) microcapsules of monolithic-type internal microstructure was studied using SEM, and a representative image is presented in Figure 11(a,b). It can be observed that the microcapsules are well dispersed throughout the matrix [Figure 11(a)]. Figure 11(b) shows the average microdroplet dimension which was found to be ~10 μm. With these microdroplet dimensions, and monolith-type morphology (diameter ~150 μm), the healant released in the crack plane is anticipated to be ~10.96% of the amount released in case of reservoir-type of microcapsules. This model prediction is in accordance with the experimental data for healing efficiency, which was found to be ~10% of the reservoir-type of microcapsules, comparison of the data is presented in Supporting Information Table S3.

The morphology of the cracked surface of neat epoxy and composite samples post-impact tests was studied by SEM, and representative fractured surfaces are presented in Figure 12(a–c). It can be seen that the surface of epoxy is rather smooth and featureless [Figure 12(a)], which are characteristics of brittle fracture. In comparison, the surface of the fractured crack plane of microcapsule-loaded samples was relatively rough throughout. "Broken" microcapsules were observed on the surface of crack plane. The monolithic- and reservoir-type internal structure of microcapsules prepared using solvent evaporation and dispersion polymerization technique respectively is clearly evident from the SEM images [Figure 12(b,c)].

CONCLUSIONS

In this article, we report the results of a comparative evaluation of the effectiveness of "reservoir" and "monolithic" microcapsules toward self-healing of thermosetting epoxy. An analytical model was developed to predict the effect of increasing microcapsule volume fraction and micro-droplet dimensions on the healant delivered in the crack plane for different internal microstructures. The difference in the predictions on healant delivery for reservoir and monolithic morphology is more pronounced at low core contents and smaller micro-droplet dimensions. Theoretical predictions reveal that with increasing core content, the monolithic morphology essentially converges to that of a reservoir (core content >50%). The developed model clearly highlights the advantages associated with the "reservoir" morphology for self-healing applications. For validation of the model, separate methodologies were used for the preparation of USP encapsulated microcapsules i.e., dispersion polymerization and solvent evaporation, with the former route leading to the formation of a "reservoir" type microcapsules and the latter leading to monolithic microcapsules. Introduction of either type of microcapsule endowed epoxy with self-healing functionality, the efficiency of which was quantified using established procedures. The healing efficiency is much higher in samples prepared using reservoir-type microcapsules, which reached a maximum (100% ± 2%) at 20% loading. For samples prepared using monolith-type structure, the healing efficiency was markedly lower, (~10% for similar loadings), which was attributed to the relatively smaller amount of healant being delivered to the crack plane in the event of microcapsule rupture.

ACKNOWLEDGMENTS

The authors gratefully acknowledge DRDO for funding this research work through grant ST/14–15/CFE-1311.

REFERENCES

1. Dry, C. *Compos. Struct.* **1996**, *35*, 263.
2. Toohey, K. S.; Sottos, N. R.; Lewis, J. A.; Moore, J. S.; White, S. R. *Nat. Mater.* **2007**, *6*, 581.
3. White, S. R.; Sottos, N. R.; Geubelle, P. H.; Moore, J. S.; Kessler, M. R.; Sriram, S. R.; Brown, E. N.; Viswanathan, S. *Nature* **2001**, *409*, 794.
4. Brown, E. N.; White, S. R.; Sottos, N. R. *J. Mater. Sci.* **2004**, *39*, 1703.
5. Yuan, L.; Gu, A.; Liang, G. *Mater. Chem. Phys.* **2008**, *110*, 417.
6. Moll, J. L.; Jin, H.; Mangun, C. L.; White, S. R.; Sottos, N. R. *Compos. Sci. Technol.* **2013**, *79*, 15.
7. Freiberg, S.; Zhu, X. X. *Int. J. Pharm.* **2004**, *282*, 1.
8. Gu, J.; Yang, X.; Li, C.; Kou, K. *Ind. Eng. Chem. Res.* **2016**, *55*, 10941.
9. Li, Q.; Siddaramaiah; Kim, N. H.; Hui, D.; Lee, J. H. *Compos. Part B: Eng.* **2013**, *55*, 79.
10. Li, Q.; Mishra, A. K.; Kim, N. H.; Kuila, T.; Lau, K-T.; Lee, J. H. *Compos. Part B: Eng.* **2013**, *49*, 6.
11. Zhao, Y.; Fickert, J.; Landfester, K.; Crespy, D. *Small* **2012**, *8*, 2954.
12. Sharma, P.; Shukla, S.; Lochab, B.; Kumar, D.; Roy, P. K. *Mater. Lett.* **2014**, *133*, 266.

13. Rule, J. D.; Sottos, N. R.; White, S. R. *Polymer* **2007**, *48*, 3520.
14. Roy, P. K.; Rawat, A. S.; Choudhary, V.; Rai, P. K. *Indian J. Chem. Technol.* **2004**, *11*, 51.
15. Tripathi, M.; Dwivedi, R.; Kumar, D.; Roy, P. K. *J. Intell. Mater. Syst. Struct.* **2016**, *27*, 1650.
16. Yuan, C. E.; Zhang, M. Q.; Rong, M. Z. *J. Mater. Chem.* **2014**, *2*, 6558.
17. Chaudhary, S.; Surekha, P.; Kumar, D.; Rajagopal, C.; Roy, P. K. *Polym. Compos.* **2015**, *36*, 174.
18. Yang, Y. Y.; Chung, T. S.; Ping Ng, N. *Biomaterials* **2001**, *22*, 231.
19. Tripathi, M.; Kumar, D.; Rajagopal, C.; Kumar Roy, P. *J. Therm. Anal. Calorim.* **2015**, *119*, 547.
20. Tripathi, M.; Rahamtullah; Kumar, D.; Rajagopal, C.; Kumar Roy, P. *J. Appl. Polym. Sci.* **2014**, *131*, 40572.

eJBio

Electronic Journal of Biology, 2017, Vol.13(2): 125-130

Vulnerability to a Bioterrorism Attack and the Potential of Directed Evolution as a Countermeasure

Ritwik Raj[1], Ananya Saxena[2], Ritcha Saxena[3,*]

1 Department of Biomedical Engineering, Delhi Technological University, Delhi, India;
2 All Saints University School of Medicine, Roseau, Dominica;
3 Department of Pathology, All Saints University School of Medicine, Roseau, Dominica.

*Corresponding author. Tel: +1-767-2777036; E-mail: ritcha.saxena@yahoo.in

Citation: Raj R, Saxena A, Saxena R. Vulnerability to a Bioterrorism Attack and the Potential of Directed Evolution as a Countermeasure. Electronic J Biol, 13:2

Received: September 20, 2016; **Accepted:** April 14, 2017; **Published:** April 21, 2017

Review Article

Abstract

This article has two purposes. First, the consolidation of facts about natural and manmade biological warfare agents as well as the use of technology in the development of biological weapons, including near-term future evolvement in genetic engineering to prevent the potential havoc-wreaking events using biological warfare. Secondly, this article aims to suggest steps to tackle the vulnerability of humans to biological warfare and to provide a potential solution that incorporates Directed Evolution to attack the problem of biological weapon proliferation.

Keywords: Biotechnology; Biological weapons; Poisons.

1. Introduction

Technology, it appears, is a double-edged sword. Express advancements in biotechnology, biomedical sciences and bioinformatics are incontrovertibly solving a multiplicity of human problems, but it seems that they are leading to as many troubles as they are intending to settle. At one extremity is the conception of breakthrough drugs against cancer and HIV, and at the other is the inadvertent or on purpose development of more dangerous biological poisons. While problems like non-uniform access and availability of food is gradually giving way to food security all over the world, nonetheless, the term 'Biological Weapon' creates insecurity in the mind of any well-aware human.

Biological warfare can be seen in very simple terms as the use of bacteria, viruses or other toxins (biological warfare agents or biohazards) to cause destruction of men, or animals and plants that serve as source of food for them, or even to the ecology of a place. If this is done in order to cause a devastating feeling of anxiety or fear among sufficiently large populations, generally occupying contiguous areas, it is called biological terrorism or bioterrorism. The biological warfare agents (biological weapons) fall under the category of weapons of mass destruction (WMD), under which also fall the dreaded chemical weapons and nuclear weapons. However, not all reported events involving biological agents are linked with a WMD that is being intended to target larger areas and intended to cause many casualties [1].

2. A Brief History

The use of biological agents as weapons to create disorder in established organizations and institutions, such as administrations and particularly bigger metropolitan populations, has been prevalent all through the course of history. In the last few decades, a noteworthy rise has been seen in the use of biological agents in order to disseminate such substances to cause detriment to the public [1]. The potential danger of *Bacillus anthracis* as a biological weapon was vividly demonstrated after an unintentional release of its spores in 1979 from a military microbiology facility in Sverdlovsk, Union of Soviet Socialist Republics, that resulted in at least 77 cases of human anthrax and 68 deaths [1,2]. Although commonly termed "bioterrorism", the purposes of such attacks are not necessarily intended to intimidate established government structures, but can also be motivated by religious, political, or ecological ideologies [2].

3. Potential of Naturally Occurring Biological Poisons

Potential natural biological warfare agents are the Ebola, the Venezuelan Equine Encephalitis virus, the Marburg virus and the like [3-5]. In more recent times, the primary biological warfare agents that have been reported are anthrax (*Bacillus anthracis*) and ricin, a non-contagious agent that is an albumin, and which can be obtained from *Ricinus communis* (castor bean plant) [2]. Not all of them cause a fatal infection, and are fought and defeated by the natural immune system of the human body. And many of these that have the potential to cause fatal infections have a fixed cure, though, for biological agents like ricin, there are no available treatment options with the exception of supportive care [1]. Obtaining the rarely occurring ones among the rest is difficult. Access to natural and highly virulent agents and strains like the smallpox virus, has become much regulated

and exceedingly restricted even for laboratories [3]. In spite of all these facts, groundwork for a possible bioterrorism attack is tremendously necessary. In case of Woolsorter's Pneumonia (anthrax), because there is no mechanism to discover a premeditated, secret release of the spores of the causative agent, the only step that appears to be exceedingly effective in combating the disease is increased clinical suspicion and vigilance based on knowledge of and insight into the pathogenesis of *B. anthracis* [6]. Similarly, in case of most of the other potential biological weapons, ways of knowing about a deliberate attack have not been identified. Indeed, the real attribute of a potential biological weapon that is chiefly utilized by those who wish to clandestinely undertake an offensive using such agents is the mortality caused by the agent before it is known that the populace has been infected with it. The fact that almost all of the potential biological weapons have the attribute of quick extensive spread among populations and utter difficulty of diagnosis of the diseases they cause makes them the most deleterious among the Weapons of Mass Destruction.

Ebola, if contracted naturally, can cause death in not a very large number of cases [4]. And those who can afford and access available treatments have a creditable probability of survival, assuming that the aim of terrorists is to create terror through deaths since treatment of the diseases they intend to cause may almost negate their efforts to create affright, even though no specific treatment for Ebola in the form of a vaccine is available, and drugs for the same are in the process of being tested or are being considered for being tested, as per WHO [7]. Similar is the case of H1N1 (swine flu), the difference there being that a potential cure was developed within a short time of the occurrence of the pandemic [8, 9], due to the rapid evolution of medical science, where new discoveries are made in the wink of an eye, and inventions occur in leaps and bounds. But susceptible groups like children, already sick people, old persons, pregnant ladies, etc., have amply heightened risk of mortality in both the cases, and that is sufficient to cause terror.

3.1 Comparison between the homicidal potential of chemical and biological poisons

A biological weapon cannot kill as quickly as a chemical weapon can. A biological warfare agent encounters huge resistance from the human body's immune system before being able to work its way out to cause its impact by means of appearance of symptoms. A chemical agent, developed to destroy life, can cause deaths in seconds by the annihilation of living cells or other subtler means. In contrast, a biological poison like the HIV, a retrovirus, can stay in the body for as long as twenty years before its first symptoms appear. Even after that, a person may live for up to a few decades by the way of using anti-retroviral therapy, even though no specific cure for its infection exists [8-11].

But HIV is not a potential biological weapon. It cannot spread extensively in a short duration. The potential biological warfare agents are chosen such that they act very rapidly. Even though, as per what is known about them their lethality pales before that of chemical weapons and they cannot act as quickly as a chemical warfare agent, difficulties related to identification of the diseases they cause render them much more detrimental than the latter.

3.2 Artificial biological weapons and modified natural agents

The creation of a biological weapon in a lab may turn out to be an absolutely onerous task. Not only that, even the most unknown microorganisms, be them viruses or bacteria, may respond to random anti-viral therapies or anti-bacterial therapies as may be the case, including broad-spectrum antibiotics or a combination of them [12]. It is known how to decipher genomic sequences (owing to the research efforts of double Nobelist Sanger), and we also possess the capability to synthesize these chains (through the studies conducted by Nobelist Hargobind Khorana) [13-15]. Thus, creating or altering the DNA or RNA sequences at will, or to perform site-specific mutagenesis (because of the research carried out Nobel Prize winning scientist Michael Smith in this direction) are tasks that can now be easily realized [13,14]. Beginning with the genome of an existing virus, in order to recreate the genome of a microbe, the sequence needs to be available on the internet or, at least, in the books. If it is available, it needs to be assembled by combining shorter DNA or RNA sequences that are already accessible, like what was done by Cello et al. in 2002 [16]. It then has to be modified to add the required genes in it, so that one can get a virus of the particular characteristics demanded by an efficient biological weapon. The protein coat may also be created in the lab, beginning with a prototype and evolving it bit by bit. These stepwise modifications may be carried out with the assistance of Directed Evolution [17,18]. Creating long genome sequences from scratch is burdensome.

If a protein coat with the required characteristics cannot be fashioned in the laboratory, the genome created can be inserted into the protein coat of a virus with its own genetic material removed. One has to take under consideration that environmental robustness depends on the genes of the pathogen along with the attributes of the protein coat. Procurement of suitable virulent strains of micro-organisms, easy multiplication and mass production without the loss of potency and pathogenicity (which also depends on the genes of the microbe), and the development of an effective medium of delivery are quite complex problems, the last one particularly so, and, on the top of that, only few well-trained and commendably proficient experts can bring about the successful culmination of the process [3,12]. Even then, credible evidence suggests that sufficiently erudite scientists are available in plenty numbers,

and are lending their support, to both the clandestine biological weapons programmes of various nations and to the terrorist organizations of the world who wish to succeed in this arena of mayhem.

The USSR's 'invisible anthrax' can be quoted as a crude example of a genetically modified biological warfare agent. The causative vehicle of the disease, *Bacillus anthracis,* had a foreign gene imbibed into its genome which modified its immunological properties, rendering existing vaccines ineffective against its infection [19]. It fulfilled almost every one of the requisite specifications of a biological weapon, viz., being able to be cultured and thus created in large quantities, being able to act rapidly, being environmentally hardy and being vaccine treatable in such a manner that an affected enemy does not have a vaccine or a drug as a cure, but the allies can be protected from its overwhelming impact [3].

But, since the early years of the current millennium, several ways to combat anthrax have been developed. Therefore, for the present generation of terrorists and other such antihuman agents, deploying anthrax as a bioterrorism instrument is almost infeasible since there are limitations with the anthrax bacteria, if the actual aim is to terrorize through destruction of human lives. Its victims can be treated using antibiotics several days after the infection, and only a small minority of those infected die, and those deaths too, can be prevented if awareness has been generated among people about this [3,6,20]. So, it could be seen as a bioweapon that is bound to fail as a mass killing machine, but might be employed to weaken enemy troops during combat.

Despite that, the development of modified anthrax bacteria can be seen as a breakthrough in the evolution of modern biological weaponry, for its creation has showed that a simple genetic change can produce very deadly results indeed. Since *Bacillus anthracis* has the potential to be employed as an agent of bioterrorism, an agent which can swiftly develop into systemic anthrax with high mortality among the persons who are exposed and go untreated, clinical guidance that can be rapidly disseminated as well as implemented must be prepared to preempt any deliberate release of the bacterium [21]. Further, to combat genetically engineered anthrax bacteria, gene therapy shows immense potential. The Human Genome Project shows great promise to have an intense constructive impact on the way biomedical research is conducted and aid in explaining the most unexplained and intricate life processes [22]. Latest techniques in biotechnology should permit the scrutiny of the events that occur in a human cell following the infection by a pathogen or the ingestion of a toxic molecule, which can help clarify the situations that result in individual vulnerability to infectious diseases [12]. Studies in functional genomics may be able to explicate mysteries that still exist regarding the human genome and facilitate the excogitation of new strategies for prevention and treatment of infectious diseases in the form of vaccines and anti-microbial drugs, respectively [12]. Researchers have developed innovative and exceedingly promising gene-editing tools like CRISPR, which can potentially be used to build up effective techniques to combat this disease by rendering humans immune to it through genetic enhancement or eugenics. Immunotherapy, also, can be possibly utilized for this purpose in combination with gene-editing, or otherwise in a standalone manner.

Molecular biology can enable the scientists and the researchers to develop new weapons, which are much more efficient and useful than the natural ones or previously existing ones [3,12]. One of the techniques under the subject can involve the introduction of antibiotic resistance by altering genomes of microbes, which can render them typically noxious biological warfare agents [3]. The human smallpox virus has the potential of being a model bio-weapon, especially for terrorists, because it is decidedly infectious and has the capability to bring about tremendous fatality rates. Nearly the whole of the latest generation among the existing population of the world has not been vaccinated against that particular virus [3,23]. It is horrible to imagine what will happen to the youth if criminals get hold of the smallpox virus somehow and hand it over to wretchedly desperate non-state actors who are capable of turning it into a catalyst of cataclysm.

Fortunately, nowhere is its genome existent except in two chosen highly secure laboratories around the world. Its genome is very big, having about a hundred and ninety thousand base pairs, and is not easily available in the form of data or diagrams [24]. Then again, the polio virus has a much smaller genome (about seven thousand and five hundred nucleotides) which been whose assembly has already been demonstrated by Cello et al. [16] and Kitamura [25]. All the same, a related virus like cowpox, or even monkey pox, can be utilized so that one can mutate and thereby change certain bases and sequences in their genome in such a way that they resemble the human smallpox [3]. Rosengard et al. [26] in 2002, have shown that the sequence of a gene related to the vaccinia virus's pathogenicity, can be transformed with the help of targeted mutation of a small number of base pairs into the sequence of the corresponding gene in the human smallpox virus [26]. But trying to repeat the process to obtain the whole genome of the smallpox virus can possibly only be fantasized, for it sounds to be an exceptionally complex, lengthy and cumbersome process.

3.3 Tackling the vulnerability through socio-political means

We have divided the focus areas for tackling vulnerability, where we particularly lack in skills at the present moment, and those which we deem to be most important, into two: those which should be undertaken during the preparation phase and those which need to be carried out during emergency response phase.

In order to reduce the vulnerability of the society to the biological agents during the preparation phase, among whatever others have prescribed, the focus must be on developing techniques regarding:

1. Quicker and simpler diagnosis of diseases, particularly the emerging infectious and non-infectious diseases, and
2. Immunization against probable natural as well as, to whatever extent possible, manmade or genetically modified biological warfare agents

During the emergency response phase, among whatever other studies have recommended, the focus must be on developing methods for:

1. discovery of the dissemination of a biological agent, if at all an agent has been disseminated, and the method as well as the agency of its dissemination
2. speedy detection of the attributes of the agent deployed, and, utilizing that information, quickly formulating a treatment protocol, and, if possible, a cure
3. inoculating the maximum possible people in the region affected in the minimum possible time through deployment and redisposition of well-protected response teams

These areas have been listed as per what we find significant to ponder. There may be more as well, that need to be taken care of. The ways of discovering the dissemination of an agent have been given a status of critical importance by Grundmann, who suggests that sentinel air-measurement devices in large urban areas or epidemiological investigations are necessary to kick off the response against a bioterrorism attack [1].

The most likely perpetrators of a possible bioterrorism attack are radicalized groups that aim to utilize biological warfare agents to cause mass casualties. Preventative measures by agencies involved against them concentrate on hunting down and disrupting the activities of such groups from obtaining and using biological agents. Most international groups are likely to collaborate with insiders who may at as their proxies, though, who are located in the country where the attack is aimed to take place [1,2]. With the knowledge that the powerful militant groups like the Daesh (ISIS), some of whose intentions include destruction of those who do not conform to their professed ideologies, have transnational targets and multinational presence, and also the existence of the rumors as well as reliable intelligence reports suggesting their setting up of new, and gaining control of old, laboratories– the armories of death – that are suitably equipped to create artificial biological weapons – the ammunition of upheaval, it is high time that the government takes a proactive stance against the possible deployment of these weapons against civilians and food supplies, along with environmental concerns that arise from the use of such weapons.

Funding towards the development of vaccines and treatment options to tackle the dissemination of a biological agent has till now chiefly been concentrated on anthrax, tularemia, smallpox, and ricin, while research projects that aim to create a remedy against a possible bioterrorism attack by other biowarfare agents have been relatively underfunded [1]. The next section shows how a new scientific technique can potentially be employed to counter bioterrorism using vaccines and drugs.

In guise of preparing anti-weed and anti-pest biological agents that might help in agriculture, or developing drugs for the benefit of humans, or under other aliases, various labs in different countries are reportedly developing biological weapons of varying modes of operation and dissemination, and deadliness. Strengthening the various safeguards like those provided under the Biological and Toxin Weapons Convention can go a long way in tackling dangers posed by these weapons.

4. Suggestions

Advances in creation of new vaccines against a number of meningitis and pneumonia bacteria have been made possible as a result of the availability of genome sequences of numerous pathogens [27]. Attempting to develop new vaccines that would stimulate immunity against a multiplicity of diseases with a single treatment, scientists have genetically engineered viruses [12,27].

There are possibilities of developing completely novel weapons on the basis of knowledge provided by biomedical research [3]. But, at the same time, Directed Evolution, in a manner similar to the way it can be involved in the development of bioterrorism agents, can be used to efficaciously counter their impact. For example, in case of genetically modified *Bacillus anthracis,* the analysis of the DNA sequences of all available variants of the anthrax pathogen may be carried out with the help of a database library, and particular regions in the sequence that are preserved in most of the variants may be found. (The sequences in the DNA that are common among all individuals of the species of the bacterium, that in fact render it a unique species, may be considered here.) Such regions can thus be considered as the chink in the armor of most, even if not all, of the *Bacillus anthracis* subtypes known hitherto. Using Directed Evolution, if, step by step, the scientists are able to develop some biocatalyst that has the capability to sever any of these regions from the genome of the bacterium, or, utilizing one or more of those regions, develop a vaccine in stages, where each iteration results in a betterment over the preceding stage, success might be achieved against possible anthrax outbreaks that may occur in case of an attempted bioterrorism assault [28].

But a problem may occur if very limited number of such regions is discovered and these regions are known by those who are in process of developing

biological weapons using genetically modified *Bacillus anthracis,* for they may modify these regions as well; rendering newly developed and effective drugs or vaccines that target those regions, useless. Further, the mission might encounter a roadblock in another way – if no such region is found out. But there is still hope, since the genome of the bacterium is long, the probability of finding such regions increases, and the chance that all these regions are discovered and modified by those who intend to create biological weapons decreases.

In case of those microorganisms whose genome has suspected to have been completely artificially synthesized in the laboratory, genes – that can possibly be incorporated in its genome to render to it the attributes that make the microbe a potential biological weapon – can be brainstormed by scientists and vaccines can be developed against such genes in the manner already described above. Although this mechanism has its drawbacks, it has the potential to not only help in battling the threat of biological warfare agents to a substantial extent but also increase the security and confidence of the people at large. This technique can be replicated against known and possible causative agents of zoonotic diseases. Such researches must be carried out covertly.

5. Conclusion

Directed Evolution may prove to be a decidedly potent tool to counter biological weapons in the near future. Reemphasizing the axiom that technology has its light and dark faces, any technology that is used for a maleficent purpose, the same technology can also be used to counteract against the ghosts which it created. To prevent the technology from being applied for injurious purposes, ultimately, it is the human outlook that has to change from self-centered to selfless, from artificiality to innateness, and from inhumane to compassionate.

References

[1] Grundmann O. (2014). The current state of bioterrorist attack surveillance and preparedness in the US. *Risk Manag Healthc Policy.* 177-187.

[2] Carus SW. (2002). Bioterrorism and biocrimes: The illicit use of biological agents since 1900. Amsterdam: Fredonia. *Federation of American Scientists.* 1-219.

[3] Aken JV, Hammond E. (2003). Genetic engineering and biological weapons. *EMBO Rep.* **4:** S57–S60.

[4] WHO Ebola Response Team. (2014). Ebola virus disease in West Africa-the first 9 months of the epidemic and forward projections. *N Engl J Med.* **371:** 1481.

[5] Feldmann H, Geisbert TW. (2011). Ebola haemorrhagic fever. *Lancet.* **377:** 849.

[6] Binkley CE, Cinti S, Simeone DM, et al. (2002). *Bacillus anthracis* as an agent of bioterrorism: A review emphasizing surgical treatment. *Ann Surg.* **236:** 9-16.

[7] http://www.who.int/medicines/emp_ebola_q_as/en/

[8] Garten RJ, Davis CT, Russell CA, et al. (2009). Antigenic and genetic characteristics of swine-origin 2009 A (H1N1) influenza viruses circulating in humans. *Science.* **325:** 197-201.

[9] US Department of Health and Human Services. (2009). H1N1 influenza improvement plan. Washington, DC: US Department of Health and Human Services 2012.

[10] http://apps.who.int/iris/handle/10665/198066

[11] Kiweewa FM, Wabwire D, Nakibuuka J, et al. (2013). Non-inferiority of a task-shifting HIV care and treatment model using peer counselors and nurses among Ugandan women initiated on ART: Evidence from a randomized trial. *J Acquir Immune Defic Syndr.* **63:** e125–132.

[12] Ainscough MJ. (2002). Next generation bioweapons-genetic engineering and biological warfare, extracted from 'The Gathering Biological Warfare Storm'.

[13] http://www.whatisbiotechnology.org/science/sequencing

[14] Smit MH. (1993). Synthetic DNA and biology-Nobel lecture. *Chemistry.* 1-18.

[15] Wells RD. (2009). Discovery of the role of non-B DNA structures in mutagenesis and human genomic disorders, The American Society for Biochemistry and Molecular Biology. *J Biol Chem.* **284:** 8997–9009.

[16] Cello J, Paul AV, Wimmer E. (2002). Chemical synthesis of poliovirus cDNA: Generation of infectious virus in the absence of natural template. *Science.* **297.** 1016–1018.

[17] Kotterman MA, Schaffer DV. (2014). Engineering adeno-associated viruses for clinical gene therapy. *Nat Rev Genet.* **15:** 445–451.

[18] Nonnenmacher M, Bakel HV. (2015). High capsid–genome correlation facilitates creation of AAV libraries for directed evolution. *Mol Ther*. **23:** 675-682.

[19] Pomerantsev AP, Staritsin NA, Mockov YV, et al. (1997). Expression of cereolysine ab genes in *Bacillus anthracis* vaccine strain ensures protection against experimental hemolytic anthrax infection. *Vaccine.* **15:** 1846–1850.

[20] http://www.cdc.gov/anthrax/medical-care/treatment.html

[21] Bradley JS, Peacock G, Krug SE, et al. (2014). Pediatric anthrax clinical management. *Am Acad Pediatr*. 0563.

[22] http://www.tigr.org/tdb/mdb/mdbcomplete.html

[23] Henderson DA, Inglesby TV, Bartlett JG, et al. (1999). Smallpox as a biological weapon: Medical and public health management. Working Group on Civilian Biodefense. *JAMA*. **281:** 2127–2137

[24] Massung RF, Liu L, Qi J, et al. (1994). Analysis of the complete genome of smallpox variola major virus strain Bangladesh-1975. *Virology*. **201**: 215-240.

[25] Kitamura N. (1981). Primary structure, gene organization and polypeptide expression of poliovirus RNA. *Nature*. **291:** 547–553.

[26] Rosengard AM, Liu Y, Nie Z, et al. (2002). Variola virus immune evasion design: Expression of a highly efficient inhibitor of human complement. *Proc Natl Acad Sci*. **99:** 8808–8813.

[27] Fraser CM, Dando MR. (2001). Genomics and future biological weapons: The need for preventive action by the biomedical community. *Nat Genet.* **29:** 253-256.

[28] http://www.thehindu.com/todays-paper/tp-features/tp-educationplus/directed-evolution-to-solve-a-drugresistant-disease/article8430069.ece

Received: 22 April 2017 | Revised: 6 June 2017 | Accepted: 7 June 2017

DOI: 10.1002/mmce.21140

RESEARCH ARTICLE

INTERNATIONAL JOURNAL OF WILEY RF AND MICROWAVE COMPUTER-AIDED ENGINEERING

Wideband high gain conical dielectric resonator antenna: An experimental study of superstrate and reflector

Sounik Kiran Kumar Dash | **Taimoor Khan**

Department of Electronics and Communication Engineering, National Institute of Technology Silchar, Silchar, Assam, India

Correspondence
Taimoor Khan, Department of Electronics and Communication Engineering, National Institute of Technology Silchar, Silchar, Assam, India
Email: ktaimoor@gmail.com

Abstract

A physically feasible new hybrid geometry for gain improvement with wideband characteristics is designed and validated experimentally in a conical shaped dielectric resonator antenna (DRA). It comprises of one superstrate and one reflector, with simple slot coupling technique for excitation. The reflector beneath the ground plane mainly accountable for significant gain improvement (∼109%) by reducing the back radiation, whereas the superstrate dedicates for maintaining wideband (12.65%). The demonstrated result shows $S_{11} < -10$ dB band between 7.4 GHz and 8.4 GHz, with 11.25 dBi peak gain, which are well matched with their simulated counter parts. It also gives high co-pol to cross-pol difference (∼40 dB) in broadside direction. This new geometry can be eligible for X-band applications as well can usher the DRA researchers for further innovations.

KEYWORDS

conical DRA, dielectric resonator antenna, high gain, reflector plane, superstrate, wideband

1 | INTRODUCTION

As per current scenario, dielectric resonator antenna (DRA) has been got extreme attention from antenna researchers in microwave as well as millimeter wave engineering.[1–3] The extraordinary techno facts like; wideband, high radiation efficiency, low loss, ease of excitation, and 3-dimensional design flexibility than conventional antennas,[1,4,5] have made it an unique radiator in electromagnetic (EM) field. In practical, most of the modern long distance communications need high gain ∼10 dBi, which are generally not reachable by standard DRAs because of medium gain ∼ 5dBi.[5,6] So engineers use traditional array concept[1] for gain improvement in bore sight direction. But always this is not convenient because of large deployment size as well volume. So researchers across the world have developed several other techniques. For example, recently Fakhte et al. have proposed anisotropic stacking of rectangular[7] and cylindrical[8] dielectric resonator (DR) for > 8 dBi gain in both the cases, by means of improving the radiation intensity from the side walls. A surface mounted horn has been designed by Nasimuddin and Esselle[9] to improve the gain up to 9 dBi. The suppression of surface wave by electromagnetic bandgap (EBG) technique has also been found productive for gain improvement up to ∼8 dBi.[10] Excitation of different higher order modes for exhibiting high gain, like $TE_{\delta 15}$ for ∼ 10 dBi,[11] $HEM_{12\delta}$ for 8 dBi,[12] for 10.3,[13] for 7.5 dBi,[14] for 8.5 dBi.[15] The feeding techniques used for revealing these modes are quite significant. Utilization of metal sheet as superstrate over DRA has also been resulted a high gain of 12 dBi[16] and 11.5 dBi.[17] Moreover, the use of reflector in some cases[18–21] has also been improved the gain and directivity by means of reducing the unnecessary back-side radiation. Nikkhah et al.[18] have used reflector as well as array of DR (4 element and 16 element) to improve the gain beyond 10 dBi. Ali and his research group[19–21] have reported different kind of reflectors to improve the directivity up to 4.5dB,[19] 5.5dB,[20] and 7.5 dB,[21] respectively. Although some of these approaches are effective in gain improvement but sometime deal with some issues, like chances of unnecessary air-gap during series of stacking, which also needs high precision fabrication,[7,8] modification of DR may affect the aspect ratio as well as the Q-factor,[9] fabrication difficulty during incorporation of EBG and vias,[10] large size, and volume.[11,12]

 © 2017 Wiley Periodicals, Inc.

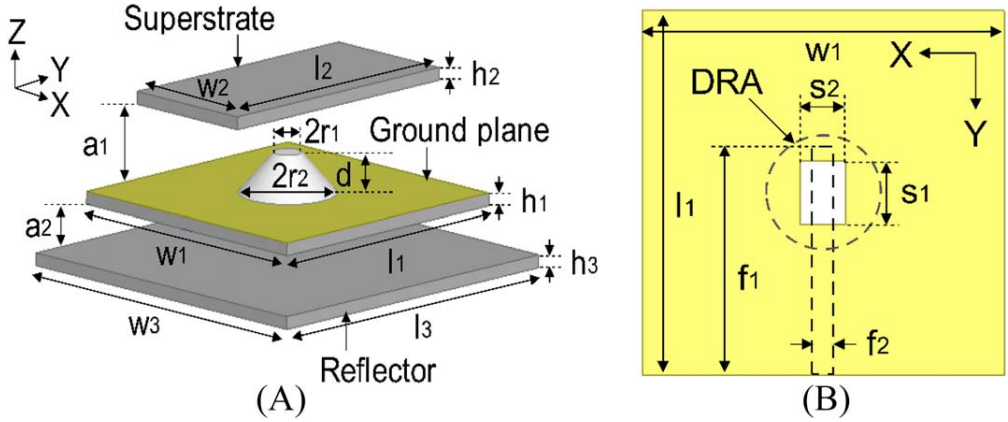


FIGURE 1 Perspective view of the slot coupled conical DRA with dielectric superstrate and reflector plane, A, 3D view, B, top view

However, from above techniques,[7–21] the superstrate[16,17] and reflector[18–21] techniques are quite adept, as they have the potential to improve the radiation characteristics without much affecting DR/substrate geometry and other performances. But the investigations on these techniques are quite limited, which need to be extended further. Moreover, no experimental analysis of reflector-based DRA is done so far. In view of those above findings, here an experimental analysis is done in DRA, which reveals that, the combined effect of superstrate and reflector is quite better than their individual effect. The fringe characteristics of this developed geometry can be considered as; (i) wide impedance matching up to 1 GHz (ie, 12.65% BW), (ii) peak gain of >11.25 dBi, (iii) comparatively compact radiator than[7–18] (also not array like[18]), (iv) introduction of light weight dielectric superstrate and reflector as well as hybrid combination of both, (v) economic fabrication and maintenance less structure. The organization of this article is done as follows: sections 2 and 3 describe the antenna configuration and design flow, respectively. The validation followed by a qualitative comparison is projected in sections 4 and 5, respectively. The final concluding remark is given in section 6.

2 | ANTENNA CONFIGURATION

Figure 1 shows the schematic diagram of the proposed antenna. Here, a conical dielectric resonator (CoDR) ($\varepsilon_r = 10$) is placed above a rectangular slot. A 50Ω microstrip line is made on the opposite side of the slot to couple the DR through it. Three FR-4 sheets of respective dimensions are used for ground plane, superstrate, and reflector substrate. In superstrate, the copper is removed from both side of the substrate and in reflector only from one side (which faces $+Z$). The superstrate and reflector are placed on top side and bottom side, respectively, with corresponding optimized air-gaps. This geometry is first modelled and optimized using commercial finite element solver Ansys HFSS v13.0[22] followed by fabrication and measurement. All the parameters of the proposed antenna are listed in Table 1.

3 | PARAMETRIC STUDY AND DESIGN ANALYSIS

The study begins with a standalone conical dielectric resonator antenna (CoDRA) ($r_1 = 1$ mm, $r_2 = 8$ mm, and $d = 5$ mm) excited by slot ($s_1 = s_2 = 5$ mm) coupling technique. The specific reason of considering CoDRA over cylindrical DRA can be attributed to more number of design parameters/flexibility, that is, upper radius, lower radius, and height than the cylindrical one. Hence, in this regard all dimensions of CoDR (r_1, r_2, and d) were taken into account to operate/resonate the antenna in the specified region. The upper radius (r_1) was varied between 1 mm and 3 mm and the respective $|S_{11}|$ and gain total responses with respect to frequency are plotted in Figure 2A. This indicates that, the resonance deep as well as frequency varies from −18dB (at 9.4 GHz) to −28dB (at 8.4 GHz) as a function of r_1; however, the peak gain remains almost same, that is, ∼4.85 dBi for all the

TABLE 1 Optimized dimensions of the proposed DRA

Name of the part	Optimized value (in mm)
Conical DR	$r_1 = 2$, $r_2 = 7$, $d = 5$
Slot	$s_1 = 7$, $s_2 = 5$
Feed line	$f_1 = 24$, $f_2 = 2$
Ground plane substrate	$l_1 = 40$, $w_1 = 40$, $h_1 = 1.6$
Superstrate substrate	$l_2 = 40$, $w_2 = 20$, $h_2 = 1.6$
Reflector substrate	$l_3 = 50$, $w_3 = 50$, $h_3 = 1.6$
Air-gaps	$a_1 = 7$, $a_2 = 4$

[a]Relative permittivity of conical DR and all substrates are 10 and 4.4, respectively.

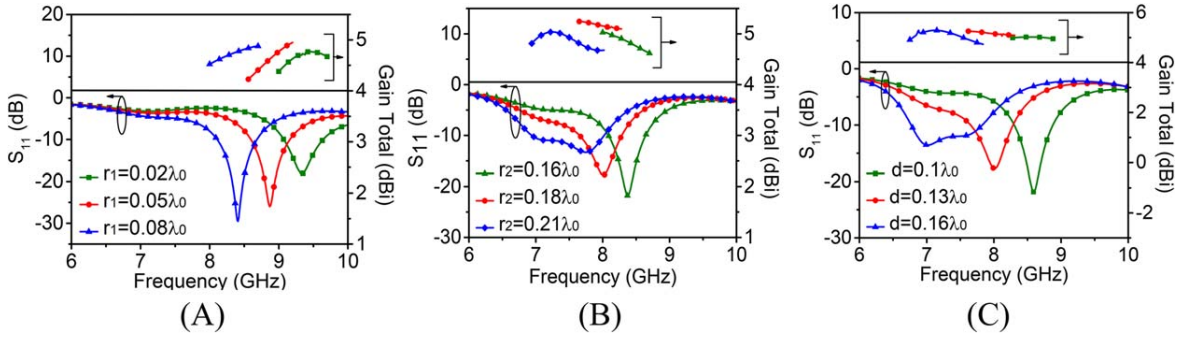


FIGURE 2 Variation of S_{11} and gain total for different DR dimensions, A, upper radius, B, lower radius, and C, height

cases. Then r_1 was fixed at 2 mm for 6.9% bandwidth and 4.88 dBi peak gain. Similarly, the lower radius (r_2) was also varied between 6 mm and 8 mm (Figure 2B) and fixed at 7 mm to achieve a bandwidth of 7.91% with 5.32 dBi peak gain. Next, the cone height (d) was taken into consideration and varied between 4 mm and 6 mm (Figure 2C). However, the existing height $d = 5$ mm, was comparatively good in terms of bandwidth and gain; hence, it was kept as it. Next, the authors conduct parametric study of slot-length and width to observe their effect. As already stated, the CoDR was placed over an $s_1 \times s_2 = 5 \times 5$ mm^2. To understand the effect of slot dimension, length (s_1), and width (s_2) were varied between 6 mm and 8 mm and 4 mm and 6 mm, respectively (Figure 3). Here, the resonance deep varies but not the frequency, which clearly interprets that, mode of operation of DR depends on its dimension but not on the slot dimensions.[1,2] It is also observed that, when length/width increases in Figure 3A,B, means the coupling area for the DRA bottom surface also increases, which creates a situation to couple maximum energy in to the DR and causes a little gain variation. However, considering the matching and bandwidth in to account the slot dimensions were fixed at $s_1 = 7$mm, $s_2 = 5$mm for 5.38 dBi peak gain over 9.20% bandwidth. It can be noted that, the dimensions in the result analysis figures are given in terms of wavelength with respect to resonating frequency 8.08 GHz. Further, at this stage the matching of the feed network is cross verified by varying the feed length (f_1) between 23 mm and 25 mm and width (f_2) between 1.5 mm and 2.5 mm (Figure 4). Considering ~50 Ω resistance and ~0 Ω reactance feed length is fixed at $f_1 = 24$ mm. In case of feed width (f_2), all values between 1.5 mm and 2.5 mm shows almost same behavior. It indicates that, the feed dimension taken initially, fits to the best. However, $f_2 = 2$ mm has been considered to smoothly accommodate the feed probe during fabrication the prototype especially for mounting the SMA connector through soldering.

In next phase, the CoDRA is partially covered by placing a dielectric superstrate ($l_2 = 40$ mm, $w_2 = 40$ mm, $h_2 = 1.6$ mm) at $a_1 = 8$ mm. Some sequential observations were made to improve the performance. The bandwidth and gain are observed for different length (l_2) and width (w_2) of superstrate. However, S_{11} is found to be less sensitive to these parameters; hence, the results are not accounted here for brevity. Moreover, their effect on gain is quite potential and plotted in Figure 5A, which is the ultimate interest of analysis here. An in-detail observation is done by examining sixteen sets of l_2, w_2 combination to find the best performance. Qualitatively, it can be stated that, gain of the antenna increases more while the bigger side of the superstrate is oriented in the direction of antenna polarization (i.e., along Y-axis). The length (l_2) and width (w_2) are fixed at 40 mm and 20 mm, respectively, for a peak gain of 6.65 dBi. It is known that, the placement of superstrate act as a cavity

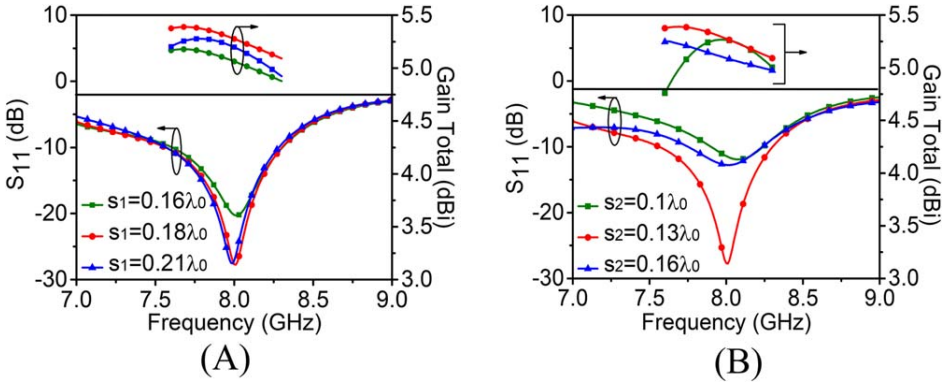


FIGURE 3 Variation of S_{11} and gain total for different, A, slot-length, B, slot-width

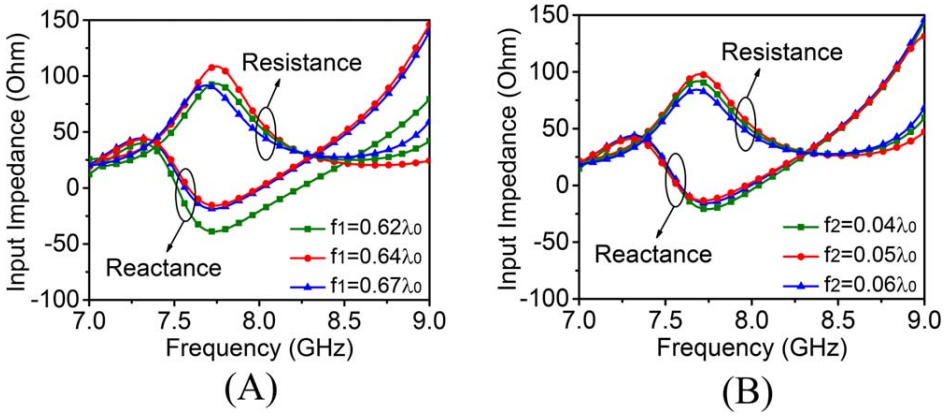


FIGURE 4 Variation of imput impedance for different, A, feed-length, B, feed-width

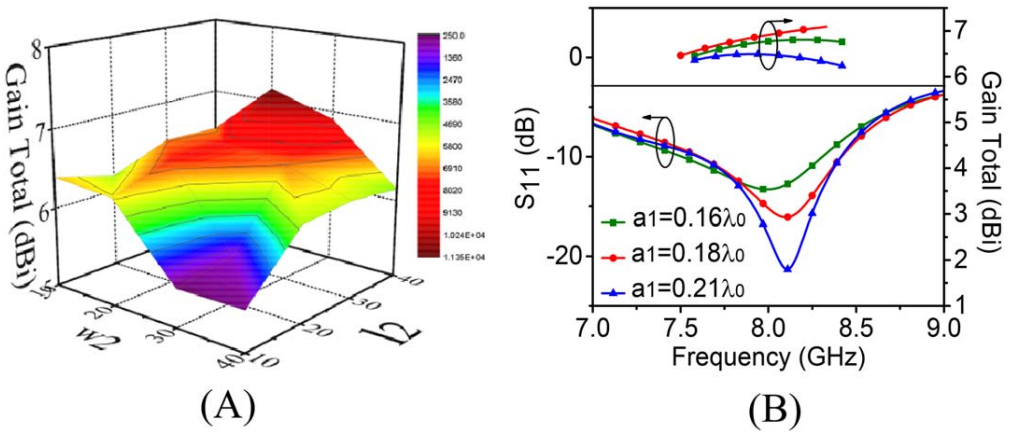


FIGURE 5 Performance of DRA with superstrate A, Gain total response for different combinations of l_2 and w_2. (at $a_1 = 8$ mm and $\theta = 0° = \phi = 0°$), B, S_{11} and gain total response for different air-gap height

resonator.[16] Hence, the effect of air-gap (a_1) is verified in Figure 5B and optimized at 7 mm in view of optimum gain 6.65 dBi over 10.73% bandwidth. Here, it can be noted that a_1 affects the resonance deep but not the frequency (Figure 5B), because it does not have control over the mode of operation of DRA.

In third phase, instead of using superstrate, a dielectric reflector is used below the ground plane substrate. The practical reason behind this approach is stated here. It is well-known that, when EM wave travel from one permittivity to another permittivity area then reflection as well as refraction occurs.[20,21] For clear understanding, the EM field sketch of Figure 1 (only reflector and ground plane substrate part is considered here) is shown in Figure 6.

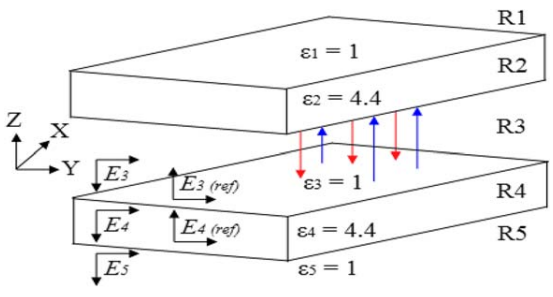


FIGURE 6 Working sketch of reflector plane

In Figure 6, R2, R4 are ground plane and reflector plane substrate ($\varepsilon_r = 4.4$), respectively, and the rest parts like R1/R3/R5 are air filled portion ($\varepsilon_r = 1$). As per Figure 1, when feed line excites the DR, some EM wave (say E_3) (Figure 6) releases to R3 and directly strikes on R4. As $\varepsilon_3 \neq \varepsilon_4$, then there must be some reflection and this can be estimated mathematically by Equation 1,[23] This indicates that the interface (of R3 and R4) reflects back ~40% of the incident wave and the rest 60% are refracted to R4. Similarly, in the very next interface (of R4 and R5) the reflection can be calculated by Equation 2, which reflects ~35% of the incident wave. These results also indicate that, the higher is the dielectric constant, the higher is the reflected wave. It can be noted that, the negative symbol indicates the 180° phase shift only.

$$E_{3(\text{ref})} = E_3\left[\frac{\sqrt{\varepsilon_3/\varepsilon_4}-1}{\sqrt{\varepsilon_3/\varepsilon_4}+1}\right] = E_3\left[\frac{\sqrt{1/4.4}-1}{\sqrt{1/4.4}+1}\right] = -0.36E_3 \quad (1)$$

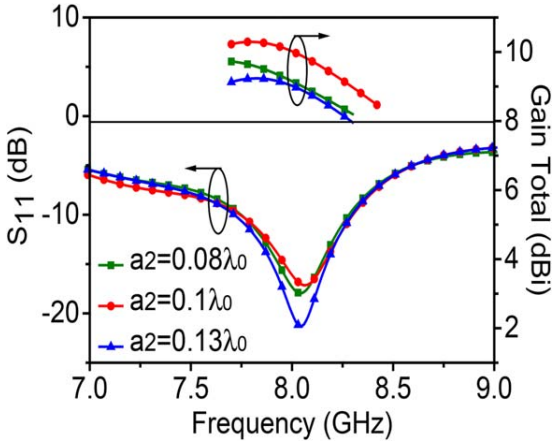


FIGURE 7 Effect of reflector airgap on S_{11} and gain total

$$E_{4(\text{ref})} = E_4\left[\frac{\sqrt{\varepsilon_4/\varepsilon_5}-1}{\sqrt{\varepsilon_4/\varepsilon_5}+1}\right] = E_4\left[\frac{\sqrt{4.4/1}-1}{\sqrt{4.4/1}+1}\right] = 0.35E_4 \quad (2)$$

Equations 1 and 2, verifies the reflective characteristics of those interfaces. Similarly, the transmitted wave to incident wave ratio can be calculated as Equations 3 and 4

$$\frac{E_4}{E_3} = \left[\frac{2}{\sqrt{\varepsilon_4/\varepsilon_3}+1}\right] \quad (3)$$

and

$$\frac{E_5}{E_4} = \left[\frac{2}{\sqrt{\varepsilon_5/\varepsilon_4}+1}\right] \quad (4)$$

Hence, the validation of our objective can be done by examining the ratio of power in R5 and R3. It is known that, for a plane wave, Poynting's vector can be stated as[23,24]:

$$P_3 = \frac{E_3^2}{\eta_3} \text{ and } P_5 = \frac{E_5^2}{\eta_5} \quad (5)$$

where,

$\eta_3 = \sqrt{\mu_0/\varepsilon_0\varepsilon_3}$ and $\eta_5 = \sqrt{\mu_0/\varepsilon_0\varepsilon_5}$

Hence, the desired ratio (P_5: P_3) can be written as[24,25]:

$$\frac{P_5}{P_3} = \left(\frac{4}{\left[\sqrt{(\varepsilon_3/\varepsilon_4)}+1\right]\left[\sqrt{(\varepsilon_4/\varepsilon_5)}+1\right]}\right)^2 \cdot \sqrt{\frac{\varepsilon_5}{\varepsilon_3}}$$

$$\Rightarrow \frac{P_5}{P_3} = \left(\frac{4}{\left[\sqrt{(1/4.4)}+1\right]\left[\sqrt{(4.4/1)}+1\right]}\right)^2 \cdot \sqrt{\frac{1}{1}} \quad (6)$$

$$\Rightarrow \frac{P_5}{P_3} = 0.77 \Rightarrow P_5 : P_3 = 0.77 : 1$$

This ratio can be decreased by choosing high dielectric materials (i.e., high ε_4). Hence, it is proved that the proposed arrangement is productive in terms of checking the back-propagating energy. It can be noted that, when $\varepsilon_4 \to \infty$ then the interface of R3 and R4 becomes "short circuit" and interface of R4 and R5 becomes "dielectric open,"[23] as all the waves are reflected at the first interface only. As in the present study, a lossy FR-4 sheet is taken a copper coating of 0.035 mm is given at the bottom side of the reflector sheet to avoid the losses.

Post theoretical analysis, this model (DRA with reflector) is simulated using finite element solver HFSS v13.0.[22] A one side etched FR-4 sheet of $l_3 = w_3 = 50$ mm, $h_3 =$ 1.6 mm, that is, slightly larger than ground plane substrate is taken to cover its bottom side properly. Initially this was placed at a random distance $a_2 = 5$ mm. Hence, as the third important parameter air-gap (a_2) is tuned between 3 mm and 5 mm (Figure 7). This indicates, the air-gap does not affect resonance frequency but the gain. It is because of high trapping and reflection of EM wave with decrease in air-gap. Apart from this, some minimum distance is needed for accommodation of the coaxial feed. So keeping all these aspects in mind, a_2 is fixed at 4 mm, which gives ~6.4% bandwidth and 9.9 dBi gain (Figure 7).

In the final and fourth stage, the authors have conceived a hybrid DRA geometry (Figure 1), that is, the combination of discussed standalone DRA with superstrate and reflector simultaneously. This comes out with an improved performance in terms of bandwidth and gain. The comparison of bandwidth, gain, and radiation pattern of all antenna types are depicted in Figure 8, which is self-explanatory. The

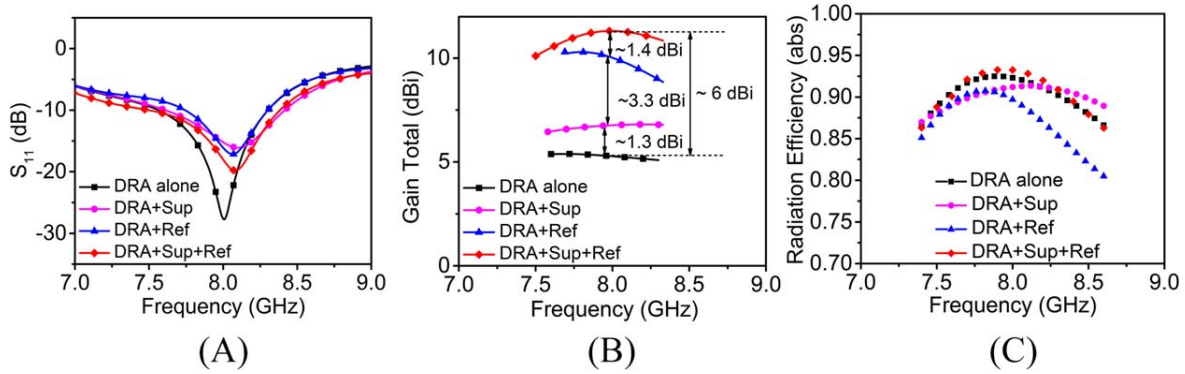


FIGURE 8 Performance comparison of all four configurations, A, S_{11}, B, gain total, C, radiation efficiency

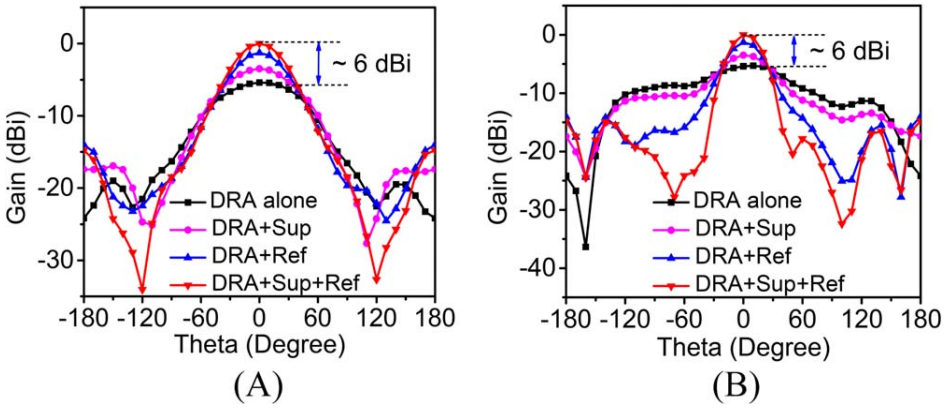


FIGURE 9 Comparative graph of gain total with and without reflector, A, $\phi = 0°$, B, $\phi = 90°$

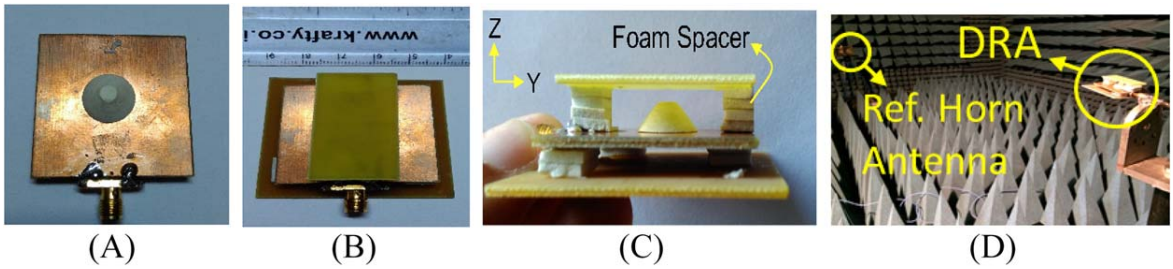


FIGURE 10 Fabricated prototype, A, top view of conical DRA only, B, C, top view and side view of the optimized DRA, respectively, D, measurement setup inside anechoic chamber

optimized antenna operates over 12.65% bandwidth (7.4 GHz-to-8.4 GHz) (Figure 8A) with 11.31 dBi peak gain (Figure 8B), and ~90% radiation efficiency (Figure 8C). The gain total radiation pattern of all antennas is compared in Figure 9. It confirms that, the antenna becomes more directional in broad side ($-30° \leq \theta \leq 30°$) by redirecting the back radiation ($-150° \leq \theta \leq 150°$). From these curves (Figure 9), it can be confirmed that, the improved performance is achieved, because of the superstrate which creates/generates high magnitude EM waves because of multiple reflection, and the reflector which redirects the back directed EM waves in broad side direction. In the hybrid combination, both superstrate and reflector effects are combined to reach peak 11.31 dBi gain. It can be noted that, even after such significant improvement of gain (Figure 8B), efficiency (Figure 8C), and far-field characteristics, the resonant frequency and band width (Figure 8A) remains almost unchanged which is rarely available in case of other techniques for gain improvement. Hence, this configuration is taken as the optimized one and considered for fabrication followed by characterization.

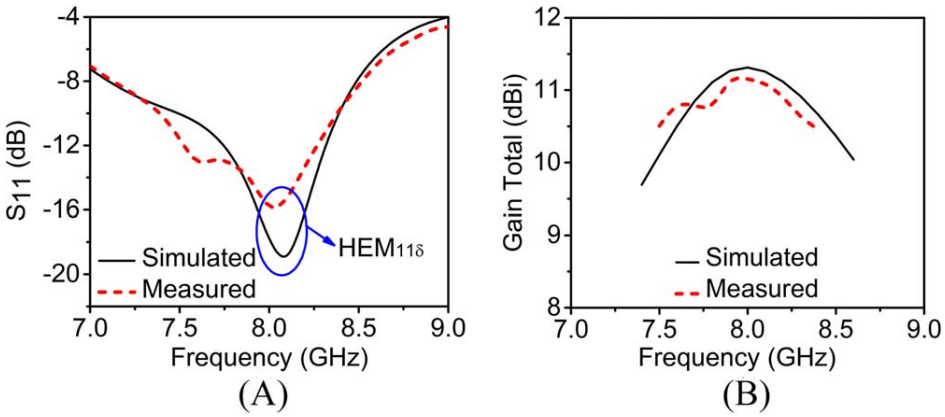


FIGURE 11 Measured and simulated results, A, S_{11}, B, gain total, and simulated radiation efficiency

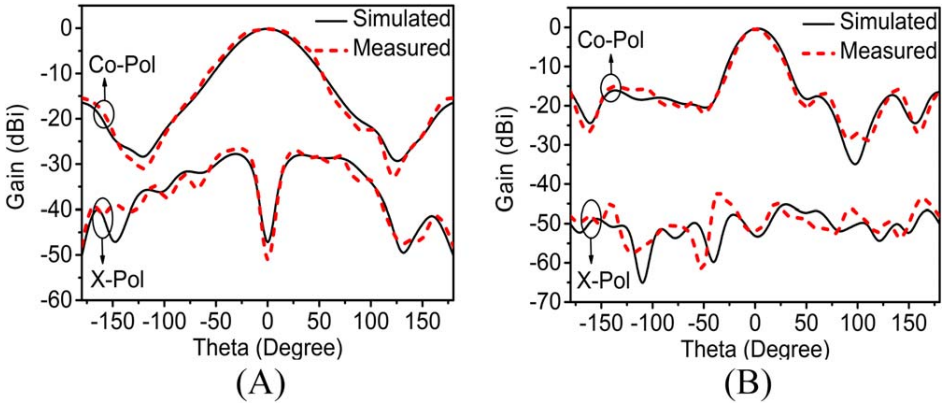


FIGURE 12 Measured and simulated radiation characteristics at 8.08 GHz, A, φ = 0°, B, φ = 90°

4 | PROTOTYPE, CHARATERIZATION, AND VALIDATION

A prototype fabricated through ordinary economic process as shown in Figure 10 has been used for characterization and validation of the proposed model. The DR is shaped out from a dielectric rod of $\varepsilon_r = 10$ and $\tan\delta = 0.003$ (Eccostock HiK) by diamond cutting tool. The three sets of dielectric sheets (ground plane substrate, superstrate, and reflector) of appropriate dimension have been made from FR-4 sheet ($\varepsilon_r = 4.4$ and $\tan\delta = 0.02$) and photolithography etching process has been applied for copper removal as required as per section 2. Vector Network Analyzer and an automatic anechoic chamber have been used for characterization of different performance parameters of the prototype.

Figure 11A reveals very good agreement between the simulated and measured reflection coefficient. The slight deviation of measured S_{11} resonance deep (due to $HEM_{11\delta}$ mode[26]) than the predicted one near 8 GHz can be attributed to unconsidered air-gap between DR and ground plane due to the roughness of DR surface. The comparison of measured and simulated gain is depicted in Figure 11B, and the attained values over the matching bandwidth is good. The peak gain is observed to be 11.25 dBi near 8 GHz compared to minimum 10.6 dBi near 7.4 GHz. It can be noted that, the peak gain is attained in higher order $HEM_{11\delta}$ mode. The simulated radiation efficiency (Figure 11B), shows ~90% efficiency almost over the entire band. The far-field radiation patterns are examined inside anechoic chamber with the help of Analog Signal Generator and Spectrum Analyzer. The spherical scanning surface technique has been used to measure the radiation. During measurement, the reference horn antenna alignment and position are kept fixed. The test antenna (DRA) is placed at a distance of 1 meter from the horn antenna and rotated along *E*-plane and *H*-plane to measure the received power (in dBm) at different angles. Later, using Friis Transmission formula, the gain value (in dBi) has been extracted. Then, the respective gain values have been normalized and compared with simulated values in Figure 12. The measurement setup of the antenna inside Anechoic chamber is depicted in Figure 10D. Both *E*-plane ($\phi = 0°$) and *H*-plane ($\phi = 90°$) radiation characteristics at 8.08 GHz show good agreement almost in all angles except little variation near $-50° \leq \theta \leq -90°$ in *H*-plane (Figure 12B). This might be due to either the misalignment between the reference

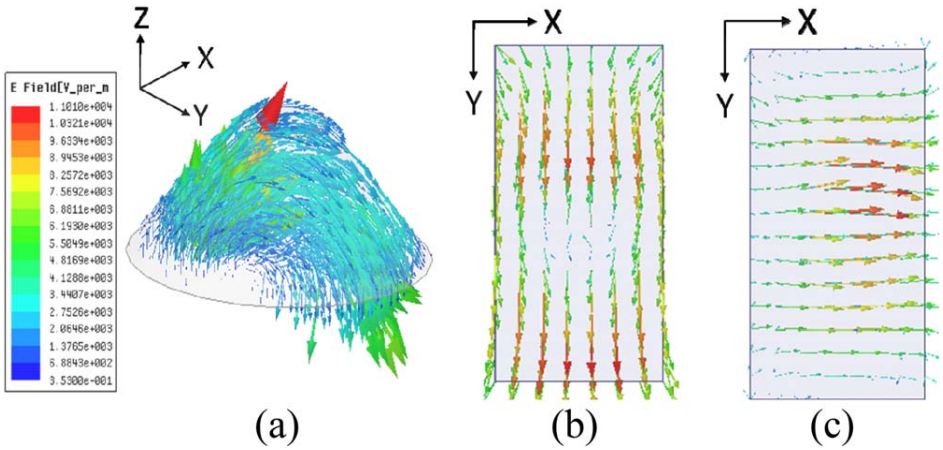


FIGURE 13 Field distribution at 8.08 GHz; A, *E*-field in DR indicates $HEM_{11\delta}$ mode, B, *E*-field in superstrate, and C, *H*-field in superstrate

TABLE 2 Comparison of developed model with published models

Ref.	DR volume (mm^3)	Bandwidth (GHz)	Peak gain (dBi)	Technique
[7]	8430.3	(3.4–4.13)	8.1	Multiple stacking + Higher order mode
[8]	15033.43	(3.15–3.72)	8.4	16 layers of stacking
[9]	614.68	(6.07–7.52)	9.0	Surface mounted horn
[10]	8128.87	6% (at 2.31)	~8.5	EBG on substrate
[11]	750	~(10.35–11.05)	10.2	Excitation of $TE_{\delta 15}$ mode
[12]	3141.59	$f_1 = 5.98$, $f_2 = 7.4$	8.5	Excitation of $HEM_{12\delta}$ mode
[13]	3098	12%, centered at 7.4 GHz	10.3	Excitation of $HEM_{12\delta}$ mode
[14]	9057.21	(4.38–4.72)	~7.5	Excitation of $HEM_{12\delta}$ mode
[15]	6902.08	(4.9–5.98)	~8.5	Excitation of $HEM_{12\delta}$ mode
[16]	3141.59	(3.6–4.35)	12.2	Metal superstrate
[17]	925.83	(3.75–6.75)	11.5	Metal superstrate
[18]	2890.08[a]	Multiple	[b]	Reflector + Array
[19]	307.88	(5.32–5.44)	~4.5[c]	Metal parabolic reflector[d]
[20]	307.88	(5.32–5.52)	~5.5[c]	Metal plate reflector[d]
[21]	307.88	(5.32–5.52)	~7.5[c]	Metal box reflector[d]
This one	**350.81**	**(7.4–8.4)**	**11.25**	**Superstrate + Reflector**

[a]Volume with respect to single element, [b]10.2 dBi → 4 element (*E*-plane array), 12.5 dBi → 4 element (*H*-plane array), and 17.6 dB → 16 element (Planar arrays), [c]Directivity, [d]Simulated.

horn antenna and the testing prototype or placement of foam spacers. The difference between the measured co-pol and cross-pol is adequately high (nearly 40 dB in both planes), which indicates application versatility. The field distribution in DR as well superstrate is shown in Figure 13.

5 | COMPARATIVE ANALYSIS

This section is dedicated for a qualitative comparison of the proposed antenna with some earlier variants. Some of the collected DRA literature(s) based on different techniques are tabulated in Table 2. The comparison is done on the basis of physical compactness, performances, and techniques. Table 2 clearly reveals that, the proposed configurations (DRA with superstrate and reflector) tops in terms of compactness and fabrication viability as well as performances.

6 | CONCLUSION

This study has targeted the gain improvement of a CoDRA by means of a compact hybrid arrangement of superstrate and reflector. The concept of multiple reflection and redirection/reduction of back radiation in broadside direction have been united for actualization of ~109% improved gain, without much altering the operating bandwidth. The proposed antenna operates over 12.65% bandwidth (7.4 GHz-to-8.4 GHz) with 11.25 dBi peak gain compared to 9.20% (7.57 GHz-to-8.3 GHz) with 5.38 dBi in case of conventional one. It can be noted that, although this arrangement looks little tricky, it is not at all a difficult task for professional manufacturer. This proposed hybrid configuration should usher the DRA investigators in developing high gain DRA with simple DR shapes. From application point of view, this can be used at the ground station for uplink data transmission (in X-band) to satellite. However, commercial antenna manufacturer should create proper scope for its optimum utilization.

REFERENCES

[1] Luk KM, Leung KW. Dielectric Resonator Antennas. Hertford-Shire, UK: Research Studies; **2002**.

[2] Petosa A. Dielectric Resonator Antenna Handbook. Norwood, MA: Artech Publication House; **2007**.

[3] Long SA, Mcallister MW, Shen LC. The resonant cylindrical dielectric cavity antenna. *IEEE Trans Antennas Propag.* **1983**; 31:406–412.

[4] Dash SKK, Khan T, De A. Modelling of dielectric resonator antennas using numerical methods: a review. *J Microwave Power Electromagn Energy.* **2016**;50(4):269–293.

[5] Dash SKK, Khan T, De A. Dielectric resonator antennas: an application oriented survey. *Int J RF Microwave Comput Aided Eng.* **2017**;27(3):e12069.

[6] Pan YM, Zheng SY. A low-profile stacked dielectric resonator antenna with high-gain and wide bandwidth. *IEEE Antennas Wireless Propag Lett.* **2016**;15:68–71.

[7] Fakhte S, Oraizi H. Compact uniaxial anisotropic dielectric resonator antenna operating at higher order radiating mode. *Electron Lett.* **2016**;52:1579–1580.

[8] Fakhte S, Oraizi H, Matekovits L, Dassano G. Cylindrical anisotropic dielectric resonator antenna with improved gain. *IEEE Trans Antennas Propag* **2017**;65(3):1404–1409.

[9] Nasimuddin, Esselle KP. A low-profile compact microwave antenna with high gain and wide bandwidth. *IEEE Trans Antennas Propag.* **2007**;55:1880–1883.

[10] Denidni TA, Coulibaly Y, Boutayeb H. Hybrid dielectric resonator antenna with circular mushroom-like structure for gain improvement. *IEEE Trans Antennas Propag.* **2009**;57:1043–1049.

[11] Petosa A, Thirakoune S. Rectangular dielectric resonator antennas with enhanced gain. *IEEE Trans Antennas Propag.* **2011**;59: 1385–1389.

[12] Guha D, Banerjee A, Kumar C, Antar YMM. Higher order mode excitation for high-gain broadside radiation from cylindrical dielectric resonator antennas. *IEEE Trans Antennas Propag.* **2012**;60:71–77.

[13] Gupta P, Guha D, Kumar C. Dielectric resonator working as feed as well as antenna: new concept for dual-mode dual-band improved design. *IEEE Trans Antennas Propag.* **2016**;64:1497–1502.

[14] Sharma A, Das G, Ranjan P, Sahu NK, Gangwar RK. Novel feeding mechanism to stimulate triple radiating modes in cylindrical dielectric resonator antenna. *IEEE Access.* **2016**;4:9987–9992.

[15] Sharma A, Ranjan P, Gangwar RK. Multiband cylindrical dielectric resonator antenna for WLAN/WiMAX application. *IET Electron Lett.* **2017**;53:132–134.

[16] Dutta K, Guha D, Kumar C, Antar YMM. New approach in designing resonance cavity high-gain antenna using nontransparent conducting sheet as the superstrate. *IEEE Trans Antennas Propag.* **2015**;63:2807–2813.

[17] Fakhte S, Oraizi H, Samiei MHV. A high gain dielectric resonator loaded patch antenna. *Prog Electromagn Res C.* **2012**;30: 147–158.

[18] Nikkhah MR, Kishk AA, Mohassel J-R. Wideband DRA array placed on array of slot windows. *IEEE Trans Antennas Propag.* **2015**;63:5382–5390.

[19] Ali SM, Zakariya MA, Baharudin Z, Adz JJ. Parabolic reflector fed by cylindrical dielectric resonator antenna with high gain. In: Proceedings of IEEE 5th International Conference on Intelligent and Advanced Systems; June **2014**.

[20] Ali SM, Zakariya MA, Baharudin Z, Md. Khir MH, Baba AA, Adz JJ. Aperture coupled dielectric resonator antenna design with flat reflector. In: Proceedings of IEEE Business Engineering and Industrial Applications Colloquium; April **2013**.

[21] Ali SM, Zakariya MA, Baharudin Z Z, Khir MHM, Baba AA, Adz JJ. Aperture coupled cylindrical dielectric resonator antenna design with box shaped reflector. In: Proceedings of IEEE Symposium on Wireless Technology and Applications; September **2013**.

[22] Ansoft. High Frequency Structure Simulator (HFSS) v13.0; **2013**.

[23] Rao Q, Denidni TA, Johnston RH. Dielectric reflector backed aperture-coupled antennas for reduced back radiation. *IEEE Trans Electromagn Compat.* **2006**;48:287–291.

[24] Harrington RF. Field Computation by Moment Methods. New York: Macmillan; **1968**.

[25] Bladel JV. Electromagnetic Fields. New York: Hemishpere, **1985**.

[26] Kishk AA, Yin Y, Glisson AW. Conical dielectric resonator antennas for wide-band applications. *IEEE Trans Antennas Propag.* **2002**;50:469–474.

AUTHOR BIOGRAPHIES



Sounik Kiran Kumar Dash received his Bachelor Degree in Electronics and Communication Engineering from Centurion University of Technology and Management, Bhubaneswar, India in 2013 and Master Degree with specialization in Communication System Engineering from KIIT University, Bhubaneswar, India in 2015. At present he is towards his doctorate degree at National Institute of Technology Silchar, India. His current research interests include Dielectric Resonator Antennas (DRAs), MIMO Antennas, Reconfigurable Antennas, Numerical Methods in Electromagnetics, and Neural Networks. He has authored or co-authored eleven research articles in international journals and international conference/symposium proceedings of repute. He is an active student member IEEE, member IAENG (China).



Taimoor Khan obtained his Polytechnic Diploma in Electronics Engineering, Bachelor Degree in Electronics and Communication Engineering and Master Degree with specialization in Communication Engineering from Board of Technical Education, Lucknow, India, The Institution of Engineers (India) Kolkata, India and Shobhit University Meerut, India in the year 2001, 2005 and 2009,

respectively. He was Awarded PhD Degree in Electronics and Communication Engineering from National Institute of Technology Patna, India in the year 2014. Presently, he is working as an Assistant Professor at National Institute of Technology Silchar, India. Prior to joining this institute, he served Delhi Technological University, Delhi, India as Assistant Professor more than two years, Netaji Subhas Institute of Technology Patna for more than a year and more than nine years to Shobhit University, Meerut as Lab Instructor, Lecturer and Assistant Professor. During September-December 2016, Dr. Khan has also worked as Visiting Assistant Professor in Telecommunication Field of Study at Asian Institute of Technology Bangkok, Thailand. His research interest includes Microstrip Antennas, Microwave Engineering and Electromagnetics and Artificial Intelligence Paradigms in Electromagnetics. He has published over twenty five research papers in international journals and international/national conference proceedings of repute. He is an active senior member of IEEE (USA), life member of IEI (India) and life member of IAE (USA).

How to cite this article: Dash SKK, Khan T. Wideband high gain conical dielectric resonator antenna: An experimental study of superstrate and reflector. *Int J RF Microw Comput Aided Eng*. 2017;e21140. https://doi.org/10.1002/mmce.21140