APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY JOURNAL

August 2021
Vol. 36 No. 8
ISSN 1054-4887

The ACES Journal is abstracted in INSPEC, in Engineering Index, DTIC, Science Citation Index Expanded, the Research Alert, and to Current Contents/Engineering, Computing & Technology.

The illustrations on the front cover have been obtained from the research groups at the Department of Electrical Engineering, The University of Mississippi

AUGUST 2021 REVIEWERS: REGULAR PAPERS

Mehmet Belen
Guan-Yu Chen
Pasquale Dottorato
Mohsen Eslami Nazari
Andrey Grigoriev
Mohammed Hadi
Julie Huffman
Tianqi Jiao
Zhu Jun
Long Li
Yingsong Li
Shengyuan Luo
Michel Ney
Mahdi Oliaei
Balachandra Pattanaik
Andrew Peterson
Mohammd Pourbagher
Dudla Prabhakar
Azhagumurugan R.
Luis Ramirez
Imaculate Rosaline
Abas Sabouni
Tale Saeidi
Luca Sani
Stefano Selleri
Yuewu Shi
Hamidreza Siampour
Yi Sui
Prabhu Sundaramoorthy
Mehmet Tabakcioglu
Junwu Tao
Christopher Trueman
Jiangong Wei
Julia Wolff
Huanhuan Yang
Hao Zhang
Mohammad Zoofaghari
Le Zuo

AUGUST 2021 REVIEWERS: EXPRESS PAPERS

Ercument Arvas
Quang Nguyen
Huthaifa Obeidat
Shinichiro Ohnuki
Hassan Ragheb
Mariselvam V.

THE APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY JOURNAL

Vol. 36 No. 8 August 2021

TABLE OF CONTENTS – REGULAR PAPERS

TABLE OF CONTENTS – EXPRESS PAPERS

A Novel Design of Aperiodic Arrays for Ultrawideband Beam Scanning and Full Polarization Reconfiguration

Ziyu Zhang [1], Jia Liu [1], Jianxun Su [1,*], and Jiming Song [2]

[1] State Key Laboratory of Media Convergence and Communication
Communication University of China, Beijing, 100024, China
*sujianxun_jlgx@163.com

[2] Electrical and Computer Engineering
Iowa State University, Ames, IA 50011, USA

Abstract ─ In this letter, a multifunction aperture array is proposed for ultrawideband (UWB) scanning and polarization reconfiguration. The UWB array consisted of linearly polarized elements is capable of operating in four polarization modes (+45° linear polarization (LP), -45° linear polarization, left-hand circular polarization (LHCP) and right-hand circular polarization (RHCP)). This work involves two essential techniques: (a) A new beam-scanning UWB array synthesis approach. An iterative convex optimization strategy is utilized to determine the element locations and obtain the minimum sidelobe level (SLL) for multiple patterns. (b) The polarization reconfigurable technique for beam-scannable arrays. In this part, a sequential rotation and excitation compensation (SR-EC) technique provides polarization reconfiguration for a beam-scannable array consisting of linearly polarized elements. A beam-scanning UWB array is designed by using the proposed UWB array synthesis approach and the SR-EC polarization reconfigurable technique. The Feko numerical result shows 0°-60° beam peak steering, a 4:1 bandwidth (1-4 GHz), and four-polarization reconfigurability.

Index Terms ─ Array synthesis, beam scanning, polarization reconfiguration, ultrawideband.

I. INTRODUCTION

With the development of wireless systems, ultrawideband (UWB) array antennas have attracted significant interests due to their potential of realizing multiple functions within one radiating aperture. In addition to the bandwidth, other requirements are needed to meet in practical multifunction systems, such as wide-angle beam steering, and multi-polarization, etc. It is a challenge to design a single radiating aperture that fulfills all the requirements.

In the literature, many methods have been proposed to develop a wideband array antenna with beam scanning. One way is to use a tightly coupled array antenna to provide ultrawideband (UWB) characteristics and beam scanning, such as in [1-4]. Although tightly coupled array antennas have many advantages, such as high aperture efficiency and wide-angle beam scanning, some drawbacks, such as sophisticated design techniques and high system cost, are also inevitable. Another way of designing UWB array antennas is to adopt nonuniformly spaced techniques. In [5-6], the stochastic optimization algorithms are used to design array antennas. In [7-8], a general SA design approach has been proposed based on polynomial model of an array. An l_0-norm constrained normalized least-mean-square adaptive beamforming algorithm for controllable sparse antenna arrays was presented in [9]. Many studies have been done to develop aperiodic array configurations. However, most aperiodic array synthesis methods focus on the problem of suppressing the grating-lobe and sidelobe levels (SLLs) at a single frequency. In [10], by virtue of the concept of design frequency, the proposed method transforms the beam-scannable UWB design to a problem of synthesizing a broadside fixed-beam array at a single frequency. In [11], the SLL is optimized at the highest operating frequency of the designed array. In the above method, although the sidelobe level is lower at the frequency, it will be uncontrolled at other frequencies. Thus, choosing the best common element positions for meeting lower SLL in the whole frequency band of interest is challenging. A novel array synthesis approach is proposed to achieve it. The optimization method realizes accurate control of the SLL for multiple patterns and guarantees the minimum element spacing in the optimization process. For the multiple constraint problem, an iterative convex optimization strategy is utilized to solve the problem. Specific principles will be introduced in Section II.

In recent years, many polarization reconfigurable antennas have been proposed [12-16]. However, few researches have been done on polarization reconfigurable and UWB beam-steering antenna arrays. An antenna

Submitted On: May 14, 2021
Accepted On: July 6, 2021

https://doi.org/10.47037/2021.ACES.J.360801

array that provides wideband frequency agility with simultaneous polarization reconfiguration is presented in [17]. The antenna has the ability of ±28° beam peak steering at 2.4 GHz for both the linear and circular polarizations. The bandwidth and scanning range are narrow.

This paper presents an ultrawideband and polarization reconfigurable electric scanning array, which is constructed by sequential rotating linearly polarized elements. The sequential rotation and excitation compensation (SR-EC) technique is an attractive approach to retain multiple polarization purity for beam-scannable arrays. The specific method will be discussed in Section III. And, the simulated results are given in Section IV.

II. A NOVEL BEAM-SCANNING UWB ARRAY SYNTHESIS APPROACH

For a linear array consisting of N elements whose elements are located at [x_0, x_1, . . . , x_{N-1}] along with x-axis. Assume that this array works at the frequency f, with a focused beam scanned in θ_o. The array factor can be expressed as the following:

$$AF(\beta,u)=\sum_{n=0}^{N-1}\exp(j\beta x_n u), \tag{1}$$

where $\beta=2\pi f/c$, $u=\sin\theta-\sin\theta_o$, and $\theta\in[-\pi/2, \pi/2]$ is the wave propagation direction measured from the z-axis. For a beam-scanning UWB array, the highest and lowest frequency is f_H and f_L, with a focused beam scanned within the range of [$-\theta_{max}$, θ_{max}]. From (1), the beam-scannable UWB array factor has the following property:

$$\{AF([\beta],[u])\}=\sum_{n=0}^{N-1}\exp(j\beta x_n u), \tag{2}$$

where $\{AF([\beta], [u])\}$ denote the range of the function $AF(\beta, u)$ varying with $\beta\in[\beta_L, \beta_H]$ and $u\in[-1-\sin\theta_{max}, 1+\sin\theta_{max}]$, $\beta_L=2\pi f_L/c$, $\beta_H=2\pi f_H/c$. This means that the design of a beam-scannable UWB array can be equivalent to a design of fixed equation. In general, one should check the scannable beam performance in the whole frequency band of interest for the beam-scannable UWB array design problem. The increased computational and storage loads is necessary to deal with requirements of multiple beam scanning angles at multiple frequencies for the beam-scannable UWB array design.

$\{AF([\beta], [u])\}$ indicates the array factor which included all the beam characteristics of the UWB array in the whole frequency band $[f_L, f_H]$ for all beam scanning cases for $\theta_o\in[-\theta_{max}, \theta_{max}]$. Thus, we can consider designing the array factor $\{AF([\beta], [u])\}$ to check all the beam performance of a scannable UWB array. For most of UWB array synthesis cases, the element spacing and the element count N can be determined by considering the antenna element structure and gain requirement in applications.

In order to achieve better UWB array performance, an iterative convex optimization strategy is developed to successively optimize the UWB array SLL. The basic idea is introducing the elememt spacing variables d_m (m = 1, . . . , N-1) to the original positions x_n (n = 0, 1, . . . , N-1) and then find a way to obtain the optimal d_m. The element position and the element spacing can be shown as:

$$\begin{cases} x_0=0 \\ x_n=\sum_{m=1}^{n} d_m \ (n=1,2,\cdots,N-1)\cdot \end{cases} \tag{3}$$

Consider the array factor $\{AF([\beta], [u])\}$ associated with varying positions x_n (n = 0, 1, . . . , N-1). That is given by

$$\{AF([\beta],[u])\}=1+\sum_{n=1}^{N-1}\exp(j\beta\sum_{m=1}^{n} d_m u). \tag{4}$$

Now the variables d_m can be optimized by using convex optimization to reduce the overall SLL. However, for a practical UWB array design problem, the following two kinds of constraints should be considered. Typically, the minimum element spacing should be constrained to avoid strong mutual coupling, which can have a detrimental effect on the driving port impedances. In addition, the array aperture size is very significant in practical applications. Thus, the maximum element spacing should also be constrained:

$$a\le d_m\le b \quad m=1,2,\cdots N-1\ , \tag{5}$$

where, a is the minimum element spacing, and b is the maximum element spacing. The optimal element spacing variables can be found by solving the following constrained optimization problem:

$$\min\ \varepsilon$$
$$\text{s.t.}\begin{cases} \max(1+\sum_{n=1}^{N-1}\exp(j\beta\sum_{m=1}^{n} d_m u_o))\le\varepsilon \\ u_o\in U_{sl} \\ \beta\in[\beta_L,\beta_H] \\ a\le d_m\le b \quad m=1,2,\cdots,N-1 \end{cases}, \tag{6}$$

where U_{sl} denotes the sidelobe region in u_o-space. Clearly, the above problem can be efficiently solved by using the iterative convex optimization strategy. Each iteration finds locally optimal element spacing and thus the element spacing can be successively updated until the obtained SLL maintains the same for many times. Besides, it should be mentioned that all the excitation amplitudes remain one in the element spacing optimization process.

The proposed beam-scannable UWB array design procedure is given in Algorithm 1. In this procedure, the

initial element spacing is randomly generated. Then a number of constrained convex optimizations given in (6) are performed to successively update the element spacing, such that the obtained maximum SLL can be reduced as much as possible.

Algorithm 1:

1: Set the frequency band $[f_L, f_H]$, beam scanning range $[-\theta_{max}, \theta_{max}]$, and element number parameter N.

2: Depending on the application requirement, set the range of element spacing $[a, b]$.

3: Set k = 0. Sample the space of u_o, frequency band, and beam scanning range, and initialize element spacings d_m (m = 1,2, . . . , N-1).

4: k = k + 1.

5: Solve the constrained convex optimization problem (6) to find the better element spacings d_m for m = 1, 2, . . . , N-1.

6: Update the element spacings, compute the pattern $\{AF([\beta], [u_o])\}$ in (4) with the new element spacings and obtain the maximum SLL ε.

7: Repeat Steps 4 to 6 until ε remains the same for multiple times or k reaches the allowable maximum number of iterations.

8: Output the obtained element spacings d_m for m = 1,2, . . . , N-1 and calculate the beam-scannable wideband pattern $\{AF([\beta],[u])\}$ in (4).

We consider synthesizing a UWB array consisted 51-element (N=51) occupying the frequency band from 1 to 4 GHz with beam scanning from 0° to 60°. That is $[\beta] = [2\pi f_L/c, ..., 2\pi f_H/c]$, $[\theta_o] = [\ 0°, ..., 60°]$. Set $[\theta] = [\ -90°, ..., 90°]$. Then $[u] = [u_{min}, ..., u_{max}]$, where $u_{min} = -1-\sqrt{3}/2$, $u_{min} = 1$. We set the minimum element spacing is $\lambda_L/2$ and the maximum element spacing is λ_L, where λ_L is the wavelength at the lowest frequency. Set the initial element spacings $\boldsymbol{d} = [\lambda_L/2, \lambda_L/2, ..., \lambda_L/2]$.

The antenna pattern is calculated according to (4) and the above data. The pattern contains both operating frequencies and beam-scanning information. Extract the maximum value SLLmax from the sidelobe information of the pattern. Change the element spacing $\boldsymbol{d}$ within the constraint range to obtain anthoer maximum value SLLmax. If the maximum value reduces, the element spacings will continue to change in this direction, otherwise the other direction will be chosen.

Repeat the above steps until SLLmax remains the same for multiple times or iterations reaches the allowable maximum number.

To show the effectiveness and robustness of the proposed method, the comparisons with the UWB array and a uniformly spaced array are presented here. The uniformly spaced array has the same number of element and aperture size as the UWB array. The optimized element locations and uniformly element locations are illustrated in Fig. 1. The minimum spacing of the optimized array is $0.5043\lambda_L$, and the aperture size is $33.3\lambda_L$. The element spacing of the uniformly spaced linear arrays is $0.6656\lambda_L$. Figure 2 shows that the maximum SLL of the optimized array scanned patterns is -11.4 dB. However, the maximum SLL of the uniformly spaced array scanned patterns is 0 dB. We can see that compared to the uniformly spaced array, the proposed UWB array can significantly reduce the maximum SLL while maintaining almost the same array aperture and element numbers.

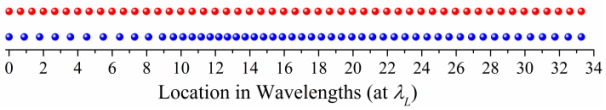


Fig. 1. Element locations of the uniformly spaced array and the optimized UWB array.

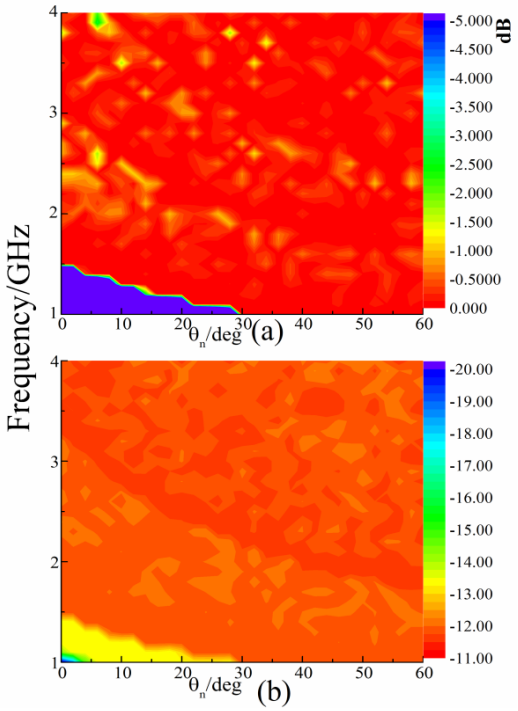


Fig. 2. The peak sidelobe levels of (a) the uniformly spaced array, and (b) the optimized UWB array in the frequency band, with steered from 0° to 60°.

III. POLARIZATION RECONGFIGURABLE TECHNOLOGY

The sequential rotation scheme was introduced by Huang in 1986 [18] to increase the axial ratio bandwidth of the circular polarization array. When arranging linearly polarized elements as shown in Fig. 3 (a), two equal orthogonal electric fields will be generated at broadside. And then, the circular polarization can be obtained with the phase excitations in 0°, 90°, 180°, 270°. Inspired by the above techniques, a novel polarization reconfigurable technique is proposed. For an array arranged as Fig. 3 (a), the four polarization (+45° LP, -45° LP, RHCP, and LHCP) can be obtained at broadside with unique phase arrangement as in Table 1 (Assuming that the co-polarized of the element is consistent with the arrow direction).

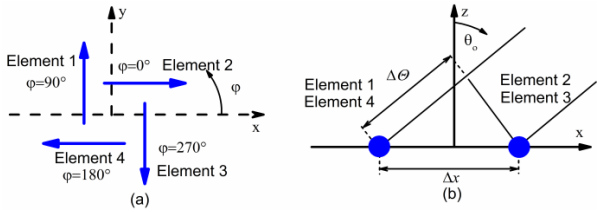


Fig. 3. (a) 2 × 2 array with elements in 0°, 90°, 180°, 270° directions. (b) Its spatial phase delay.

Table 1: The phase arrangements for four polarizations

	+45° LP	**-45° LP**	**RHCP**	**LHCP**
Element 1	0°	0°	0°	0°
Element 2	0°	180°	90°	270°
Element 3	180°	180°	180°	180°
Element 4	180°	0°	270°	90°

However, the approach is not suitable for scanning arrays. When the beam is steered away from the broadside, the polarization characteristics will severely degrade. One reason is the differences between the E-plane and H-plane patterns of linearly polarized elements. The other is that the spatial phase delay disturbs the required phase differences. The practical problem can be solved by compensating for the amplitudes and phases of the individual element [19]. We will illustrate the necessary procedure using a 2 × 2 array arranged as Fig. 3 (a). All the cases we discussed as followed are in φ=0° plane.

Suppose the polarization reconfigurability in the θ_o direction is needed, two equal orthogonal electric fields would be required in the direction of θ_o. However, the differences between the two principal patterns of linearly polarized elements contribute to two orthogonal electric fields are unequal in the direction of θ_o. The basic idea of amplitude compensation is to pre-simulate each element patterns in the array. From the simulated patterns, the required amplitude correction values can be calculated.

$E_{1\theta}(\theta_o)$ and $E_{1\varphi}(\theta_o)$ are the θ-component and φ-component due to Element 1 for the direction of θ_o in φ=0° plane, respectively. A similar explanation holds for the other field components. They are obtained by pre-simulating. $E_\theta(\theta_o)$ and $E_\varphi(\theta_o)$ represent the θ- and φ-components of the total electric field in the direction of θ_o, and can be expressed in (7) and (8). To obtain two equal orthogonal electric fields, a_j is introduced here, which represents the amplitude excitation of Element j:

$$E_\theta(\theta_o) = a_1E_{1\theta}(\theta_o) + a_2E_{2\theta}(\theta_o) + a_3E_{3\theta}(\theta_o) + a_4E_{4\theta}(\theta_o), \quad (7)$$

$$E_\varphi(\theta_o) = a_1E_{1\varphi}(\theta_o) + a_2E_{2\varphi}(\theta_o) + a_3E_{3\varphi}(\theta_o) + a_4E_{4\varphi}(\theta_o). \quad (8)$$

Set a_1=a_3 and a_2=a_4, and the amplitudes will be obtained as shown in (9) by solving $E_\theta(\theta_o)$= $E_\varphi(\theta_o)$:

$$\frac{a_2}{a_1} = \frac{\left[E_{1\theta}(\theta_o) + E_{3\theta}(\theta_o)\right] - \left[E_{1\varphi}(\theta_o) + E_{3\varphi}(\theta_o)\right]}{\left[E_{2\varphi}(\theta_o) + E_{4\varphi}(\theta_o)\right] - \left[E_{2\theta}(\theta_o) + E_{4\theta}(\theta_o)\right]}. \quad (9)$$

In addition, the spatial phase delay would disturb the required phase differential, and contribute to the poor polarization quality in the direction of θ_o as shown in Fig. 3 (b). The space phase compensation value required for Element 2 and Element 3 can be expressed as:

$$\Delta\Theta = -\beta\Delta x \sin\theta_o, \quad (10)$$

where Δx is the spacing between the two elements. For a large scanning array consisting of 2 × 2 sequential rotation subarray, the amplitude compensation values and phase arrangements required for polarization of the element are the same in different 2 × 2 subarray. However, the spatial phase delay compensation values vary from the locations of element in φ=0° plane.

IV. THE DESIGN OF A RECONFIGURABLE ANTENNA ARRAY AND SIMULATED VALIDATIONS

An aperiodic array consisting of linearly polarized elements is placed on x-y plane as Fig. 4. The ideal dipoles are used to provide linearly polarized elements. The element positions in the direction of x-aixs are derived from the optimization results in Section II. Thus, the antenna array will have the ability to accurately control the sidelobe over a 4:1 bandwidth. The element spacing in the direction of y-axis is $\lambda_L/2$, where λ_L is the wavelength at the lowest frequency.

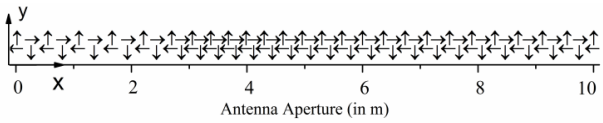


Fig. 4. The structure of the designed antenna.

According to (9), the amplitude compensation values of the ideal dipoles can be calculated, and the values are the same in different subarray as:

$$\frac{a_2}{a_1}=\frac{a_4}{a_3}=\frac{1}{\cos\theta_o}. \tag{11}$$

The phase excitation of the element consists of two parts, one is the space phase delay, and the other is the phase differential required by four polarizations. It is worth noting that the phase arrangements required for four polarizations are the same in different subarray, and it has been shown in Table 1. The space phase delay compensation values are $-\beta x_n \sin\theta_o$, where x_n is the locations of element in the direction of x-axis.

Figure 5 (a) shows that the patterns of the designed reconfigurable array antenna steered to 15°, 30°, 45°, and 55° with +45° LP, -45° LP, RHCP, and LHCP at 1 GHz. Figure 5 (b) shows that the patterns of the array antenna steered to 0°, 20°, 40°, and 60° with +45° LP, -45° LP, RHCP, and LHCP at 4 GHz. Figure 6 shows axial ratios of the circular polarized scanned patterns.

From Figs. 5-6, it can be observed that the four polarizations can be obtained at different scanning angles. The simulation results show that the method can be used to design the UWB, wide-angle scanning, and polarization reconfigurable antenna arrays. When considered a more realistic wideband linearly polarized element, we need to account for variations element pattern over bandwidth. A more excellent polarization performance will be obtained by calculating the element amplitude compensation values at different operating frequencies.

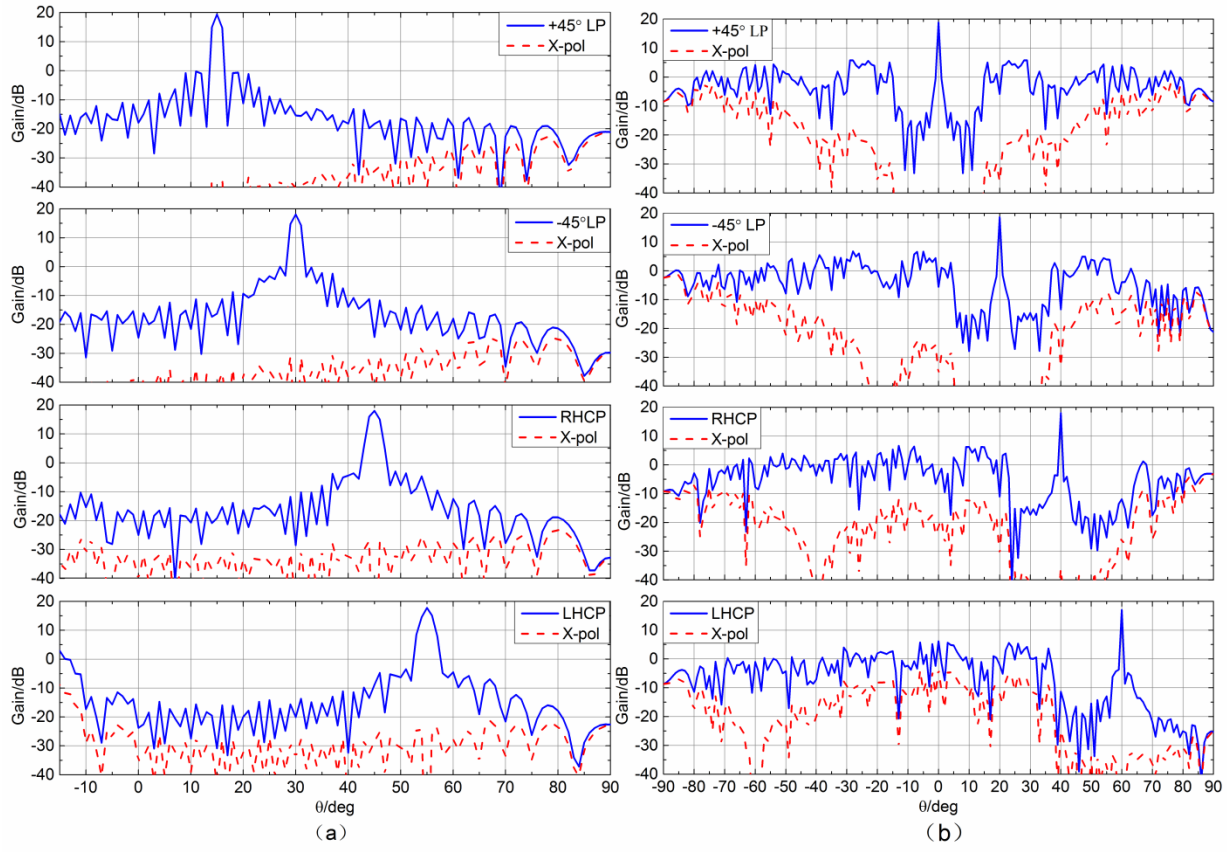


Fig. 5. (a) The radiation pattern at 1 GHz. The main beam is steered to 15° with +45° LP, 30° with -45° LP, 45° with RHCP, and 55° with LHCP. (b) The radiation pattern at 4 GHz. The main beam is steered to 0° with +45° LP, 20° with -45° LP, 40° with RHCP, 60° with LHCP.

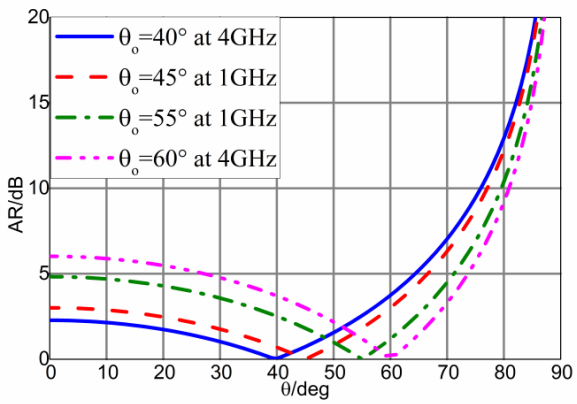


Fig. 6. Axial ratios for the circular polarized scanning.

V. CONCLUSION

A beam-scanning UWB array consisting of linearly polarized elements is proposed for polarization reconfiguration. Through optimizing multiple radiation patterns simultaneously, the UWB array synthesis approach can control the sidelobe level in the ultrawide frequency band. An iterative convex optimization strategy is used to successively optimize the element spacing for SLL reduction. The sequential rotation and excitation compensation (SR-EC) technique provides polarization reconfigurable electrical scanning ability for UWB aperiodic arrays. This paper presents an effective approach of designing multifunctional arrays.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China (61701448 and 62071436).

REFERENCES

[1] B. Wang, S. Yang, Y. Chen, S. Qu, and J. Hu, "Low cross-polarization ultrawideband tightly coupled balanced antipodal dipole array," *IEEE Trans. Antennas Propag.*, vol. 68, no. 6, pp. 4479-4488, June 2020.

[2] W. F. Moulder, K. Sertel, and J. L. Volakis, "Superstrate-enhanced ultrawideband tightly coupled array with resistive FSS," *IEEE Trans. Antennas Propag.*, vol. 60, no. 9, pp. 4166-4172, Sep. 2012.

[3] E. Yetisir, N. Ghalichechian, and J. L. Volakis, "Ultrawideband array with 70° scanning using FSS superstrate," *IEEE Trans. Antennas Propag.*, vol. 64, no. 10, pp. 4256-4265, Oct. 2016.

[4] H. Zhang, S. Yang, S.-W. Xiao, Y. Chen, S.-W. Qu, and J. Hu, "Ultrawideband phased antenna arrays based on tightly coupled open folded dipoles," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 2, pp. 378-382, Feb. 2019.

[5] R. L. Haupt, "Genetic algorithm applications for phased arrays," *Applied Computational Electromagnetics Society Journal*, vol. 21, no. 3, pp. 1054-4887, Nov. 2006.

[6] M. H. Rahmani and A. Pirhadi, "Optimum design of conformal array antenna with a shaped radiation pattern and wideband feeding network," *Applied Computational Electromagnetics Society Journal*, vol. 29, no. 1, pp. 1054-4887, Jan. 2014.

[7] W. Shi, Y. Li, and S. A. Vorobyov, "Low mutual coupling sparse array design using ULA fitting," in *Proc. ICASSP. IEEE*, pp. 4610-4614, May 2021.

[8] W. Shi, S. A. Vorobyov, and Y. Li "ULA fitting for sparse array design," arXiv e-prints, arXiv: 2102.02987, Feb. 2021.

[9] W. Shi, Y. Li, L. Zhao, and X. Liu, "Controllable sparse antenna array for adaptive beamforming," *IEEE Access*, vol. 7, pp. 6412-6423, Jan. 2019.

[10] Y. Liu, Y. Yang, F. Han, Q. H. Liu, and Y. J. Guo, "Improved beam-scannable ultra-wideband sparse antenna arrays by iterative convex optimization based on raised power series representation," *IEEE Antennas Wireless Propag. Lett.*, vol. 68, no. 7, pp. 1-1, 2020.

[11] M. D. Gregory and D. H. Werner, "Ultrawideband aperiodic antenna arrays based on optimized raised power series representations," *IEEE Trans. Antennas Propag.*, vol. 58, no. 3, pp. 756-764, Mar. 2010.

[12] Y. J. Sung, T. U. Jang, and Y. S. Kim, "A reconfigurable microstrip antenna for switchable polarization," *IEEE Antennas Wireless Propag. Lett.*, vol. 14, no. 11, pp. 534-536, Nov. 2004.

[13] C. Wenquan, Z. Bangning, L. Aijun, Y. Tongbin, G. Daosheng, and P. Kegang, "A reconfigurable microstrip antenna with radiation pattern selectivity and polarization diversity," *IEEE Antennas Wireless Propag. Lett.*, vol. 11, pp. 453-456, Feb. 2012.

[14] N. N.-T. Muhammad Ikram and Amin Abbosh, "A simple single-layered continuous frequency and polarization-reconfigurable patch antenna array," *IEEE Trans. Antennas Propag.*, vol. 10, no. 1109, pp. 1-6, 2019.

[15] Y. Fan, R. Li, and Y. Cui, "Development of polarisation reconfigurable omnidirectional antennas using crossed dipoles," *IET Microw. Antenna P*, vol. 13, no. 4, pp. 485-491, Nov. 2019.

[16] J.-S. Row and Y.-H. Wei, "Wideband reconfigurable crossed-dipole antenna with quad-polarization diversity," *IEEE Trans. Antennas Propag.*, vol. 66, no. 4, pp. 2090-2094, Apr. 2018.

[17] B. Babakhani, S. K. Sharma, and N. R. Labadie, "A frequency agile microstrip patch phased array antenna with polarization reconfiguration," *IEEE Trans. Antennas Propag.*, vol. 64, no. 10, pp. 4316-4327, Oct. 2016.

[18] J. Huang, "A technique for an array to generate circular polarization with linearly polarized elements," *IEEE Trans. Antennas Propag*., vol. 34, no. 9, pp. 1113-1124, Sep. 1986.
[19] A. B. Smolders and U. Johannsen, "Axial ratio enhancement for circularly-polarized millimeter-wave phased-arrays using a sequential rotation technique," *IEEE Trans. Antennas Propag.*, vol. 59, no. 9, pp. 3465-3469, Sep. 2011.

Fast Range Decoupling Algorithm for Metamaterial Aperture Real-time Imaging

Yuteng Gao [1], Wencan Peng [2], Min Wang [3], Chenjiang Guo [1], and Jun Ding [1]

[1] School of Electronics and Information, Northwestern Polytechnical University, Xi'an, 710129, China
ytgao2002@163.com, cjguo@nwpu.edu.cn, dingjun@nwpu.edu.cn

[2] Xi'an Institute of Space Radio Technology, (CASC, Xi'an), Xi'an, 710100, China
wcanpeng@163.com

[3] National Lab of Radar Signal Processing, Xidian University, Xi'an, 710071, China
wangmin@xidian.edu.cn

Abstract — While metamaterial aperture imaging systems do not require mechanical scanning equipment or complex components by employing a spatially variant radiation field, they require large amount of data and many computations. In this paper, we deduce the contribution of the resonator to the radiation fields of the metamaterial aperture. We propose a fast range decoupling algorithm that can improve the data processing speed and obtain real-time images of far-field scenes. The algorithm decomposes the scene into numerous range cells, drastically reduces the range of interest, and reconstructs the scene in parallel. Simulation results show that computational cost is significantly decreased and image quality is maintained.

Index Terms — millimeter-wave imaging, metamaterial apertures, compressed sensing, decoupling

I. INTRODUCTION

Millimeter-wave imaging techniques can provide high-quality imagery with nonionizing radiation and are promising in a variety of applications, such as medical diagnosis [1], bio-imaging [2], security screening [3] and through-the-wall imaging [4],[5]. Conventional millimeter-wave imaging systems usually rely on mechanical equipment to scan a static aperture. Phased arrays can generate different imaging models without mechanical scanning, yet they need many phase shifts, attenuators and other components which results in increased cost and size. In recent research, frequency-diverse aperture imaging has been shown to be a possible solution to the above problems. By generating a group of different spatial radiation fields that change rapidly with frequency, the frequency-diverse aperture antenna can encode the scene information. Many frequency-diverse antenna architectures have been demonstrated, such as metamaterial apertures [6]-[11], cavity apertures [12]-[14], and dynamic metasurface apertures [15]-[18]. Compared with dynamic metasurface apertures and cavity apertures, metamaterial apertures have the advantages of simple structure and low cost.

Although simple and inexpensive, metamaterial apertures also have flaws. First, a metamaterial aperture antenna is not easy to design, especially for millimeter wave imaging applications. The radiation patterns change with respect to frequency, and cannot be predicted in an exact way. Since Hunt et al. [6] first used waveguide-fed metamaterial apertures that consist of complementary electric-field-coupled (cELC) elements in imaging experiments, many studies have focused on new structures of metamaterial apertures. To obtain a uniform excitation, Guy [10] designed a sparse aperture in which they removed elements randomly. Na [19] designed apertures with cELC and complementary Jerusalem cross (CJC) units to raise radiation efficiency. In [20], Zhao designed a bunching metamaterial antenna that could generate frequency-diverse bunching random radiation pattern. In these literatures, cELC elements and other units were distributed randomly, and entire designs relied on the full-wave electromagnetic simulation. Some traditional modeling methods could also be used for metamaterial aperture modeling and designing, such as fast multipole method (FMM) [21], series expansion method [22], and boundary element method (BEM) method [23]. Recently years, some new technics have been developed in electromagnetic modeling. In [24], a full-wave finite-element-based method was introduced. Feng [25] used combined neural networks for parametric modeling of electromagnetic (EM) behavior of microwave components. Calik [26] proposed a fully-connected regression model based on Bayesian optimization for frequency selective surfaces. These methods could be considered in the design of the metamaterial apertures antenna. Second, in the metamaterial aperture imaging model, discretization of the whole imaging zone will inevitably result in the

Submitted On: May 16, 2021
Accepted On: July 12, 2021

https://doi.org/10.47037/2021.ACES.J.360802

particularly high dimensionality of the measurement matrix and the huge computational cost in further imaging processing. In order to get real-time images, reducing the scale of the inverse problem is necessary. A feasible approach is to constrain the imaging volume and thus decrease the corresponding number of discretization voxels. It is mentioned in [27] that the Microsoft Kinect sensor working at infrared wavelength can potentially be used for reducing reconstruction volume, whereas this approach would be problematic if the targets span across many far different range bins. To enhance the capability of the system, multi-GPU parallel processing architectures were studied in [28]. In traditional SAR imaging, there is coupling between the range dimension and azimuth dimension. The range-Doppler (RD) algorithm [29] decomposes the coupling two-dimensional scene focusing problem into two one-dimensional focusing problems, which is also the most popular decoupling method. In [30], a range decoupling method was applied in metamaterial aperture imaging to reduce the dimensionality of the measurement matrix and accelerate computation, but the running time still could not be neglected.

In this work, we deduce the contribution of the cELC resonator to the radiation fields of the metamaterial aperture, and propose a fast range decoupling method for metamaterial aperture-based imaging systems. The algorithm allows the scene to be decomposed into a set of range cells, the range of interest to be drastically reduced according to the IFFT results, and the scene to be reconstructed in a parallel way. Compared with the conventional approach, the proposed algorithm requires less computational cost and running time, and produces real-time images. Multiple simulations verify the effectiveness of the proposed algorithm.

The rest of this paper is structured as follow. In Section II, we briefly introduce the general forward metamaterial imaging model. Section III deduces the contribution of the cELC resonator to the radiation fields of the metamaterial aperture, and discusses the character of the measurement matrix. Section IV devotes the derivation fast range decoupling approach. In section V, we present the imaging simulation to demonstrate the effectiveness of the proposed method. Conclusions are presented in Section VI.

II. MATHEMATICAL IMAGING MODEL

The mathematical model of the metamaterial aperture forward imaging system will be introduced in this section. As shown in Fig. 1, a dielectric-filled parallel-plate waveguide is employed as the transmitter. An array of subwavelength cELC elements with different physical geometries are randomly distributed on the upper plane of the waveguide. Different subsets of elements are excited with changing frequency. Backscatters from the objects in the scenes are received by an open-ended waveguide probe, which is on the antenna plane and close to the transmitter plane. The vector $\mathbf{r}=(x,y,0)$ represents the points on the metamaterial aperture antenna, which is on the reference plane, and the vector $\mathbf{r}'=(x',y',z')$ indicates the points in the imaging scene.

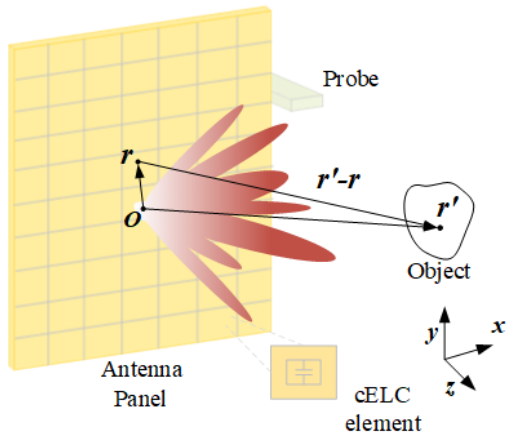


Fig. 1. Panel to probe imaging system.

Suppose $T_{tx}(\mathbf{r})$ is the antenna impulse response. The field in the reference plane could also be illustrated by this response. According the Helmholtz equation, the field in the imaging scene can be given as:

$$U_{0TX}(\mathbf{r}')=\int_{\text{scene}} T(\mathbf{r})\frac{\partial}{\partial z}G(\mathbf{r}',\mathbf{r})\mathrm{d}^2\mathbf{r}\,, \tag{1}$$

where

$$G(\mathbf{r}',\mathbf{r})=\frac{\exp(-jk_0|\mathbf{r}'-\mathbf{r}|)}{4\pi|\mathbf{r}'-\mathbf{r}|}, \tag{2}$$

is Green's function, which describes the solution for impulse response from a source in free space. Taking z as the propagation axis, $(\partial/\partial z)G(\mathbf{r}',\mathbf{r})$ represents the scalar field propagator from the location $\mathbf{r}$ on the reference plane to the location $\mathbf{r}'$ on a parallel plane in imaging scene [30]. k_0 is the wavenumber, given as $k_0=2\pi f/c$, where f denotes the spatial frequency and c denotes the speed of light.

Assuming that the incident radiation field irradiates the scene target and the target scatters the incident electromagnetic wave. $U_S(\mathbf{r}')$ is used to represent the target scattering field, and the total field of the imaging scene can be obtained as follows:

$$U_{To}(\mathbf{r}')=U_{oTX}(\mathbf{r}')+U_S(\mathbf{r}')\,. \tag{3}$$

The wave equation in the case of uniform medium scattering is expressed as:

$$\nabla^2U_{To}(\mathbf{r}')+k(\mathbf{r}')^2U_{To}(\mathbf{r}')=0\,, \tag{4}$$

where

$$k(\mathbf{r}')=k_0n(\mathbf{r}')=k_0[1+n_\delta(\mathbf{r}')]\,, \tag{5}$$

is the refractive index of the scattering medium, and

$n_\delta(\mathbf{r}') = [\mu(\mathbf{r}')\varepsilon(\mathbf{r}')/\mu_0\varepsilon_0]^{1/2}$.

Considering that the incident radiation field $U_{oTX}(\mathbf{r}')$ must satisfy the Helmholtz equation, it can be obtained:

$$\nabla^2 U_{oTX}(\mathbf{r}') + k_0^2 U_{oTX}(\mathbf{r}') = 0 . \quad (6)$$

Substituting (5) and (6) into (4), the target scattering field $U_S(\mathbf{r}')$ can be written as:

$$\nabla^2 U_S(\mathbf{r}') + k_0^2 U_S(\mathbf{r}') = -(k(\mathbf{r}')^2 - k_0^2)U_{To}(\mathbf{r}') . \quad (7)$$

It should be noted that there is a nonlinear relationship between the scattering field and the target scatter function. Without any approximate treatment of the scattering field, the upper differential equation cannot be solved. In this section, the first order Born approximation assumption is used to ignore the second-order terms [31]. Then, the Helmholtz equation could be represented as:

$$\nabla^2 U_S(\mathbf{r}') + k_0^2 U_S(\mathbf{r}') = -\sigma(\mathbf{r}')U_{To}(\mathbf{r}') , \quad (8)$$

where

$$\sigma(\mathbf{r}') = 2k_0 n_\delta(\mathbf{r}') , \quad (9)$$

donates the reflectivity of scene target.

Treating the scene as the radiation source, the source-field propagation Green's function is used to solve the above equation. In this case, the scattering field at the transmit aperture (reference plane) $U_S(\mathbf{r}')$ is represented as:

$$U_S(\mathbf{r}') = \int_V -G(\mathbf{r},\mathbf{r}')\sigma(\mathbf{r}')U_{To}(\mathbf{r}')\mathrm{d}^3\mathbf{r}' . \quad (10)$$

Considering the effect of weak perturbation from the scattered field on the total field $U_{To}(\mathbf{r}')$, we can rewrite (10) as:

$$U_S(\mathbf{r}') = \int_V -G(\mathbf{r},\mathbf{r}')\sigma(\mathbf{r}')U_{0TX}(\mathbf{r}')\mathrm{d}^3\mathbf{r}' . \quad (11)$$

Revisiting (2), we obtain the relationship between the derivative of Green's function $(\partial/\partial z)G(\mathbf{r}',\mathbf{r})$ and the Green's function G yield itself [32]:

$$\frac{\partial}{\partial z}G(\mathbf{r},\mathbf{r}') = \cos(\mathbf{z},\mathbf{r_P})(jk_0 - \frac{1}{|\mathbf{r_P}|})G(\mathbf{r},\mathbf{r}') , \quad (12)$$

where $\cos(\mathbf{z},\mathbf{r_P})$ denotes the cosine of the angle between $\mathbf{z}$ and $\mathbf{r_P}$ joining $\mathbf{r}$ to $\mathbf{r}'$. Assuming that the field of view (FOV) is narrow, we can simplify the above equation as:

$$\frac{\partial}{\partial z}G \approx jk_0 G . \quad (13)$$

Therefore, the scattering field at the transmit aperture $U_S(\mathbf{r}')$ can be represented as:

$$U_S(\mathbf{r}') = -\frac{1}{jk_0}\int_V \sigma(\mathbf{r}')U_{oTX}(\mathbf{r}')\frac{\partial}{\partial z}G(\mathbf{r},\mathbf{r}')\mathrm{d}^3\mathbf{r}' . \quad (14)$$

Assume that the receiving probe antenna is located at the same plane as the transmitting metamaterial aperture antenna, and the pulse response is $T_{rx}(\mathbf{r})$. Since the scattering field on the receiving aperture is known, the measured value g of the receiving probe antenna can be calculated as [33]:

$$g = \int_S T_{rx}U_S(\mathbf{r}')d^2\mathbf{r} . \quad (15)$$

Substituting (12) and (14) into (15), g can be rewritten as:

$$g = \int_V U_{oTX}(\mathbf{r}')U_{oRX}(\mathbf{r}')\sigma(\mathbf{r}')d^3\mathbf{r} , \quad (16)$$

where $U_{oRX}(\mathbf{r}')$ represents the field produced by the receiving probe antenna at location $\mathbf{r}$.

Considering that the metamaterial aperture antenna can generate frequency-diverse random radiation fields to detect the scene, a set of measurement echo vectors indexed by frequency can be calculated as:

$$\mathbf{g}(f) = \int_V U_{oTX}(\mathbf{r}', f)U_{oRX}(\mathbf{r}', f)\sigma(\mathbf{r}')d^3\mathbf{r} . \quad (17)$$

Assuming the reflectivity of objects in scene is isotropic, since the imaging resolution is limited by the antenna aperture, the scene can be decomposed to 3D discrete voxels and encoded by a finite number of frequency diverse models. Then (17) can be rewritten as a finite-dimensional matrix equation:

$$\mathbf{g} = \mathbf{H}\boldsymbol{\sigma} + \mathbf{n} , \quad (18)$$

where $\mathbf{g}$ denotes the $M \times 1$ measurement vector, $\mathbf{n}$ is an $M \times 1$ noise vector, $\boldsymbol{\sigma}$ is the $N \times 1$ scatting coefficient vector across N pixies of the imaging scene, and $\mathbf{H}$ denotes the $M \times N$ measurement matrix, which is the dot product of the incident and scatted fields. The above equation clearly describes the receiving and processing of scattering electromagnetic waves in the scene.

In order to get a unique linear solution for the scatting information $\boldsymbol{\sigma}$ in the scene, the rank of $\mathbf{H}$ must equal the dimension of $\boldsymbol{\sigma}$. In another word, $M = N$. However, in the application of imaging scenarios, the rank of $\mathbf{H}$ is much less than the dimension of scatting coefficient $\boldsymbol{\sigma}$. According the compress sensing theory, the underdetermined scenes could be reconstructed by solving the optimization problem:

$$\boldsymbol{\sigma}_{\mathbf{est}} = \arg\min_\sigma \left\|\mathbf{g} - \mathbf{H}\boldsymbol{\sigma}_{\mathbf{est}}\right\|_2^2 + \gamma\left\|\boldsymbol{\sigma}_{\mathbf{est}}\right\|_1 , \quad (19)$$

where γ is the regularization parameter.

Numerous algorithms exist for solving this optimization problem, such as greedy algorithm, convex optimization algorithm and Bayesian algorithm. Here we use the orthogonal matching pursuit (OMP) algorithm, which is a classic greedy algorithm, to solve this problem. The ensemble measurement matrix is used directly during this processing, which is called brute force method. Since $M \ll N$, the brute force method needs amount of computational cost and relative long running time. In order to get new algorithms which could accelerate the calculation, we will analysis the measurement matrix in following section.

III. APERTURE MODEL AND MEASUREMENT MATRIX CHARACTER

For the frequency-diverse metamaterial aperture employed in our imaging system, the field pattern **T** consists of a relatively random set of nulls and lobes that change with frequency. At a given angle, the magnitudes and phase of the field vary irregularly over the bandwidth. The frequency diverse radiation field of the metamaterial aperture is generated by an array of cELC elements with various resonance frequencies. Generally, a cELC element can be modelled as a polarizable radiation magnetic dipole. The polarization of it can be described according to the Lorentzian dispersion:

$$\alpha(\omega) = -\frac{F\omega^2}{\omega^2 - \omega_0^2 + j\omega\gamma}, \tag{20}$$

where $\omega = 2\pi f$ is the angular frequency, ω_0 is the angular resonance frequency of the resonator, $\gamma = \omega_0 / 2Q$ is the damping factor, and F is proportional to the oscillator strength and coupling, which is assumed to be 1. The dipole moment of the dipole at location $\mathbf{r}$ can be calculated as:

$$\mathbf{m}(\mathbf{r},\omega) = \alpha(\mathbf{r},\omega)\cdot U_G(\mathbf{r},\omega), \tag{21}$$

where $U_G(\mathbf{r},\omega)$ is the local guided field. Assuming that these dipoles are all y-polarized, the radiation pattern of location $\mathbf{r}'$ from a dipole at location $\mathbf{r}$ can be approximated as:

$$U(\mathbf{r}',\mathbf{r},\omega) \approx \mathbf{m}(\mathbf{r},\omega)\frac{Z_0 k_0}{4\pi R'}\exp(-jk_0R')\sin\theta, \tag{22}$$

where R' is the distance between $\mathbf{r}'$ and $\mathbf{r}$, Z_0 is the impedance of free space, and θ is the angle between $\mathbf{r}'$ and $\mathbf{m}(\mathbf{r})$. Compared with the standard approximation for the far-field radiation of a dipole, we ignore all higher-order terms that vary as $1/R'^2$ or $1/R'^3$. The total transmit field at location $\mathbf{r}'$ in the scene can be calculated by superposing radiation fields from all dipoles in the aperture:

$$U_{oTX}(\mathbf{r}', f) = U_{oTX}(\mathbf{r}',\omega) = \sum\nolimits_{\mathbf{r}} U(\mathbf{r}',\mathbf{r},\omega). \tag{23}$$

The above deduction shows that by sweeping frequency ω, the polarizability and local guided field of each dipole are changed, and then the dipole moment is modified. Moreover, the radiation fields of the resonator and entire array are affected. Hence the radiation fields of the metamaterial aperture change with respect to ω. ω serves as a parameter to index the measurement modes.

In the derivation of Section II, **H** is the dot product of transmitting fields in the scene of the transmitting metamaterial aperture antenna and the receiving probe. In the far field, the distance from the metamaterial aperture antenna to location $\mathbf{r}'$ and the distance from the receiving probe to location $\mathbf{r}'$ can be approximately equal. The relevant term $1/|\mathbf{r}'-\mathbf{r}|$ in the Green's function could be ignored. The transmitting field on $\mathbf{r}'$ in the scene can be written as:

$$U_{oTX}(\mathbf{r}') = Te^{-jk_0R}, \tag{24}$$

where R is the distance between the center of the transmitting aperture and position $\mathbf{r}'$ in the scene, and T denotes the radiation field of the metamaterial aperture panel.

For convenient analysis, we ignore the radiation pattern from the open waveguide probe, and treat it as an omnidirectional receiving probe. The product of incident and scattering fields at position $\mathbf{r}'$ in the scene (corresponding range R), can be expressed as:

$$\mathbf{h}_R = \mathbf{t}e^{-j2k_0R} = \mathbf{t}e^{-j4\pi\mathbf{f}R/c}, \tag{25}$$

where **t** and **f** denote the radiation field pattern and the variational frequency, respectively. $e^{-j4\pi\mathbf{f}R/c}$ is the propagation factor and is a linear phase term related to R. It is obvious that $\mathbf{h}_R$ is a column entry of measurement matrix **H**, and **H** can be expressed as:

$$\mathbf{H} = \mathbf{TG}(\mathbf{r}',\mathbf{r},\mathbf{f}). \tag{26}$$

Hence a row entry of **H** is the product of the radiation field pattern **T** and appropriate Green's function. When the frequency changes, a distinct spatial radiation pattern **T** is generated by the metamaterial aperture and a new row entry of **H** is obtained.

For a given metamaterial aperture antenna, the frequency sampling interval Δf needs to be determined. However, when we choose Δf, it is important to note that the maximum unambiguous range (MUR), represented by $c/2\Delta f$, is closely related to image quality, and Δf directly influences the frequency points. After determining Δf, the limited frequency bandwidth (named subband) B_Δ has to be ascertained. By the IFFT process for the radiation field within a proper B_Δ, we could get a sequence with a sinc-like magnitude distribution. Then, the limited frequency bandwidth B_Δ can be ascertained experimentally.

Table 1: Antenna parameters

Bandwidth	18.5~25GHz
Frequency sampling interval	8.125MHz
Field of view	-70°~70°
Azimuth sampling interval	0.5°
Panel length	0.4m
Dimension of **T**	801×281

Here we employed measured radiation field data of the antenna from [6]. The parameters of antenna are listed in Table 1. The prototype is a 1D leaky waveguide, formed by patterning the top conductor of a standard microstrip line with cELC elements. This antenna could

explore 2D (range and azimuth) sparse scenes.

We randomly select an azimuth angle within the field of view. For the radiation fields **t** at this angle within the subband B_Δ, the results of IFFT operation are shown in Fig. 2 (a). The magnitudes take the form of a sinc-like distribution. Recalling (20), the measurement vector $\mathbf{h}_R$ is the product of the radiation fields **t** and the propagation factor. This propagation factor, as a linear phase term related to R, corresponds to the shift of the sequence in the time domain, according to the frequency-shift property of the IFFT. The results of IFFT with respect to $\mathbf{h}_R$, which is calculated from **t**, are also shown in Fig. 2 (a). We can find that the IFFT results of $\mathbf{h}_R$ shift to the right compared to that of radiation field **t**. The shift units represent range R corresponding to $\mathbf{h}_R$. This behavior would still exist if the IFFT is performed over entire bandwidth, as shown in Fig. 2 (b).

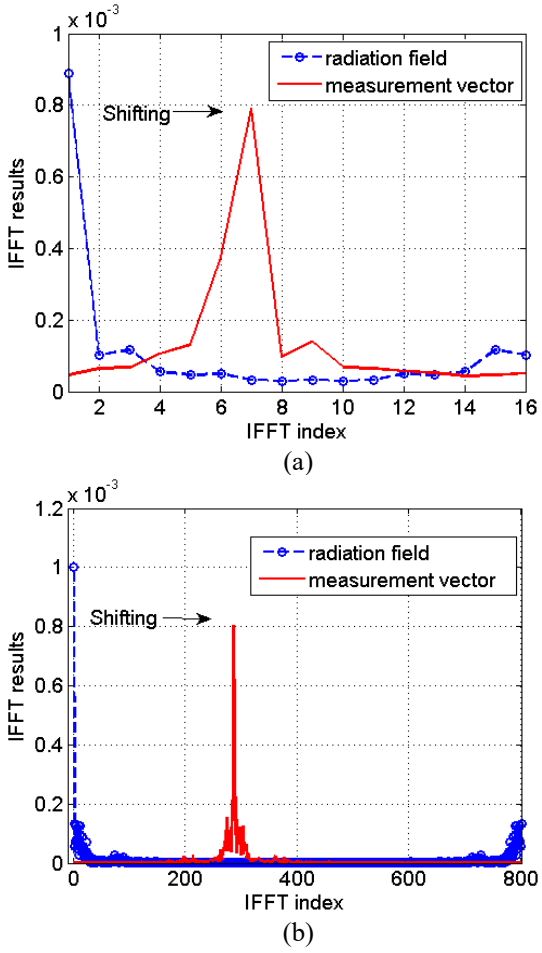


Fig. 2. IFFT results of radiation fields **t** and the measurement vector $\mathbf{h}_R$.

Considering that the echo data measured by the receiving antenna are the product of $\mathbf{h}_R$ and reflectivity of scatters, if we perform IFFT with respect to echo data **g**, the frequency-shift property still exists. Hence, we can deduce the approximate ROI according to the shift units of maxima and the range resolution corresponding to B_Δ. Since the resolution is inversely proportional to the bandwidth, if we perform IFFT with respect to **g** over the entire working bandwidth, the detailed range position of scatters could be obtained, and the ROI would be reduced. This behavior will be verified in the simulation of a later section.

The above analysis focuses on arbitrarily chosen azimuth angle radiation fields within B_Δ. Now, we use a sliding window that has the same size as B_Δ to carry out IFFT operation on the entire bandwidth, while the shift length is one frequency point. The results of the sliding-windowed IFFT are shown in Fig. 3. It is obvious that the position of maxima barely moves. Hence the Sinc-like distributions remain on entire band. Considering that the azimuth angle is randomly chosen, this behavior is also suitable for the ensemble azimuth field, which is the foundation for the further deduction of our algorithm.

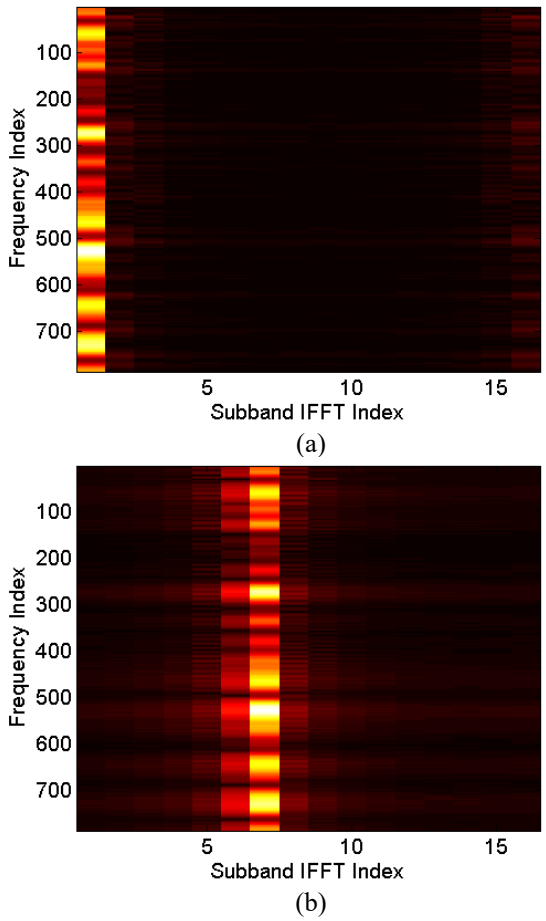


Fig. 3. Sliding-windowed IFFT results of (a) the azimuth fields, and (b) measurement vector over the entire frequency band.

In the above discussion, the targets contained in the imaging space are concentrated at the relatively close range bins. Then the reflectivity coefficients of the voxels outside the ROI are equal to 0. The measured echo data $\mathbf{g}$ are purely the product of constrained $\mathbf{H}$ and $\boldsymbol{\sigma}$ corresponding to the ROI. If we apply IFFT with respect to $\mathbf{g}$, the ROI can be obtained, and then $\boldsymbol{\sigma}$ can be retrieved with the constrained $\mathbf{H}$. However, if the targets contained in the imaging space span some far different range bins, this method will be intractable rapidly. This issue will be discussed in follow section.

IV. FAST RANGE DECOUPLING ALGORITHM

Let $\boldsymbol{\sigma}$ denote all reflection coefficients contained in the imaging scene, and $\mathbf{g}_\Delta \in \mathbb{C}^{P\times 1}$ denote P receiving measurements of the whole scene within subband B_Δ, which contains P frequency points. For an extracted range area $\mathbf{R_1}$, the corresponding measurement matrix is $\mathbf{H}_{\mathbf{R}_1}$; then, the relative received frequency measurement can be represented as $\mathbf{g}_{\mathbf{R}_1} = \mathbf{H}_{\mathbf{R}_1}\boldsymbol{\sigma}_{\mathbf{R}_1}$, which constitutes a portion of $\mathbf{g}_\Delta$. The IFFT results with respect to $\mathbf{g}_{\mathbf{R}_1}$ can be expressed as matrix equation

$$\boldsymbol{\psi} = \mathbf{D}_P\mathbf{g}_{\mathbf{R}_1}, \tag{27}$$

where $\boldsymbol{\psi}$ denotes the IFFT return and $\mathbf{D}_P$ is the inverse discrete Fourier transform matrix, which can be written as:

$$\mathbf{D}_p = \frac{1}{P}\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_P^1 & W_P^2 & \cdots & W_P^{P-1} \\ 1 & W_P^2 & W_P^4 & \cdots & W_P^{2(P-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W_P^{P-1} & W_P^{2(P-1)} & \cdots & W_P^{(P-1)(P-1)} \end{bmatrix}, \tag{28}$$

where $W_p = \exp(j2\pi/P)$. Since $\mathbf{g}_{\mathbf{R}_1} = \mathbf{H}_{\mathbf{R}_1}\boldsymbol{\sigma}_{\mathbf{R}_1}$, $\boldsymbol{\psi}$ can be written as:

$$\boldsymbol{\psi} = \mathbf{D}_P\mathbf{H}_{\mathbf{R}_1}\boldsymbol{\sigma}_{\mathbf{R}_1} = \mathbf{D}_P\begin{bmatrix}\mathbf{h_1}\\ \mathbf{h}_2\\ \vdots\\ \mathbf{h}_P\end{bmatrix}\boldsymbol{\sigma}_{\mathbf{R}_1}. \tag{29}$$

Processing IFFT with respect to $\mathbf{g}$ on all frequency points, we get the detailed range index of objects, and the corresponding area $\mathbf{R}_2$, where $\mathbf{R}_2 \subset \mathbf{R}_1$. Since all the received waves corresponding to $\mathbf{R_1}$ are scattered by the objects in $\mathbf{R}_2$, $\mathbf{g}_{\mathbf{R}_1}$ can be expressed as:

$$\mathbf{g}_{\mathbf{R}_1} = \mathbf{H}_{\mathbf{R}_1}\boldsymbol{\sigma}_{\mathbf{R}_1} = \mathbf{H}_{\mathbf{R}_2}\boldsymbol{\sigma}_{\mathbf{R}_2} \tag{30}$$

and $\boldsymbol{\psi}$ can be written as:

$$\boldsymbol{\psi} = \mathbf{D}_P\mathbf{H}_{\mathbf{R}_1}\boldsymbol{\sigma}_{\mathbf{R}_1} = \mathbf{D}_P\mathbf{H}_{\mathbf{R}_2}\boldsymbol{\sigma}_{\mathbf{R}_2} = \mathbf{D}_P\begin{bmatrix}\mathbf{h_1}\\ \mathbf{h}_2\\ \vdots\\ \mathbf{h}_P\end{bmatrix}\boldsymbol{\sigma}_{\mathbf{R}_2}. \tag{31}$$

For the case in Fig. 3, the corresponding maximum can be expressed as:

$$\psi_{\max} = \frac{1}{P}\begin{bmatrix}1 & W_P^4 & \cdots & W_P^{4(P-1)}\end{bmatrix}\begin{bmatrix}\mathbf{h_1}\\ \mathbf{h}_2\\ \vdots\\ \mathbf{h}_P\end{bmatrix}\boldsymbol{\sigma}_{\mathbf{R}_2}. \tag{32}$$

Henceforth, the maximum $\psi_{\max}$ can be considered as a new measurement ψ. The inner product between the Fourier coefficient and original measurement matrix $\mathbf{H}_{\mathbf{R}_2}$, denoted by $\boldsymbol{\varphi}$, is the new sensing vector.

In the above derivation, the measurement vector $\mathbf{g}_{\mathbf{R}_1}$ turns to a single measurement $\psi_{\max}$, and the measurement matrix $\mathbf{H}_{\mathbf{R}_1}$ turns to a sensor vector $\boldsymbol{\varphi}$. We get a new measurement equation:

$$\boldsymbol{\psi} = \boldsymbol{\Phi}\boldsymbol{\sigma}. \tag{33}$$

The new measurement matrix is still underdetermined, so we can get reflection coefficients $\boldsymbol{\sigma}$ by solving an optimization problem:

$$\boldsymbol{\sigma}_{\mathbf{est}} = \arg\min_\sigma \left\|\boldsymbol{\Psi} - \boldsymbol{\Phi}\boldsymbol{\sigma}_{\mathbf{est}}\right\|_2^2 + \gamma\left\|\boldsymbol{\sigma}_{\mathbf{est}}\right\|_1. \tag{34}$$

Comparing (34) with (19), the dimension of the new measurement matrix $\boldsymbol{\Phi}$ is much less than the dimension of $\mathbf{H}$. If we use same algorithm, such as OMP, (34) would require less computational cost and running time.

If the targets are located in far more range cells in the scene, we can obtain multiple equations, and the above reconstruction can be performed in parallel.

Our proposed fast range decoupling algorithm can be summarized as follows:

Step 1. Initialize the received frequency measurements $\mathbf{g}$ and radiation field $\mathbf{T}$.

Step 2. Confirm sub-bandwidth B_Δ from $\mathbf{T}$, and record the index of maximum column of the IFFT return.

Step 3. Process $\mathbf{g}$ with sliding-window IFFT to ascertain the range region $\mathbf{R_1}$ that contains targets by the index of maxima. And record the corresponding new measurements $\boldsymbol{\Psi}$.

Step 4. Perform IFFT with respect to $\mathbf{g}$ and confirm the range location $\mathbf{R_2}$ of targets according to the peak index of results.

Step 5. Build the new measurement matrix $\boldsymbol{\Phi}$ with respect to $\mathbf{R_2}$, and get new measurement equation $\boldsymbol{\Psi} = \boldsymbol{\Phi}\boldsymbol{\sigma}_{\mathbf{R}_2}$.

Step 6. Reconstruct $\boldsymbol{\sigma}_{\mathbf{R}_2}$ from the measurement

equation and rebuild the ensemble $\boldsymbol{\sigma}$.

V. SIMULATIONS AND ANALYSIS

In this section, the imaging capabilities of the proposed fast range decoupling algorithm are demonstrated. Radiation field data of the measured antenna are employed in multiple imaging simulations. The parameter of antenna is listed in Table 1.

The azimuth resolution and range resolution can be obtained from:

$$\theta_a = \frac{\lambda_c}{L_{\text{eff}}} \qquad \delta_r = \frac{c}{2\text{BW}}, \tag{35}$$

where λ_c is the wavelength of the center frequency and L_{eff} denotes the effective size of the aperture. For the antenna we used in the simulation, the azimuth resolution is 1.7°, and the range resolution is 2.3 cm.

The imaging scene is discretized into a dense grid with range and azimuth dimensions, and the scattering coefficient of the targets obeys $\mathbb{CN}(0,1)$. All the simulations are processed in the MATLAB environment and on a computer with an Intel Xeon CPU and 6 GB RAM.

The results of sliding-window IFFT processing with randomly chosen azimuth field data are shown in Fig. 4(b). The length of window P is set as 16 experimentally, as well as the number of frequencies.

The first imaging scene has three scatters located in the same range but different azimuth angles. The measurement vector $\mathbf{g}$ can be calculated according to equation (18). Then we perform $\mathbf{g}$ with sliding-window IFFT, as shown in Fig. 4 (b). The length of the window is P, and the corresponding subband width is $\text{B}_\Delta = (P-1)\Delta f$. The range resolution of results is given by $\Delta R = c/2\text{B}_\Delta$. The maximum column of results corresponds to the new measurement vector $\boldsymbol{\Psi}$. According to the index of maxima, the range of interest is [4.34, 5.54] m. The IFFT results with respect to g are shown in Fig. 4 (c). We can reduce ROI to [4.48, 4.52] m, and reconstruction areas are limited to a very narrow band. Then, we update the measurement matrix $\boldsymbol{\Phi}$ by using the inner product of the corresponding Fourier coefficients and constrained measurement matrix $\mathbf{H}_{\text{ROI}}$.

Since the updated measurement matrix $\boldsymbol{\Phi}$ is still underdetermined, we need a compress sensing reconstruction algorithm to solve the modified equation. Here, the orthogonal matching pursuit algorithm is employed to retrieve the scene. The reconstructed reflectivity coefficients of $\boldsymbol{\sigma}_{\mathbf{R2}}$ are shown in Fig. 4 (d). It is obvious that all three scatters are preserved clearly. Moreover, to compare the performance of the proposed method with that of existing algorithms, we reconstruct the scene with the algorithm in [30], as shown in Fig. 4 (e), and the brute force method, which uses the ensemble measurement matrix directly, as shown in Fig. 4 (f). We can see that in both images, all three scatters are well estimated. However, the consumed times in Fig. 4 (d), (e) and (f) are 0.35 s, 2.7 s and 9.4 s, respectively. Hence the acceleration ratio of the proposed fast range decoupling algorithm is calculated to be 26.8, while that of the algorithm in [30] is 3.5.

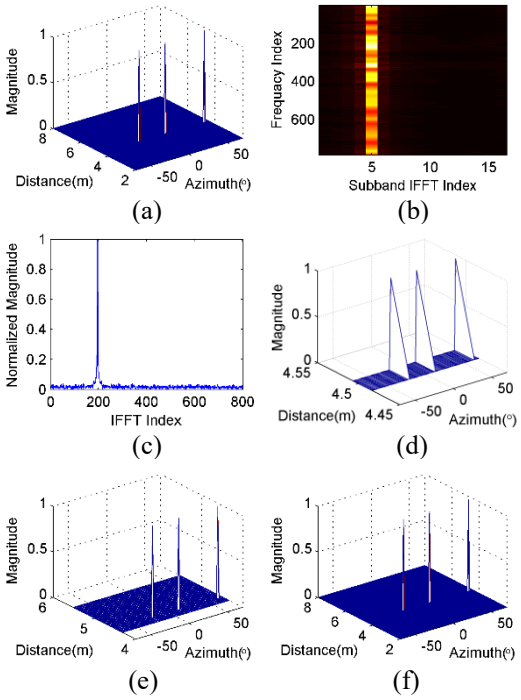


Fig. 4. 2D scene imaging results with 3 point-like scatters. (a) True scene with 3 scatters located in the range of approximately 4.5 m. (b) Result of slight-windowed IFFT with respect to $\mathbf{g}$ with a window length of 16. (c) Results of IFFT operation with respect to $\mathbf{g}$ with all frequency points. (d) Images obtained with proposed method. (e) Images obtained with method of [30]. (h) Image obtained by ensemble H corresponding to [2,8] m.

For the problem of that targets located in far more range cells in scenes, the scale of the measurement matrix would increase rapidly. It is seen from Fig. 5 (a) that targets lie in the range of around 4.5 and 6.9 m from the antenna panel, and the entire range area is [2,8] m. The result of the sliding-windowed IFFT to the received data $\mathbf{g}$ is plotted in Fig. 5 (b). We can see from Fig. 5 (b) that the targets are well separated in two different range cells. Fig. 5 (c) shows the IFFT results to $\mathbf{g}$ with all frequency points, and the size of the ROI could be slumped. Then the modified measurement equation corresponding to each ROI could be built and the reconstruction results are shown in Fig. 5 (d) and Fig. 5 (e). We can observe that the scatters at different locations are estimated accurately.

Similarly, we reconstruct the scene with the algorithm in [30], as shown in Fig. 5 (f) and Fig. 5 (g), and the brute force method, as shown in Fig. 5 (h). The consumed times in Figs. 5 (d)-(h) are 0.32 s, 0.26 s, 3.76 s, 3.70 s and 9.61 s, respectively. Since the reconstruction could be processed in parallel, the acceleration ratio of the proposed method can be calculated to be 30, while that of the algorithm in [30] is 2.6. The imaging efficiency is enhanced obviously.

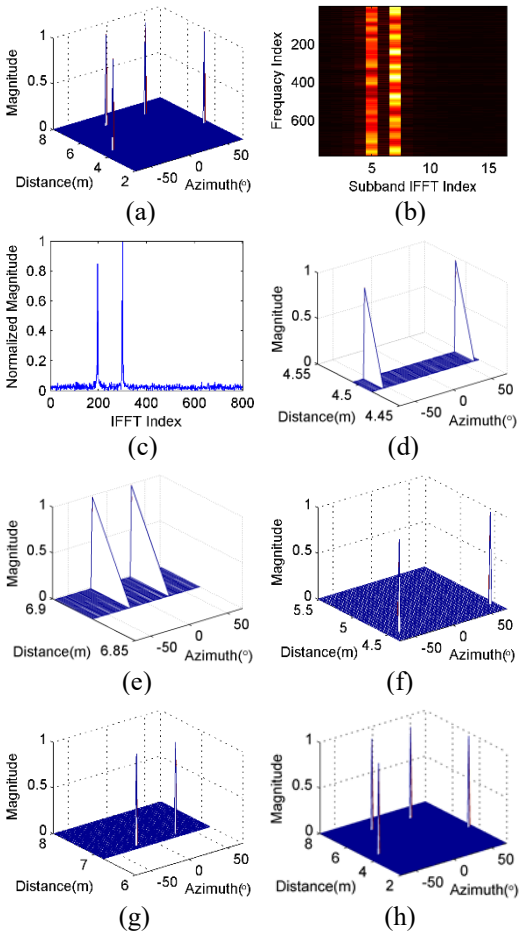


Fig. 5. 2D scene imaging results with scatters located at different range positions. (a) True scene with 4 point-like scatters located in the range of approximately 4.5 m and 6.9 m. (b) Result of slight-windowed IFFT with respect to **g** with a window length of 16. (c) Result of IFFT operation with respect of **g**. (d), (e) Images obtained with the proposed method. (f), (g) Images obtained with the method of [30]. (h) Is an image obtained by ensemble H corresponding to [2,8] m.

Let us analyze the computational complexity of the proposed algorithm. The reconstructed approach employed in this paper is the OMP algorithm, whose computational complexity is $O(LMN)$ [34],[35]. L is the number of iterations, M is the number of frequency points, and N is the number of unknown voxels in the ROI. It should be noted that $L < M \ll N$. The computational cost of the proposed algorithm is mainly from step 3 to step 6 of the summarization. Assuming that the length of the sliding window is m, then the scene could be decomposed to m range cells. In step 3, the cost of IFFT operations is $O(m\log(m))$. K denotes the number of IFFT operation, given by $K = M - m + 1$. Then, the total computation cost is $O(Km\log(m))$. In step 4, the computational cost is $O(M\log(M))$. Define n as the $\mathbf{R}_1$ to $\mathbf{R}_2$ ratio. The computational cost of calculating measurement matrix $\mathbf{\Phi}$ is $K \times \frac{N}{mn} \times m\log(m) = \frac{KN}{n}\log(m)$ in step 5. In step 6, since the different range cells could be processed in parallel, the computational cost is $O(LKN/mn)$. Therefore, the total computational cost of the proposed algorithm is:

$$O(Km\log(m) + M\log(m) + K\frac{N}{n}\log(m) + LK\frac{N}{mn}). \quad (36)$$

Although the computational complexities of both the proposed algorithm and the brute force method are $O(N)$, the memory requirements of our approach for calculating and storing the measurement matrix are decreased sharply. The cost reduction benefits from the sharp drop of the ROI and the parallel processing of partitioned range cells. Furthermore, considering the high dimensionality feature of imaging information, the total computational costs are lowered significantly.

VI. CONCLUSIONS

In this paper, we propose a fast range decoupling algorithm that can reconstruct the partitioned range area in real-time. This proposed algorithm has a lower computational cost and a higher imaging efficiency, due to the ROI reduction and the parallel way of reconstructing the scene. Moreover, the method could produce real-time images of far-field scenes. Further research will concentrate on improving the algorithm in practical applications.

ACKNOWLEDGMENT

We would like to thank Dr. Guy Lipworth and Center for Metamaterials and Integrated Plasmonics, Duke University for providing us the data.

REFERENCES

[1] J. A. Ribeiro, C. M. Pereira, A. F. Silva, and M. G. F. Sales, "Disposable electrochemical detection of breast cancer tumour marker CA 15-3 using poly (Toluidine Blue) as imprinted polymer receptor," *Biosensors and Bioelectronics*, vol. 109, pp. 246-

254, June 2018.

[2] S. S. Gorthi, D. Schaak, and E. Schonbrun, "Fluorescence imaging of flowing cells using a temporally coded excitation," *Opt. Exp.*, vol. 21, no. 4, pp. 5164-5170, Feb. 2013.

[3] X. Zhuge and A. G. Yarovoy, "A sparse aperture MIMO-SAR-based UWB imaging system for concealed weapon detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 1, pp. 509-518, July 2011.

[4] U. Alkus, A. B. Sahin, and H. Altan, "Stand-off through-the-wall W-band millimeter-wave imaging using compressive sensing," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 7, pp. 1025-1029, Apr. 2018.

[5] Y. Gao, W. Peng, Y. Qu, and J. Ding, "Through-the-wall imaging based on modified compressive sampling matching pursuit," *2017 Sixth Asia-Pacific Conference on Antennas and Propagation (APCAP)*, Xi'an, China, pp. 1-3, Oct. 2017.

[6] J. Hunt, T. Driscoll, A. Mrozack, G. Lipworth, M. Reynolds, D. Brady, and D. Smith, "Metamaterial apertures for computational imaging," *Science*, vol. 339, no. 6117, pp. 310-313, Jan. 2013.

[7] J. Hunt, J. Gollub, T. Driscoll, G. Lipworth, A. Mrozack, M. Reynolds, D. Brady, and D. Smith, "Metamaterial microwave holographic imaging system," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 31, no. 10, pp. 2109-2119, Oct. 2014.

[8] J. Gollub, O. Yurduseven, K. P. Trofatter, D. Arnitz, M. F. Imani, T. Sleasman, M. Boyarsky, A. Rose, A. Pedross-Engel, H. Odabasi and T. Zvolensky, "Large metasurface aperture for millimeter wave computational imaging at the human-scale," *Scientific Reports*, vol. 7, no. 1, pp. 1-9, Feb. 2017.

[9] G. Lipworth, A. Mrozack, J. Hunt, D. L. Marks, T. Driscoll, D. Brady, and D. R. Smith, "Metamaterial apertures for coherent computational imaging on the physical layer," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 30, no. 8, pp. 1603-1612, Aug. 2013.

[10] G. Lipworth, A. Rose, O. Yurduseven, V. R. Gowda, M. F. Imani, H. Odabasi, P. Trofatter, J. Gollub, and D. R. Smith, "Comprehensive simulation platform for a metamaterial imaging system," *Appl. Opt.*, vol. 54, no. 31, pp. 9343-9353, Nov. 2015.

[11] G. Lipworth, J. Hunt, A. Mrozack, D. Brady, and D. R. Smith, "Simulations of 2D metamaterial apertures for coherent computational imaging," *2013 IEEE International Conference on Microwaves, Communications, Antennas and Electronic Systems (COMCAS 2013)*, Tel Aviv, Israel, pp. 1-4, Oct. 2013.

[12] T. Fromenteze, O. Yurduseven, M. F. Imani, J. Gollub, C. Decroze, D. Carsenat, and D. R. Smith, ''Computational imaging using a mode-mixing cavity at microwave frequencies,'' *Appl. Phys. Lett.*, vol. 106, no. 19, pp. 9343-53, May 2015.

[13] T. Fromenteze, O. Yurduseven, M. Boyarsky, J. Gollub, D. L. Marks, and D. R. Smith, "Computational polarimetric microwave imaging," *Opt. Exp.*, vol. 25, no. 22, pp. 27488-27505, Oct. 2017.

[14] D. L. Marks and D. R. Smith, "Mode diversity of weakly modulated cavity antennas," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 35, no. 1, pp. 135-147, Jan. 2018.

[15] T. Sleasman, M. F. Imani, J. N. Gollub, and D. R. Smith, "Dynamic metamaterial aperture for microwave imaging," *Appl. Phys. Lett.*, vol. 107, no. 20, pp. 204104, Nov. 2015.

[16] T. Sleasman, M. Boyarsk, M. F. Imani, J. N. Gollub, and D. R. Smith, "Design considerations for a dynamic metamaterial aperture for computational imaging at microwave frequencies," *J. Opt. Soc. Amer. B, Opt. Phys.*, vol. 33, no. 6, pp. 1098-1111, June 2016.

[17] A. V. Diebold, M. F. Imani, T. Sleasman, and D. R. Smith, "Phaseless computational ghost imaging at microwave frequencies using a dynamic metasurface aperture," *Appl. Opt.*, vol. 57, no. 9, pp. 2142-2149, Mar. 2018.

[18] M. F. Imani, T. Sleasman, and D. R. Smith, "Two-dimensional dynamic metasurface apertures for computational microwave imaging," *IEEE Antennas and Wireless Propagation Letters*, vol. 17, no. 12, pp. 2299-2303, Oct. 2018.

[19] K. Na, L. Li, S. Tian, and Y. Li. "Measurement matrix analysis and radiation improvement of a metamaterial aperture antenna for coherent computational imaging," *Applied Sciences*, vol. 7, no. 9, pp. 933-942, Sep. 2017.

[20] M. Zhao, S. Zhu, J. Li, H. Shi, J. Chen, Y. He, and A. Zhang, "Frequency-diverse bunching metamaterial antenna for coincidence imaging," *Materials*, vol. 12, no. 11, pp. 1817-1828, Jan. 2019.

[21] G. Antonini, "Fast multipole formulation for PEEC frequency domain modeling," *Applied Computational Electromag. Society Journal*, vol. 17, no. 3, pp. 1-17, Nov. 2002.

[22] S. Kahng, "Predicting and mitigating techniques of the PCB rectangular power/ground planes' resonance modes," *Applied Computational Electromagnetics Society Newsletter*, vol. 22, no. 3, pp. 15-23, Nov. 2007.

[23] S. Patil, M. Y. Koledintseva, and R. W. Schwartz, "Modeling of field distribution and energy storage in diphasic dielectrics," *2006 15th IEEE International Symposium on the Applications of Ferroelectrics*, Sunset Beach, NC, USA, pp. 307-310, July 2006.

[24] L. Xue and D. Jiao, "Fast and rigorous method for solving low-frequency breakdown in full-wave finite-element-based solution of general lossy problems," *2018 International Applied Computational Electromagnetics Society Symposium (ACES)*, Denver, CO,

USA, pp. 1-2, Mar. 2018.
[25] F. Feng, C. Zhang, J. Ma, and Q. Zhang, "Parametric modeling of EM behavior of microwave components using combined neural networks and pole-residue-based transfer functions," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 1, pp. 60-77, Jan. 2016.
[26] N. Calik, M. A. Belen, P. Mahouti, and S. Koziel, "Accurate modeling of frequency selective surfaces using fully-connected regression model with automated architecture determination and parameter selection based on Bayesian optimization," *IEEE Access*, vol. 9, pp. 38396-38410, Mar. 2021.
[27] O. Yurduseven, J. N. Gollub, A. Rose, D. L. Marks, and D. R. Smith, "Design and simulation of a frequency-diverse aperture for imaging of human-scale targets," *IEEE Access*, vol. 4, pp. 5436-5451, July 2016.
[28] D. L. Marks, O. Yurduseven, and D. R. Smith, "Fourier accelerated multistatic imaging: A fast reconstruction algorithm for multiple-inputmultiple-output radar imaging," *IEEE Access*, vol. 5, pp. 1796-1809, Feb. 2017.
[29] C. G. Walter, R. M. Majewski, and R. S. Goodman, *Spotlight Synthetic Aperture Radar: Signal Processing Algorithms*. Boston, MA, USA: Artech House, 1995.
[30] Z. Wu, L. Zhang, H. Liu, and N. Kou, "Range decoupling algorithm for accelerating metamaterial apertures-based computational imaging," *IEEE Sensors Journal*, vol. 18, no. 9, pp. 3619-3631, Mar. 2018.
[31] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics*. Cambridge University, 1995.
[32] J. W. Goodman, *Introduction to Fourier Optics*. 3rd ed., Englewood, CO, USA: Roberts and Company, 2005.
[33] F. C. Lin and M. A. Fiddy, "Image estimation from scattered field data," *Int. J. Imag. Syst. Technol.*, vol. 2, no. 2, pp. 76-95, June 1990.
[34] B. L. Sturm and M. G. Christensen, "Comparison of orthogonal matching pursuit implementations," *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, Bucharest, Romania, pp. 220-224, Oct. 2012.
[35] S. Hsieh, C. Lu, and S. Pei, "Fast OMP: Reformulating OMP via iteratively refining ℓ2-norm solutions," *2012 IEEE Statistical Signal Processing Workshop (SSP)*, Ann Arbor, MI, USA, pp. 189-192, Aug. 2012.

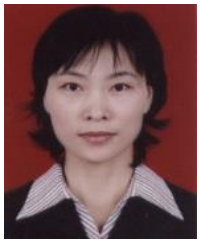
Yuteng Gao was born in Shannxi Province, China, in 1988. He received the B.S. and M.S degree in School of Electronics and Information, Northwestern Polytechnical University in Xi'an City, China, in 2010 and 2013 respectively. He is presently working on his Ph.D. degree in School of Electronics and Information, Northwestern Polytechnical University in Xi'an City, China. His research interests include milimeter wave, antenna design and radar imaging.

Wencan Peng was born in Hubei Province, China, in 1987. She received the B.S. degree in School of Measuring and Optical Engineering, Nanchang Hangkong University in Nanchang City, China, in 2010. Then, she received the M.S. and Ph.D. degree in School of Electronics and Information, Northwestern Polytechnical University in Xi'an City, China, in 2013 and 2020 respectively. Now, she is working with Payload Research Center, Academy of Space Information System (CASC Xi'an). Her research interests include: array signal processing and array calibration.

Min Wang (IEEE Member), received the B.S. degree from Xidian Univ. China in 2000; M.S. and Ph.D. degrees in Signal and Information Processing from Xidian Univ., Xi'an China, in 2003 and 2005 respectively. He is working with National Lab of Radar Signal Processing in Xidian University. His research interests include sparse signal processing, miliwave/Terahertz radar and high resolution radar imaging.

Chenjiang Guo was born in Shannxi Province, China, in 1963. CIE Senior Member, Antenna Society Committee Member. He received the B.S., M.S. and Ph.D. in School of Electronics and Information, Northwestern Polytechnical University in Xi'an City, China, in 1984, 1987 and 2007 respectively. He is a Professor in School of Electronics and

Information NWPU. His research interests include: array signal processing, theory and design of antenna.



Jun Ding, received her B.Eng. degree in Electronic Engineering in 1986 from Northwestern Polytechnical University (NWPU). She obtained her M.S. degree in Electromagnetic Fields and Microwave Techniques in 1989 from the NWPU. In 2005, she received her Ph.D. degree in Circuits and Systems from the NWPU. She is now a Professor of Electromagnetic Fields and Microwave Techniques in the NWPU. Her research interests include electromagnetic calculation, antenna theory and design, microwave circuit design, and electromagnetic compatibility (EMC).

Time-dependent Schrödinger Equation based on HO-FDTD Schemes

M. Zhu, F. F. Huo, and B. Niu

School of Electronic and Information Engineering
Jinling Institute of Technology, Nanjing, 211169, China
zomi@jit.edu.cn, huofeifei4750@jit.edu.cn, niuben@jit.edu.cn

Abstract ─ A high order finite-different time-domain methods using Taylor series expansion for solving time-dependent Schrödinger equation has been systematically discussed in this paper. Numerical characteristics have been investigated of the schemes for the Schrödinger equation. Compared with the standard Yee FDTD scheme, the numerical dispersion has been decreased and the convergence has been improved. The general update equations of the methods have been presented for wave function. Numerical results of potential well in one-dimension show that the application of the schemes is more effective than the Yee's FDTD method and the higher order has the better numerical dispersion characteristics.

Index Terms ─ HO-FDTD, numerical dispersion, stability, the Schrödinger equation.

I. INTRODUCTION

With the devices size shrinks, quantum effects can not to be ignored in the nano-scale material and semiconductor in equipment system. In some certain circumstances, like in a quantum-dot structure, the quantum effects are remarkable and the problem requires new mathematical method to investigate. The Maxwell's equations are the basic theory for researching the electromagnetic phenomenon in macroscopic scale. The Schrödinger equations are employed to solve the nano-scale components in microscopic scale electromagnetic domain. There are many numerical methods are used to solve the Schrödinger equations [1]. The Yee's FDTD [2] algorithm is conservative and limitation with the second-order accuracy both in time and space. The numerical dispersion and computational memory of the FDTD are the dominant restriction to the application in the electronically large electromagnetic targets. The FDTD [3-5] scheme has been chosen to deal with the quantum problem for its convenient and simple. The absorbing boundary conditions for the FDTD method of the time-dependent one-dimensional Schrödinger equation are proposed in [6]. The absorbing boundary conditions are considered by using a rational-function approximation for the one- and two-dimensional Schrodinger equations on a finite interval [7]. A new FDTD simulation scheme for coupled Maxwell-Schrödinger system is presented in [8] and the computation has been made more efficient.

The dimensional time domain which using the FDTD method to describe the simulation of the quantum magnetic susceptibility for a quantum toroid [9].

In recent years, the high order finite-difference time-domain (HO-FDTD) [10] method and multiresolution time-domain (MRTD) [11] scheme are outstanding technique in the quantum computational field. The HO-FDTD schemes is a little more complicate than the FDTD but is less complex than the MRTD method.

In this paper, we give HO-FDTD methods to deal with the time-dependent Schrödinger equation. The simulation results of the method are testified by one dimensional potential well problem. The paper is organized as follows. The basic theory of the method for the time-dependent Schrödinger equation is stated in Section II. The numerical characteristics including the dispersion and stability are discussed in Section III. Numerical simulation is given in Section IV. Consequently, the HO-FDTD algorithms is more accuracy but more computational time than that of FDTD method.

II. THEORY AND ALGORITHM

A. The Schrödinger equation

The time-dependent Schrödinger equation is the governing equation for system's quantum effects can be written as follows:

$$\mathrm{i}\hbar\frac{\partial\psi(\vec{\mathrm{r}},t)}{\partial t}=-\frac{\hbar^2}{2m}\nabla^2\psi(\vec{\mathrm{r}},t)+\mathrm{V}(\vec{\mathrm{r}})\psi(\vec{\mathrm{r}},t), \qquad (1)$$

where ψ is a particle's wave function, the mass of the electron m is 9.109e-31kg, the Planck constant $\hbar$ is $1.0546\times10^{-34}J\cdot\mathrm{s}$, the imaginary unit "i" is the square root of minus one, $\mathrm{V}(\vec{r})$ is the potential energy, $-(\hbar^2/2m)\nabla^2$ is the kinetic energy operator, and $-(\hbar^2/2m)\nabla^2+\mathrm{V}$ is the Hamiltonian operator. The wave function $\psi(\vec{\mathrm{r}},t)$ is a complex which can be written as follows:

Submitted On: February 21, 2021
Accepted On: July 10, 2021
https://doi.org/10.47037/2021.ACES.J.360803

$$\psi(\vec{r},t)=\psi_{re}(\vec{r},t)+\mathrm{i}\,\psi_{im}(\vec{r},t), \quad (2)$$

where ψ_{re} and ψ_{im} is the real part and imaginary part of wave function ψ.

The equation (1) can be rewritten as:

$$\frac{\partial\psi_{re}(\vec{r},t)}{\partial t}=-\frac{\hbar}{2m}\frac{\partial^2\psi_{im}(\vec{r},t)}{\partial x}+\frac{1}{\hbar}\mathrm{V}(\vec{r})\psi_{im}(\vec{r},t), \quad (3)$$

$$\frac{\partial\psi_{im}(\vec{r},t)}{\partial t}=\frac{\hbar}{2m}\frac{\partial^2\psi_{re}(\vec{r},t)}{\partial x}-\frac{1}{\hbar}\mathrm{V}(\vec{r})\psi_{re}(\vec{r},t). \quad (4)$$

B. The HO-FDTD for Schrödinger equation

The wave function ψ can be written as follows:

$$\psi^{n+1/p}(i,j,k)=\psi^{n+1/p}(i\Delta x, j\Delta y, k\Delta x). \quad (5)$$

The wave function relevant in (5) is expand in a Taylor series of basis function, the update equations of the HO-FDTD(2,2*n*) methods for the real part of the method can be written as follows:

$$\begin{aligned}\psi_{re}^{n+1}(i,j,k)&=\psi_{re}^{n}(i,j,k)+\frac{\Delta t}{\hbar}\mathrm{V}(i,j,k)\psi_{im}^{n}(i,j,k)\\&-\sum_{v=1}^{\infty}a(v)\frac{\Delta t}{\Delta x^2}\frac{\hbar}{2m}\left[\psi_{im}^{n}(i+v,j,k)\text{-}2\psi_{im}^{n}(i,j,k)+\psi_{im}^{n}(i\text{-}v,j,k)\right]\\&-\sum_{v=1}^{\infty}a(v)\frac{\Delta t}{\Delta y^2}\frac{\hbar}{2m}\left[\psi_{im}^{n}(i,j+v,k)\text{-}2\psi_{im}^{n}(i,j,k)+\psi_{im}^{n}(i,j\text{-}v,k)\right],\\&-\sum_{v=1}^{\infty}a(v)\frac{\Delta t}{\Delta z^2}\frac{\hbar}{2m}\left[\psi_{im}^{n}(i,j,k+v)\text{-}2\psi_{im}^{n}(i,j,k)+\psi_{im}^{n}(i,j,k\text{-}v)\right]\end{aligned} \quad (6)$$

where Δx, Δy, Δz is the cell size and *i*, *j*, *k* is the spatial grid index in *x*, *y*, *z* direction, respectively, Δt is the time step size, $\alpha(v)$ are the coefficients as shown in Table 1.

Table 1: Coefficients of spatial difference

(2, 2*p*)	$a(1)$	$a(2)$	$a(3)$	$a(4)$	$a(5)$
FDTD	1				
(2,6)	97/84	-1/84	-1/84		
(2,8)	353/324	-1/324	-1/324	-1/324	
(2,10)	489/462	-1/924	-1/924	-1/924	-1/924

III. NUMERICAL CHARCTERISTICS

A. Stability condition of the HO-FDTD method for the Schrödinger equation

On the basis of the von Neumann stability, the solution of the wave function can be written as superpositions of the plane wave [19] as:

$$\psi(x,y,z,t)=A\exp(-\mathrm{i}(i\Delta_x k_x+j\Delta_y k_y+k\Delta_z k_z\text{-}\omega n\Delta t)), \quad (7)$$

where $k_x=k_0\sin\theta\cos\varphi$, $k_y=k_0\sin\theta\sin\varphi$, $k_z=k_0\cos\theta$, k_0 is the wave number, $k_0=\mathrm{p}_m/\hbar$, p_m is the momentum, (θ, φ) is the wave propagation angle in spherical coordinate, *n* is the quantity of the time steps, ω is the frequency of the incident wave.

In one dimension and V = 0 for simplicity, the general stability condition of the HO-FDTD methods for the Schrödinger equation [12]-[18] can be given as:

$$\Delta t\le\frac{m\Delta^2}{\hbar\sqrt{d}\sum_{v=1}^{\infty}|a(v)|}=s_{\max}\frac{m\Delta^2}{\hbar}, \quad (8)$$

$$s_{\max}=\frac{1}{\sqrt{d}\sum_{v=1}^{\infty}|a(v)|}, \quad (9)$$

where $\Delta x=\Delta y=\Delta z=\Delta$, *d* is the number of spatial dimensions, 1, 2 or 3, the Courant-Friedrichs-Lewy (CFL) stability condition is in equation (8), $s_{\max}$ is the maximum CFL stability factor that is shown in Table 2. The investigations are that the HO-FDTD methods have the more stricter stability condition than FDTD. From the comparisons among the HO-FDTD, it is obvious that the higher order has the looser stability condition.

Table 2: The maximum stability factor $s_{\max}$ for HO-FDTD methods

	FDTD	(2, 6)	(2, 8)	(2,10)
$s_{\max}$	$1/\sqrt{d}$	$0.848/\sqrt{d}$	$0.910/\sqrt{d}$	$0.941/\sqrt{d}$

B. Numerical dispersion of HO-FDTD methods for the Schrödinger equation

Using the 2*p*th-order difference to discretize the spatial derivatives as follows:

$$\frac{\partial^2\psi}{\partial x^2}\approx\sum_{v=-p}^{p}C_v\frac{\psi(i+v,j,k)}{(\Delta x)^2}=\sum_{v=-p}^{p}C_v\frac{\exp(-j_0vk_x\Delta x)}{(\Delta x)^2}\psi(i,j,k), \quad (10)$$

$$\frac{\partial^2\psi}{\partial y^2}\approx\sum_{v=-p}^{p}C_v\frac{\psi(i,j+v,k)}{(\Delta y)^2}=\sum_{v=-p}^{p}C_v\frac{\exp(-j_0vk_y\Delta y)}{(\Delta y)^2}\psi(i,j,k), \quad (11)$$

$$\frac{\partial^2\psi}{\partial z^2}\approx\sum_{v=-p}^{p}C_v\frac{\psi(i,j,k+v)}{(\Delta z)^2}=\sum_{v=-p}^{p}C_v\frac{\exp(-j_0vk_z\Delta z)}{(\Delta z)^2}\psi(i,j,k), \quad (12)$$

where C_v is the coefficients of 2*p*th-order spatial difference as shown in Table 3.

In Maxwell's equation and free space, the dispersion relation of the free electron can be written as

$$\omega=\left(\frac{\hbar}{2m}\right)|\vec{\mathrm{k}}|^2, \quad (13)$$

where $\vec{\mathrm{k}}$ is the wave vector with the amplitude *k*, *c* is velocity of the light. Similarly, the theoretic velocity of the Schrödinger equation is $v=(\hbar/2m)$ and the numerical velocity is $v_p=\omega/k_0^2$. Substituting the equations (7) and

(10), (11), (12) in to the wave equation to get the relative numerical dispersion error $\eta=20\log10|v_p/v-1|$ shows in the Fig. 1 with the stability factor $(\hbar\Delta t/\Delta 2m) = 0.25$ and wave propagation angle $\theta = 0°$ and $\phi = 0°$.

Table 3: Coefficients of 2*p*th-order spatial difference

(2,2*p*)	C-5	C-4	C-3	C-2	C-1	C0	C1	C2	C3	C4	C5
FDTD					1	-2	1				
(2,4)				-1/12	4/3	-5/2	4/3	-1/12			
(2,6)			-1/84	-1/84	97/84	-95/42	97/84	-1/84	-1/84		
(2,8)		-1/324	-1/324	-1/324	353/324	-175/81	353/324	-1/324	-1/324	-1/324	
(2,10)	-1/924	-1/924	-1/924	-1/924	489/462	-487/231	489/462	-1/924	-1/924	-1/924	-1/924

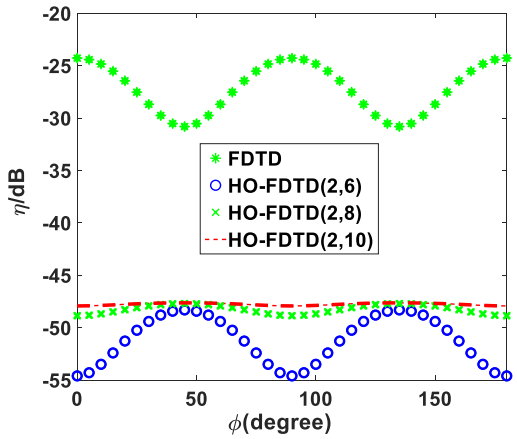


Fig. 1. Dispersion errors against ϕ with the incident angle θ = 30°of different HO-FDTD methods.

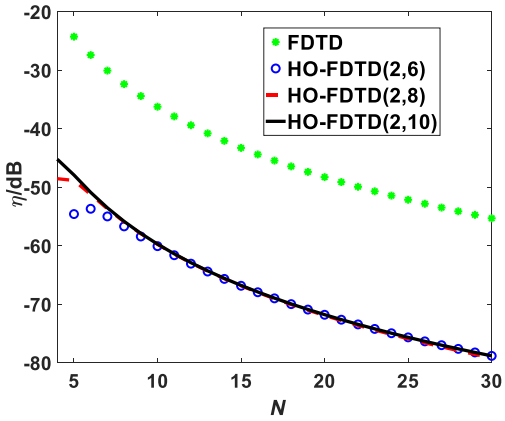


Fig. 2. Numerical dispersion against *N* of different HO-FDTD methods.

Figure 1 shows that HO-FDTD methods present the better numerical dispersion than FDTD. The higher order of the methods the more accuracy, but the HO-FDTD (2, 6) performs better. The relative numerical dispersion error η versus the number of cells per wavelength *N* for different HO-FDTD methods are illustrated in Fig. 2. The comparisons show that the HO-FDTD methods have the lower numerical dispersion and the higher order allows coarser cells per wavelength and less memory required.

IV. NUMERICAL SIMULATION

A particle trapped in an infinite potential well in one dimension are discussed as follows:

$$V(x)=\begin{cases}0, & 0<x<l\\ \infty, & \text{otherwise}\end{cases}, \qquad (14)$$

where the length of the well l is $0<l<1nm$, x is the position of the particle in axis. The cell size Δx is $0.015nm$. The total time step is 3000. The expression of energy for level L of the quantum well is quantized as:

$$E_L=\frac{1}{2m}\left(\frac{\hbar\pi L}{l}\right)^2, \quad L=1,2,3... \qquad (15)$$

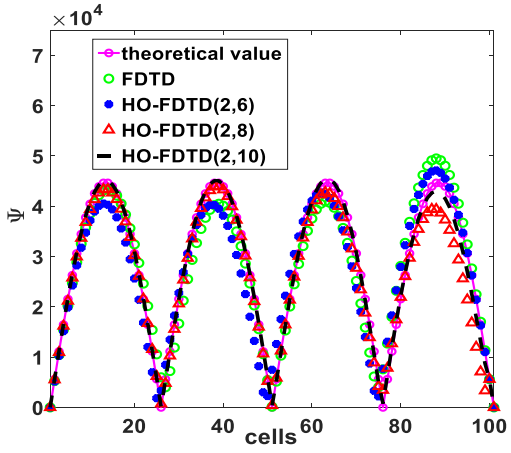


Fig. 3. The comparisons of simulations for the wavefunction with different methods and theoretical value in one dimensional infinite well.

Figure 3 shows the simulation of the FDTD and HO-FDTD algorithms for wavefunction with the energy level L = 4. The results show that the HO-FDTD more similar to the theoretical value than FDTD method and the HO-FDTD (2, 10) is the more closely to the theoretical value among other HO-FDTD methods. The time step size, spatial cell size, time steps, computational domain and CPU time consumption are listed in table 4.

It shows that the CPU time consumption of the HO-FDTD methods are more than FDTD. It is clear that the more complicated algorithm affords the better dispersion but need more computational time with the same computational condition.

Table 4: Computational parameters different methods

Methods	**FDTD**	**(2,6)**	**(2,8)**	**(2,10)**
Δx (*nm*)	0.015*nm*	0.015*nm*	0.015*nm*	0.015*nm*
Δt(s)	4.3188×10^{-21}	4.3188×10^{-21}	4.3188×10^{-21}	4.3188×10^{-21}
Spatial cells	150	150	150	150
Time steps	3000	3000	3000	3000
CPU time(s)	6.8750	8.6378	9.2674	9.5683

V. CONCLUSIONS

Different order of the HO-FDTD methods for solving time-dependent Schrödinger equation has been investigated in this paper. The simulation of the wave function for a particle in one dimensional infinite potential well are used to validate the accuracy of the HO-FDTD and FDTD methods. The results show that the HO-FDTD schemes afford the better numerical dispersion under the same computational condition and the higher order the more accuracy.

ACKNOWLEDGMENT

The work is supported by the High-Level Talent Foundation of Jinling Institute of Technology (Grant No. jit-b-201716, jit-b-201629, jit-b-201704).

REFERENCES

[1] S. Datta, *Quantum Transport: Atom to Transistor*, Cambridge University Press, New York, 2005.

[2] A. Z. Elsherbeni, *FDTD Course Notes*, Department of Electrical Engineering, The University of Mississippi, MS; Spring 2001.

[3] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equation in isotropic media," *IEEE Trans. Antennas Propag*. vol. AP-14, pp. 302-307, 1966.

[4] D. M. Sullivan, *Electromagnetic Simulation using the FDTD Method*, IEEE Press, New York, 2000.

[5] C. W. Manry, S. L. Broschat, and J. B. Schneider, "High-order FDTD methods for large problems," *Applied Computational Electromagnetics Society*, vol. 10, no. 2, pp.17-29, 1995.

[6] A. Goldberg, H. M. Schey, and J. L. Schwartz "Computer-generated motion pictures of one dimensional quantum mechanical transmission and reflection phenomena," *American Journal of Physics*, vol. 35, no. 3, pp.177-186, 1967.

[7] T. Shibata, "Absorbing boundary conditions for the finite-difference time-domain calculation of the one dimensional Schrödinger equation," *Physical Review B*, vol. 42, no.8, pp. 6760-6763, 1991.

[8] J. P. Kuska, "Absorbing boundary conditions for the Schrödinger equation on finite intervals," *Physical Review B*, vol. 46, no. 8, pp. 5000-5003, 1992.

[9] C. J. Ryu, A. Y. Liu, W. E. I. Sha, and C. W. Chew, "Finite-difference time-domain simulation of the Maxwell-Schrodinger system," *IEEE Journal on Multiscale and Multiphysics Comput. Techniques*, vol. 1, pp. 1-8, 2016.

[10] J. Houle, D. Sullivan, E. Crowell, S. Mossman, and M. G. Kuzyk, "Three dimensional time domain simulation of the quantum magnetic susceptibility," *IEEE Workshop on Microelectronics and Electron Devices*, 2019.

[11] K. Lan, Y. W. Li, and W. G. Lin, "A High order (2, 4) scheme for reducing dispersion in FDTD algorithm," *IEEE Trans. Electromagn. Compat*., vol. 41, no. 2, pp. 160-165, 1995.

[12] M. Krumpholz and L. P. B. Katehi, "MRTD: New time-domain schemes based on multiresolution analysis," *IEEE Trans. on Microwave Theory and Techniques*, vol. 44, no. 4, pp. 555-571, 1996.

[13] G. Xie, Z. X. Huang, and W. E. I. Sha, "Simulating Maxwell-Schrödinger equations by high-order symplectic FDTD algorithm," *IEEE Journal on Multiscal and Multiphysics Comput. Tech.*, vol. 4, pp. 143-151, 2019.

[14] Q. S. Cao, R. Kanapady, and F. Reitich, "High-order Runge-Kutta multiresolution time-domain methods for computational electromagnetics," *IEEE Trans. Microw. Theory Tech*., vol. 54, no. 8, pp. 3316-3326, 2006.

[15] G. Sun and C. W. Trueman, "Analysis and numerical experiments on the numerical dispersion of two-dimensional," *IEEE Antenna and Wireless Propag. Lett.*, vol. 2, no. 7, pp. 78-81, 2003.

[16] M. Zhu and Q. S. Cao, "Analysis for three-dimensional curved objects by Runge-kutta high order time-domain method," *Applied Computational Electromagnetics Society*, vol. 30, no. 1, pp. 86-92, 2015.

[17] M. Zhu and Q. S. Cao, "Studying and analysis of the characteristic of the high-order and MRTD and RK-MRTD scheme," *Applied Computational Electromagnetics Society*, vol. 28, no. 5, pp. 380-389, 2013.

[18] S. Gottlieb, C. W. Shu, and E. Tadmor, "Strong stability Strong stability-preserving high-order time discretization methods," *SIAM Rev*., vol. 43, no. 1, pp. 89-112, 2001.

[19] J. Shen, W. E. I Sha, Z. X. Huang, M. S. Chen, and X. L. Wu, "High-order symplectic FDTD scheme for solving a time-dependent Schrödinger equation," *Comput. Physics Comm.*, vol. 184, no. 3, pp. 480-492, 2013.

RK-HO-FDTD Scheme for Solving Time-dependent Schrodinger Equation

Min Zhu [1] and Yi Wang [2]

[1] School of Electronic and Information Engineering
Jinling Institute of Technology, Nanjing, 211169, China
zomi@jit.edu.cn

[2] College of Electronic and Information Engineering
Nanjing University of Aeronautics and Astronautics, Nanjing, 211106, China
jflsjfls@nuaa.edu.cn

***Abstract*—** The Runge-Kutta high-order finite-difference time-domain (RK-HO-FDTD) method is adopted to solve the time-dependent Schrodinger equation. The update equations of the RK-HO-FDTD method have been presented for wave function. The simulation results of the 1D potential well strongly confirm the advantages of the RK-HO-FDTD scheme over the conventional FDTD.

***Index Terms* —** Dispersion, potential well, RK-HO-FDTD, stability, the Schrodinger equation.

I. INTRODUCTION

Recently, more and more numerical solutions are employed to solve the Schrodinger equation [1]. The finite-difference time-domain (FDTD) [2] method also has been played an important role in quantum computational fields. The first attempt to use the FDTD algorithm for the Schrodinger equation was Goldberg et al. in 1967 [3], but it didn't get a lot of attention. 1990 years later, the topic began to cause greater attention [4-6], which is based on the Crank-Nicholson scheme. Soriano et al. rigorously formulated an efficient FDTD-Q algorithm and distinguished its application for quantum systems [7]. FDTD-Q method is one of the commonly adopted methods to solve the time-dependent eigenvalue problem of the Schrodinger equation, but it suffers from large dispersion errors in a long erm simulation [8-9]. The symplectic algorithm has been proposed to solve the Maxwell–Schrodinger (M-S) system for investigating light-matter interaction [11], but it suffers from intolerable dispersion errors in a long-term simulation.

The RK-HO-FDTD are proposed in [10] and has the better dispersion error and convergence. In this paper, the RK-HO-FDTD [10] scheme is used to solve the Schrodinger equation for one dimensional (1D) potential well problem. The paper is organized as follows. First, the basic theory and algorithm are introduced in Section II. The dispersion and stability for different methods are shown in Section III. Numerical example is given in Section IV. The simulation results show better numerical dispersion than the traditional FDTD and the HO-FDTD approach.

II. THEORY AND ALGORITHM

The time-dependent Schrodinger equation is written as follows:

$$i\hbar\frac{\partial\psi(\vec{\mathrm{r}},t)}{\partial t}=-\frac{\hbar^2}{2m}\nabla^2\psi(\vec{\mathrm{r}},t)+\mathrm{V}(\vec{\mathrm{r}})\psi(\vec{\mathrm{r}},t), \tag{1}$$

where ψ is wave function of a particle related to position and time t, m is the mass of the particle, $-\frac{\hbar^2}{2m}\nabla^2$ is the kinetic energy operator, $\mathrm{V}(\vec{r})$ is the time-independent potential energy, and $-\frac{\hbar^2}{2m}\nabla^2+\mathrm{V}$ is the Hamiltonian operator. The variable $\psi(\vec{\mathrm{r}},t)$ is a complex number that can be written as:

$$\psi(\vec{\mathrm{r}},t)=\psi_R(\vec{\mathrm{r}},t)+i\psi_I(\vec{\mathrm{r}},t). \tag{2}$$

Substituting (2) into (1), the equations can be obtained as:

$$\frac{\partial\psi_R(\vec{r},t)}{\partial t}=-\frac{\hbar}{2m}\frac{\partial^2\psi_I(\vec{r},t)}{\partial x}+\frac{1}{\hbar}\mathrm{V}(\vec{r})\psi_I(\vec{r},t), \tag{3}$$

$$\frac{\partial\psi_I(\vec{r},t)}{\partial t}=\frac{\hbar}{2m}\frac{\partial^2\psi_R(\vec{r},t)}{\partial x}-\frac{1}{\hbar}\mathrm{V}(\vec{r})\psi_R(\vec{r},t), \tag{4}$$

where ψ_R is the real part of wave function ψ, ψ_I is the imaginary part of wave function ψ.

The equation can be rewritten as:

$$\frac{\partial}{\partial t}\begin{pmatrix}\psi_R\\ \psi_I\end{pmatrix}=A\begin{pmatrix}\psi_R\\ \psi_I\end{pmatrix}, \tag{5}$$

$$\frac{\partial}{\partial t}F=AF, \tag{6}$$

where $A=\begin{pmatrix}0 & \kappa\\ -\kappa & 0\end{pmatrix}$, $\kappa=-\frac{\hbar}{2m}(\nabla^2)+\frac{\mathrm{V}}{\hbar}$, $F=\begin{pmatrix}\psi_R\\ \psi_I\end{pmatrix}$.

Submitted On: December 30, 2020
Accepted On: July 6, 2021

https://doi.org/10.47037/2021.ACES.J.360804

According to [10], we use the *p*th-order *p* stage SSP-RK scheme to discretized the temporal system in equation (5).

The SSP-RK scheme is written as follows:

$$F_{n+1}=\sum_{l=0}^{p}\alpha_{p,l}F^{(l)}, \tag{7}$$

where $F_n=F(t_n)$, and the coefficients $\alpha_{p,l}$ [12] are given by:

$$\alpha_{p,p}=\frac{1}{p!},\qquad l=p,$$

$$\alpha_{p,l}=\frac{1}{l}\alpha_{p-1,l-1},\quad l=1,2,\ldots p-1,\quad p\geq 2,$$

$$\alpha_{p,0}=1-\sum_{l=1}^{p}\alpha_{p,l},\quad l=0.$$

Based on the Staggered difference, the wave function ψ can be written as:

$$\psi^{n+l/p}(i,j,k)=\psi_I^{n+l/p}(i\Delta x,j\Delta y,k\Delta x). \tag{8}$$

Using the fourth order staggered difference to replace the spatial derivative in *x*-axis as follows:

$$\frac{\partial\psi^{n+l/p}(i)}{\partial x^2}=\frac{-\psi^{n+l/p}(i-2)+16\psi^{n+l/p}(i-1)-30\psi^{n+l/p}(i)}{12\Delta x^2}+\frac{+16\psi^{n+l/p}(i+1)-\psi^{n+l/p}(i+2)}{12\Delta x^2}. \tag{9}$$

Employing the SSP-RK algorithm to substitute time derivates and the Taylor series to replace spatial derivates, the update equations for the real part of the RK-HO-FDTD method can be derived as follows:

$$\begin{aligned}\psi_R^{n+l/p}(i,j,k)&=\psi_R^{n+(l-1)/p}(i,j,k)+\frac{\Delta t}{\hbar}\mathrm{V}(i,j,k)\psi_I^{n+l/p}(i,j,k)\\&-\alpha_{x1}\frac{\Delta t}{\Delta x^2}\frac{\hbar}{2m}\left[\psi_I^{n+l/p}(i+1,j,k)-2\psi_I^{n+l/p}(i,j,k)+\psi_I^{n+l/p}(i-1,j,k)\right]\\&-\alpha_{x2}\frac{\Delta t}{\Delta x^2}\frac{\hbar}{2m}\left[\psi_I^{n+l/p}(i+2,j,k)-2\psi_I^{n+l/p}(i,j,k)+\psi_I^{n+l/p}(i-2,j,k)\right]\\&-\alpha_{y1}\frac{\Delta t}{\Delta y^2}\frac{\hbar}{2m}\left[\psi_I^{n+l/p}(i,j+1,k)-2\psi_I^{n+l/p}(i,j,k)+\psi_I^{n+l/p}(i,j-1,k)\right]\\&-\alpha_{y2}\frac{\Delta t}{\Delta y^2}\frac{\hbar}{2m}\left[\psi_I^{n+l/p}(i,j+2,k)-2\psi_I^{n+l/p}(i,j,k)+\psi_I^{n+l/p}(i,j-2,k)\right]\\&-\alpha_{z1}\frac{\Delta t}{\Delta z^2}\frac{\hbar}{2m}\left[\psi_I^{n+l/p}(i,j,k+1)-2\psi_I^{n+l/p}(i,j,k)+\psi_I^{n+l/p}(i,j,k-1)\right]\\&-\alpha_{z2}\frac{\Delta t}{\Delta z^2}\frac{\hbar}{2m}\left[\psi_I^{n+l/p}(i,j,k+2)-2\psi_I^{n+l/p}(i,j,k)+\psi_I^{n+l/p}(i,j,k-2)\right],\end{aligned} \tag{10}$$

where Δx, Δy, Δz is the cell size and *i*, *j*, *k* is the spatial grid index in *x*, *y*, *z* direction, respectively, and Δt is the time step size, α_{x1}, α_{x2}, α_{y1}, α_{y2}, α_{z1} and α_{z2} are the coefficients of the HO-FDTD and $\alpha_{x1}=\alpha_{y1}=\alpha_{z1}=a(1)$, $\alpha_{x2}=\alpha_{y2}=\alpha_{z2}=a(2)$ as given in Table 1.

Table 1: Coefficients for HO-FDTD method

	$a(1)$	$a(2)$
FDTD	1	
HO-FDTD (2,4)	4/3	-1/12

III. NUMERICAL STABILITY AND DISPERSION

A. Stability of RK-HO-FDTD for the Schrodinger equation

The solution of the wave function can be represented as a superposition of plane wave on the basis of the von Neumann stability method as follows:

$$\psi(x,y,z,t)=A_m\exp(-j_0(i\Delta_x k_x+j\Delta_y k_y+k\Delta_z k_z)), \tag{11}$$

where $k_x=k_0\sin\theta\cos\varphi$, $k_y=k_0\sin\theta\sin\varphi$, $k_z=k_0\cos\theta$, k_0 is the wave number and $k_0=\mathrm{p}_m/\hbar$, p_m is the momentum, (θ, φ) is the wave propagation angle in spherical coordinate, j_0 is the imaginary unit.

For simplicity, we consider the Schrodinger equation in one dimension (1D) and set the potential energy zero (V=0). Using the m'th-order difference to discretize the spatial derivatives as follows:

$$\frac{\partial\psi}{\partial x^2}\approx\sum_{L=-m'/2}^{m'/2}C_L\frac{\psi(i,j,k+L)}{\Delta_x^2}=\sum_{L=-m'/2}^{m'/2}C_L\frac{\exp(-j_0Lk_x\Delta_x)}{\Delta_x^2}\psi(i,j,k)=T_x\psi, \tag{12}$$

where $T_x=\sum_{L=-m'/2}^{m'/2}C_L\frac{\exp(-j_0Lk_x\Delta_x)}{\Delta_x^2}$, $\mathrm{C_L}$ is the coefficients of *m*′th-order spatial difference which is listed in Table 2.

Table 2: Coefficients of *m*′th-order spatial difference [14-15]

Order (m')	C_1	C_2	C_3	C_4	C_5
FDTD	1	-2	1		
HO-FDTD (2,4)	4/3	-1/12	-5/2	4/3	-1/12

Equation (5) can be rewritten as:

$$\frac{\partial}{\partial t}\begin{pmatrix}\psi_R\\ \psi_I\end{pmatrix}=H\begin{pmatrix}\psi_R\\ \psi_I\end{pmatrix}, \tag{13}$$

$$H=\begin{pmatrix}0 & -\frac{\hbar}{2m}T_x\\ \frac{\hbar}{2m}T_x & 0\end{pmatrix}, \tag{14}$$

where *H* is spatial growth matrix, the eigen equation of *H* is written as $|\lambda\mathrm{I}-H|=0$ which can be used to solve the eigen value, the positive sign solution of the eigen value

can be derived as follows:

$$\left|\lambda \mathrm{I}\text{-}H\right| = \lambda^2 + (\frac{\hbar}{2m}T_x)^2 = 0\,, \tag{15}$$

$$\begin{aligned}\lambda &= j_0 \frac{\hbar}{2m}T_x \\ &= j_0\lambda_j\,,\end{aligned} \tag{16}$$

where λ_j is the imaginary part of λ.

Equation (7) of the SSP-RK method can be rewritten [9], [12] as:

$$F(t_{n+1}) = \sum_{l=0}^{p}\frac{1}{l!}(\Delta t H)^l F(t_n) = GF(t_n)\,, \tag{17}$$

$$\xi = \sum_{l=0}^{p}\frac{1}{l!}(\lambda\Delta t)^l\,, \tag{18}$$

where G is the total gain factor of F, ξ is the gain factor of F when $H = \lambda$. The SSP-RK algorithm is stable and convergent under the condition of $|G| \le 1$ and $|\xi| \le 1$, and the solution of the equation (17) can be written as $\lambda_j\Delta t \le C_f$, C_f is a constant. The stability condition of the RK$_3$-HO-FDTD (2, 4) when p = 3 for the Schrodinger equation can be derived as:

$$\xi = \left|\sum_{l=0}^{3}\frac{1}{l!}(\lambda\Delta t)^l\right| = \left|\sum_{l=0}^{3}\frac{1}{l!}(j_0\lambda_j\Delta t)^l\right| \le 1\,, \tag{19}$$

$$\begin{aligned}&\left|\left(1-\frac{1}{2}(\lambda_j\Delta t)^2\right) + j_0\left(\lambda_j\Delta t - \frac{1}{6}(\lambda_j\Delta t)^3\right)\right| \le 1 \\ &\left(1-\frac{1}{2}(\lambda_j\Delta t)^2\right)^2 + \left(\lambda_j\Delta t - \frac{1}{6}(\lambda_j\Delta t)^3\right)^2 \le 1 \\ &1+\frac{1}{4}(\lambda_j\Delta t)^4 - (\lambda_j\Delta t)^2 + (\lambda_j\Delta t)^2 + \\ &\qquad \frac{1}{36}(\lambda_j\Delta t)^6 - \frac{1}{4}(\lambda_j\Delta t)^4 \le 1 \\ &\qquad \frac{1}{36}(\lambda_j\Delta t)^2 \le \frac{1}{12} \\ &\qquad \lambda_j\Delta t \le \sqrt{3}.\end{aligned} \tag{20}$$

In equation (20) $C_f = \sqrt{3}$ and C_f is dependent on the order of the time discretization, then C_f of the RKp-HO-FDTD (2, 2m') can be deduced in the same way. From the equation (16), it can be found that the spatial stability condition is dependent on λ_j correspond to the spatial discretization order m. To sum up, the stability condition of the RK-HO-FDTD method for the Schrodinger equation is decided by equation (16) and (20), then the general form of stability condition for the RKp-HO-FDTD (2, 2m') [11] can be given as:

$$\Delta t \le C_f \frac{m\Delta^2}{\hbar\sqrt{d}\sum_{\mathrm{L'}=1}^{m'}|a(m')|} = \beta\frac{m\Delta^2}{\hbar}\,, \tag{21}$$

$$\beta = \frac{C_f}{\sqrt{d}\sum_{L'=1}^{m'}|a(m')|}\,, \tag{22}$$

where $\Delta x = \Delta y = \Delta x = \Delta$, L' is the number of the coefficients in Table 1, d is the spatial dimension, equation (21) is the Courant-Friedrichs-Lewy (CFL) stability condition, β is the CFL stability factor which is shown in Table 3. The investigation is that the RK$_3$-HO-FDTD (2, 4) has the looser stability condition than other methods and the SSP-RK method can increase the algorithm stability.

Table 3: Stability factor for different methods

	FDTD	HO-FDTD (2, 4)	RK$_3$-HO-FDTD (2, 4)	RK$_4$-HO-FDTD (2, 4)
β	$1/\sqrt{d}$	$0.7059/\sqrt{d}$	$1.2226/\sqrt{d}$	$1.9965/\sqrt{d}$

B. Dispersion of RK-HO-FDTD for the Schrodinger equation

In Maxwell's equation, the dispersion relation of the free electron in free space can be written as:

$$\omega = \left(\frac{\hbar}{2m}\right)\left|\vec{k}\right|^2\,, \tag{23}$$

where c is velocity of light, $\vec{\mathrm{k}}$ is the wave vector with the amplitude k, and $k_x = k\sin\theta\cos\phi$, $k_y = k\sin\theta\sin\phi$, $k_z = k\cos\theta$, (θ, ϕ) is the wave propagation angle in spherical coordinate. Similarly, the theoretic velocity of the Schrodinger equation is $v = \left(\frac{\hbar}{2m}\right)$ and the numerical velocity is $v_p = \frac{\omega}{k_0^2}$. The dispersion relation can be derived from the expansion of the plane wave and the SSP-RK theory [13] as follows:

$$\omega\Delta t = \mathrm{Arg}(\xi)\,. \tag{24}$$

According to the CFL stability condition, defining the stability factor $S = (\hbar\Delta t/\Delta 2m) = 0.25$, and with the wave propagation angle $\theta = 0°$ and $\phi = 0°$.

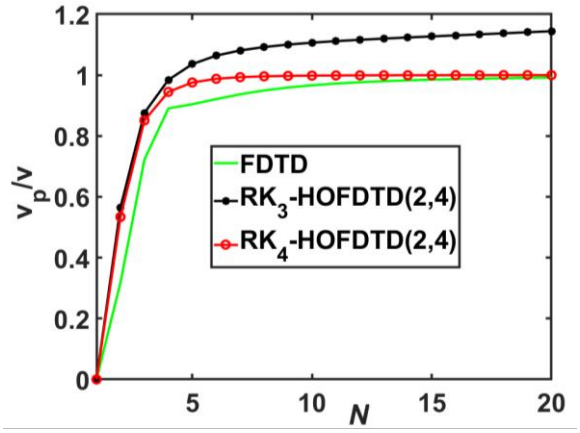


Fig. 1. Dispersion errors against N of the FDTD and RK-HO-FDTD method.

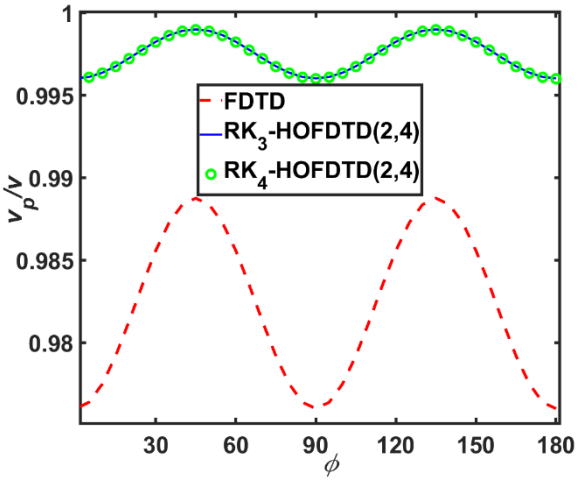


Fig. 2. Dispersion errors against ϕ with the incident angle $\theta = 30°$ of the FDTD and RK-HO-FDTD method.

Figure 1 shows the dispersion errors v_p / v versus the number of cells per wavelength N for the FDTD, RK_3-HO-FDTD(2, 4) and RK_4-HO-FDTD(2, 4) methods. The results show that the RK-HO-FDTD methods have the lower numerical dispersion and the higher order allows coarser cells within a given accuracy. The dispersion characteristic of the RK_3-HO-FDTD is not good, the reason maybe the temporal order 3 is not matching with the spatial order 4. There is no difference between RK_4-HO-FDTD (2, 4) and FDTD when N is greater than 20. Figure 2 demonstrates that RK-HO-FDTD method present the better dispersion error.

IV. NUMERICAL SIMULATION

In order to demonstrate the efficiency of the RK-HO-FDTD methods for 1D infinite potential well simulation is presented. Considering a particle trapped in an infinite potential well as:

$$V(x) = \begin{cases} 0, & 0 < x < d \\ \infty, & \text{otherwise} \end{cases}, \tag{25}$$

where d is the length of the well and x is the position of the particle in axis.

Without loss of generality, the size of the d is $0<d<10^{-9}$m and the cell size Δx is 10^{-11}m. The total time step is 2000. The eigenenergies of the quantum well are quantized as:

$$E_n = \frac{h^2\pi^2}{2ma^2}n^2, \quad n = 1, 2, 3... \tag{26}$$

The wavefunction calculated with the FDTD and RK-HO-FDTD schemes are drawn in Fig. 3. It shows that the RK-HO-FDTD schemes accord closely with the theoretical predictions than FDTD method. Table 4 gives the temporal discretization, spatial discretization, total computational domain, total time steps and CPU time. From the data, we know that the CPU times of the HO-FDTD (2,4) is more than the FDTD. The conclusion is the better dispersion need more computational time with the same computational condition.

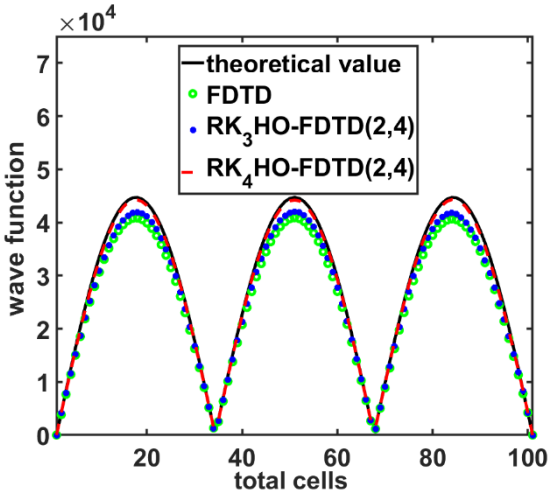


Fig. 3. The comparisons of the wavefunction calculated in different methods and theoretical value in 1D infinite well.

Table 4: Computational time for FDTD and RK-HO-FDTD method

Methods	FDTD	RK3-HO-FDTD (2,4)	RK4-HO-FDTD (2,4)
Δx (m)	10^{-11}	10^{-11}	10^{-11}
Δt	6.5039× 10^{-19}	6.5039× 10^{-19}	6.5039× 10^{-19}
Total cells	100	100	100
Total time	2000	2000	2000
CPU time(s)	5.995	13.296	18.873

V. CONCLUSIONS

The characteristics of the RK-HO-FDTD methods have been discussed in this paper for solving the Schrodinger equation. The RK-HO-FDTD and FDTD method are implemented to mode the wave function of a particle in 1D infinite potential well. The results show that the scheme can increase the accuracy of wavefunction simulation.

ACKNOWLEDGMENT

The work is supported by the High Level Talent Research start-up project of Jinling Institute of Technology under the contract jit-b-201716.

REFERENCES

[1] S. Datta, *Quantum Transport: Atom to Transistor*. Cambridge University Press, New York, 2005. A.

Z. Elsherbeni, *FDTD Course Notes*, Department of Electrical Engineering, The University of Mississippi, MS, Spring 2001.

[2] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propag*., vol. AP14, no. 3, pp. 302-307, May 1966.

[3] A. Goldberg, H. M. Schey, and J. L. Schwartz, "Computer-generated motion pictures of one dimensional quantum mechanical transmission and reflection phenomena," *American Journal of Physics*, vol. 35, no. 3, pp. 177-186, Mar. 1967.

[4] T. Shibata, "Absorbing boundary conditions for the finite-difference time-domain calculation of the one dimensional Schrodinger equation," *Physical Review B*, vol. 42, no. 8, pp. 6760-6763, Mar. 1991.

[5] J. P. Kuska, "Absorbing boundary conditions for the Schrodinger equation on finite intervals," *Physical Review B*, vol. 46, no. 8, pp. 5000-5003, Aug. 1992.

[6] D. M. Sullivan, *Electromagnetic Simulation Using the FDTD Method*. New York, NY: IEEE Press, 2000.

[7] A. Soriano, E. A. Navarro, J. A. Porti, and V. Such, "Analysis of the finite difference time domain technique to solve the Schrodinger equation for quantum devices," *Journal of Applied Physics*, vol. 95, no. 12, pp. 8011-8018, June 2004.

[8] A. Soriano, E. A. Navarro, J. A. Porti, V. Such, *J. Appl. Phys*., vol. 95, pp. 8011-8018, 2004.

[9] D. M. Sullivan and D. S. Citrin, *J. Appl. Phys*., vol. 97, 2005.

[10] M. Zhu and Q. Cao, "Study and analysis of a novel Runge-Kutta high-order finite-difference time-domain method," [J]. *IET Microw. Antennas Propag*., vol. 8, no. 12, pp. 951-958, 2014.

[11] G. Xie, Z. X. Huang, and W. E. I Sha, "Simulating Maxwell-Schrodinger equations by high-order symplectic FDTD algorithm," *IEEE Journal on Multiscale and Multiphysics Computational Techniques*, vol. 4, pp. 143-151, 2019.

[12] Q. Cao, R. Kanapady, and F. Reitich, "High-order Runge-Kutta multiresolution time-domain methods for computational electromagnetics," *IEEE Trans. Microw. Theory Tech*., vol. 54, no. 8, pp. 3316-3326, Aug. 2006.

[13] G. Sun and C. W. Trueman, "Analysis and numerical experiments on the numerical dispersion of two-dimensional," *IEEE Antenna and Wireless Propagation Lett*., vol. 2, no. 7, pp. 78-81, 2003.

[14] M. Zhu and Q. Cao, "Analysis for three-dimensional curved objects by Runge-Kutta high order time-domain method," [J]. *Applied Computational Electromagnetics Society Journal*, vol. 30, no. 1, pp. 86-92, Jan. 2015.

[15] M. Zhu and Q. Cao, "Studying and analysis of the characteristic of the high-order and MRTD and RK-MRTD scheme," *Applied Computational Electromagnetics Society Journal*, vol. 28, no. 5, pp. 380-389, May 2013.

A Recovery Performance Study of Compressive Sensing Methods on Antenna Array Diagnosis from Near-Field Measurements

Oluwole John Famoriji and Thokozani Shongwe

Department of Electrical and Electronic Engineering Science
University of Johannesburg, P.O. Box 524, Auckland Park, 2006, Johannesburg, South Africa
famoriji@mail.ustc.edu.cn, tshongwe@uj.ac.za

Abstract — Antenna testing consists locating the potential defaults from radiated field measurements. It has been established in literature, that compressive sensing methods provide faster results in failure detection from smaller number of measurement data compared to the traditional back-propagation mechanisms. Compressive sensing (CS) methods require a priori measurement of failure-free reference array and require small number of measurements for diagnosis. However, there are conflicting reports in literature regarding the choice of appropriate CS method, and there is no sufficient comparison study to justify which one is a better choice under a very harsh condition. In this study, recovery performance test of CS methods for the diagnosis of antenna array from few near-field measured data under various signal-to-noise ratios (SNRs) is presented. Specifically, we tested three prominent regularization procedures: total variation (TV), mixed l_1/l_2 norm, and minimization of the l_1 in solving diagnosis problems in antenna array. Linear system that relates the difference between near-field measured data from reference antenna (RA) array and array under test (AUT), and the difference that exist between coefficients of RA and the AUT, is solved by the three compressive sensing regularization methods. Numerical experiment of a 10×10 rectangular waveguide array under realistic noise scenario, operating at 10 GHz is used to conduct the test. Minimization l_1 technique is more robust to additive data noise. It exhibits better diagnosis at 20 dB and 10 dB SNR, making it a better candidate for noisy measured data as compared to other techniques.

Index Terms — Antenna array diagnosis, antenna testing, compressive sensing, near-field, low SNR.

I. INTRODUCTION

Nowadays, near-field equipment is used for routine array test. Apart from radiation pattern measurements, diagnosis of antenna array is another major application. Current and future technologies employ sophisticated active or phased arrays with large elements. For example, large array employed in RADAR systems, full MIMO systems, massive MIMO, and personal communication devices that require complex antenna arrays. As a result, there will always be a demand for fast and accurate complex antenna systems diagnosis, to resolve the unacceptable radiation pattern distortion caused by element (s) failure in the array.

Many antenna array diagnosis methods, based on genetic algorithms [1], [2], exhaustive search [3], matrix inversion [4], and MUSIC [5], have been developed in literature to identify faulty antenna elements in an array. All the methods compare the array under test (AUT) with the radiation pattern of an "error free" reference array. All the methods in [1-5] need big measurement samples for large antenna arrays, to get reliable diagnosis. Reducing the measurement samples, compressive sensing (CS) based techniques have been reported in [6]-[9]. Despite the compelling outcome, the methods in [1]-[9] focused on the detection of sparsity pattern of a failed array, i.e., failed elements location, not on the complex blockage, as addressed in [10]. Recently, compressive sensing approaches have shown great advantages in terms of reduced number of measurement data, reconstruction accuracy, and simulation time [11]. Compressive sensing (CS) is a method of signal processing through which it is possible to reconstruct or recover a signal from a set of linear measurements rather than the original signal itself, and the measurements set is less than the signal. Consequently, the primary signal will be reconstructed from measurement matrix, which is ill-posed as a result of decreased dimension.

In literatures, such as [7-11], differential scenario with sparse recovery algorithms have been used to diagnose antenna arrays and retrieve element excitations. This method results to a small number of unknown, however, they require a well detailed array model with the exact information of the radiation patterns of the antenna to generate appropriate results. Particularly, total variation (TV) norm, mixed ℓ_1/ℓ_2 norm and minimization of the ℓ_1 norm were tested to proffer solution to an inversion issue. However, based on the acquired information, no work has thoroughly compared the methods, particularly the recovery performance

Submitted On: June 15, 2018
Accepted On: June 10, 2021

https://doi.org/10.47037/2021.ACES.J.360805

under different signal-to-noise ratios (SNRs). For a noisy measured data, which of those methods is more suitable? i.e., which method gives better recovery performance for near-field measured data with low SNR? This paper provides answer to this question. In addition, there are contradicting reports and recommendations in literature. For instance, Fuchs *et al*. [11] recommended the use of mixed ℓ_1/ℓ_2 norm while Migliore [7] recommended minimization of the ℓ_1 norm. Again, these are conflicting results necessitating the further study.

In CS mechanisms, the number of measurements required increases logarithmically and slowly based on the number of unknowns [12-17]. Hence, field synthesis scheme benefits more in sparse recovery-based methods [11]. Modeled diagnostic issues can be evaluated by using the available customs whose calculation times are little more than the standard approaches. Therefore, it is good to show that total time required in getting the array of antenna diagnosed majorly depend on time taken in measurement, with post-processing time of higher magnitude faster. That is the reason that sparse recovery methods with few numbers of measurements will provide faster antenna array diagnosis. Researchers have proposed various compressive sensing algorithms [11], [12]-, [25] but only three of them are reported in this paper because of their robustness and efficiency [4]. Total variation (TV) norm, mixed ℓ_1/ℓ_2 norm and minimization of the ℓ_1 norm methods are adapted and applied to a simulated near-field data of a 10 GHz waveguide array with 100 elements in which failures had been added intentionally. And the recovery performance under low SNR was evaluated, to determine the best algorithm fit for such scenario.

CS methods provide good, reliable, and accurate antenna array diagnosis at low SNR compared to the conventional methods. This is a useful feature, particularly in very harsh measurement environment. Nevertheless, no work has thoroughly compared the CS methods. Furthermore, in a very harsh measurement environment (i.e., low SNR), which of these methods is preferable? That is, which of the methods provide better recovery performance for low SNR near-field measurement? This paper answers this question, specifically; total variation (TV) norm, mixed ℓ_1/ℓ_2 norm and minimization of the ℓ_1 norm were examined to solve an inversion problem. The results obtained will influence the choice of CS diagnosis method especially under very low SNR measurements.

II. DIAGNOSIS PROBLEM OF ANTENNA ARRAY

Here, we consider a rectangular radiating antenna array in space (Fig. 1). The radiated field of antenna is usually considered in phase or/and amplitude within the near-field region. The AUT is as shown in Fig. 1 (b). The associated parameters of AUT is marked with "u" as superscript. Especially, $\boldsymbol{E}^u(x,y)$ is the tangential field that is concentrated on the antenna aperture, i.e.,

$$\boldsymbol{E}^u(x,y) = E_x^u(x,y)\hat{x} + E_y^u(x,y)\hat{y}, \qquad (1)$$

where $E_x^u(x,y)\hat{x}$ and $E_y^u(x,y)\hat{y}$ are components x and y of the electric field situated on the aperture Σ respectively. Near-field $N^u(r,\theta,\emptyset)$ is field measured on part of hemispherical surface $(0 \le \theta \le \pi/2, 0 \le \phi \le 2\pi)$ at radius r from center of AUT, and $r < 2D^2/\lambda$, D is diameter of the antenna. Also, the near-field of RA is assumed available. The associated parameters are with "o" superscript. $\boldsymbol{E}^o(x,y)$ is the field on RA aperture Σ and $\boldsymbol{N}^o(r,\theta,\emptyset)$ denotes the far-field radiated. For the DA depicted in Fig. 1 (c), the tangential distribution $\boldsymbol{E}(x,y)$ on its aperture equals the difference between the RA and AUT field, and the corresponding near-field $\boldsymbol{N}(r,\theta,\phi)$ is expressed as the difference between the near field of RA and AUT as:

$$\boldsymbol{E}(x,y) = \boldsymbol{E}^u(x,y) - \boldsymbol{E}^o(x,y), \qquad (2)$$

$$\boldsymbol{N}(r,\theta,\phi) = \boldsymbol{N}^u(r,\theta,\phi) - \boldsymbol{N}^o(r,\theta,\phi). \qquad (3)$$

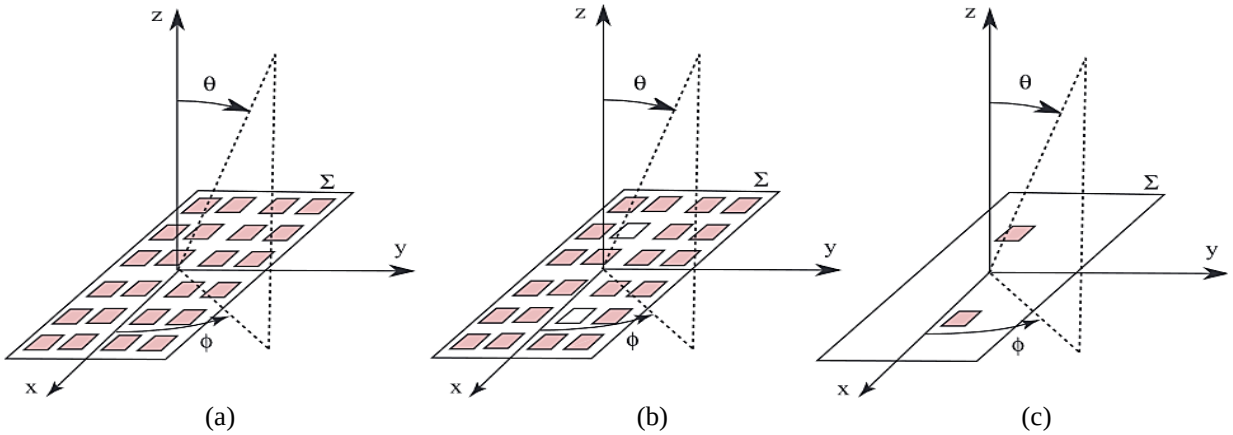


Fig. 1. Antenna array decomposition and modeling for diagnosis. (a) Reference antenna (RA), (b) antenna under test (AUT), and (c) differential antenna (DA). Number of failures here is *2* for 24 elements, and Σ is the aperture [11].

III. COMPRESSED SPARSE RECOVERY TECHNIQUES

The role of matrix inversion technique is to introduce *a priori* knowledge within the inversion. The efficient approach that is required in getting this regularization is by approximately reducing the selected norm *q* of **x** solution. Hence, the optimization problem to be solved becomes:

$$\min_{\mathbf{x}} \|X\|_q \text{ subject to } \|\boldsymbol{y} - \boldsymbol{AX}\|_2 \leq \gamma, \quad (4)$$

where $\|\cdot\|_q$ represents l_q norm, and γ is a function of noise and factors affecting the measurement. Different routines solutions are readily available to solve the convex optimization problem of Eqn. (4), such as [19-25]. The three norms l_q, chosen according to *a priori* knowledge of a differential antenna diagnosis setup, can then be explained for inversion regularization. We can consequently apply them to effectively diagnosis radiating elements.

A. TV norm

According to *a priori* knowledge that solution x has small discontinuities due to the presence of failures. Besides the failures, it is expected that the field x is to be made almost zero. Therefore, TV-norm is a smooth function to regularize x [25]. So, minimizing TV-norm is minimizing its gradient, this is the effect of smoothing. Consider 2-dimensional complex data set $X \in \mathbb{C}^{M\times N}$, TV-norm gives:

$$\|X\|_{TV} = \sum_{m,n} \left|\boldsymbol{X}_{m+1,n} - \boldsymbol{X}_{m,n}\right| + \left|\boldsymbol{X}_{m,n+1} - \boldsymbol{X}_{m,n}\right| \quad (5)$$
$$\|vec(\boldsymbol{\nabla}_x \boldsymbol{X})\|_1 + \left\|vec(\boldsymbol{X}\boldsymbol{\nabla}_y)\right\|_1,$$

vec (**X**) produces vector *N by M* that has the columns **X,** stacked beside each other. Gradient matrix $\boldsymbol{\nabla}_x$ and $\boldsymbol{\nabla}_y$ are made up of $M \times M$ and $N \times N$ size, respectively. They are computed as:

$$\boldsymbol{\nabla}_x = \begin{bmatrix} -1 & & 1 & 0 \\ & \ddots & & \ddots \\ 0 & & -1 & 1 \end{bmatrix} \text{ and}$$

$$\boldsymbol{\nabla}_y = \begin{bmatrix} -1 & & 0 \\ 1 & \ddots & \\ & \ddots & -1 \\ 0 & & 1 \end{bmatrix}.$$

Then, the problem of optimization in Eqn. (4) becomes:

$$\min_{\boldsymbol{X}} \|\boldsymbol{X}\|_{TV} \text{ subjects to } \|\boldsymbol{y} - \boldsymbol{A}vec(\boldsymbol{X})\|_2 \leq \epsilon. \quad (6)$$

B. ℓ_1 norm

Since sparse solution *X* exists, then a search space can be reduced by initiating *a priori* knowledge in inversion. Particularly, the ℓ_1-norm ($\|X\|_1 = \sum_k |x_k|$) is the leading convex surrogate of the appropriate vector estimate. That is, the quasi-norm ℓ_0 that calculates nonzero occurrences of a particular vector). Consequently, ℓ_1 norm technique is an effective method for enhancing the sparse solution [7], [8], [10], [11]. The problem of the regularization is:

$$\min_X \|X\|_1 \text{ subject to } \|\boldsymbol{y} - \boldsymbol{A}vec(\boldsymbol{X})\|_2 \leq \epsilon. \quad (7)$$

Minimization of the ℓ_1-norm forces the pointwise sparsity of solution per sample x_k of EM field on the DA aperture.

C. ℓ_1/ℓ_2-norm

The position and dimension of the radiating aperture can be taken. The solution *X* is grouped into *G* groups X^g that corresponds to each aperture of the radiating element *g*. For a faulty element, all regions of discretization x_k^g within the aperture becomes nonzero. Let vector X of dimension *MN* be divided into *G* non-overlapping groups depicted as X^g of size N_g, such as $\sum_{g=1}^{G} N_g = MN$. Then, the mixed ℓ_1/ℓ_2-norm is given as:

$$\|X\|_{1,2} = \sum_{g=1}^{G} \|X^g\|_2 = \sum_{g=1}^{G} \sqrt{\left|x_1^g\right|^2 + \cdots + \left|x_{N_g}^g\right|^2}. \quad (8)$$

Mixed ℓ_1/ℓ_2-norm have similar behavior with ℓ_1 Norm on vector$\|X^1\|_2$,…, $\|X^g\|_2$,… $\|X^G\|_2$, it therefore induces group sparsity at the radiating aperture level. The regularized inversion optimization problem is given as:

$$\min_X \|X\|_{1,2} \text{ subject to } \|\boldsymbol{y} - \boldsymbol{A}X\|_2 \leq \epsilon. \quad (9)$$

IV. NUMERICAL SIMULATIONS

Considering an open-ended waveguide array constituted by a 10×10 *WR90* waveguides that operates at 10 GHz, aperture size $22.86 \times 10.16\ mm^2$, and spaced uniformly by λ *and* $\lambda/2$ along *x*- and *y*-planes, respectively, as depicted in Figs. 2 (a-c). Because normalized patterns provide sufficient information to obtain substantial results, we considered the normalized pattern as shown in Figs. 2 (d, e). The antenna array's radiation pattern is computed using antenna toolbox in Matlab software. At first, all the elements *N* in the array are excited with the same value in order to emulate RA (i.e., array without failure). To model the AUT, *K* failures in either phase δ_Φ or amplitude δ_A are added.

Practically, noise contaminates measurements; hence, we added a Gaussian noise **n** to the radiation pattern of the reference and defaults as $\boldsymbol{y}_n^q = \boldsymbol{y}^q + \boldsymbol{n}^q$ where $q = \{r, d\}$. Noise level is computed by SNR which could be extracted from the highest magnitude of received field to fit with dynamic range of measurement. Then the noise can be given as:

$$n^q = \frac{\mathbb{N}(0,1) + j\mathbb{N}(0,1)}{\sqrt{2}}\, max|\boldsymbol{y}^q|.10^{-SNR_{dB}/20}, \quad (10)$$

where $\mathbb{N}(0,1)$ represents Gaussian random vector with 0 mean and 1 standard deviation. SNR can be varied in random near-field measurements, and subjecting the

compressive sensing methods to different scenario can provide recovery performance of each methods.

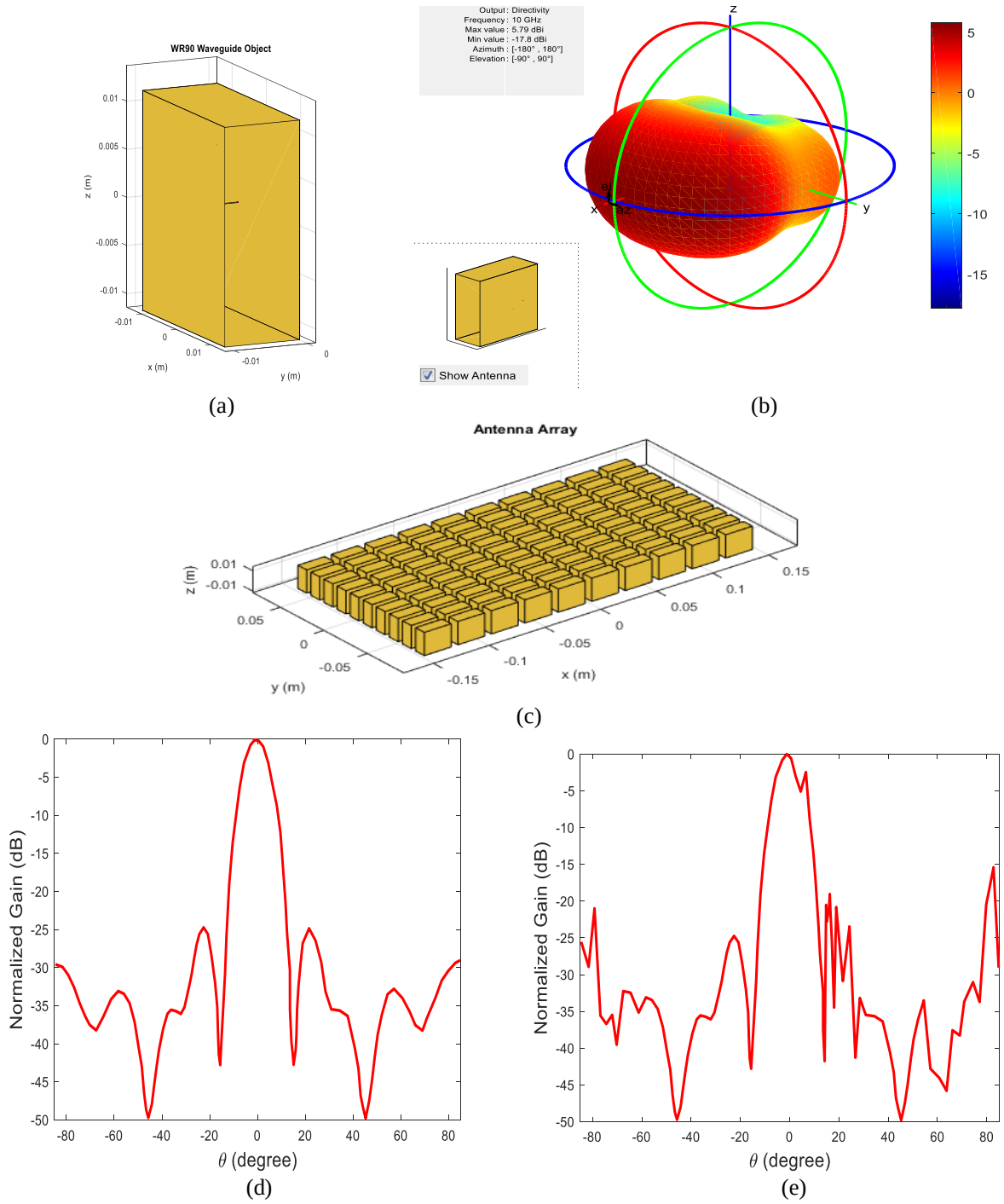


Fig. 2. Simulation setup. (a) Waveguide, (b) radiation pattern, (c) antenna array, (d) normalized radiation pattern with no element failure, and (e) normalized radiation pattern with 6 elements failures.

V. RECOVERY PERFORMANCE

Quantifying the diagnostic performances, the following indicator is introduced. Firstly, we add the field's magnitude samples x_k situated on each element's aperture g: $\alpha_g = \sum_g |x_k^g|, for\ g = 1, \dots, G$. From that, we get a positive number for each radiating element g, the difference Δ_{FA} between lowest failure level and highest false alarm is computed. Δ_{FA} value is the margin

set as threshold to differentiate between true failure and false alarm. Higher margin Δ_{FA} implies easier diagnostic, while a negative Δ_{FA} implies incorrectly performed diagnostic because false alarm value appears bigger than a failure.

For all the methods adopted, we repeated the simulation for 120 times with the incorporation of the Gaussian white noise. The simulation results given in this study are average figures over the 120 simulation times, so as to ensure appreciable results. The total number of measurement employed for investigation is $12 \times 12 = 144$ for all the methods because choosing higher measured points have no effect on the recovery performance. The associated parameter ϵ in the data-fitting of the sparse recovery techniques (Eqn. 4) is selected to be higher than noise intensity. We specifically put ϵ at $1.1\|\boldsymbol{n}\|_2$ while in measurements, an estimation of the SNR is employed to calculate ϵ. Anechoic chamber that exhibits various SNR is assumed, and we set the value of ϵ accordingly.

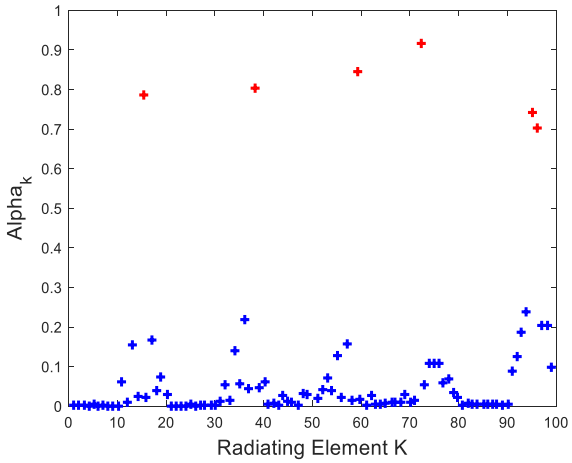


Fig. 3. Result sample of array diagnosis. $alpha_k$ is the radiating element k level. Δ_{FA} is the distance or gap between the least of the six errors (red) and the highest false alarm (black).

D. Amplitude failures

At first we consider K=6 waveguides are excited with $1 - \delta_A$ amplitude as against 1 to initiate amplitude failures. Accessing the recovery performance, the margin Δ_{FA} is computed for different SNRs. The results are reported in Fig. 3 and Tables 1 and 2 respectively.

It is simpler to conduct diagnosis when amplitude error is more important which is as expected. Generally, for a particular SNR, Δ_{FA} is higher when $\delta_A = 1$ than $\delta_A = 0.1$. The ℓ_1 Norm exhibit highest Δ_{FA} in both cases, this implies best diagnostic performance at different SNRs. Though it took longer simulation time than other compressed sensing methods.

E. Phase failures

In this case, K=6 waveguides are excited with amplitude of $e^{j\delta_\Phi}$ instead of 1 to emulate failures in phase. Δ_{FA} was computed for different SNRs to evaluate the recovery. The results obtained are shown in Tables 3 and 4. All the comments made for amplitude failures also hold here. As expected, larger phase error is much easier to diagnose than smaller one. The best diagnostic results are obtained by ℓ_1 norm approach.

Generally, the diagnosis is weak at lower SNRs for both amplitude and phase errors. Therefore, this is a gap to be filled up by new algorithms, which is open for research and development. The algorithm should exhibit higher Δ_{FA} recovery performance both at higher and lower SNRs. Because most near-field equipment exhibit different degrees of SNRs, it will make the diagnosis better and correct. But based on this study, ℓ_1 norm offer better diagnosis at lower SNRs; 20 dB and 10 dB from near-field. In all the cases considered for both amplitude and phase failures, and results obtained in Table 1-4, ℓ_1 norm has widest Δ_{FA} gap among the three regularizers considered. Next is TV, which performs better than the mixed l_1/l_2 norm. For instance, at 10 dB SNR of Table 1-4, only ℓ_1 norm gives value for Δ_{FA}. This case is the same for $\delta_A = 1$ and $\delta_A = 0.1$ amplitude failure configurations, and $\delta_\Phi = 70^0$ and $\delta_\Phi = 10^0$ phase failure configurations. Therefore, ℓ_1 norm is the best choice among the three CS methods for all SNR levels considered.

Ensuring common ground for comparison with reference [11], the diagnosing scheme and the AUT configurations are made similar to [11], but the diagnosing results are different. In [11], the mixed l_1/l_2 norm minimization indicates best performance, while in this study; the ℓ_1 norm minimization indicates best performance. The disparity might be because of the nature of field considered. In [11], the CS methods are investigated at far-field, but in this work and that of Migliore [7], minimization $l_1$technique show best performance and more robust against measurement noise. Probably because [7] and this work are investigated at near-field (i.e., measurement is taken on part of hemispherical surface $(0 \le \theta \le \pi/2, 0 \le \phi \le 2\pi)$ at radius r from center of AUT, and $r < 2D^2/\lambda$. The minimization $l_1$technique does not require any hardware modification of standard (i.e., not scattering-modulated based) near-field measurement systems allowing an increasing of the "throughput" of the array testing process at practically zero cost.

The key point regularization procedure minimizing the 1-norm of the difference vector between a failure-free excitation vector and the excitation vector of the AUT. This allows to obtain an equivalent sparse array, discarding the pieces of information not of interest for the failure identification problem.

This study centers on numerical experiments and shows that the ℓ_1 norm technique gives best performance. However, the identification of the minimum number of measurements in array diagnosis, and in general the development of precise undersampling theorems in electromagnetic theory, remains an open problem that requires further studies.

Table 1: Recovery performance for various *SNR*s; *K=6* amplitude failures and $\delta_A = 1$ amplitude failure configuration

SNR (dB)	100	80	60	40	20	10
Δ_{FA} of L_1L_2	0.4251	0.3997	0.3196	0.1982	NAN	NAN
Δ_{FA} of TV	0.9846	0.8636	0.6738	0.6321	0.3121	NAN
Δ_{FA} of L_1	0.9997	0.9892	0.9771	0.8562	0.4321	0.021

Table 2: Recovery performance for various *SNR*s; *K=6* amplitude failures and $\delta_A = 0.1$ amplitude failure configuration

SNR (dB)	100	80	60	40	20	10
Δ_{FA} of L_1L_2	0.4531	0.2310	0.0000	NAN	NAN	NAN
Δ_{FA} of *TV*	0.5312	0.3921	0.2101	0.1031	0.0010	NAN
Δ_{FA} of L_1	0.8251	0.7834	0.7321	0.6101	0.3221	0.0310

Table 3: Recovery performance for various *SNR*s; *K=6* phase failures and $\delta_\Phi = 70^0$ phase failure configuration

SNR (dB)	100	80	60	40	20	10
Δ_{FA} of L_1L_2	0.7984	0.7231	0.6321	0.1321	NAN	NAN
Δ_{FA} of TV	0.5312	0.3921	0.4210	0.4521	0.0210	NAN
Δ_{FA} of L_1	0.9876	0.9231	0.8532	0.6324	0.3871	0.1310

Table 4: Recovery performance for various *SNR*s; *K=6* phase failures and $\delta_\Phi = 10^0$ phase failure configuration

SNR (dB)	100	80	60	40	20	10
Δ_{FA} of L_1L_2	0.4421	0.3410	0.3750	NAN	NAN	NAN
Δ_{FA} of *TV*	0.6120	0.5881	0.5101	0.0000	NAN	NAN
Δ_{FA} of L_1	0.8612	0.7994	0.7821	0.2101	NAN	NAN

VI. CONCLUSION

In this paper, a recovery performance analysis test of compressive sensing methods for antenna array diagnosis from near-field measured data is presented. Particularly, we considered three prominent regularization procedures: total variation (TV), mixed l_1/l_2 norm, and minimization of the l_1 to solve diagnosis problems in antenna array at low SNR. Simulation of a 10 GHz 10×10 rectangular waveguide array under realistic noise conditions was presented and used to conduct the test for various SNRs. Generally, the diagnosis is weak at low SNRs for both amplitude and phase errors, so for data with low SNRs. Minimization l_1 technique show more robustness to additive data noise than total variation, and mixed l_1/l_2 norm. Therefore, minimization l_1 is a better choice whenever antenna array diagnosis is to be performed in a very harsh measurement environment.

ACKNOWLEDGMENT

This work is supported in part by Global Excellence Stature (GES) of the University of Johannesburg, South Africa.

REFERENCES

[1] J. A. Rodriguez, F. Ares, E. Moreno, and G. Franceschetti, "Genetic algorithm procedure for linear array failure correction," *Electron. Lett.*, vol. 36, no. 3, pp. 196-198, Feb. 2000.

[2] R. Iglesias, F. Ares, M. Fernandez-Delgado, J. A. Rodriguez, J. Bregains, and S. Barro, "Element failure detection in linear antenna arrays using case-based reasoning," *IEEE Antennas Propag. Mag*., vol. 50, no. 4, pp. 198-204, Aug. 2008.

[3] O. M. Bucci, M. D. Migliore, G. Panariello, and P. Sgambato, "Accurate diagnosis of conformal arrays from near-field data using the matrix method," *IEEE Trans. Antennas Propag*., vol. 53, no. 3, pp. 1114-1120, Mar. 2005.

[4] J. A. Rodriguez-Gonzalez, F. Ares-Pena, M. Fernandez-Delgado, R. Iglesias, and S. Barro, "Rapid method for finding faulty elements in antenna arrays using far field pattern samples," *IEEE Trans. Antennas Propag*., vol. 57, no. 6, pp. 1679-1683, June 2009.

[5] A. Buonanno, and M. D'Urso, "On the diagnosis of arbitrary geometry fully active arrays," *in Proc. Eur. Conf. Antennas Propag.* (EuCAP), Barcelona, Spain, pp. 1-4.S, Apr. 2010.

[6] S. Clauzier, S. M. Mikki, and Y. M. M. Antar, "Design of near-field synthesis arrays through global optimization," *IEEE Trans. Antennas Propag*., vol. 63, no. 1, pp. 151-165, Jan. 2015.

[7] M. D. Migliore, "A compressed sensing approach for array diagnosis from a small set of near-field measurements," *IEEE Trans. Antennas Propag.*, vol. 59, no. 6, pp. 2127-2133, June 2011.

[8] G. Oliveri, P. Rocca, and A. Massa, "Reliable diagnosis of large linear arrays—A Bayesian compressive sensing approach," *IEEE Trans. Antennas Propag.*, vol. 60, no. 10, pp. 4627-4636, Oct. 2012.

[9] O. Famoriji, Z. Zhang, A. Fadamiro, R. Zakariyya, and F. Lin, "Planar array diagnostic tool for millimeter-wave wireless communication systems," *Electronics*, vol. 7, no. 383, 2018.

[10] O. J. Famoriji, Z. Zhang, A. Fadamiro, Z. Khan, and F. Lin, "Active antenna array diagnosis from far-field measurements," *IEEE Int. Conference on Integrated Circuits, Technologies and Applications* (ICTA), Beijing, China, pp. 1-2.S, Dec. 2018.

[11] B. Fuchs, L. L. Coq, and M. D. Migliore, "Fast antenna array diagnosis from a small number of far-field measurements," *IEEE Trans. Antennas Propag.*, vol. 64, no. 6, pp. 2227-2235, June 2016.

[12] A. F. Morabito, R. Palmeri, and T. Isernia, "A compressive-sensing-inspired procedure for array antenna diagnostic by a small number of phaseless measurements," *IEEE Trans. Antennas Propag.*, vol. 64, no. 7, pp. 3260-3265, July 2016.

[13] 5G: A technological Vision, HUAWEI white paper, Huawei Technologies Co., Shenzhen [Online] Available: http://www.huawei. com/5gwhitepaper/

[14] J. J. Lee, E. M. Ferrer, D. P. Woollen, and K. M. Lee, "Near-field probe used as a diagnostic tool to locate defective elements in an array antenna," *IEEE Trans. Antennas Propag*., vol. 36, no. 3, pp. 884-889, June 1988.

[15] L. Gattoufi, D. Picard, Y. R. Samii, and J. C. Bolomey, "Matrix method for near-field diagnostic techniques of phased array antennas," *in Proc. IEEE Int. Symp. Phased Array Syst. Technol*., pp. 52-57, 1996.

[16] C. Xiong, G. Xiao, Y. Hou, and M. Hameed, "A compressive sensing-based element failure diagnosis method for phased array antenna during beam steering," *IEEE Trans. Antennas Wireless Propag. Lett.*, vol. 18, pp. 1756-1760, Sept. 2019.

[17] M. D. Migliore, "A compressed sensing approach for array diagnosis from a small set of near-field measurements," *IEEE Trans. Antennas Propag.*, vol. 59, no. 6, pp. 2127-2133, June 2011.

[18] G. Oliveri, P. Rocca, and A. Massa, "Reliable diagnosis of large linear arrays, a Bayesian compressive sensing approach," *IEEE Trans. Antennas Propag.*, vol. 60 no. 10, pp. 4627-4636, Oct. 2012.

[19] M. D. Migliore, "Array diagnosis from far-field data using the theory of random partial Fourier matrices," *IEEE Trans. Antennas Wireless Propag. Lett.*, vol. 12, pp. 745-748, July 2013.

[20] B. Fuchs and M. D. Migliore, "Accurate array diagnosis from near-field measurements using reweighted minimization," *IEEE Antennas Propag. Symp*., Orlando, FL, USA, pp. 2255-2256, 2013.

[21] B. Fuchs, L. Le Coq, L. Ferro-Famil, and M. D. Migliore, "Comparison of methods for reflectarray diagnostic from far field measurements," *in Proc. IEEE Int. Symp. Antennas Propag*., pp. 398-399, July 2015.

[22] D. L. Donoho, "Compressive sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289-1306, Apr. 2006.

[23] E. J. Candes, J. Romberg, and T. Tao "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489-509, Feb. 2006.

[24] D. L. Donoho and J. Tanner, "Precise undersampling theorems," *Proc. IEEE,* vol. 98, no. 6, pp. 913-923, June 2010.

[25] M. D. Migliore, B. Fuchs, L. Le Coq, and L. Ferro-Famil, "Compressed sensing approach for reflectarray diagnostic from far field measurements," *in Proc. Eur. Microw. Conf.*, Sep. pp. 289-292, 2015.

[26] A. Massa, P. Rocca, and G. Oliveri, "Compressive sensing in electromagnetics—A review," *IEEE Antennas Propag. Mag.*, vol. 57, no. 1, pp. 224-238, Feb. 2015.

Performance Investigations of a Quad-band Microstrip Antenna for Body Wearable Wireless Devices

Varshini Karthik [1] and T. Rama Rao [2]

[1] Dept. of Biomedical Engineering
[2] Dept. of Telecommunication Engineering
SRM University, Chennai, Tamil Nadu - 603203, India
varshinikps@gmail.com, ramaraotr@gmail.com

Abstract ─ This research work proposes a microstrip-based quad-band monopole antenna for body Wearable Wireless Devices (WWD) for Wireless Body Area Network (WBAN) applications. Ultra high frequency (UHF) and Ultra wide band (UWB) technology have been known for their efficiency to meet power, size and distance considerations for WBANs. The designed antenna resonates at four frequencies, 1.8, 2.4, 5.0 and 8.9 GHz covering licensed and license-free wireless technologies. The antenna design considers the electromagnetic (EM) effects due to the interaction of body tissues with the radio frequency (RF) waves, which are very different when compared to their interaction with free space. The performance of the antenna is investigated in terms of radiation efficiency, total gain, specific absorption rate (SAR), and thermal effects (short and long term). A simplified, human body tissue layer model is used for simulations utilizing EM computations. Simulated and experimental results are paralleled and are found to be in good agreement.

Index Terms ─ Microstrip antenna, quad-band, specific absorption rate, thermal effects, ultra wideband, WBAN, wearable wireless devices.

I. INTRODUCTION

The goal of wearable technology is to create expedient and portable access to information in real time. With the use of body Wearable Wireless Devices (WWD), Wireless Body Area Networks (WBANs) find extensive applications in health care, rescue and emergency missions, augmented reality, ambient assisted living setups, smart entertainment and much more due to its multi-frequency capability. It has been anticipated that wearable technology would be a vital supporter of consumer electronics industry. Wearable devices interconnect with the help of wearable antennas. The field of wearable antennas is a multidisciplinary one that combines electromagnetics, material science, and bioelectronics. The parameters that impact the design of in-body and on-body antennas have been enumerated in [1]. A review paper [2] has revealed research on wearable, body mounted antennas for various applications at different frequencies over a decade. A wide assortment of antennas ranging from monopole, dipole, PIFA, microstrip, U slot, coplanar, circular, triangular patch antenna, EBG textile antennas, antennas for dual bands and those mounted on a button, belt and helmet have been studied.

In recent years, antennas for health care systems, effects of planar fabric substrate antenna, dual polarized printed antenna, tripod kettle antenna, and several antennas when placed near to human body has been studied by [3-7]. Diverse antenna types working at GHz frequencies has been discussed [8-11]. Antennas that are wearable and which work at multiband are in trend. A fractal shape multiband antenna that can operate from 1-20 GHz has been proposed [12]. A WiMax and C band antenna of compact size has been designed [13] and compared with other dual and triple band antennas. A triple band fractal Koch dipole denim antenna has been studied for its performance and its suitability has been tested for wearable applications [14].

The human body is inhomogeneous with high loss and permittivity that are frequency dependent and which affect the properties of an antenna. The interactions of electromagnetic (EM) waves with the body using single and multi-layered tissue models have been discussed [15]. A good antenna design is essential for limited SAR values as the SAR depends on directivity, orientation, and gain of the antenna [16]. The effect of human body on RF transmission of biotelemetry and portable radio antennas at 10-1000 MHz has been discussed [17]. Recent research in wearable textile antennas that work under bending conditions on body has been carried out by [18]. Researchers [3] have investigated wearable antenna performance using human modelling for gain and diversity. In [19] authors have designed wearable fractal antenna and have worked on SAR as a useful parameter for wearable application. Safe use of the antenna on the body is more related to temperature

Submitted On: May 16, 2016
Accepted On: June 28, 2017
https://doi.org/10.47037/2021.ACES.J.360806

changes than power density [20]. Emphasis has been made on thermal and biological aspects when it is RF waves interactions with the body [21]. In this context, elevation of temperature measurements for the compliance evaluation of mm-wave wireless devices operating close to the body has been suggested [22].

In this work, a quad-band microstrip based monopole antenna that works efficiently at four bands within UHF (30-3000 MHz) and UWB (3.1-10.6 GHz) frequencies on the human body has been proposed for WBANs utilizing WWDs. The antenna is offered to communicate from WWD to a personal digital assistant device / smart digital device. Due to the growth of wireless technology, several wireless systems operate simultaneously in a given scenario. The four bands chosen enable the antenna to be used along with the existing wireless technologies licensed and license-free bands. The UWB sub-bands selected allow the antenna to be used for low and high data rate applications, for varying power levels of reception and realistic environments when the full UWB cannot be utilized [23-24]. The antenna designed is proposed to work in proximity to the human body. The effects of the body on the antenna performance and vice versa have to be taken into account. The performance of the designed antenna for varying skin thickness and its heating effects are studied and analyzed utilizing Ansys HFSS [25] and Sim4Life by SPEAG [26]. Sim4Life is used for simulations on-body for the designed antenna at various body sites utilizing its high-fidelity computable human phantoms with advanced tissue models, and HFSS (based on finite element method) is used for antenna design and fabrication.

II. MODELLING HUMAN BODY TISSUES

A. Effect of electromagnetic field on human tissues

The human body consists of different types of material, each differing in the way they interact with the EM fields. The geometry of the human body and its different parts has been described to be very complex [27]. The electrical properties of the biological tissues are the result of the electromagnetic radiations interacting with constituents of the body at the cellular and molecular level [28]. Understanding the interaction and the electrical properties of body tissues is the primary need in the design work of antennas for body wearable devices. The biological effects of antennas working in the UWB frequency ranges are proportional to the rate of energy absorption given in terms of SAR and the ability to heat human tissues. Both these effects can be hazardous if exposure is sufficiently intense or prolonged. Since the magnetic permeability of body tissues is same as that of free space, the electrical permittivity and conductivity are the important properties that determine the electric field distribution in the body and the power dissipated in it [29]. These properties change with frequency. Tissues with the highest water content have the highest relative permittivity (e.g., skin and muscle), decreasing with increasing frequency. The tissue's water content results in the specific permittivity value, which affects the wavelength inside tissues [30].

B. Layered model of human tissue

The antenna has been designed for wireless devices, where the influence of the outermost body tissues has to be considered. There are different simplified tissue models like homogeneous models, three layered body model with different cross sections as flat, rectangular and elliptical used in literature. There are also detailed 3D complex voxel body models that use a lot of computational resource and time. For computational simplicity and considerable accuracy, a four layered rectangular biological tissue model made up of skin, subcutaneous fat, muscle and bone has been chosen as it represents most of the body regions [31]. The designed antenna is proposed to be placed over the stacked tissues with an average thickness of 2 mm, 3.5 mm, 10 mm and 10 mm for skin, fat, muscle and bone layer respectively, as in Fig. 1.

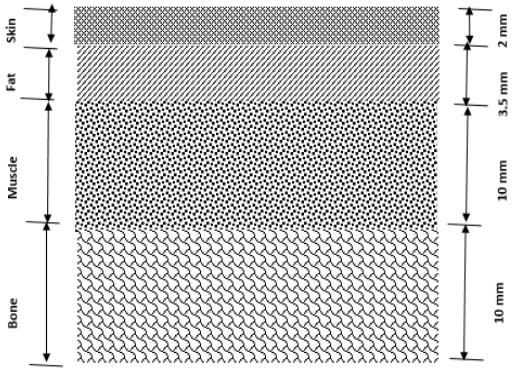


Fig. 1. Layered rectangular phantom model of human tissues.

The electrical properties of the tissues at the various resonating frequencies of the designed antenna have been based on the works of Italian National Research Council, which is available online [32] and are tabulated in Table 1. The calculation of these electrical properties have been based on the data available [28], where the variation of the electrical properties of tissues over a frequency range of 10 KHz-10 GHz around a body temperature of 20˚C to 40˚C has been considered. Skin, due to its inhomogeneous structure has inhomogeneous dielectric properties. Muscles, bones have intermediate dielectric properties when compared to fat which has poor dielectric properties.

Table 1: Electrical properties at the desired frequencies

Frequency (GHz)	Tissue	ε_r	σ (S/m)	Tan θ (Loss Tangent)
1.8	Skin	38.8	1.18	0.30
	Fat	5.30	0.07	0.14
	Muscle	53.5	1.34	0.25
	Bone	19.3	0.58	0.30
2.4	Skin	38.0	1.44	0.28
	Fat	5.28	0.10	0.14
	Muscle	52.7	1.70	0.24
	Bone	18.6	0.78	0.31
5.0	Skin	35.7	3.06	0.30
	Fat	5.02	0.24	0.17
	Muscle	49.5	4.04	0.29
	Bone	16.0	1.81	0.40
8.9	Skin	32.3	6.78	0.42
	Fat	4.68	0.50	0.21
	Muscle	44.2	9.05	0.41
	Bone	13.2	3.42	0.52

III. ANTENNA DESIGN

It is critical that the antenna has a compact design for use on the human body. The antenna design requirements differ from that of a conventional antenna that is designed for free space because of the tissue environment. Microstrip model design equation [33] has been used to design the antenna. The antenna has a dimension of 35 x 32 x1.57 mm^3 and is made of Rogers RT/ Duroid 5880 substrate with relative permittivity of 2.2 and loss tangent of $9x10^{-3}$. The selection of the substrate is due to its low electric loss along with low moisture absorption. An optimized design of the antenna with a partial ground plane and staircase rectangular shaped patch has been used. The step size of 1mm followed by 1.5mm from both the top and bottom edges is used for the staircase pattern. Horizontal slots each measuring 6 mm in length and 0.5 mm in width are used on the right and left sides of the radiating patch. A 50Ω microstrip line was used to feed the radiating patch. The power input at the port was 1 mW. Good impedance matching has been achieved by a microstrip tapered line. The prototype of the antenna along with the dimensions is shown in Fig. 2.

Strips of conducting material at top and bottom have been used to provide a defected ground plane. The design has been optimized to make the antenna work at four bands with S_{11} less than -10 dB at all the bands. The return loss plot for the various slot size of antenna design and optimization is shown in Fig. 3. It has been learnt that the defected ground plane helps to achieve the multiband operation of the antenna. The slots on the patch tend to align the shifted bands to the required resonant frequencies and the steps offer a greater value of return loss at the four bands. A quad band operation has been chosen to offer better EM compatibility with existing wireless communication technologies and unlicensed UWB and to be used for both low and high data rate applications.

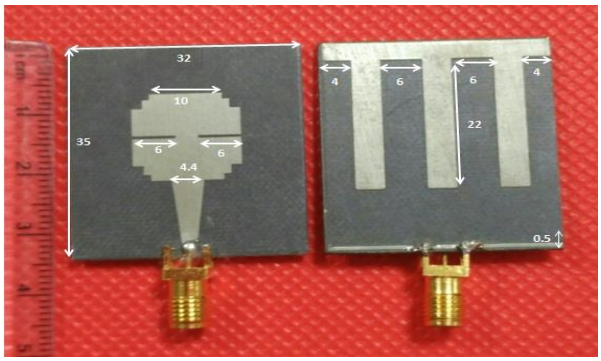


Fig. 2. Photograph of the prototype of staircase microstrip antenna with the dimensions in mm for: (a) radiating patch and (b) ground plane respectively.

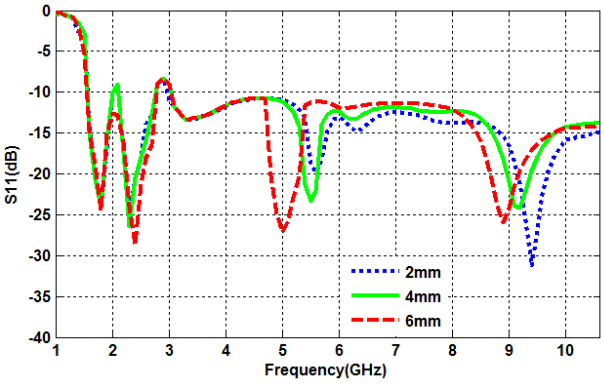


Fig. 3. S_{11} plot for various simulated length of the horizontal slot on the radiating patch of the antenna.

IV. SIMULATION AND MEASURED RESULTS

A. Antenna simulations and measurements in free space

The antenna has been simulated and measured for its performance in free space before placing it on the body to understand the influence of human body proximity on the working of the microstrip antenna. Performance metrics of the antenna considered are total gain, radiation efficiency, and SAR. Measurements of return loss were made using RS-ZVL Vector Network Analyzer of frequency range 9 KHz-13.6 GHz.

B. Antenna simulations on rectangular phantom model and measurements on human body

Placement of the antenna in direct contact with the body has resulted in decrease in gain and radiation efficiency with due effects on the impedance matching due to the dielectric loading provided by the lossy human tissues. A frequency shift in the resonating frequency of 100 MHz from 1.8 to 1.9 GHz, 80 MHz from 2.4 to 2.48 GHz, 140 MHz from 5.0-5.14 GHz, 100 MHz from 8.9 to 8.8 GHz was observed between simulated and measured results in the four bands respectively. Measurements of return loss when the antenna is placed on the body were made using RS-ZVL Vector Network Analyzer of frequency range 9 KHz-13.6 GHz. The designed antenna was kept on the human body with cotton padding of 5 mm between them. The performance of the antenna on the body has been enhanced due to the bio and electromagnetic compatibility nature of the chosen fabric which makes it realizable for real-time use when compared to the few mm of air gap used in previous designs. The measured return loss plot for free space and on body operation of the antenna is shown in Fig. 4.

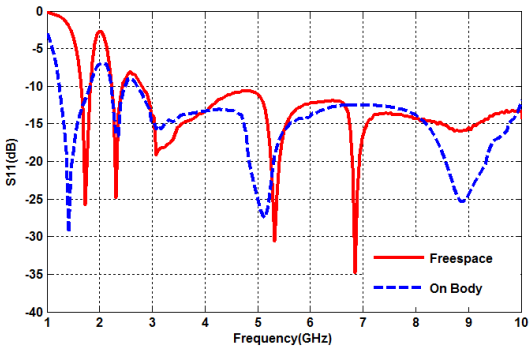


Fig. 4. Measured return loss plot for free space and on body placement of the antenna.

The performance of the quad bands has been studied and tabulated in Table 2. It is seen from the table that the upper bands perform well when compared to lower bands. This is important while considering the EM waves (thus also antennas) interaction with the body [28]. Major reflections occur due to the wave impedance mismatch between low (fat) and high water content (skin, muscle) tissues. The performance of the antenna at the four bands is dependent on the electrical properties of the tissues at those frequencies. It has been observed that the electrical permittivity decreases with increasing frequency and the electrical conductivity and loss tangent increase with increasing frequency for the tissues. Negative gain values at lower bands were due to the near omnidirectional radiation pattern of the antenna.

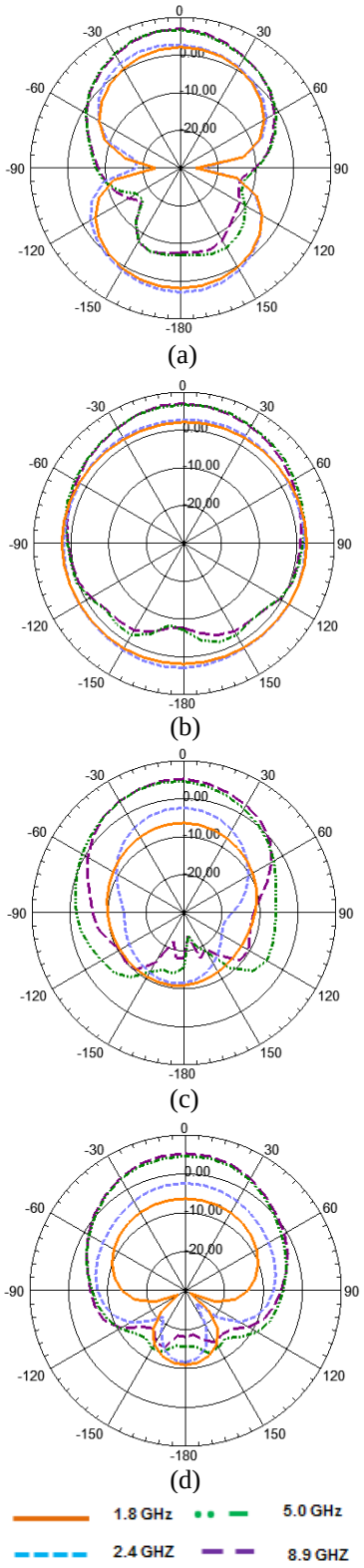


Fig. 5. Radiation pattern (simulated gain in dB): (a) free space E plane, (b) free space H plane, (c) on body E plane, and (d) on body H plane.

As frequency increases, there is less penetration into the body and hence an increase in radiation efficiency of the antenna. It is observed that there is better performance in the third band. This is due to the combination of the change in electrical properties of the tissues at this band when compared to the other bands.

The quasi-omnidirectional radiation pattern of the antenna that has been obtained considering the thickness of the layers of skin, subcutaneous fat, muscle and bone as 2, 3.5,10, 10 mm respectively at the four bands are shown in Fig. 5. The patterns have been compared with free space radiation patterns of the antenna. The radiations are seen away from the body ensuring less absorption of radiations by the body. The simulated radiation parameters for on-body placement of the antenna are tabulated in Table 3. It has been observed that there is increase in main lobe level, peak directivity at higher frequencies due to lesser penetration of RF waves into body and hence more radiations away from the body. The values of half power beam width (HPBW) and side lobe levels are also given. This makes it suitable for on-off body communication as from a wearable device to a mobile phone. In case of free space operation, the absence of body tissue influence has ensured near omni- directional radiation pattern as seen in Fig. 5.

Table 2: Performance of quad band in free space and on body

Frequencies	**Total Gain (dB)**		**Radiation Efficiency (%)**	
	Free Space	On Body	Free Space	On Body
1.8 GHz	2.7	-6.7	98	9.0
2.4 GHz	3.2	-2.3	98	14.4
5.0 GHz	7.0	5.2	99	55.9
8.9 GHz	6.8	5.0	99	48.1

Table 3: Radiation properties (simulated) for on-body placement of the antenna

Radiation Parameter	**1.8 GHz**	**2.4 GHz**	**5.0 GHz**	**8.9 GHz**
Main lobe magnitude (dB)	-6.4	-2.3	5.2	5.0
Peak directivity (dB)	2.5	4.0	5.9	6.1
Side lobe level (dB)	-4.5	-9.3	-15.3	-15.0
HPBW (degree)	130	117	79.5	81.5

C. Robustness study of the antenna on the body

The antenna has been analyzed for its robust performance on the human body by considering the scenario of varying skin thickness and placement of the antenna at random sites on the body. The thickness of skin increases with age. In that view, the thickness of the skin layer used in the model has been varied from 1-2 mm, depicting a scenario of young to aged skin. The return loss plot for the same is seen in Fig. 6. It is observed that the quad band operation is maintained for all skin thickness. With a decrease in skin thickness, for the second band, there is a right shift towards higher resonating frequencies.

The proposed antenna has been simulated at ten different anatomical sites on the body. This was to analyze the performance of the antenna on any site on the body where the thickness of the layers of tissue varies with the placement site. The body sites that were considered are forehead, neck, biceps, triceps, chest, wrist, anterior abdomen, anterior thigh, posterior thigh, calf of legs as shown in Fig. 7 and named in Table 4. They are the most familiar and comfortable sites on body for monitoring vital bio-signals. The anatomical sites have been chosen considering a scenario of sports training or general health monitoring where the sensors and transceivers can be placed for the specific application. They are best sites for picking up physiological parameters like body temperature, pulse rate, blood pressure, respiration rate, ECG, heart rate, EMG, glucose level, motion and perspiration. The thickness of skin and subcutaneous fat layer for these body sites measured using ultrasonography have been referred from literature [34-40].

The performance of the antenna for various tissue layer combinations considering the muscle and bone thickness as 10 mm is given in Table 4. As frequency increase, they have shorter wavelengths and hence smaller penetration into the tissues. This results in lesser absorption of EM waves, and low values of SAR as frequency increases.

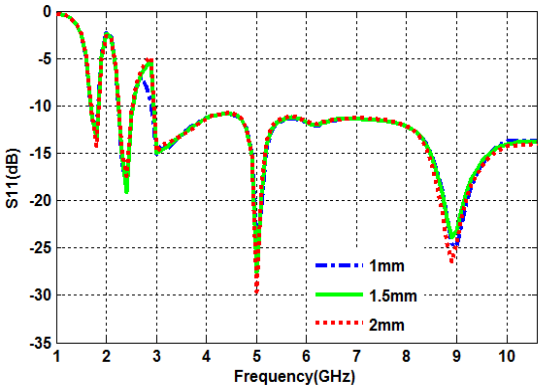


Fig. 6. Comparison of simulated S_{11} plot for varying skin thickness.

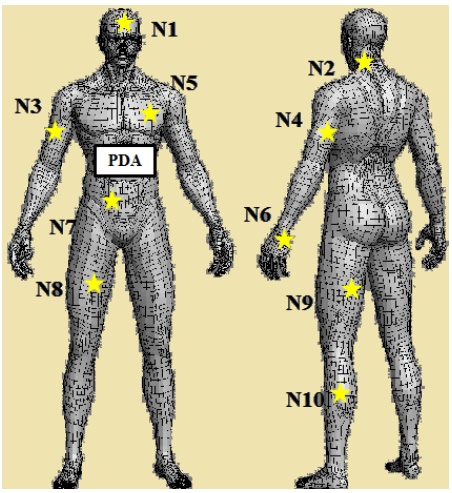


Fig. 7. Anatomical sites of the body (anterior and posterior) considered for simulations along with PDA that serves as a gateway for WBAN applications.

D. SAR and on body thermal effects

Specific Absorption Rate (SAR) is a good dosimetric quantity that measures the rate of energy absorption by human body when exposed to radio frequency EM field. SAR is calculated as:

$$SAR = \frac{\sigma.|E|^2}{\rho_m}, \quad (1)$$

where, E is the RMS value of induced field in (V/m), σ is the conductivity of tissue in (S/m), ρ_m is the mass density of tissue in kg/m^3. A high level of SAR, in any tissue, above standard limits, can prove dangerous for human use. The effect has been measured as an increase in tissue temperature related to the time of exposure of tissues to the EM field. It has been observed from literature that an increase of tissue temperature equal to or greater than 1K would ascertain danger.

The temperature rise of tissues have a linear relationship with the time of exposure for short-term exposure of a few seconds to minutes. This is because of the little significant contribution made by conductive or convective heat distribution to the temperature rise. A time duration of seconds to minutes has been taken as short term and minutes to hours has been taken as long term. The temperature rise in tissues (K) for short time exposure has been obtained from the following equation [41] by knowing the heat capacity of the tissue (J/kg/K), average SAR value for 1 gram of tissue and the time of exposure (Δt) in seconds:

$$\Delta T = SAR.\frac{\Delta t}{c}. \quad (2)$$

The short time temperature rise for a period of 60 seconds of radiation from the designed antenna on the human body has been calculated for the outer three tissues. Adiabatic conditions were assumed for heat exchange with the environment. Adiabatic conditions give the worst case temperature increase [42] in a state where no heat is exchanged with the environment.

To obtain the relation between temperature rise and time, in the case of long time exposures, the Pennes Bio Heat equation considering the role of blood in thermoregulation of the body has been used. Simplification of the Bio heat differential equation has yielded equation (3) [43] that gives the maximum temperature rise over an extended period of time in the order of minutes. In the equation, S is average SAR for 1 gram of tissue in W/kg, ρ is the mass density of tissue in kg/m^3, K is thermal conductivity of the tissue in W/m/K, and w is blood perfusion rate in ml/g/min, c_b is the heat capacity of blood in J/kg/K, λ is the wavelength of the EM wave in m. The parameter values for the considered body tissues and the average SAR values for the tissues at the four bands considered are given in Table 5.

Table 4: Performance of the quad bands at different body sites using simulations

Performance Metrics	**Total Gain (dB)**				**Efficiency (%)**				**SAR (W/Kg)**			
Site (Node)	**I**	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**
Forehead (N1)	-4	0.5	4.6	3.7	12.5	25.5	48.0	47.8	0.23	0.20	0.17	0.15
Neck (N2)	-4.3	-0.7	4.7	3.5	13.3	21.9	47.7	46.8	0.22	0.20	0.16	0.13
Forearm (N3)	-6	-2.1	5.2	3.8	9.3	15.3	53.4	50.7	0.21	0.20	0.14	0.13
Triceps (N4)	-6.7	-2.7	5.3	3.9	9.1	14.2	55.6	50.6	0.20	0.17	0.14	0.12
Chest (N5)	-6.9	-1.7	4.8	3.7	8.5	15.3	53.9	50.1	0.17	0.14	0.13	0.10
Wrist (N6)	-5.3	-1.4	4.2	2.5	10.5	16.4	16.4	39.6	0.25	0.19	0.15	0.14
Abdomen (N7)	-9	-6.6	5.1	3.5	6.1	7.4	58.1	46.4	0.21	0.17	0.16	0.09
Ant. Thigh (N8)	-6.8	-3.2	5.2	3.9	9.1	13.3	55.9	50.0	0.20	0.17	0.13	0.10
Posterior thigh (N9)	-7.9	-4.7	5.3	3.8	7.6	10.4	57.3	49.3	0.18	0.16	0.10	0.09
Calf (N10)	-7.3	-3.0	5.7	3.9	8.7	13.5	57.8	50.6	0.14	0.17	0.11	0.11

Table 5: Parameter values of body tissues used for temperature calculations

Tissue/ Parameters	Mass Density (kg/m^3)	Thermal Conductivity (W/m/K)	Heat Capacity (J/kg/K)	Average SAR (W/kg)			
				1.8 GHz	2.4 GHz	5.0 GHz	8.9 GHz
Skin	1109	0.37	3391	0.22	0.17	0.15	0.11
Fat	911	0.21	2348	0.03	0.02	0.01	0.01
Muscle	1090	0.49	3421	0.09	0.05	0.02	0.02
Bone	1908	0.32	2065	0.0003	0.0002	0.0001	0.0001

The maximum SAR values among them has been used to calculate the maximum temperature rise for the particular tissue and the temperature rise are tabulated in Table 6 using (2) and (3). It is observed that the highest temperature increase has been found in the skin, with 0.004 K for short term and 0.134 K for long term exposure. The temperature increase has been found to be indeed less of about fraction of a kelvin. The value of blood parameters considered are blood perfusion rate (w) and heat capacity of blood (c_b) whose values are taken as 0.5 ml/g/min and 3617 J/kg/K respectively:

$$Tmax = \frac{q}{\lambda'}[1 - \left(\sqrt{\lambda'}\,A + 1\right)e^{-\sqrt{\lambda' A}}\ ], \qquad (3)$$

where, $q = \frac{\rho S}{K}$; $\lambda' = \frac{w\,c_b}{K}$; $A = \frac{\lambda}{4}$.

Table 6: Maximum SAR values at different tissues and their heating effects

Tissue	Average SAR (Max.) (W/Kg)	Short Term Temp Increase (K)	Long Term Temp Increase (K)
Skin	0.22	**0.004**	**0.134**
Fat	0.03	0.001	0.016
Muscle	0.09	0.002	0.056
Bone	0.0003	0.0001	0.0003

V. DISCUSSIONS AND CONCLUSIONS

The performance of the quad-band based microstrip monopole antenna for WBAN applications has been studied in free space and on the body for the use of WWDs. The effects of the human body on the performance of the proposed antenna and the effects of the performing antenna on the human body in terms of thermal effects were analyzed. The antenna's S_{11} satisfies less than -10 dB in free space and on body simulations at all the four bands. Less value of gain at lower bands is due to the omnidirectional radiation pattern at these bands. Investigations have been made on the performance of the antenna at various human body sites and for varying skin thickness. It was observed from Table 4 that the location of placement of the antenna on the body has an effect on the performance of the antenna. Sites with lesser fat thickness (wrist, forehead, neck) had less reflections, greater absorption by the body, hence higher SAR values. Higher bands performed better due to a change in permittivity leading to less penetration into the tissues as frequency increases. Radiation absorption is less and efficiency of the antenna is more at higher frequencies. However, the performance of the antenna at all the considered body sites was reasonably good regarding gain and radiation efficiency.

It is also analyzed from Table 6 that the short term and long term exposure of the body to the EM field due to the irradiating antenna was found to produce a temperature increase of a fraction of a kelvin which is indeed a small value. The thermal effect was found to be maximum on the skin which is the outermost tissue with a value of 0.134 K temperature rise under long term exposure. The maximum SAR value was observed to be 0.22 W/kg over 1 g of tissue. SAR values were found to be far less than the maximum allowable SAR values limits 2 W/kg averaged over 10 g of tissue exposed to EM radiations given by International Council on Non-Ionizing radiation protection [44] and 1g averaged SAR value of 1.6 W/kg provided by IEEE/ICES C95.1-2005 [45]. These observations indicate that the proposed antenna module can be used as a candidate for WWDs for WBAN applications which will proliferate very widely in near future as a promising way of monitoring signals in all walks of life where the antennas performance is greatly affected by the dielectric properties of body tissues.

REFERENCES

[1] S. Movassaghi, M. Abolhasan, J. Lipman, D. Smith, and A. Jamalipour, "Wireless body area networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 16.3, pp. 1658-1686, 2014.

[2] B. Gupta, S. Shankaralingam, and S. Dhar, "Development of wearable and implantable antennas in the last decade: A review," *MMS'10*, Guzelyurt, 2010.

[3] E. S. Florence, M. Kanagasabai, and G. N. M Alsath, "An investigation of a wearable antenna using human body modelling," *Applied Computational Electromagnetics Society Journal*, vol. 29, no. 10, 2014.

[4] C. H. Lin and K. Ito, "A compact dual-mode

wearable antenna for body-centric wireless communications," *Electronics*, vol. 3.3, pp. 398-408, 2014.

[5] D. D. Cara, J. Trajkovikj, R. Torres-Sánchez, J. F. Zürcher, and A. K. Skrivervik, "A low profile UWB antenna for wearable applications: The tripod kettle antenna (TKA)," In *Antennas and Propagation (EUCAP), IEEE 2013 7th European Conference on*, pp. 3257-3260, 2013.

[6] A. Sabban, "Comprehensive study of printed antennas on human body for medical applications," *International Journal of Advance in Medical Science*, 2013.

[7] J. Anguera, A. Andújar, C. Picher, L. González, C. Puente, and S. Kahng, "Behavior of several antenna topologies near the human head at the 2.4–2.5 GHz band," *Microwave and Optical Technology Letters*, vol. 54, no. 8, pp. 1911-1916, 2012.

[8] T. Tuovinen, K. Y. Yazdandoost, and J. Iinatti, "Monopole ultra-wideband antenna for on-body communication in wireless body area network," *In Antennas and Propagation Conference (LAPC)*, Loughborough, pp. 1-4, 2011.

[9] A. A. Kishk and K. F. Lee, "Ultrawide-band coplanar waveguide-fed rectangular slot antenna," *IEEE Antennas and Wireless Propagation Letters*, vol. 1.3, pp. 227-229, 2004.

[10] Z. N. Chen, T. See, and X. Qing, "Small printed ultrawideband antenna with reduced ground plane effect," *IEEE Transactions on Antennas and Propagation*, vol. 55.2, pp. 383-388, 2007.

[11] M. A. Osman, M. K. A. Rahim, K. Kamardin, F. Zubir, and N. A. Samsuri, "Design and analysis UWB wearable textile antenna," *In Antennas and Propagation (EUCAP), IEEE Proceedings of the 5th European Conference on*, pp. 530-533, 2011.

[12] S. A. Ali, U. Rafique, U. Ahmad, and M. A. Khan, "Multiband microstrip patch antenna for microwave applications," *IOSR Journal of Electronics & Communication Eng.*, vol. 3.5, pp. 43-48, 2012.

[13] M. R. Ahsan, M. T. Islam, M. Habib Ullah, H. Arshad, and M. F. Mansor, "Low-cost dielectric substrate for designing low profile multiband monopole microstrip antenna," *The Scientific World Journal*, 2014.

[14] M. E. B. Jalil, M. K. Abd Rahim, N. A. Samsurin, N. A. Murad, H. A. Majid, K. Kamardin, M. Azfar Abdullah, "Fractal Koch multiband textile antenna performance with bending, wet conditions and on the human body," *Progress in EM Research*, vol. 140, pp. 633-652, 2013.

[15] F. B. Saadi and M. R. Kamarudin, "EM wave effects upon the human body using UWB antennas," *Blood*, vol. 57.3, pp. 149209, 2014.

[16] K. M. S. Thotahewa, J. M. Redoute, and M. R. Yuce, "SAR, SA and temperature variations in the human head caused by IR-UWB implants operating at 4 GHz," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61.5, pp. 2161-2169, 2013.

[17] P. A. Neukomm, "Body-mounted antennas," *Doctoral dissertation*, Swiss Federal Institute of Technology, ETH, Zurich, 1979.

[18] S. Sankaralingam and B. Gupta, "Development of textile antennas for body wearable applications and investigations on their performance under bent conditions," *Progress in Electromagnetics Research B*, vol. 22, pp. 53-71, 2010.

[19] S. Velan, E. F. Sundarsingh, M. Kanagasabai, A. K. Sarma, C. Raviteja, R. Sivasamy, and J. K. Pakkathillam, "Dual-band EBG integrated monopole antenna deploying fractal geometry for wearable applications," *IEEE Antennas and Wireless Propagation Letters*, vol. 14, pp. 249-252, 2015.

[20] T. Wu, T. S. Rappaport, and C. M. Collins, "The human body and millimeter-wave wireless communication systems: Interactions and implications," *ICC'2015 IEEE International Conference on Communications*, pp. 2423-2429, 2015.

[21] Z. Maxim, N. Chahat, R. Sauleau, C. Le Quement, and Y. Le Drean, "Millimeter-wave interactions with the human body: State of knowledge and recent advances," *Intern. Journal of Microwave & Wireless Technologies*, vol.3.2, pp. 237-247, 2011.

[22] T. Wu, S. T. S. Rappaport, and C. M. Collin, "Safe for generations to come: Considerations of safety for millimeter waves in wireless communications," *IEEE Microwave Magazine*, vol. 16. 2, pp. 65-84, 2015.

[23] H. Yamamoto and T. Kobayashi, "Ultra-wideband propagation loss around a human body in various surrounding environments," Ultra-*Wideband, Short Pulse Electromagnetics*, vol. 9, pp. 11-17, 2010.

[24] A. Batra, J. Balakrishnan, G. R. Aiello, J. R. Foerster, and A. Dabak, "Design of a multiband OFDM system for realistic UWB channel environments," *IEEE Transactions on Microwave Theory and Techniques*, vol. 52.9, pp. 2123-2138, 2004.

[25] http://www.ansys.com/

[26] http://www.speag.com/

[27] N. Siauve, R. Scorretti, N. Burais, L. Nicolas, and A. Nicolas, "Electromagnetic fields and human body: a new challenge for the electromagnetic field computation," *The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, vol. 22.3, pp. 457-469, 2003.

[28] C. Gabrie, S. Gabriely, E. Corthout, "The dielectric properties of biological tissues: I. Literature survey," *Phys. Med. Biol.*, vol. 41, pp. 2231-2249, 1996.
[29] S. Hernandez and A. David, *High Frequency Electromagnetic Dosimetry*. Artech House, 2009.
[30] T. Tuovinen, M. Berg, K. Y. Yazdandoost, and J. Iinatti, "Ultra wideband loop antenna on contact with human body tissues," *IET Microwave and Antennas Propagation*, vol. 7.7, pp. 588-596, 2013.
[31] M. Klemm and G. Troester, "EM energy absorption in the human body tissues due to UWB antennas," *Progress In Electromagnetics Research*, vol. 62, pp. 261-280, 2006.
[32] Italian National Research Council, Institute for Applied Physics, homepage on Dielectric properties of body tissues. [Online]. Available: http://niremf.ifac.cnr.it
[33] R. Garg, P. Bhartia, I. Bahl, and A. Ittipiboon, *Microstrip Antenna Design Handbook*. Artech House, 2001.
[34] R. Y. Ha, K. Nojima, P. William, and A. S. Brown, "Analysis of facial skin thickness: Defining the relative thickness index," *Plastic and Reconstructive Surgery*, vol. 115.6, pp. 1769-1773, 2005.
[35] W. L. Weiss and C. F. Clark "Ultrasonic protocols for separately measuring subcutaneous fat and skeletal muscle thickness in the calf area," *Physical Therapy Journal*, vol. 65.4, pp. 477-481, 1985.
[36] S. P. Jerrold, J. D M. Prowse, and E. Lohman, "The influence of aging and diabetes on skin and subcutaneous fat thickness in different regions of the body," *The Journal of Applied Research*, vol. 8.1, pp. 55, 2008.
[37] F. J. Menick, "A 10-year experience in nasal reconstruction with the three-stage forehead flap," *Plastic and Reconstructive Surgery*, vol. 109.6, pp. 1839-1855, 2002.
[38] R. Moller, R. Horejsi, E. Tafeit, M. Fraidl, G. Dietmaier, U. Anegg, H. Klemen, and S. Wallner-Liebmann, "Effects of hyperbaric environment on SAT-Top," *Coll. Antropol*, vol. 34.4, pp. 1309-1313, 2010.
[39] O. Akkus, A. Oguz, M. Uzunlulu, and M. Kizilgul, "Evaluation of skin and subcutaneous adipose tissue thickness for optimal insulin injection," *Journal on Diabetes Metab*, 2012.
[40] A. Akesson, H. R. A. Schema, and M. Wildt, "Longitudinal development of skin involvement and reliability of high frequency ultrasound in systemic sclerosis," *Ann Rheum Diseases*, vol. 63.7, pp. 791-796, 2004.
[41] F. S. González, "Radiation effects of wearable antenna in human body tissues," *University of Colorado Springs*, 2014.
[42] T. Samaras, A. Christ, A. Klingenbock, and N. Kuster, "Worst case temperature rise in a one-dimensional tissue model exposed to radio frequency radiation," *IEEE Transactions on Biomedical Engineering*, vol. 54.3, pp. 492-496, 2007.
[43] H. N. Kritikos and H. P. Schwan, "Potential temperature rise induced by electromagnetic field in brain tissues," *IEEE Transactions on Biomedical Engineering*, vol. 1, pp. 29-34, 1979.
[44] ICNIRP Guidelines for limiting to time varying electric, magnetic, and electromagnetic fields (up to 300GHz), *Health Physics*, vol. 74.4, pp. 494-522, 1998.
[45] IEEE Standard for Safety Levels with Respect to Human Exposure to Radio Frequency Electromagnetic Fields, 3 kHz to 300 GHz, IEEE Standard C 95.1, 2005.

Varshini Karthik is presently working as an Assistant Professor, Department of Biomedical Engineering, SRM University, India. She received a Bachelor of Engineering degree in Electronics and Communication Engineering from the Bharathidasan University in 2002 and received her Master of Engineering degree in Medical Electronics from Anna University in 2009. Her main research interests are design and optimization of Body wearable wireless devices and their Bio electromagnetic effects.

T. Rama Rao is a Professor of Telecommunication Engineering at SRM University, India, and has a long-standing research experience on radiowave propagation studies for wireless communications. In the past, he worked at Aalborg University, Denmark as a Research Professor in an EU funded project on cellular networks capacity enhancement, at Universidad Carlos III de Madrid, Spain, and at the University of Sydney, Australia, as a Visiting Professor. He is a member of IEEE, WWRF, IET, ACM, and IETE Fellow.

Design of Compact Reconfigurable Antenna Array for WLAN Applications

Yazeed M. Qasaymeh

Department of Electrical Engineering
Majmmah University, Majmmah, 11952, Kingdom of Saudi Arabia
y.qasaymeh@mu.edu.sa

Abstract ─ In this communication, a compact design of a reconfigurable antenna array operating in the band IEEE 802.11a is presented. The proposed antenna array contains four radiating elements excited by a hybrid parallel-series-feed network. The hybrid feed technique is used to avoid the main beam squint due to frequency changes compared to series and parallel array feed topologies. Each of the four radiating elements consists of straight narrow strip inductor in parallel with an interdigital capacitor. The antenna resonant frequency is electronically controlled by placing PIN diodes switches at the resonant element's inputs. The antenna permits reconfigurable switching frequency bands between 5.25 and 5.82 GHz. The results of the return loss and pattern radiation are shown. The size of the whole antenna structure is about 64×18 mm^2 and can potentially be used in wireless systems.

Index Terms ─ Antenna array, hybrid feed, reconfigurable, miniaturized.

I. INTRODUCTION

Recent advancements in wireless technology components have resulted in their increased use in gadgets meant for numerous wireless applications and use cases. Even when is a constant push for miniaturization of gadgets, reducing the sizes of antennas and batteries have been historically difficult, owing to performance constraints [1].

For instance, the principal limitation of the microstrip patch antenna is its narrow impedance bandwidth which is usually a few percent. Various procedures have been proposed to increase this bandwidth [2-3]. The techniques include using a parasitic patch stacked, a substrate of a low dielectric constant, as well as employing matching structures [4-6]. However, these techniques can only achieve limited degrees of bandwidth advancement.

The frequency agile antennas were introduced to overcome the microstrip antenna narrow resonance [7]. These antennas can alter their resonant frequency to function at a specific frequency with the multi-serviced frequency bands [8]. As such, they can be built in lower sizes, and finds multipurpose uses for several wireless equipments. The feature of the antenna's ability to be built in lower sizes, is highly demanded for cognitive radio systems [9-10]. Numerous methods were reported to achieve frequency tunable antennas, and they can be fundamentally categorized into three main types namely mechanical tuning [11,12], using tunable materials [13,14] and electronics switches [15,16]. Mechanical tuning involves a physical displacement, which needs some time suspension. It is difficult to attain consistent performance with mechanical tuning. Using tunable material also have drawbacks such as high bias voltages, temperature sensitivity and dc power consumption. To address these shortcomings, electronic switches are preferred in designing frequency tunable antennas. Additionally, electronic elements provide the benefit of packing efficiency, size reduction, and the ready availability of miniaturized commercial components that easily combined with the antennas.

A few rounded electronically reconfigurable arrays were reported in the literature to operate in the band IEEE 802.11a. Byford et al., (2015) presents a 2×2 reconfigurable array consisting of interconnected radiating elements which is controlled by switches. If the switch is ON state, a larger radiating element with lower resonating frequencies than the individual elements is formed [17]. Li et al. (2014) introduced an array that consists of four double-sided bow-tie elements as well as an ultra-wideband feed network. The effective electrical lengths of the elements can be altered, once the state of the PIN diodes loaded on the radiators is controlled. Hence, the array resonant frequency can be electronically switched [18].

The purpose of this paper is threefold. The first is to propose a way for antennas to acquire frequency agile resonance, by using the PIN diodes as ON/OFF switches at each of the four array element inputs. Next, the paper proposes a way to avoid radiation pattern squint due to frequency change by using a hybrid feeding technique. The third is to present a miniaturized array real estate compared with the ones reported in the literature. If the radiating elements are compact, then the array size would be reduced.

The remainder of this paper is structured as follows.

Submitted On: December 22, 2020
Accepted On: August 15, 2021

https://doi.org/10.47037/2021.ACES.J.360807

In Section 2, the single element resonator is explained. Description of the array is presented in Section 3. The PIN diode integration at array feeding arms is investigated in Section 4. Section 5 gives a comparison between the proposed antennas's simulated measured results for the reconfigurable reflection coefficient and radiation pattern. The conclusions are drawn in Section 6.

II. SINGLE ELEMENT ANTENNA GEOMETRY

Quasi lumped resonators (QLR) are widely discussed in literature owing to its attractive feature of reduced size. Its intrinsic circuit components determine the resonance frequency, and its periodic interdigital lumped element geometry enables it to have miniature size, which is its premium advantage. Quasi lumped elements are microwave elements whose physical lengths are very small compared to their respective wavelength. The microstrip lines used in microwave realization with lengths are shorter of quarter wavelength at the operating frequency, are generally denoted by the QLR.

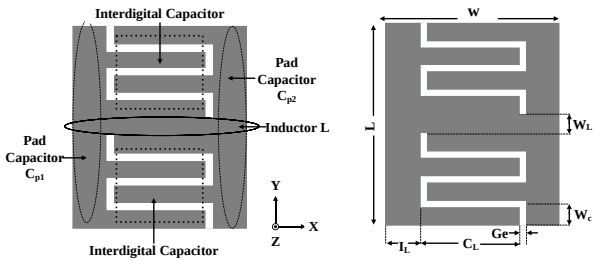


Fig. 1. Quasi-lumped element layout.

For the QLR shown in Fig. 1, the inductance that at the center and can be obtained by solving equation (1):

$$L = 200 \times 10^{-9} I_L \left(ln\left(\frac{2I_L}{W_L+h}\right) + 0.50094 + \frac{W_L}{3I_L} \right), \quad (1)$$

where *h* is the substrate thickness

In Fig. 1, the periodic multifinger structure represents the interdigital capacitor. The capacitance exists between the slits between the conductors. Naturally, these slits are very long. Hence, for reducing the QLR real estate, they are bent. Equation (2) reported by [19,20] is used to determine the interdigital capacitance:

$$C = \varepsilon_0 \left(\frac{\varepsilon_r+1}{2}\right)(N-\Delta)C_L\,, \quad (2)$$

where Δ is the correction factor Δ= (w_{eff}-w), w_{eff} is the effective width, w finger is the width:

N is the fingers number.

The pad capacitors C_{p1} and C_{p2} are situated at the QLR sides are used to fine tune the resonant frequency. equation (3) given by [21], is used to calculate the pad capacitors:

$$C_p = \left(\frac{2.85\varepsilon_{eff}}{ln\left(1+(0.5)\times\left(\frac{8h}{w_{eff}}\right)\right)\times\left(\frac{8h}{w_{eff}}+\sqrt{\left(\frac{8h}{w_{eff}}\right)^2+\pi^2}\right)} \right) \times \left(\frac{1}{25.4\times10^{-3}}\right). \quad (3)$$

The resonant frequency of proposed QLR antenna can be estimated using equation (4) reported by [22,23]. Table 1 shows the dimensions of a single quasi lumped element at a resonant frequency of 5.8 GHz:

$$f = \frac{1}{2\pi\sqrt{L\left(\frac{C_{p1}C_{p2}}{C_{p1}+C_{p2}}+C\right)}}. \quad (4)$$

Table 1: The proposed quasi lumped antenna element parameters

Parameter	Dimensions [mm]
W_e	0.35
I_L	3.35
C	3.05
N	8
$I_L^{\backslash}$	1.23
g_e	1.23
W_L	1.2
L	5.4
W	5.8
h	0.813

III. DESCRIPTION OF THE ARRAY

Two kinds of array feeding topologies are frequently used, parallel topology and series topology geometries [24,25]. The series fed structures are mainly classified into resonant and traveling wave feeds [26,27]. The arrays for which the impedances at the junctions of the feeding line and the resonant elements are not matched are called the resonant arrays. In the traveling-wave array topology, the feeding line impedances and the radiating elements are commonly all matched. The advantages of series topology is that it primarily reduces the feed lines' dielectric insertion losses. Furthermore, it decreases the radiation leakage of the feeding network compared to the parallel feed network. Broadly, the series feed arrangements are more compact as they need shorter transmission lengths, lower insertion loss and fewer junctions. However, it mainly produces narrow bandwidth and suffers of inherent phase difference produced by lengths differences of the feed lines [28]. The parallel feed, also called the corporate feed [29], are common and versatile as it enables controlling the feed of each element. However, corporate feeds need a relatively larger size, and it can be complicated to integrate them into an array environment. Also, due to the supplementary line lengths, losses are increased [30].

The hybrid feed is a combination of series and

parallel feed lines, and its design can eliminate the disadvantages of each of the parallel and the series feed networks. It provides a wider bandwidth over the series feed and can be more miniaturized than the corporate feed design. For purely series-feed array, the antenna input power is transmitted from one end of the array. The main beam angle will be quite sensitive to frequency variation because of the phase variation of the series-fed elements. A combination of parallel-series feed geometry can be employed to prevent the main beam squint due to frequency variations. For the hybrid feed, the beam phase of the series array will squint away from broadside as frequency varies, but, with parallel integration, the beam of the entire array will sustain pointed in the broadside direction. However, the hybrid array insertion loss is higher than purely series-fed array due to its fractional parallel feed. Nonetheless, the hybrid technique provides the designer with the opportunity to trade-offs between insertion loss and bandwidth [31].

The proposed center-fed series-parallel array is employed to design the array as shown in Fig. 2. The power distribution and phase arrangement are accomplished by means of quarter-wavelength lines in the feeding network. The figure shows the array sub-element dimensions and impedances calculated at 5.8 GHz. The quarter-wavelength segment is employed for several reasons as the shortest length for a transformation, reduced size and overall loss. The use of multi-sections of quarter-wavelength transmission lines to gradually transform an impedance will result in lowering dissipation compared to a single quarter-wavelength transmission line [32].

IV. TUNING TECHNIQUE

Different electronic components, such as PIN diodes, FETs, and MEMS, are used to build an RF switch for frequency reconfigurable antennas [33,34]. The MEMS-based RF switches are useful in cases where low loss and high isolation are required, although they are costlier and require higher operating voltage. The FETs have low power consumption, yet suffer from higher loss and poor linearity. The PIN diode-based RF switches are cheaper and offer low-loss operation. The limitation of a PIN diode switch is that a forward DC current is required to switch the diode in the ON state, leading to poor radiation efficiency.

The Schematic diagram of the diode coupling circuit is depicted in Fig. 5. The appropriate value of the RF choke can be estimated to get a very high impedance (approaching infinity) at the desired frequency. However, it's important to refer to the inductor's datasheet and make sure it does not resonate at the designed frequency of the antenna in which it is deployed.

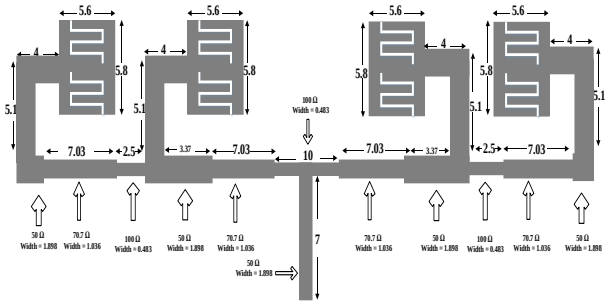


Fig. 2. The proposed hybrid series-parallel array.

The RF PIN diodes is the most prominent in designing reconfigurable antenna among other switches. The PIN diode has several advantages. It requires less complicated biasing components, and its cost makes it a preferable choice for researchers. However, PIN diodes have several weaknesses such as high insertion loss, low efficiency, and requires an additional RF extension. The direct current blocks is mandatory and must be taken into account during the design. The PIN diode equivalent circuit of forward bias and reverse bias is shown in Fig. 3. The allocation of the PIN diodes on the array is depicted in Fig. 4.

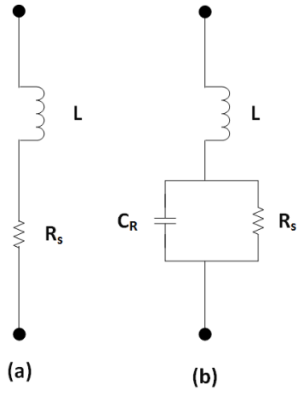


Fig. 3. PIN diode equivalent circuit. (a) Forward bias and (b) reverse bias.

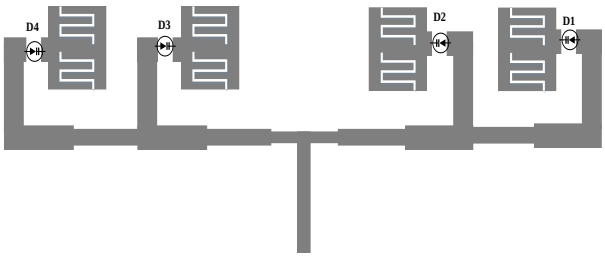


Fig. 4. The proposed array with PIN diodes located at QLR inputs.

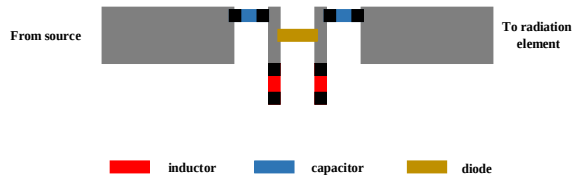


Fig. 5. The schematic coupling circuit of the PIN diode.

V. RESULTS AND DISCUSSIONS

Figure 6 depicts the proposed antenna prototype. The simulation has been executed by 3D Simulation Technology (CST) microwave studio software. The BAR63-02V PIN diode from Infineon Technologies is used as a switch for the reconfigurable antenna design. The diode used provide the advantage of very low capacitance that offers high isolation. Figure 6 (a) shows the discrete ports allocation coupled to resonating elements. The PIN diode equivalent circuit model in the ON and OFF states are shown in Figs. 6 (b) and 6 (c) respectively.

The assembled antenna with the integrated PIN diodes switches is depicted in Fig. 7. Surface mount component (SMC) were integrated to PIN diodes circuits. The circuit consists of one PIN diode, two DC block capacitors and two RF choke inductors. The DC source is applied to the inductor to perform as a short circuit. On the other side, the capacitor only permits the AC current to flow through. For that, the DC run though the other route and ON the PIN diode. A 5 volt DC supply is applied through the red wire. Whilst, the ground is connected to the green wire.

The simulated and measured S_{11} parameters of the proposed antenna are shown in Fig. 8 and Fig. 9, respectively. For the simulated ones, when all diodes are ON, a resonating notch occurs at 5.82 GHz (mode 1). In mode 2, frequency is shifted to between 5.762 to 5.558 GHz by putting one PIN to OFF condition. In mode 3 the resonance is moved between 5.6 to 5.29 GHz when two PIN are OFF. Finally, with three diodes at OFF state, the antenna resonates at 5.107 to 5.298 GHz. From the measured results, at mode 1 if all diodes are ON state the resonating notch occurs at 5.836 GHz. For mode 2 at which three diodes is ON state is observed that the maximum resonance occurs at 5.758 GHz whilst the minimal 5.532 GHz. For mode 3, when two diodes are OFF state, the maximum resonance obtained at 5.586 GHz and the minimal is 5.27 GHz. Lastly, if three diodes are OFF state the resonance occurs in between 5.276 GHz and 5.1 GHz. Different modes for PIN diode modes and corresponding resonance frequencies are given in Table 2.

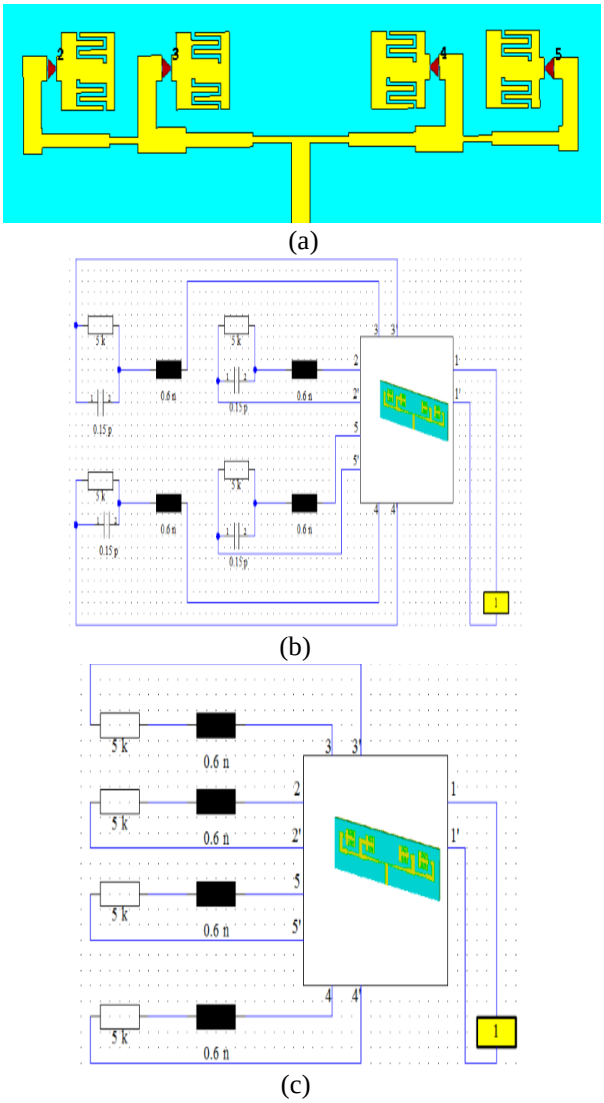


Fig. 6. The discrete port to connect the main radiation plane to the additional plane. (b) Equivalent circuit model of theBAR63-02V PIN diode in the ON state. (c) Equivalent circuit model of theBAR63-02V PIN diode in the OFF state.

In Fig. 8, the reflection coefficients shifts further to the left if more diodes are in OFF mode state. Generally,

the matching impedance is decreased if more PIN Diodes are OFF state. The Simulated S_{11} results are also in good agreement with the measured ones. This is due to the accuracy level of the adjustment of both simulation and measurement.

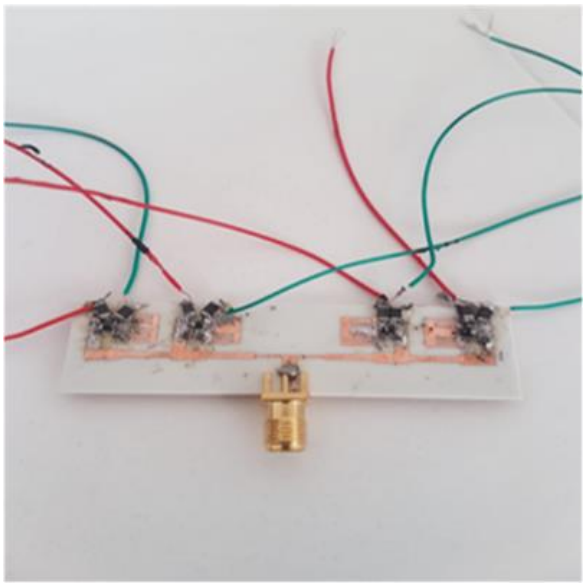
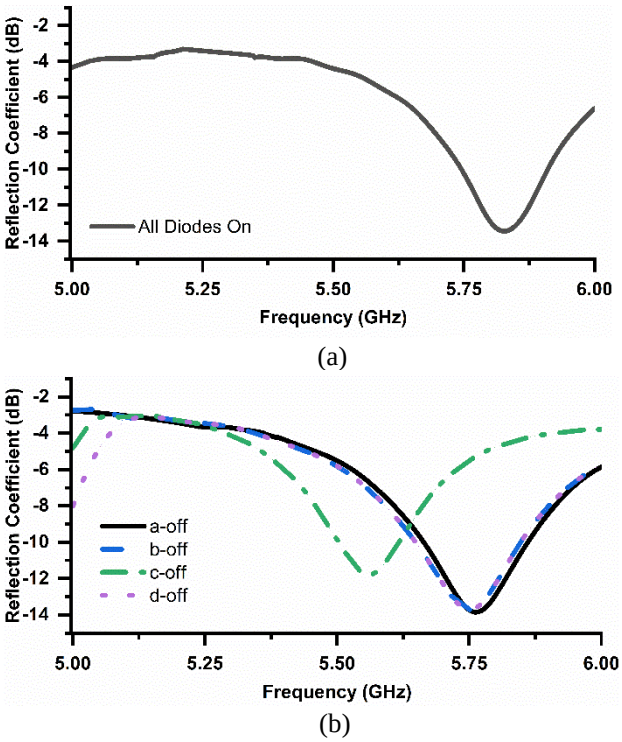


Fig. 7. Geometry of the fabricated antenna structure.



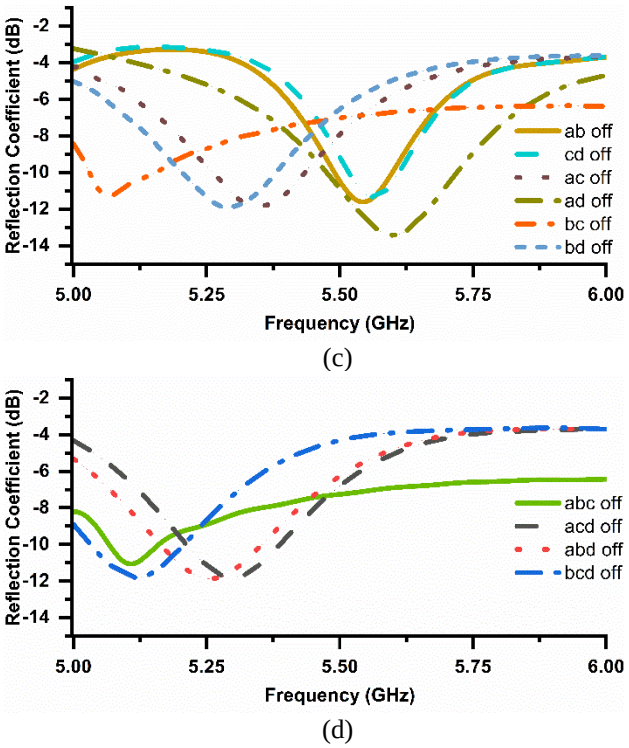
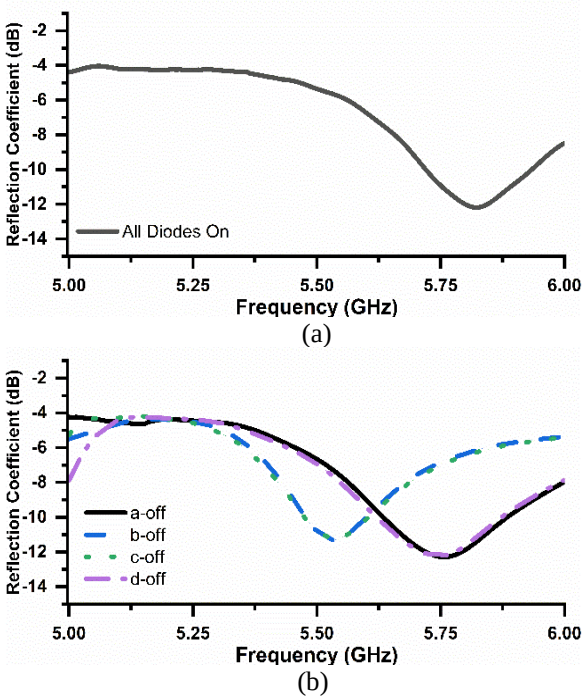
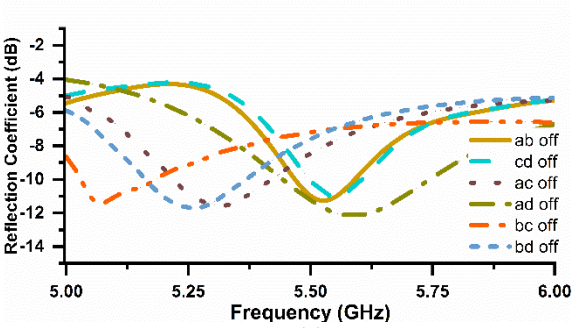
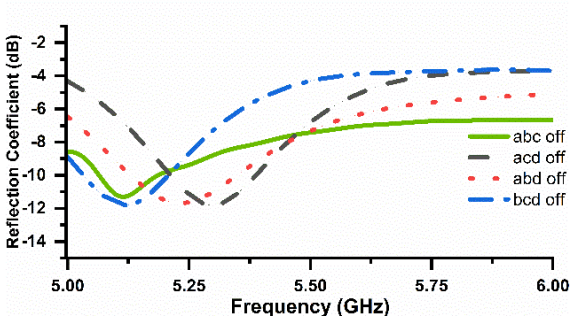


Fig. 8. Simulated reflection coefficient of the proposed antenna. (a) All Diodes are ON. (b) One diode is OFF. (c) Two Diodes are OFF. (D) Three Diodes are OFF.





(c)

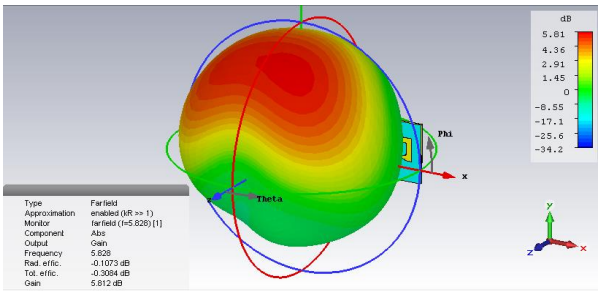


(d)

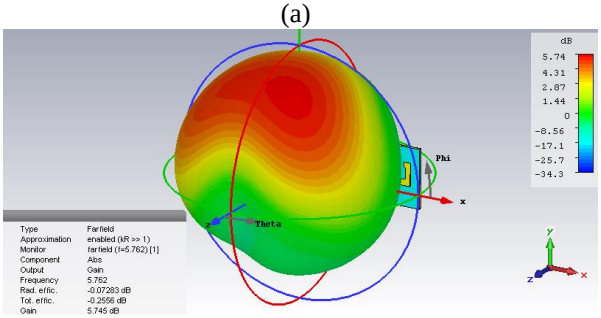
Fig. 9. Measured reflection coefficient of the proposed antenna. (a) All Diodes are ON. (b) One diode is OFF. (c) Two Diodes are OFF. (d) Three diodes are OFF.

Figure 10 depicts the simulated 3D radiation pattern of the PIN diode switches configurations. The figure reveals that the radiation pattern is maintained despite the frequency variation. This result coincides with the assumption that the hybrid feeding topology results in no alteration on the radiation pattern orientation. The figure shows the radiation pattern with all diodes ON state, one diode (a) is ON state, two diodes (a) (b) and (b) are OFF state and three diodes (a) (c) (d) are OFF state. Also, other the diodes states produce the same radiation orientation, i.e., no beam squint due to frequency changes. Figure 11 depicts the polar simulated, and measured radiation pattern of the PIN diode switches configurations for both elevation and azimuth plane. The agreement between the simulated and measured pattern was fairly good.

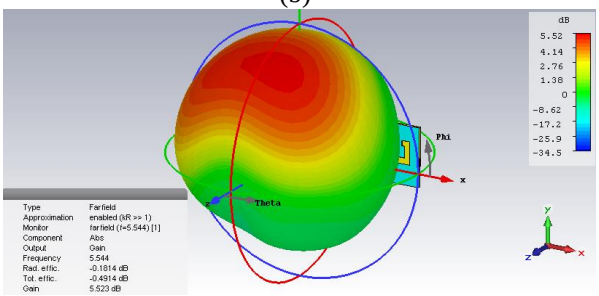
Figure 12 depicts the simulated and measured realized gains of the proposed antenna. The simulated gain ranges from 5.16 to 5.87 dBi. While the measured gain is maintained between 5.14 and 5.84 dBi. Good correspondence exists between these two results. The simulated total efficiency ranges from 0.75% to 0.82.



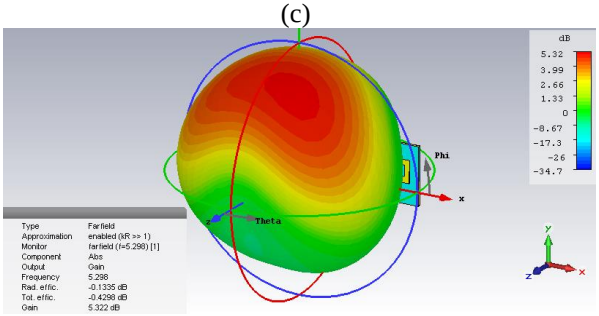
(a)



(b)



(c)



(d)

Fig. 10. 3D simulated radiation pattern of the proposed antenna. (a) All Diodes are ON. (b) One diode is OFF. (c) Two Diodes are OFF. (d) Three Diodes are OFF.

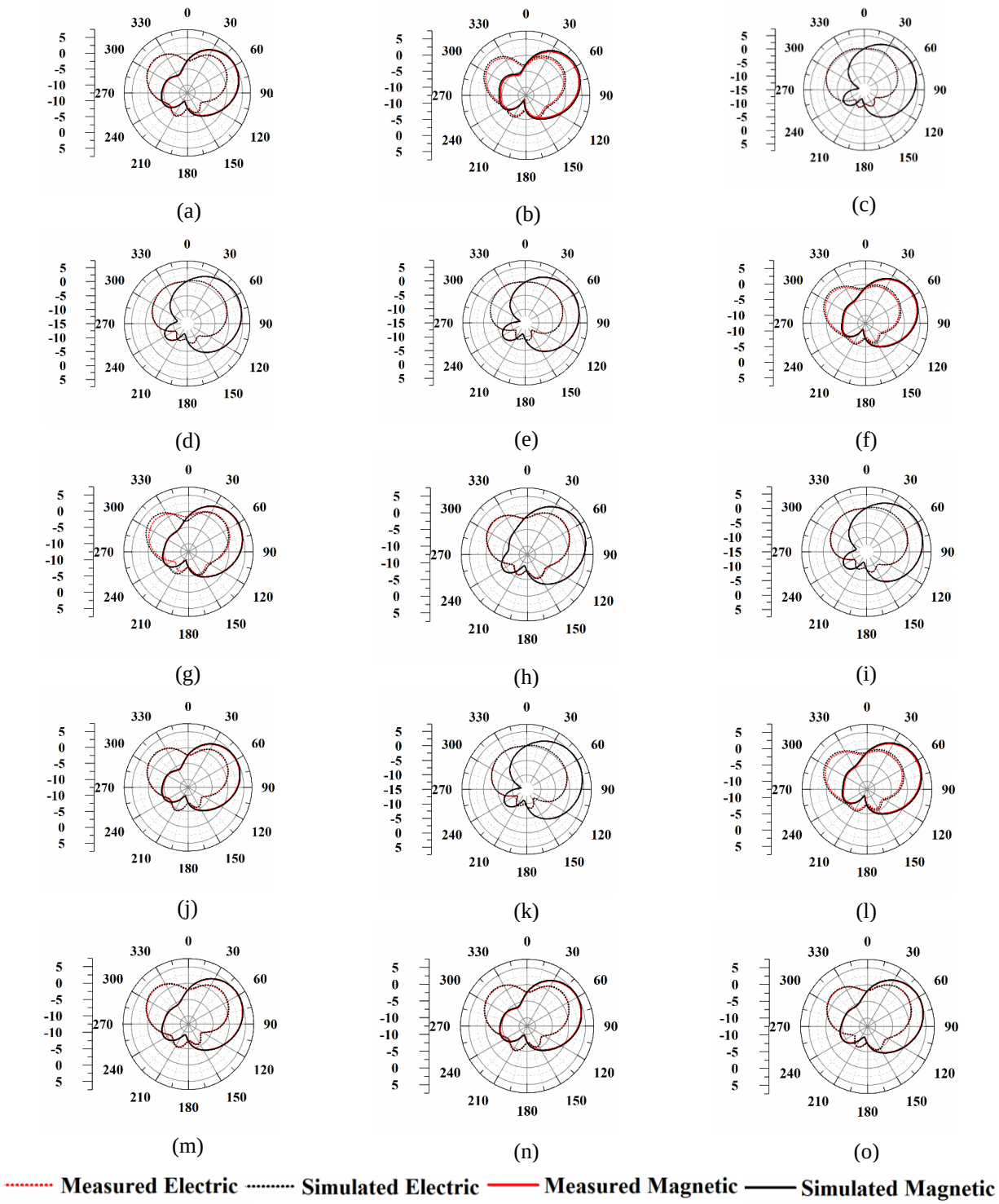


Fig. 11. 2D simulated radiation pattern of the proposed antenna. (a) All Diodes are ON. (b) Diode a is OFF. (c) Diode b is OFF. (d) Diode c is OFF. (e) Diode d is OFF. (f) Diodes ab are OFF. (g) Diodes ac are OFF. (h) Diodes ad are OFF. (i) Diodes bc are OFF. (j) Diodes bd are OFF. (k) Diodes cd are OFF. (l) Diodes abc are OFF. (m) Diodes abd are OFF. (n) Diodes bcd are OFF. (o) Diodes acd are OFF.

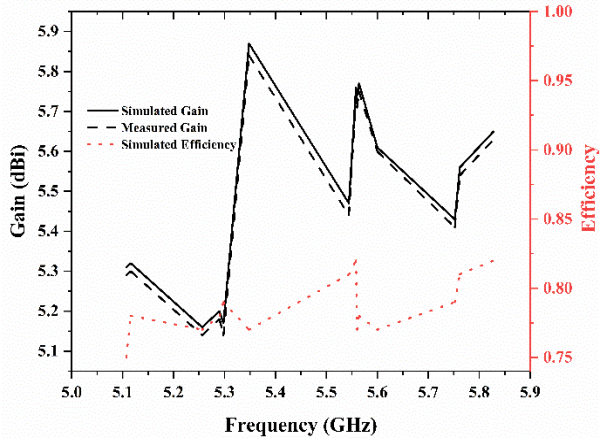


Fig. 12. Simulated and measured realized gain and simulated efficiency.

VI. CONCLUSION

A novel hybrid frequency agile array antenna design with directional and main beam squint immunity operating in the band IEEE 802.11a is proposed. The antenna applies the miniaturized quasi- lumped elements to reduce the array real estate. The array shows the validity of frequency reconfigurability through the range of 5.1 - 5.836 GHz. The proposed antenna with abundant diode modes finds potential use as a portable WiFi application. Moreover, the array has attained a compact physical measurement of by 64×18 mm^2. Table 3 shows the noteworthy size mitigation once compared with other designs reported in literature employed diodes operating in the band IEEE 802.11a.

Table 3: A size comparison between the proposed design and other designs reported in the literature meant for IEEE 802

Reference	Size [mm^2]	Tuning Technique	Number of Elements	Freq [GHz]
17	130x80	PIN diode	4	4.14 – 6
18	68x68	Capacitor switches	4	5.4 - 5.6
Proposed work	64x18	PIN diodes	4	5.1 - 5.836

ACKNOWLEDGMENT

The author extends his appreciation to the Deanship of the Scientific Research at Majmaah Univeristy for funding this work under project number RGP-2019-20.

Table 2: Summary of PIN diode status and corresponding simulated and measured resonance

Pin Diode Status					Simulated		Measured	
Mode	A	B	C	D	Frequency [GHz]	Reflection Coefficient [dB]	Frequency [GHz]	Reflection Coefficient [dB]
All on	On	On	On	On	5.828	-13.44	5.836	-12.11
One Off	Off	On	On	On	5.762	-13.85	5.758	-12.28
	On	off	On	On	5.558	-11.87	5.532	-11.3
	On	On	Off	On	5.56	-11.78	5.538	-11.2
	On	On	On	Off	5.752	-13.71	5.744	-12.23
Two off	Off	Off	On	On	5.544	-11.61	5.526	-11.26
	On	On	Off	Off	5.564	-11.34	5.54	-10.98
	Off	On	Off	On	5.348	-11.83	5.32	-11.55
	Off	On	On	Off	5.6	-13.43	5.586	-12.16
	On	Off	Off	On	5.29	-11.25	5.27	-11.42
	On	Off	On	Off	5.29	-11.91	5.264	-11.67
	Off	Off	Off	On	5.107	-11.06	5.115	-11.29
Three Off	Off	On	Off	Off	5.298	-11.8	5.276	-11.47
	Off	Off	On	Off	5.256	-11.87	5.232	-11.68
	On	Off	Off	Off	5.116	-11.08	5.1	-11.45

REFERENCES

[1] A. Vijayan, L. R. KarlMarx, K. J. Jegadish Kumar, and C. C. Vimlitha, "N-shaped frequency reconfigurable antenna with auto switching unit," *Applied Computational Electromagnetics Society Journal*, vol. 33, no. 6, pp. 710-713, June 2018.

[2] J. Jame and P. Hall, *Handbook of Microstrip Antennas*. Peregrinus on behalf of the Institution of Electrical Engineers, London, 1989.

[3] F. Zurcher and E. Gardioi, *Broadband Patch*

Antennas. Artech House, London, 1995.

[4] G. Kumer and C. Gupta, "Broadband microstrip antennas using additional resonators gap coupled to the radiating edges," *IEEE Trans. Antennas and Propagation*, vol. 32, no. 12, pp. 1375-1379, Dec. 1984.

[5] K. Smith and E. Mayes, "Stacking resonators to increase the bandwidth of low-profile antennas," *IEEE Trans. Antennas and Propagation*, vol. 47, no. 2, pp. 1473-1476, Dec. 1987.

[6] M. Pozar and B. Kaufman, "Increasing the bandwidth of a microstrip antenna by proximity coupling," *Electronics Letters*, vol. 23, no. 8, pp. 368-369, Apr. 1987.

[7] Y. Li, W. Li, and Q. Ye, "A reconfigurable triple-notch-band antenna integrated with defected microstrip structure band-stop filter for ultra-wideband cognitive radio applications," *International Journal of Antennas and Propagation*, vol. 2013, pp. 1-13, May 2013.

[8] Y. Li, W. Li, and W. Yu. "A switchable UWB slot antenna using SIS-HSIR and SIS-SIR for multi-mode wireless communications applications," *Applied Computational Electromagnetics Society Journal*, vol. 27, no. 4, pp. 340-351, Apr. 2012.

[9] Y. Li, W. Li, and Q. Ye "A reconfigurable wide slot antenna integrated with sirs for UWB/multiband communication applications," *Microwave and Optical Technology Letters*, vol. 55, no. 1, pp. 52-55, Jan. 2013.

[10] Y. Tawk, J. Costantine, K. Avery, and G. Christodoulou, "Implementation of a cognitive radio front-end using rotatable controlled reconfigurable antennas," *IEEE Transactions on Antennas and Propagation*, vol. 59, no. 5, pp. 1773-1778, Mar. 2011.

[11] H. Haertling, "Rainbow actuators and sensors: A new smart technology," *Proc. SPIE 3040, Smart Structures and Materials 1997: Smart Materials Technologies*, San Diego, CA, USA, pp. 81-92, June 1997.

[12] A. Bokhari, F. Zurcher, R. Mosig, and E. Gardiol, "A small microstrip patch antenna with a convenient tuning option," *IEEE Transactions on Antennas and Propagation*, vol. 44, no. 11, pp. 1521-1528, Nov. 1996.

[13] J. Rainville and J. Harackewiez, "Magnetic tuning of a microstrip patch antenna fabricated on a ferrite film," *IEEE Microwave and Guided Wave Letter*s, vol. 2, no. 12, pp. 483-485, Dec. 1992.

[14] K. Mishra, S. Pattnaik, and N. Das, "Tuning of microstrip antenna on ferrite substrate," *IEEE Transactions on Antennas and Propagation*, vol. 41, no. 2, pp. 230-233, Feb. 1993.

[15] H. Al-Charchafchi, and M. Frances, "Electronically tunable microstrip patch antennas," *IEEE Antennas and Propagation Society International Symposium*, Atlanta, GA, USA, pp. 304-307, June 1998.

[16] K. Guney, "Resonant frequency of a tunable rectangular microstrip patch antenna," *Microwave and Optical Technology Letters*, vol. 7, no. 12, pp. 581-585, Aug. 1994.

[17] Z. Li, D. Rodrigo, L Jofre, and A. Cetiner, "A new class of antenna array with a reconfigurable element factor," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 4, pp. 1947-1955, Dec. 2013.

[18] A. Byford, K. Park, P. Chahal, and J. Rothwell, "Frequency reconfigurable patch antenna array," *Electronics Letters*, vol. 51, no. 21, pp. 1628-1630, Oct. 2015.

[19] F. Huang, B. Avenhaus, and M. J. Lancaster, "Lumped-element switchable superconducting filters," *IEE Proceedings - Microwaves, Antennas and Propagation*, vol. 146, no. 3, pp. 229-233, June 2013.

[20] E. Bogatin," Design rules for microstrip capacitance," *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, vol. 11, no. 3, pp. 253-259, Sep. 1988.

[21] G. Alley, "Interdigital capacitors and their application to lumped-element microwave integrated circuits," *IEEE Transactions on Microwave Theory and Techniques*, vol. 18, no. 12, pp. 1028-1033, Dec. 1970.

[22] I. Bahl, *Lumped Elements for RF and Microwave Circuits*. Artech House, Boston, 2003.

[23] H. T. Su, M. J. Lancaster, F. Huang, and F. Wellhofer, "Electrically tunable superconducting quasilumped element resonator using thin-film ferroelectrics," *Microwave and Optical Technology Letters*, vol. 24, no. 3, pp. 155-158, Feb. 2000.

[24] A. Abbaspour-Tamijani and K. Sarabandi "An affordable millimeter-wave beam-steerable antenna using interleaved planar subarrays," *IEEE Transactions Antennas and Propagation*, vol. 51, no. 9, pp. 2193-2202, Nov. 2003.

[25] J. R. James, P. S. Hall, and C. Wood, *Microstrip Antenna: Theory and Design*. Peter Peregrinus, Ltd., London, 1981.

[26] R James, S. Hall and C. Wood, *Microstrip Antenna: Theory and Design*. Peter Peregrinus: London, UK, 1981.

[27] Q. Zhu, J. Zhu, W. Lu, L. Wu, and S. Xu, "Dual-linearly polarized microstrip array based on composite right/left-handed transmission line," *Microwave and Optical Technology Letters*, vol. 48, no. 7, pp. 1366-1369, July 2006.

[28] S. Horng and G. Alexopoulos, "Corporate feed design for microstrip arrays," *IEEE Transactions on Antennas and Propagation*, vol. 41, no. 14, pp. 1615-1624, Dec. 1993.

[29] Petosa, A. *Dielectric Resonator Antenna Handbook*. Artech House Inc: Boston (MA), London, 2007.
[30] J. Huang, "A parallel-series-fed microstrip array with high efficiency and low cross-polarization," *Microwave and Optical Technology Letters*, vol. 5, no. 5, pp. 230-233, May 1992.
[31] C. Li and S. Ricketts, "Loss minimisation in λ/4 impedance transformers using multiple λ/4 segments," *Electronics Letters*, vol. 49, no. 4, pp. 274-276, Feb. 2013.
[32] B. Mun, C. Jung, M. Park, and B. Lee, "A compact frequency-reconfigurable multiband LTE MIMO antenna for laptop applications," *IEEE Antennas and Wireless Propagation Letters*, vol. 13, pp. 1389-1392, July 2014.
[33] X. Yang, J. Lin, G. Chen, and F. Kong, "Frequency reconfigurable antenna for wireless communications using GaAs FET switch," *IEEE Antennas and Wireless Propagation Letters*, vol. 14, pp. 807-810, Dec. 2014.
[34] H. Rajagopalan, M. Kovitz and Y. Rahmat-Samii, "MEMS reconfigurable optimized E-shaped patch antenna design for cognitive radio," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 3, pp. 1056-1064, Nov. 2013.

Yazeed Mohammad Qasaymeh received B.Sc. Eng. degree from Mutah University, Karak, Jordan, in 2006, M.Sc. Eng. degree in Engineering from College of Electrical and Computer Engineering, University Sains Malaysia, Pinang, Malaysia, in 2009, and Ph.D. degree in the College of Electrical and Electronic Engineering, Universiti Sains Malaysia, Pinang, Malaysia. He is currently Assistant Professor in College of Engineering, Al-majmaah University at Saudia Arabia. His research interest is antennas array, wireless communication MIMO, distributed antenna systems.

Compact High Gain Multiband Antenna Based on Split Ring Resonator and Inverted F Slots for 5G Industry Applications

Ranjan Mishra [1], Rajeev Dandotiya [1], Ankush Kapoor [2], and Pradeep Kumar [3]

[1] Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies, Dehradun, India
rmishra@ddn.upes.ac.in, rajeevdandotiya@gmail.com

[2] Department of Electronics and Communication Engineering, Jawaharlal Nehru Government Engineering College
Sundernagar, Mandi, India
ankush8818@yahoo.com

[3] Discipline of Electrical, Electronic and Computer Engineering, University of KwaZulu-Natal
Durban-4041, South Africa
pkumar_123@yahoo.com

Abstract ─ This paper presents the design, optimization, fabrication, and measurement of the compact high gain microstrip antenna with a split ring resonator and set of inverted-F slots along with a matching stub for sub-6 GHz5G applications. In this investigation, different iterations are visualized by incorporating inverted F slots, a split ring resonator, and a matching stub in the transmission line. The advantages of each incorporated structure are analyzed, and a hybrid antenna consisting of the combination is proposed as a final antenna configuration with the optimum results. The proposed final design attains compactness and multi-band operation. Impedance matching is improved by using the stub matched technique at the feed line. The designed antenna shows the resonances at precisely 2.1 GHz, 3.3 GHz, and 4.1GHz. The proposed antenna is suitable for mobile cellular communication such as the LTE band (2.1 GHz), n78 band (3.3 GHz), and n77 band (4.1 GHz) of 5G bands. The gain retrieved from each band attains more than 5 dB value.

Index Terms ─ Fifth generation (5G), inverted F-slot, Split Ring Resonator (SRR), stub matching.

I. INTRODUCTION

In the 21st century, wireless communication and networking devices require multiple operating frequencies due to the rapid increase in the demand of users. This demand for multi-band operated antennas needs to cater to compactness in dimensions while maintaining the performance characteristics. These reasons are sufficient for going to the design of a multi-band antenna. Multiband antennas are becoming an eye of sight for researchers in applications with a need for downward compatibility and adequate facilities. With the rise of LTE and demand for 5G technology, there is always a need for multi-band antennas. High data rate and increasing users for data transmission systems have motivated researchers to develop multi-band antennas. In recent years, researchers have made various efforts to get the multi-band operation by using patch antennas. A patch antenna with a T-shaped slit on the radiating aperture has shown a multi-band operation with operating frequencies at 3.92 GHz and 5.82 GHz of C band, and 7.88 GHz and 11.35 GHz of X band of microwave spectrum [1]. A planar MIMO antenna structure integrated with mm-wave has been discussed [2] for future V2X applications. This phased array antenna structure is used in LTE/sub-6 GHz 5G and the miniaturization was achieved using a gap capacitor and inter digital capacitors. Further, a multi-band antenna array structure suitable for smart phones operating in LTE bands 42/43/46 has been presented using two diverse open slots on a T-shaped slot antenna [3]. Sub-6 GHz has taken the interest of researchers in designing the antenna, radio frequency circuits, and spatial filters [4]. A simple single feed design is proposed for a multi-band antenna with defected ground substrate structure [5]. A compact slot antenna with the 3 L-shaped slots and the ground plane for three operating frequency bands is proposed in [6]. A compact F-shaped slot antenna showing multi-band frequency is proposed, and its application area is focused on WLAN, Wi-MAX, and X band applications [7]. In [8], the design of a four-band antenna is presented, and the structure involves a T-shaped feed patch, an inverted T-shaped stub, and two E-shaped stubs. Further investigations have been done in designing a partial slotted ground antenna for wideband applications [9]. Further in [10], patch antennas for sub-6 GHz 5 G communications have a T slot on a

Submitted On: March 23, 2021
Accepted On: July 7, 2021
https://doi.org/10.47037/2021.ACES.J.360808

rectangular patch and a defective ground structure. A wideband antenna, with gain enhancement using FSS based spatial filters, is presented for operation in less than 6 GHz of frequency [11]. A compact reconfigurable 3-D slot antenna suitable for 5G mobile application is offered with a metal casing [12]. In [13], a dual multi-band planar inverted F antenna system is proposed in which the bandwidth is enhanced by using a specific configuration of slots. A design of an inverted F slot antenna structure has claimed to be easily integrated into handheld devices or printed on WLAN card is proposed [14]. In [15], a new planar inverted-F antenna with a very large bandwidth starting from 817 MHz to 11.5 GHz is proposed as an alternative for high performance mobile phones and also intended to cover the major part of the mobile phone frequencies as well as the ultra-wideband (UWB) frequency range. The research on multi-band antennas has been extended in which a dual-band antenna consisting of a rectangular patch fed by the coaxial probe feeding technique and a ground plane loaded with two rectangular strip slots and one elliptical slot is proposed [16]. Further, a compact triple-band microstrip antenna, with a rectangular stub connected to the feed line, is proposed for tri-band applications [17]. The proposed antenna covers 2.1-2.8 GHz, 3.3-4.0 GHz, and 5.5-5.8 GHz. A triple-band microstrip antenna consisting of a slotted rectangular patch with the defected ground is designed for triple-band operation at 1.2 GHz, 2.45 GHz, and 5.6372 GHz [18]. In [19], the PIFA is proposed employing a rectangular split-ring resonator structure exhibiting multi-band characteristics and suitable for mobile applications.

In this paper, a design of multi-band antenna is recommended by incorporating fractal slots, inverted F-slots, and split ring resonator. The antenna shows multi-band behavior with an adequate value of gain and finds its applications in sub-6 GHz applications.

The manuscript is organized as follows. The antenna geometry is described in Section II. Section III is devoted to measuring and validating the results, and finally, the conclusion is briefed in Section IV.

II. ANTENNA GEOMETRY

This section presents the design and the optimization of an inverted-F slot microstrip antenna for sub-6 GHz band in 5G application. In this investigation, an inverted-F slot on the antenna patch is designed to get the multi-band operations with a higher order of miniaturization. Impedance matching is improved by using a stub matched technique at the feed line. The designed antenna shows a perfect resonance at 2.1 GHz, 3.3 GHz, and 4.1 GHz of the frequency with a good reflection coefficient. The gain of each band is more than 5dB. These three frequencies are chosen for LTE (2.1 GHz), n77 band of 5G (3.3 GHz), and n78 band of 5G (4.1 GHz), respectively, as per ITU-NR and department of telecommunications (DOT), Government of India proposal for 5G applications [20]-[21]. The proposed antenna geometry considers a square fractal aperture with microstrip line feeding. The feeding is optimized with stub matching to improve the reflection coefficient characteristics and provide operability in the 5G applications. The square aperture of the antenna is first optimized by the combination of inverted-F slots at the diagonal corner to improve the performance of the antenna. A split ring slot is dogged just after the feed line, which cooperatively improves the radiation characteristics of the designed antenna. The proposed antenna demonstrates three resonance bands, the first one is at 2.1 GHz, and the second one is at 3.3 GHz, while the third one is at 4.1 GHz of frequency. Typically gain at each resonance is enhanced to more than 4 dB and with a good radiation pattern. The combined fractal and slotted configuration realize multi-band operation applicable for the mobile communication system. The proposed antenna is suitable for mobile cellular communications such as LTE band (2.1 GHz) and 5G-NR (3.3 GHz and 4.1 GHz) communication bands. A compact microstrip inverted-F slotted fractal antenna for mobile and 5G-NR applications is presented. In this investigation, the inverted-F slot on the antenna aperture is analyzed to demonstrate multi-band mobile operations with a higher order of miniaturization. The impedance matching is improved by using a stub-matched technique with offset feeding. The substrate chosen for this design is RT-duroid (RT 5880) substrate, having a dielectric constant value of 2.2, and a loss tangent of 0.009. Substrates of low tangent losses will minimize dielectric losses, but these substrates are costly. The chosen height of the substrate is 1.57 mm and using this moderate thicker substrate, conductor losses may also be minimized, thereby influencing the antenna's performance. This RT-duroid substrate is although expensive, but at the same time, enhances the performance in terms of reflection coefficient and gain. The aim in choosing this specification is to design a multi-band antenna with good gain, reflection coefficient, and directive radiation property. The design is proposed to find its scope in the modern wireless mobile communication system. Primarily the physical dimensions are determined using the equations given below [22].

The width of the patch, W_p, is given as:

$$W_P = \frac{c}{2f}\sqrt{\frac{2}{(\varepsilon_r+1)}}, \quad (1)$$

where c, f and $\boldsymbol{\varepsilon}_r$ denote the velocity of light, the resonant frequency andthe relative permittivity, respectively.

The length of the patch, L_p, is given as:

$$L_P = L_{eff} - 2\Delta L, \quad (2)$$

where ΔL denotes the extended effective length. The effective length, L_{eff}, is given as below:

$$L_{eff} = \frac{c}{2f\sqrt{\varepsilon_{eff}}}. \quad (3)$$

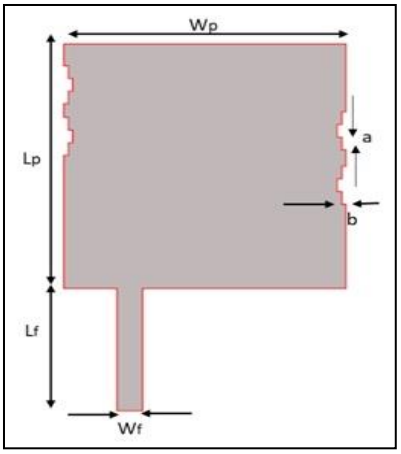
The effective dielectric constant, ε_{eff}, of the substrate, is given as:

$$\varepsilon_{eff} = \frac{\varepsilon_r+1}{2}+\frac{\varepsilon_r-1}{2}\sqrt{\left[1+\frac{12h}{W_p}\right]}, \quad (4)$$

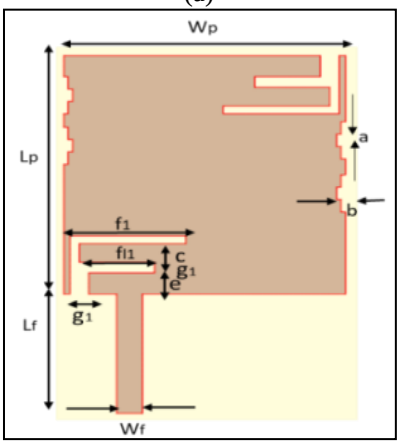
where h denotes the thickness of the substrate. By considering the normalized extension, the actual length of patch, L_p, is given as:

$$L_P = \frac{c}{2f}\left(\frac{\varepsilon_r+1}{2}+\frac{\varepsilon_r-1}{2}\sqrt{\left[1+\frac{12h}{W}\right]}\right)^{-\frac{1}{2}} - 2\Delta L. \quad (5)$$

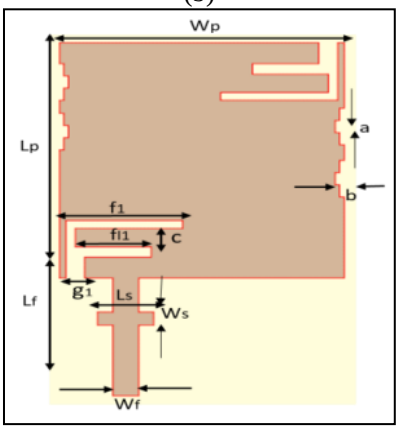
The design flow of the proposed antenna has gone through little iteration. A conventional rectangular patch is implemented on the top surface of dielectric substrate wherein a fully conductive ground is considered. In the first iteration, fractal slots are engraved at the opposite boundaries along the length of the patch for making the design a fractal slot based rectangular patch antenna (FSRPA). Fractal slot is derived from the self-identical and broken irregular pieces, belonging to a family of intrinsic geometrical structures [23]. The geometrical layout along with the dimensional parameters is illustrated in Fig. 1 (a). A fractal slot carved on the edge along the length of the antenna produces different frequency bands owing to its repeating geometry structure. The microstrip patch antenna embedded with the modified fractal design exhibited a dual frequency response. This antenna provides resonance at various frequencies in the band due to its fractal design. In the next step of the design iteration, a dual inverted plane F slot is introduced on the radiating patch (DIFRPA). The DIFRPA consists of F-slots structures that lie opposite to each other as demonstrated in Fig. 1 (b). However, the design of Fig. 1 (a) produces the multi resonance characteristic, but the performance is not good in terms of reflection coefficient for the specified bands. This shortcoming leads to the second design. The second design of Fig. 1 (b) also illustrates the three well distinguishable bands, but the second band doesn't give a good reflection coefficient. Moreover, the three bands are still not falling precisely as per the specified frequency. The objective of exact resonance at a specific frequency enables the structure of the third design. The third design of Fig. 1 (c) consists of stub matching (SMRPA) helps to achieve triple-band operation with the narrow bandwidth. The improvement is required in the third design for better matching impedance between the source and the antenna. This inverted F structure is also used in controlling the frequency and to enhance the gain by improving impedance matching and reducing the backward radiation [24]-[26].



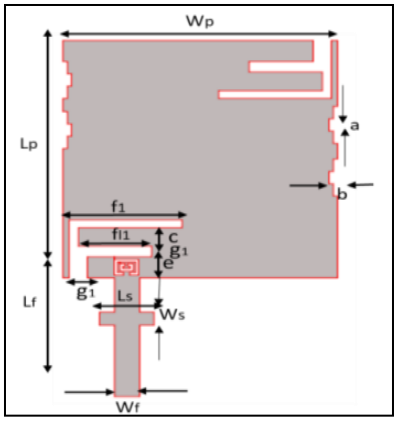
(a)



(b)



(c)



(d)

Fig. 1. Design evolution steps of the proposed antenna, (a) FSRPA, (b) DIFRPA, (c) SMRPA, and (d) SRRRPA.

In short, this structure consists of radiating patches parallel to a ground that is joined to the ground by a shorting plate to resonate at a quarter wavelength. The impact of the inverted-F slot over the fractal slot-loaded square patch antenna is to bring the first resonance at 2.2 GHz and the second resonance at 3.0 GHz. The designed antenna also demonstrates the third band around 4.2 GHz of frequency and thus results in multiple resonances in sub-6 GHz band, as shown in the reflection coefficient plot depicts in Fig. 2. This design brings the first band at 2.2 GHz. The second band also comes to be at 3.3 GHz, but its resonance is not good, and the radiation is near to -10 dB. The third band is exciting at 4.2 GHz with a poor reflection coefficient. Implementation of multiple slots improves the multi-band response, but it degrades the impedance characteristics. Stub matching technique has been investigated to improve the impedance response at multiple resonances. Analysis has been carried out using different stub size and position. It has been observed that the change in stub dimension gives considerable order flexibility in impedance tuning without disturbing the original resonances. It is realized the stub length (Ls) of 8.9 mm and width (Ws) of 2.5 mm meet the appropriate reflection coefficient response. The comparison of the reflection coefficients obtained from all four designs is shown in Fig. 2. This design brings the first band at 2.2 GHz. The second band also comes to be at 3.0 GHz, but its resonance is not good and the radiation is near to -10 dB. The third band is exciting at 4.2 GHz with a reflection coefficient near -10 dB. The introductions of stub matching improve the three bands' overall reflection coefficient and precisely position the three bands at the desired and specified frequency.

In the last step, a tiny split ring slot is implemented on the patch of the antenna aperture making 4th iteration as split ring resonator based rectangular patch antenna (SRRRPA). This is done to improve the impedance matching of the antenna further to get it excited at the specified frequencies. The position of this slot is just at the junction point of feed line and radiating patch. The final proposed antenna design**,** i.e.**,** SRRRPA is loaded with fractal slots, inverted F slot and split ring, and a stub matching technique to achieve a fine tuning in the impedance profile. All these structures are plane in nature with easily etched on the patch, and the structure of the final designed antenna is shown in Fig. 1 (d). As per Fig. 2, the reflection coefficient curve of the SRRRPA demonstrates that all the three bands are well distinguished, and they are resonating strictly at 2.1 GHz, 3.3 GHz, and 4.1 GHz of the frequency with a reflection coefficient of -28 dB, -21 dB, and -24 dB respectively. The dimensional parameters for the proposed SRRRPA are enumerated in Table 1.

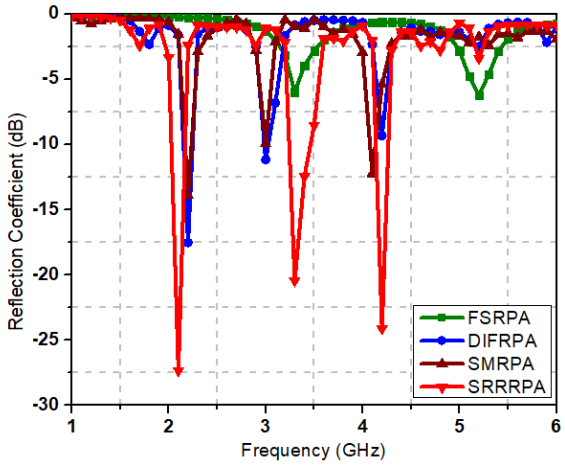


Fig 2. Reflection coefficient of various antenna configurations.

Table 1: Dimensional parameters of the SRRRPA

Parameter	Value
Substrate height (hs)	0.01 λ_{mid}
Patch length, Lp	0.46 λ_{mid}
Patch width, Wp	0.46 λ_{mid}
Feed length (Lf)	0.23 λ_{mid}
Feed width (Wf)	0.04 λ_{mid}
Upper arm of F slot, fl	0.19 λ_{mid}
Lower arm of F slot, fl1	0.12 λ_{mid}
Fractal slot, a	0.02 λ_{mid}
Fractal slot, b	0.007 λ_{mid}
Gap between F slot, c	0.03 λ_{mid}
Distance of f slot, e	0.04 λ_{mid}
Width of base of F slot, g1	0.03 λ_{mid}
Width of arm of F slot, g2	0.02 λ_{mid}
Length of stub, Ls	0.09 λ_{mid}
Width of stub, Ws	0.02 λ_{mid}

Where λ_{mid} is the middle wavelength corresponding to 3.1 GHz. The comparative analysis of the variation of the peak gain along the operating frequency bands is illustrated in Fig. 3. Figure 3 depicts that the antenna has a good gain response of more than 5 dB in the entire zone of its operating frequency for the final design of SRRRPA. The values of the peak gain of SRRRPA at the three bands are 5.4 dB at 2.1 GHz, 6.4 dB at 3.3 GHz, and 5.1 dB at 4.1 GHz.

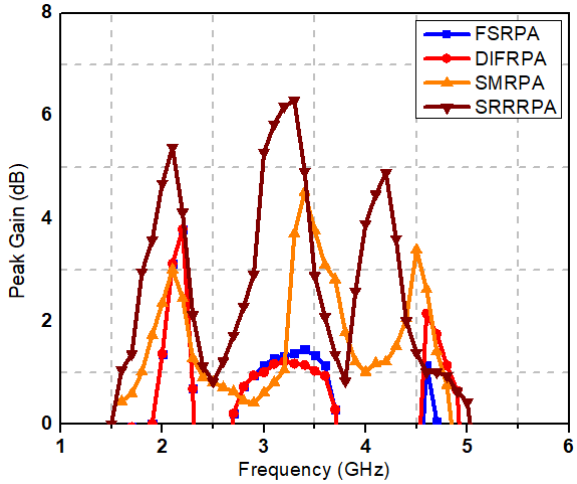


Fig. 3. Peak gain of various antenna configurations.

III. MEASUREMENT VALIDATION

The final antenna design of SRRRPA is fabricated using RT-duroid (Rogers 5880) as dielectric substrate with a height of 1.57 mm. In the fabrication process, a photographic film of the antenna is carved out initially, and then it is glued to the substrate. The patch and the ground of the substrate are copper cladding. In the next stage, the antenna is etched out with a chemical solution. Finally, the fabrication of the antenna is done by a computerized mechanical etching process. In this process, necessary care was taken to align the ground slit exactly below the stub-matched transmission line. Figure 4 shows the photograph of the fabricated antenna.

The reflection coefficient of the fabricated antenna is measured using a vector network analyzer and is validated by comparing it with the simulated results. The gain and the radiation pattern are measured in an anechoic chamber. The measured values of the reflection coefficient with frequency and the simulated results are shown in Fig. 5. The reflection coefficient of the fabricated antenna is measured using a vector network analyzer and validated by comparing it with the simulated results. The gain and the radiation are measured in an anechoic chamber. As demonstrated, the experimental analysis follows the resonance characteristics precisely the same as that of the simulation work with a very nominal deviation. The experiment analysis validates the claimed resonances at 2.1 GHz, 3.3 GHz, and 4.1 GHz of the frequency with reflection coefficient -19 dB, -17 dB, and -18 dB, respectively. The plot for gain response is shown in Fig. 6, and it depicts that the measured peak gain at three bands is 4.6 dB, 5.5 dB, and 4.1 dB, respectively. The E-plane and H-plane plot for all the resonance frequencies (both measured and simulated) are shown in Fig. 7, respectively. The radiation pattern of the E-plane demonstrates that the measured results are in accordance with the simulated design work. The maximum radiation intensity is towards the principal lobe of the antenna, and the back lobe is almost null. It depicts the directive nature of the radiation. The measurement of H-plane results reveals that the radiation intensity is concentrated more towards the main lobe. The main lobe is dominating with almost no presence of back lobe radiation. The radiation pattern also clearly shows the directive nature of the radiation.

Fig. 4. Prototype of the SRRRPA.

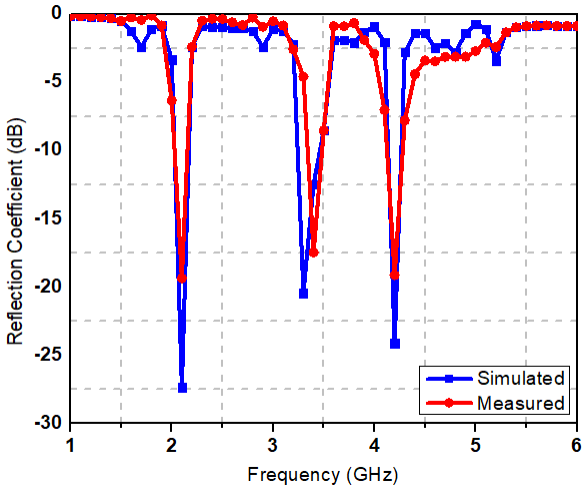


Fig. 5. Simulated versus measured reflection coefficient of the SRRRPA.

With these viable performances, it claims its efficient application in the three chosen bands of frequencies. Table 2 summarizes the results, which show the matching between the simulated and experimental results. In this section, a comparison of the converged design is presented with a couple of other research work on sub-6 GHz of frequency, summarized in Table 3. where λ_{mid} is the wavelength at the middle of the frequency, i.e., between the highest and the lowest frequency of operation in the multi-band operation.

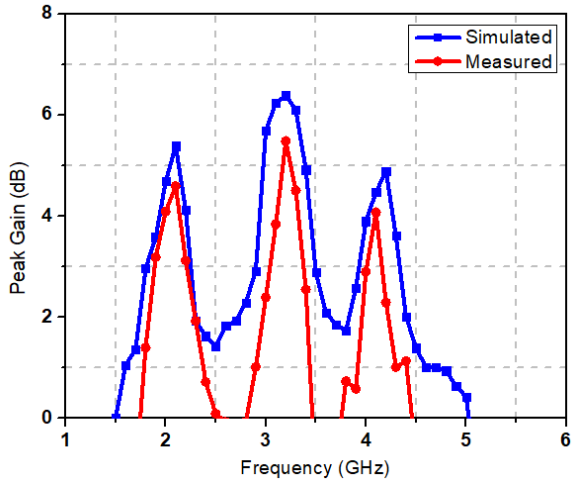


Fig. 6. Simulated and measured gain of the SRRRPA.

Table 2: Comparison between simulated and measured results

Freq.	Parameter	Simulated	Measured
2.1 GHz	Reflection coefficient	-27.5 dB	-19.7 dB
	VSWR	1.1	1.09
	Gain	4.5 dB	4.5 dB
	Bandwidth	200 MHz	230 MHz
3.3 GHz	Reflection coefficient	-20.5 dB	-17.5 dB
	VSWR	1.28	1.21
	Gain	6.4 dB	5.5 dB
	Bandwidth	170 MHz	190 MHz
4.1 GHz	Reflection coefficient	-24.1dB	-19.6 dB
	VSWR	1.2	1.18
	Gain	4.3 dB	4.1 dB
	Bandwidth	200 MHz	220 MHz

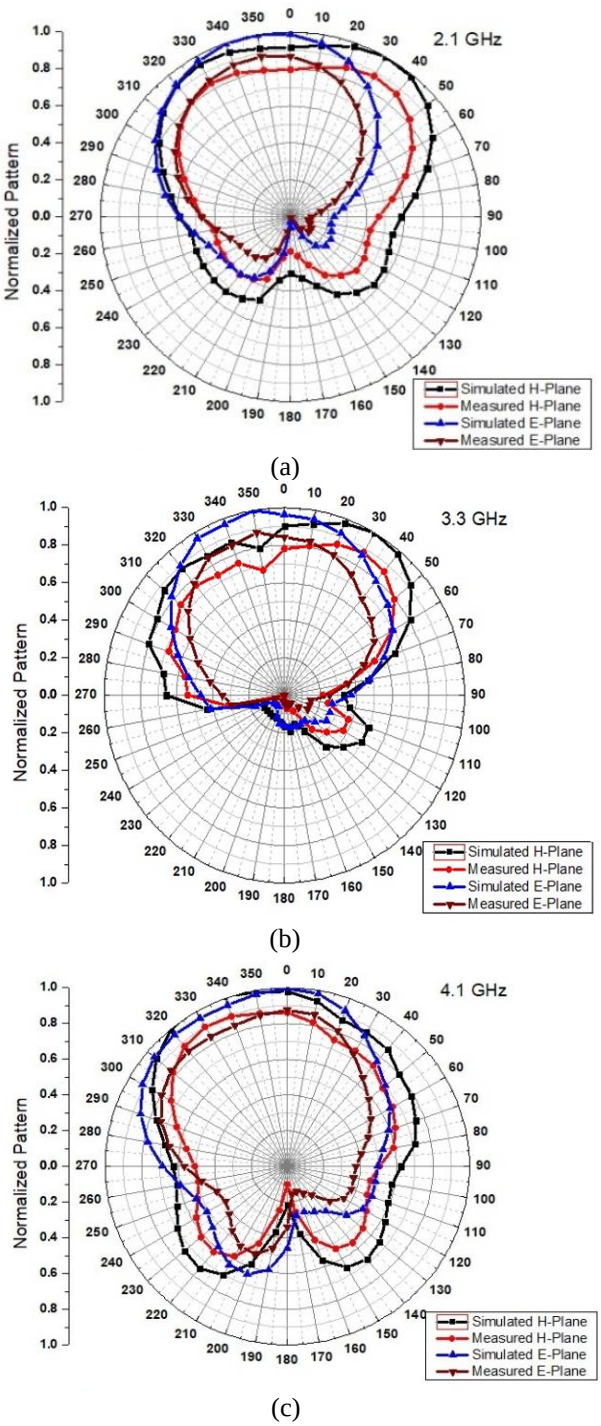


Fig. 7. Normalized radiation patterns of the SRRRPA: (a) 2.1 GHz, (b) 3.3 GHz, and (c) 4.1 GHz.

IV. CONCLUSION

In this manuscript, the design of the multi-band antenna, to be operated in sub-6 GHz 5G spectrum, is presented consisting of three bands of operating frequency. The compact microstrip antenna loaded with an inverted-F slotted fractal stub matching structure is suitable for applications in LTE-A and 5G-NR mobile technology. In the design process, investigation on fractal slots and inverted- F slot on the antenna patch is analyzed to demonstrate multi-band behavior with a higher order of precision. The symmetrical fractal lies at the edge of the length, whereas inverted-F slots are opposite to each other. The exact frequency resonance is controlled by these specified structures. A stub-matched technique along with SRR slot in the feeding is used to improve the impedance matching. The three resonating bands are precisely at 2.1 GHz, 3.3 GHz, and 4.1 GHz. The proposed antenna configuration shows a good gain response up to 6.4 dB. The antenna exhibits good impedance bandwidth up to 400 MHz and also shows the enhanced gain of more than 5 dB at all three bands. The proposed antenna is suitable for future mobile communication systems to cover up the integration of 5G with the existing LTE technology. The precise resonating frequency is highly recommended for the upcoming industry application of 5G technology.

Table 3: Comparison of the proposed SRRRPA with the state of art literature

Ref.	Resonant Frequencies (GHz)	Substrate Material	Bandwidth (GHz)	Peak Gain	Size
[3]	3.5, 5.5	FR4 with ε_r=4.4	3.400-3.800, 5.150-5.925	13 dB	$2.30\lambda_{mid} \times 1.22\lambda_{mid} \times 0.01\lambda_{mid}$
[5]	2.4, 5.2, 5.8, 26	Rogers RT5880 with ε_r = 2.2	2.450-2.495, 5.0-6.3, 23-28	4.72 dB at 5.8 GHz	$1.52\lambda_{mid} \times 1.52\lambda_{mid} \times 0.02\lambda_{mid}$
[8]	1.57, 2.45, 3.5, 5.2	FR4 with ε_r=3.5	1.575-1.665, 2.4-2.545, 3.27-3.97, 5.17-5.93	5.02 dB at 3.5 GHz	$0.6\lambda_{mid} \times 0.2\lambda_{mid} \times 0.01\lambda_{mid}$
[9]	3.3, 3.8	FR4 with ε_r=4.4	3.3-4.2, 3.3-3.8	2.5 dB at 3.5 GHz	$0.42\lambda_{mid} \times 0.18\lambda_{mid} \times 0.01\lambda_{mid}$
[10]	6, 10, 15	FR4 with ε_r=4.4	5-7, 9-10.8, 14-15	5.85 dB at 15 GHz	$0.61\lambda_{mid} \times 0.4\lambda_{mid} \times 0.1\lambda_{mid}$
[27]	0.9, 1.8, 2.4	FR4 with ε_r=4.4	0.88-0.91, 1.79-1.83, 2.39-2.42	2.12 dB at 1.8 GHz	$0.23\lambda_{mid} \times 0.08\lambda_{mid} \times 0.0.008\lambda_{mid}$
[28]	2.45, 3.5, 4.65	FR4 with ε_r=4.4	2.42-2.48, 3.3-3.7, 4.45-4.85	3.68 dB at 4.65 GHz	$0.38\lambda_{mid} \times 0.55\lambda_{mid} \times 0.01\lambda_{mid}$
[29]	1.422, 1.791, 2.467	FR4 with ε_r=4.4	1.409-1.437, 1.723-1.911, 1.858-2.045	3.276 dB at 1.422 GHz	$0.21\lambda_{mid} \times 0.16\lambda_{mid} \times 0.009\lambda_{mid}$
[30]	2.43, 3.3, 6.1	FR4 with ε_r=4.4	2.39-2.48, 3-3.7, 5-7	5 dB at 3.5 GHz	$0.40\lambda_{mid} \times 0.50\lambda_{mid} \times 0.02\lambda_{mid}$
Proposed antenna	2.1, 3.3, 4.1	Rogers RT5880 ε_r = 2.2	2-2.4, 3.3-3.6, 4-4.4	6.4 dB at 3.3 GHz	$0.46\lambda_{mid} \times 0.46\lambda_{mid} \times 0.01\lambda_{mid}$

REFERENCES

[1] R. Kumari and M. Kumar, "Design of multiband antennas for wireless communication," *International Conference on Communication Systems and Network Technologies*, Gwalior, pp. 1-6, 2013.

[2] M. Ko, H. Lee, and J. Choi, "Planar LTE/sub-6 GHz 5G MIMO antenna integrated with mmWave 5G beamforming phased array antennas for V2X applications," *IET Microwaves, Antennas & Propagation*, vol. 14, no. 11, pp. 1283-1295, 2020.

[3] Y. Li, C. Sim, Y. Luo, and G. Yang, "Multiband 10-antenna array for sub-6 GHz MIMO applications in 5G smartphones," *IEEE Access*, vol. 6, pp. 28041-28053, 2018.

[4] A. Kapoor, R. Mishra, and P. Kumar, "Analysis and design of passive spatial filter for sub-6 GHz 5G communication systems," *Journal of Computational Electronics*, 2021.

[5] Z. Khan, M. H. Memon, S. U. Rahman, M. Sajjad, F. Lin, and L. Sun, "A single-fed multiband antenna for WLAN and 5G applications," *Sensors*, vol. 20, no. 21, pp. 6332, Nov. 2020.

[6] J. Bao, Q. Huang, X. Wang, and X. Shi, "Compact multiband slot antenna for WLAN/WiMAX operations," *International Journal of Antennas and Propagation*, vol. 2014, 7 pages, 2014.

[7] B. Maity and S. K. Nayak, "Compact CPW-fed multiband F-shaped slot antenna for wireless communications," *International Conference on Wireless Communications Signal Processing and*

Networking (WiSPNET), Chennai, India, pp. 92-96, 2020.

[8] Y. F. Cao, S. W. Cheung, and T. I. Yuk, "A multiband slot antenna for GPS/WiMAX/WLAN systems," *in IEEE Transactions on Antennas and Propagation*, vol. 63, no. 3, pp. 952-958, 2015.

[9] A. Kapoor, R. Mishra, and P. Kumar, "Compact wideband printed antenna for sub-6-GHz fifth-generation applications," *International Journal on Smart Sensing and Intelligent Systems*, vol. 13, no. 1, pp. 1-10, 2020.

[10] T. O. Olawoye and P. Kumar, "A high gain microstrip patch antenna with slotted ground plane for sub-6 GHz 5G Communications," *International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems*, pp. 1-6, South Africa, 2020.

[11] A. Kapoor, R. Mishra, and P. Kumar, "A compact high gain printed antenna with frequency selective surface for 5G wideband applications," *Advanced Electromagnetics*, 2021, (In press).

[12] P. Yang, "Reconfigurable 3-D slot antenna design for 4G and sub-6G smart phones with metallic casing," *Electronics*, vol. 9, no. 2, p. 216, Jan. 2020.

[13] R. Addaci, F. Ferrero, D. Seetharamdoo, T. Le Huy, R. Staraj, and M. Berbineau, "Multiband multi-antenna system with new approach of PIFA bandwidth enhancement," *The 8th European Conference on Antennas and Propagation (EuCAP 2014*, pp. 2880-2883, 2014.

[14] M. Gallo, O. Losito, V. Dimiccoli, D. Barletta, and M. Bozzetti, "Design of an inverted F antenna by using a transmission line model," *Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP)*, Rome, pp. 635-638, 2011.

[15] R. G. Villanueva, R. L. Miranda, J. A. Tirado-Mendez, and H. Jardon Aguilar, "Ultra-wideband planar inverted-F antenna (PIFA) for mobile phone frequencies and ultra-wideband applications," *Progress In Electromagnetics Research C*, vol. 43, pp. 109-120, 2013.

[16] M. Mabaso and P. Kumar, "A dual band patch antenna for bluetooth and wireless local area network applications," *International Journal of Microwave and Optical Technology*, vol. 13, no. 5, pp. 393-400, 2018.

[17] K. Yu, Y. Li, and W. Yu, "A compact triple band antenna for bluetooth, WLAN and WiMAX applications," *Applied Computational Electromagnetics Society Journal*, vol. 32, no. 5, pp. 424-429, 2017.

[18] M. Mabaso and P. Kumar, "A microstrip patch antenna with defected ground structure for triple band wireless communications," *Journal of Communications*, vol. 14, no. 8, pp. 684-688, 2019.

[19] V. G. Raviteja, V. R. Lakshmi, and S. N. Reddy, "Multiband planar inverted-F antenna employing rectangular SRR for UMTS and WiMax/WiFi applications," *International Journal of Computer Applications*, vol. 182, no. 36, pp. 36-39, 2019.

[20] IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond, ITU, https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!!PDF-E.pdf, Feb. 2014, Accessed on June 27, 2021.

[21] White paper on Enabling 5G in India, https://www.trai.gov.in/sites/default/files/White_Paper_22022019.pdf. Feb. 2019, Accessed on June 27, 2021.

[22] R. Garg, *Microstrip Antenna Design Handbook*. Artech house, Inc., 2001.

[23] A. Azari, "A new ultra wideband fractal monopole antenna," *Applied Computational Electromagnetics Society Journal*, vol. 26, no. 4, pp. 348-352, 2011.

[24] D. M. Nashaat, H. A. Elsadek, and H. Ghali, "Single feed compact quad-band PIFA antenna for wireless communication applications," *IEEE Transactions on Antennas and Propagation*, vol. 53, no. 8, pp. 2631-2635, 2005.

[25] M. Singh, V. Marwaha, A. Thakur, H. S. Saini, and N. Kumar, "Design of a low return loss planar inverted F antenna (PIFA) for 4G & WLAN," *3rd International Conference on Signal Processing and Integrated Networks (SPIN)*, pp. 539-543, 2016.

[26] Z. Lin and H. Tun, "Design and fabrication of a planar inverted-F antenna (PIFA) for LEO satellite application," *American Journal of Electromagnetics and Applications*, vol. 8, no. 1, pp. 28-32, 2020.

[27] X. Sun, "Design of a triple-band antenna based on its current distribution," *Progress In Electromagnetics Research Letters*, vol. 90, pp. 113-119, 2020.

[28] Pragati, S. L. Tripathi, S. R. Patre, S. Singh, and S. P. Singh, "Triple-band microstrip patch antenna with improved gain," *International Conference on Emerging Trends in Electrical, Electronics and Sustainable Energy Systems (ICETEESES–16)*, pp. 106-110, 2016.

[29] R. K. Verma and D. K. Srivastava, "Design and analysis of triple band rectangular microstrip antenna loaded with notches and slots for wireless applications," *Wireless Personal Communications*, pp. 1-18, 2020.

[30] D. L. Jin, T. T. Bu, J. S. Hong, J. F. Wang, and H. Xiong, "A tri-band antenna for wireless applications using slot-type SRR," *Applied Computational Electromagnetics Society Journal*, vol. 29, no. 1, pp. 47-53, 2014.

Ranjan Mishra is working as an Associate Professor in the Department of Electrical and Electronics Engineering, UPES, Dehradun, India. He received his Ph.D. in Microstrip Antenna Design from University of Petroleum and Energy Studies, Dehradun in the year 2016. He is the author of 20 Scopus indexed refereed journal and international conference papers and 4 edited books.

Rajeev Dandotiya is an engineering graduate in Electronics. He has received Ph.D. from UPES Dehradun. He has a 13 years of experience in Research and Development in the area of sensor, wireless communication, antenna design, cyber platform, and developing communication systems for industry.

Ankush Kapoor is currently pursuing his Ph.D. degree from UPES Dehradun. He has around 8 years of experience in academics. Currently, he is appointed as an Assistant Professor in Department of ECE, JNGEC Sundernagar, India. His research interests include frequency selective surfaces, design and analysis of microstrip antennas and metamaterials.

Pradeep Kumar received his Bachelor's degree, M.Eng. and Ph.D. in Electronics and Communication Engineering in 2003, 2005 and 2009, respectively. He completed his postdoctoral studies from Autonoma University of Madrid, Spain. At present, he is working with University of KwaZulu-Natal, South Africa. His current research areas include design and analysis of microstrip antennas, antenna arrays, wireless communications etc.

Design of Cylindrical Conformal Array Antenna based on Microstrip Patch Unit

Tianming Bai, Di Jiang*, Sha Luo, and Kai Zhu

Information and Communication Engineering
University of Electronic Science and Technology of China, Chengdu, 611731, China
tianyumingfeng@163.com, *dijiang@uestc.edu.cn, luosha109@163.com, zhukai1278@163.com

Abstract ─ Based on microstrip patch antenna, a cylindrical conformal 8×8 array antenna is designed, which uses a T-shaped power divider to realize 64 feed channels. The simulation results show that the peak gain of planar array antenna can reach 24.8dB, while the peak gain of cylindrical conformal antenna decreases by 1.2dB and 1.7dB in phi=0° and phi=90° respectively. And the main beam direction deflects by 0° and 4° respectively. The measurement results show that the performance of the processed object is close to the simulation. After conformal with cylinder, the peak gain is 23.5dB, and the beam deflection is 4°, which verifies the feasibility of the designed cylindrical conformal array antenna.

Index Terms ─ Array antenna, conformal, microstrip patch.

I. INTRODUCTION

With the development of electronic information technology, the requirements for the antenna are developing towards miniaturization and integration. And the demands for the beam scanning range and gain of the antenna are also higher and higher. The emergence of planar array antenna solves the problems of limited gain and difficult wave scanning of single antenna to a certain extent [1-2]. However, with the expansion of application scope of antenna, planar array antenna is gradually difficult to adapt to the needs of complex occasions.

Conformal array antenna can keep consistent with its bearing surface, that is, it fits the surface of the carrier completely. If it is attached to the cylindrical surface, it is called cylindrical array. The beam direction of the array element can basically achieve complete coverage along the curved surface [3-4], which can save space and avoid the fast change of the inherent beam scanning performance of planar array with the increase of scanning angle. It can be applied in satellite, airborne, missile and other complex occasions [5-7].

Compared with the planar array antenna, conformal array antenna has some unique advantages in practical applications. It is suitable for highly integrated electronic design, and the conformal array antenna can effectively reduce the weight of integrated electronic system. Using conformal array antenna can reduce the air resistance of the aircraft in flight, improve its aerodynamic performance, and effectively reduce the radar cross section (RCS) of the aircraft and improve its stealth performance.

In [8], a cylindrical conformal circularly polarized (CP) series-fed microstrip array design for broadside radiation is presented. They primarily focusing of the CP major lobe of the conformal array by proper dimensioning of the aperture spacings. And the direction-of-arrival (DOA) estimation of a conformal antenna array with directive elements is studied in [9]. The simulation results prove that this conformal array achieves better DOA estimation performance than that of the planar array antenna. In addition, a cylindrical conformal transmission array has been developed [10], which has a peak measured gain of 19.6 dB and an aperture efficiency of 25.1%. In the aspect of beam scanning, [11] designed a dual-layer multibeam conformal slot array antenna, which has a scanning range of ±46°. And [12] proposed a conformal phased array antenna for unmanned aerial vehicle (UAV) radar, which can scan ±60° in E-plane. It shows a wide range of common antenna application scenario. Conformal array antenna is naturally suitable for wide-angle beam scanning and integrated design. It can provide more comprehensive field of view, better aerodynamic performance and reduce the volume and weight of aircraft as much as possible. Conformal array antenna is an important branch of antenna development in military and civil fields in the future.

Compared with other common microwave antennas, microstrip patch antenna has the characteristics of small size, light weight, easy integration, easy processing, and easy carrier conformal [13-15], etc. It has been widely apply in electronic information, wireless communication and other fields [16-18]. In this paper, a conformal array antenna is designed based on microstrip patch antenna unit.

Ku-band has low ground interference, low resistance to electromagnetic interference of other frequencies, and low receiving environment. The antenna of this frequency band is small in diameter and easy to integrate.

Submitted On: May 1, 2021
Accepted On: July 6, 2021

https://doi.org/10.47037/2021.ACES.J.360809

The design of the Ku-band antenna can effectively alleviate the shortage of spectrum resources and improve the utilization rate of spectrum. There have been many studies on Ku band patch antennas. In [19], a S-shape microstrip patch antenna that works from 15.35 GHz to 19.65 GHz is designed. In [20], a single-layer multi-band reflectarray antenna in X bands (10.8 GHz to 12.8 GHz), Ku bands (15.3 GHz to 17.3 GHz) K bands (24 GHz to 26 GHz) bands is proposed. In [21], a Ku-band patch antenna with Enhanced bandwidth is designed (14.58 GHz to 16.33 GHz).

II. UNIT DESIGN

In this paper, the array element is designed based on the microstrip rectangular patch antenna, the substrate is Rogers 5880, and the designed center frequency is 15GHz. The initial size of the patch is calculated by the theoretical formula. The width of the rectangular patch can be calculated by equation (1):

$$W = \frac{c_0}{2 f_0} \sqrt{\frac{2}{\varepsilon_r + 1}}, \tag{1}$$

where c_0 is the speed of light in free space, f_0 is the center frequency, and ε_r is the dielectric constant of the substrate.

The effective permittivity of microstrip patch antenna is obtained by equation (2):

$$\varepsilon_e = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2} \left(1 + 12 \frac{h}{W} \right)^{-\frac{1}{2}}. \tag{2}$$

Because of the edge shortening effect of rectangular radiation patch, ΔL is the length of equivalent radiation gap, which can be obtained from equation (3):

$$\Delta L = 0.412h \frac{(\varepsilon_e + 0.3)(\mathrm{W}/\mathrm{h} + 0.264)}{(\varepsilon_e - 0.258)(\mathrm{W}/h + 0.8)}. \tag{3}$$

The actual length of the radiation patch should be:

$$L = \frac{c}{2 f_0 \sqrt{\varepsilon_e}} - 2\Delta L. \tag{4}$$

After getting the initial size, the central embedded structure is used for unit feeding and the structure is shown in Fig. 1. The radiation impedance of traditional microstrip antenna usually adds an impedance converter between the antenna and the feeder to achieve smaller reflection. However, this kind of patch antenna loaded with impedance converter has complicated structure and needs to provide loading space for impedance converter, which means it is difficult to achieve miniaturization. Based on the principle of impedance loading, the central embedded feed patch antenna can realize impedance matching in a small space by opening two slots between the feeder and the antenna and loading a reactance on the feeder of the antenna.

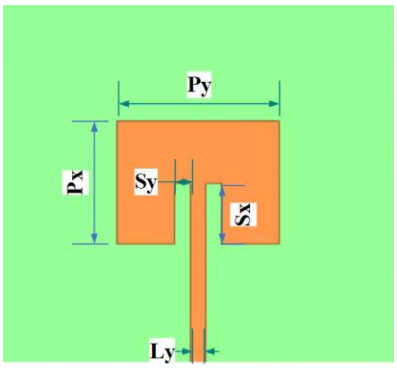


Fig. 1. Microstrip patch antenna unit.

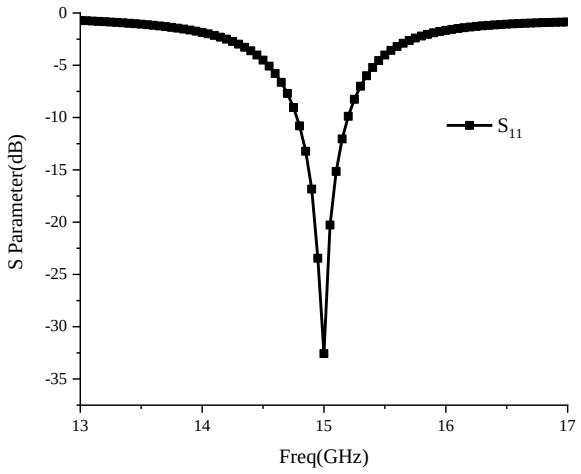
The optimized parameters are shown in Table 1, and the simulation results of S parameters and gain are shown in Fig. 2. At the center frequency of 15GHz, S_{11} is less than -30dB and the peak gain is 6.7dB.

Table 1: Design parameters and their values

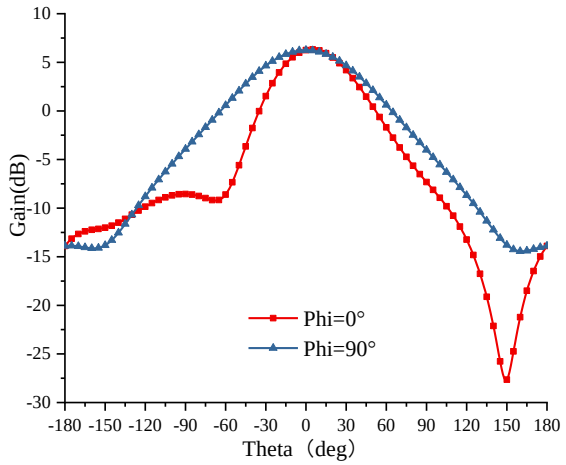
Parameters	Value (mm)
Px	6.4
Py	8
Sx	1.5
Sy	0.9
Lx	6.5
Ly	0.4
h	0.508

III. ARRAY SIMULATION

In order to reduce the mutual coupling between array elements, the array spacing is 0.7λ= 14mm. The structure model of power divider is shown in Fig. 3. The designed planar array antenna is 8×8 uniform array, and the improved T-type power divider structure is still used to realize 64 channels feeding.



(a)



(b)

Fig. 2. Unit simulation results: (a) S_{11} and (b) gain.

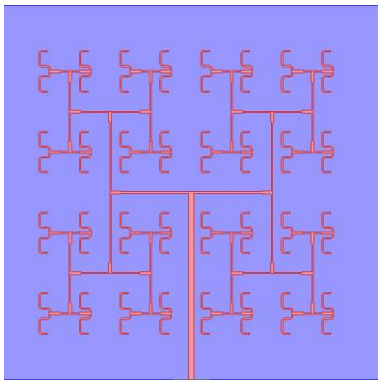


Fig. 3. 64 channels power divider.

After adding radiation patch, the structure of array antenna is shown in Fig. 4, which is 8×8 uniform array.

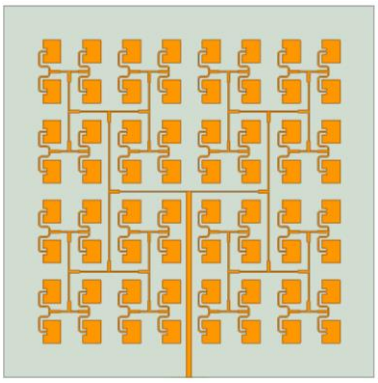
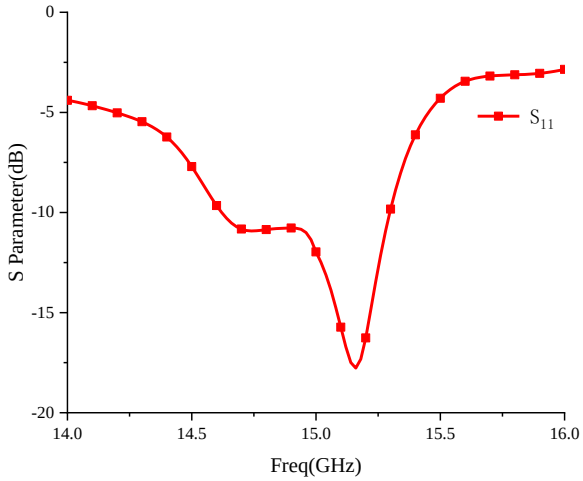
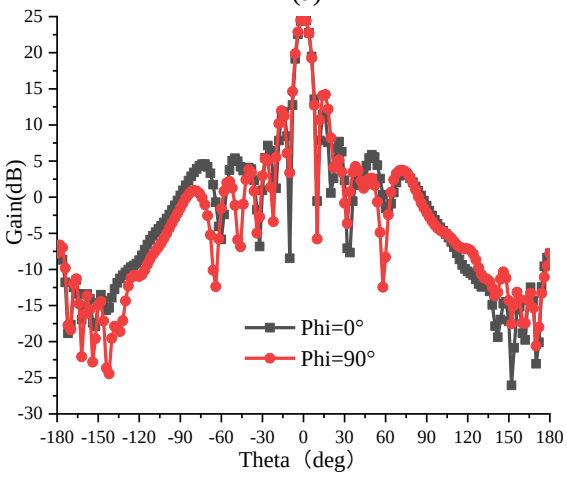


Fig. 4. Schematic diagram of planar array antenna model.

The simulation results of the array antenna gain are shown in Fig. 5. The S_{11} of the planar array antenna is -12.7dB at 15GHz, and the working bandwidth of -10dB is 640MHz. The peak gain of the planar array antenna can reach 24.8dB.



(a)



(b)

Fig. 5. Simulation results of planar array antenna: (a) S_{11} and (b) gain.

The common array structure of cylinder based on the designed array antenna is shown in Fig. 6. According to the different radius of the conformal cylinder, the gain of the antenna will be attenuated and the beam direction will be deflected.

When the radius of the conformal cylinder is 600mm (30λ), the simulation results of the conformal array antenna pattern are shown in Fig. 7. The S_{11} of the planar array antenna is -12.3dB at 15GHz, and the working bandwidth of -10dB is 2GHz (14.2GHz-16.2GHz), which is nearly three times larger than before. The peak gain of phi=0° and phi=90° is reduced to 23.6dB and 23.1dB respectively, which is 1.2dB and 1.7dB lower than that of the planar array, and the main beam direction is deflected by 0° and 4° respectively.

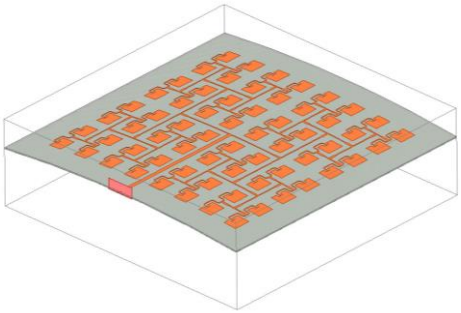


Fig. 6. Cylindrical conformal array antenna model.

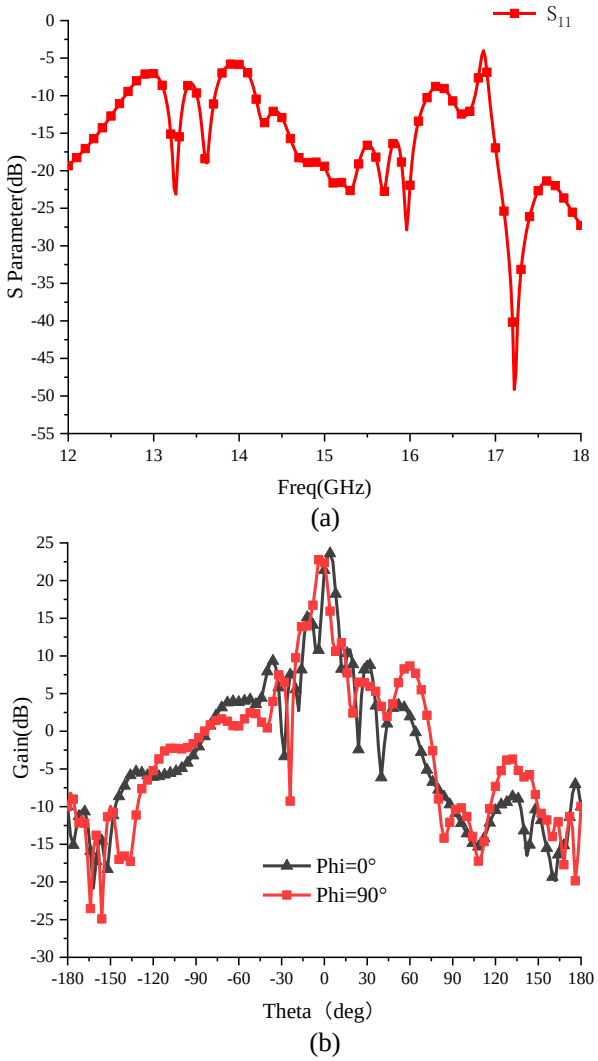


Fig. 7. Simulation results of cylindrical conformal array antenna: (a) S_{11} and (b) gain.

IV. ARRAY MEASUREMENT

Based on the designed antenna model, the prototype is fabricated. The prototype of the designed array antenna is shown in Fig. 8.

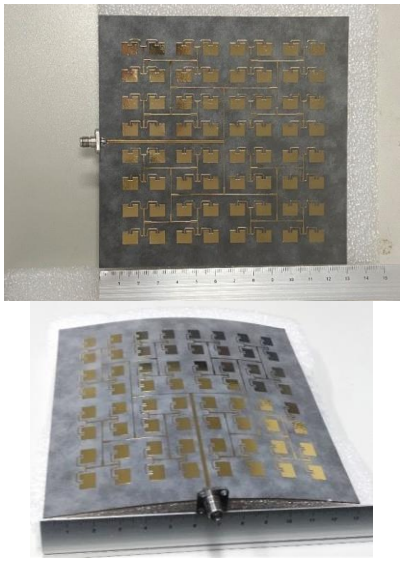


Fig. 8. Array antenna prototype.

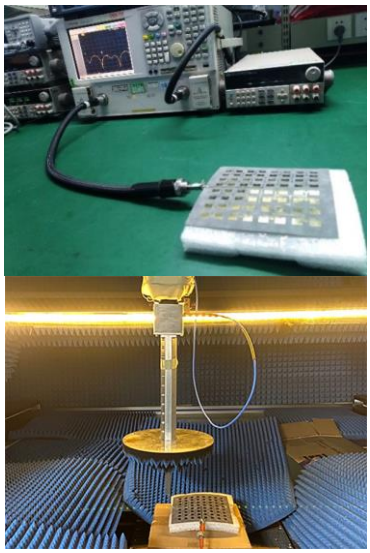
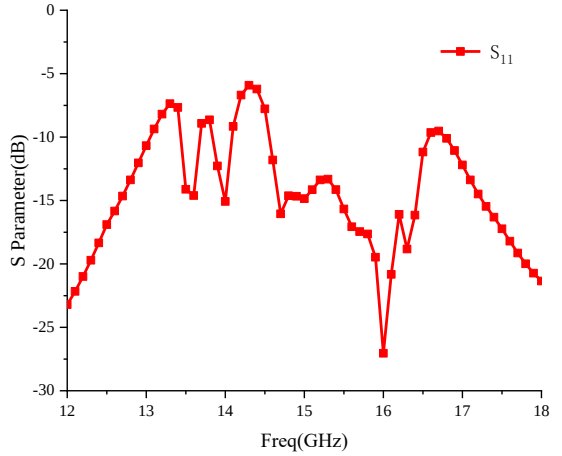
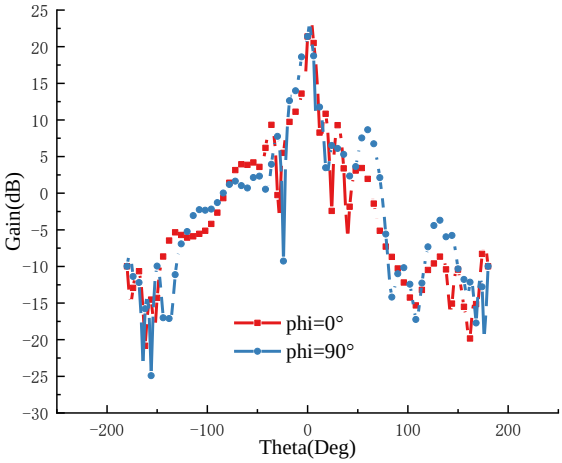


Fig. 9. The measurement of designed antenna.

As shown in Fig. 9, the fabricated antenna is tested to verify the effectiveness of the conformal array antenna design. The antenna test environment is shown in Fig. 9. The S_{11} and Gain measurement of the conformal antenna is shown in Fig. 10. The working bandwidth of -10dB is 1.9GHz(14.6GHz-16.5GHz), and the peak gain is 23.5dB, the beam direction is deflected by 4° in the measurement. The comparison results of measurement and simulation are shown in Fig. 11. The comparison of single antenna, planar antenna and conformal antenna is shown in Table 2. The comparison of proposed work with recent antennas is shown in Table 3. These results shows that the designed conformal array antenna has certain practical significance.

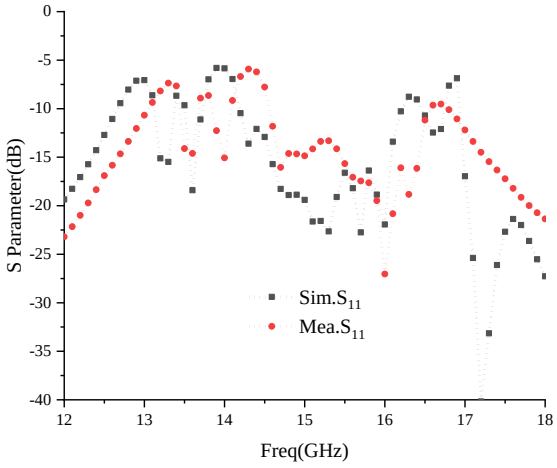


(a)

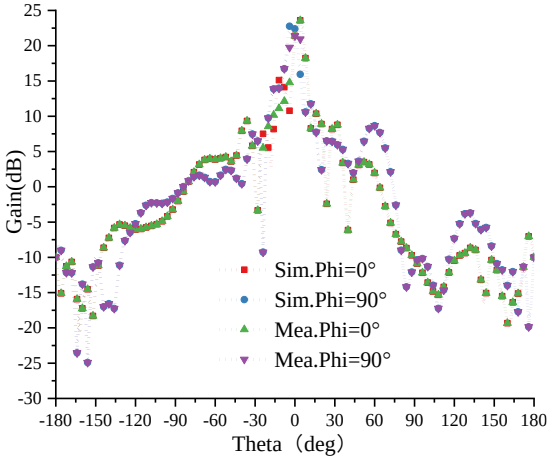


(b)

Fig. 10. Measurement results of designed antenna: (a) S_{11} and (b) gain.



(a)



(b)

Fig. 11. Comparison of simulation and measurement results: (a) S_{11} and (b) gain.

Table 2: Comparison of single, planar and conformal antenna

Antenna	-10dB Bandwidth (GHz)	Peak Gain (dB)	Array Size
Single	0.4	6.7	-
Planar	0.7	24.8	8×8
Conformal	1.9	23.5	8×8

Table 3: Comparison of proposed work with recent antennas

Ref.	Bending Radius (mm)	-10dB Bandwidth (GHz)	Peak Gain (dB)	Array Size
[8]	50	＜0.5	10	1×10
[11]	90	0.6	19.3	10×10
[14]	-	0.25	4.5	1×4
[15]	304.8	0.6	15.36	4×4
This work	600	1.9	23.5	8×8

V. CONCLUSION

In this paper, a cylindrical conformal array antenna is designed based on microstrip patch antenna unit, which can achieve a gain of 24.8dB. Based on the model, a cylindrical conformal array antenna prototype is fabricated, and the measurement is carried out. The results show that the peak gain of the conformal antenna prototype is 23.5dB and the beam deflection is 4°. The feasibility of the designed cylindrical conformal array antenna is close to the simulation results, which can

verify the rationality of the conformal antenna design.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (Grant No. 61871086); the Science and Technology Support Project of Sichuan Province (2020YFG0231, 2021YFG0017, 2021YFG0015, 2020YFG0043, 2021YFG0023), Sichuan Province Outstanding Young Science and Technology Talent Project (2021JDJQ0026); Sichuan Science and Technology Service Industry (2020ZHCG0065); the Central University Basic Research Funding (ZYGX2020ZB006, ZYGX2019J007); Foundation of GF Key Laboratory 612502200303.

REFERENCES

[1] Z. Hao, X. Liu, X. Huo, and K. Fan, "Planar high-gain circularly polarized element antenna for array applications," in *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 5, pp. 1937-1948, May 2015.

[2] A. Pal, A. Mehta, R. Lewis, and N. Clow, "A wide-band wide-angle scanning phased array with pattern reconfigurable square loop antennas," *2015 9th European Conference on Antennas and Propagation (EuCAP)*, pp. 1-4, Apr. 2015.

[3] P. Sanchez-Olivares, R. V. Haro-Baez, P. P. Sanchez-Dancausa, J. L. Masa-Campos, and J. A. Ruiz-Cruz, "Conformal array antenna fed by radial-waveguide divider for omnidirectional coverage at Ku band," *2016 10th European Conference on Antennas and Propagation (EuCAP)*, Davos, pp. 1-5, Apr. 2016.

[4] Y. Xia, B. Muneer, and Q. Zhu, "Design of a full solid angle scanning cylindrical-and-conical phased array antennas," in *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 9, pp. 4645-4655, Sept. 2017.

[5] M. G. M. Hussain, "Cylindrical array beamforming based on ultra-wideband signals," *IEEE International Radar Conference*, pp. 618-622, May 2005.

[6] J.-J. Peng, S.-W. Qu, M. Xia, and S. Yang, "Wide-scanning conformal phased array antenna for UAV radar based on polyimide gilm," in *IEEE Antennas and Wireless Propagation Letters*, vol. 19, no. 9, pp. 1581-1585, Sept. 2020.

[7] H. Zhu, X. Liang, S. Ye, R. Jin, and J. Geng, "A cylindrically conformal array with enhanced axial radiation," in *IEEE Antennas and Wireless Propagation Letters*, vol. 15, pp. 1653-1656, Jan. 2016.

[8] S. Ogurtsov and S. Koziel, "A conformal circularly polarized series-fed microstrip antenna array design," in *IEEE Transactions on Antennas and Propagation*, vol. 68, no. 2, pp. 873-881, Feb. 2020.

[9] S. Mohammadi, A. Ghani, and S. H. Sedighy, "Direction-of-arrival estimation in conformal microstrip patch array antenna," in *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 1, pp. 511-515, Jan. 2018.

[10] P.-Y. Qin, L.-Z. Song, and Y. J. Guo, "Beam steering conformal transmitarray employing ultra-thin triple-layer slot elements," in *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 8, pp. 5390-5398, Aug. 2019.

[11] Y. Liu, H. Yang, Z. Jin, F. Zhao, and J. Zhu, "A multibeam cylindrically conformal slot array antenna based on a modified Rotman lens," in *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 7, pp. 3441-3452, July 2018.

[12] J.-J. Peng, S.-W. Qu, M. Xia, and S. Yang, "Wide-scanning conformal phased array antenna for UAV radar based on polyimide film," in *IEEE Antennas and Wireless Propagation Letters*, vol. 19, no. 9, pp. 1581-1585, Sept. 2020.

[13] S. F. Jilani, M. O. Munoz, Q. H. Abbasi, and A. Alomainy, "Millimeter-wave liquid crystal polymer based conformal antenna array for 5G applications," in *IEEE Antennas and Wireless Propagation Letters*, vol. 18, no. 1, pp. 84-88, Jan. 2019.

[14] W. T. Li, Y. Q. Hei, P. M. Grubb, X. Shi, and R. T. Chen, "Inkjet printing of wideband stacked microstrip patch array antenna on ultrathin flexible substrates," in *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 8, no. 9, pp. 1695-1701, Sept. 2018.

[15] A. T. Castro and S. K. Sharma, "Inkjet-printed wideband circularly polarized microstrip patch array antenna on a PET film flexible substrate material," in *IEEE Antennas and Wireless Propagation Letters*, vol. 17, no. 1, pp. 176-179, Jan. 2018.

[16] B. D. Braaten, S. Roy, I. Irfanullah, S. Nariyal, and D. E. Anagnostou, "Phase-compensated conformal antennas for changing spherical surfaces," in *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 4, pp. 1880-1887, Apr. 2014.

[17] X. Dai, T. Zhou, and G. Cui, "Dual-band microstrip circular patch antenna with monopolar radiation pattern," in *IEEE Antennas and Wireless Propagation Letters*, vol. 15, pp. 1004-1007, Oct. 2016.

[18] Y. Cao, Y. Cai, W. Cao, B. Xi, Z. Qian, T. Wu, and L. Zhu, "Broadband and high-gain microstrip patch antenna loaded with parasitic mushroom-type structure," in *IEEE Antennas and Wireless Propagation Letters*, vol. 18, no. 7, pp. 1405-1409, July 2019.

[19] M. H. Ullah, M. T. Islam, and J. S. Mandeep, "Printed prototype of a wideband S shape microstrip

patch antenna for Ku/K band applications," *Applied Computational Electromagnetics Society Journal*, vol. 28, no. 7, pp. 307-313, Apr. 2013.

[20] Z. Hamzavi-Zarghani and Z. Atlasbaf, "A single-layer multi-band reflectarray antenna in X/Ku/K-bands," *Applied Computational Electromagnetics Society Journal*, vol. 31, no. 5, pp.555-561, May 2016.

[21] M. M. Islam, R. I. Faruque, R. Azim, M. T. Islam, and T. Alam, "Design of a Ku-band patch antenna with enhanced bandwidth," *Review of Progress in Applied Computational Electromagnetics*, pp.89-94, Mar. 2014.

Tianming Bai Graduate student of Information and Communication Engineering, University of Electronic Science and Technology of China. Main research direction is conformal antenna, reflection array antenna, phased array antenna, etc.

Di Jiang Associate Professor and Doctoral Advisor of Information and Communication Engineering, University of Electronic Science and Technology of China. Mainly engaged in broadband reconfigurable microwave devices, array antennas and other research.

Sha Luo Graduate student of Information and Communication Engineering, University of Electronic Science and Technology of China. I love learning new knowledge, like to contact new things, willing to communicate with people, have a sense of responsibility. In the graduate stage, I studied in the direction of RF and microwave. My main research topics are reconfigurable antenna and passive components. Designed passive phased array antenna, holographic antenna, phase shifter, filter power divider, microstrip log periodic antenna, waveguide slot antenna, etc.

Kai Zhu Graduate student of Information and Communication Engineering, University of Electronic Science and Technology of China. Main research direction is terahertz array antenna, reflection array antenna, liquid crystal phase shifter, etc.

Sub-6 GHz Quad-Band Reconfigurable Antenna for 5G Cognitive Radio Applications

Sivakumar Ellusamy [1] and Ramachandran Balasubramanian [2]

[1] Department of Electronics and Communication Engineering
SRM Institute of Science and Technology, Chennai, Tamilnadu 603203, India
sivakume@srmist.edu.in

[2] Department of Electronics and Communication Engineering
SRM Institute of Science and Technology, Chennai, Tamilnadu 603203, India
ramachab@srmist.edu.in

Abstract — This article describes a quad-band frequency tunable antenna for 5G applications that operates in the sub-6 GHz frequency range. On the top side, a stub-loaded square patch is printed on an inexpensive glass epoxy substrate, and on the opposite side, a C-shaped slot embedded into the partial ground plane. To achieve frequency reconfigurability with a consistent radiation pattern, the C-shaped slot and matching stub are used. The antenna is electronically frequency tunable by placing two positive intrinsic negative diodes (PIN) in the C-shaped slot and one PIN diode between the stub and feed line. The frequency is tunable between one ultra-wideband (UWB) and three communication bands based on the switching conditions of the PIN diodes. The proposed configuration is small, with a substrate dimension of 25×25×1 mm^3. The antenna was fabricated, tested and found measurements result back up the simulation; it can switch between UWB (3.31.0-6.03 GHz) and three communication modes (3.31-4.32, 3.78-4.98, 4.98-5.96 GHz). The antenna has peak gains of 1.91, 1.86, 2.0 and 2.0 dB, and radiation efficiencies of 80, 78, 83 and 86%, respectively, in the four frequency bands. The developed antenna is ideally suited for multi-functional wireless systems and cognitive radio applications since it covers the frequency bands below 6 GHz and is tunable between wide and narrow bands.

Index Terms — Narrow band, PIN diode, quad-band, reconfigurable, square patch, stub, tunable, ultra wide band.

I. INTRODUCTION

Implementation of the future communication system (5G) is probably either in sub-6GHz or in the mm-wave region. The research in 5G is growing since the demand for high data rates and better performance are ever-increasing. Coverage, adequate signal quality, low latency, and so on, are just a few of the benefits sparked interest in 5G networks [1,2]. The 5G spectrum are divided into three operating region. They are below 1 GHz (Low-Bands), between 1-6 GHz (Mid-Bands), and above 24 GHz (mm-wave). For 5G deployment, several countries are mainly concentrating on the mid-band (3-6GHz) and mm-wave (24 GHz). Higher frequencies have higher propagation losses, so the mid-band frequency was chosen over the mm-wave frequency [3]. Nowadays, wireless systems are becoming more multifunctional, necessitating the use of mobile devices that operate on various standards or applications. Reconfigurable antennas are advanced antenna that can adjust its resonant frequency, impedance bandwidth, radiation pattern, and polarization to achieve multi-functionality. The reconfigurable antenna is a single element that supports multiple wireless standards in a single unit, as opposed to traditional antennas. However, more conventional antennas are required to maintain the various wireless standards [4].

Multiple-input-multiple-output (MIMO), cognitive radio (CR), and beamforming technologies are all essential in the implementation of 5G systems. The cognitive radio is a cutting-edge technology that improves spectrum efficiency in modern communication systems. Furthermore, the cognitive radio requires a wideband antenna for scanning the whitespace within the finite frequency spectrum of interest. A reconfigurable communicating antenna to transmit/receive the information in the unoccupied frequency bands [5]. For the past few years, antenna design for CR has been a hot research subject. The researchers are interested in developing frequency tunable antennas as well as a sensing antenna. The authors of [6] discussed cognitive radio architecture, working process, as well as possible antenna configurations and design approaches. Cognitive radio are drawn to frequency reconfigurable antennas because they can dynamically change their resonant frequency to meet device requirements. They have various features

Submitted On: May 14, 2021
Accepted On: August 7, 2021
https://doi.org/10.47037/2021.ACES.J.360810

such as efficient utilization of spectrum, compact size, easy integration, operation in wide and narrowbands, and frequency selectivity characteristics to avoid co-site interference. Electrical, mechanical, material change, and optical techniques can all be used to achieve reconfigurability in an antenna. Electrical and electro-mechanical devices such as varactor diodes [7], PIN diodes [8-11], and RF MEMS are the most widely used reconfiguration techniques [12].

Several prototypes for CR applications are available in the literature [13-17]. A wide-band antenna and slot radiator were used in the design mentioned in [13]. The UWB antenna operates between 3 and 11 GHz, while a slot-radiator serves as a reconfigurable communicating antenna between 5 and 6 GHz. The mechanical rotation is used in the antenna structure in [14] to achieve frequency tunability. A wide-band antenna for monitoring the spectrum between 2 and 6 GHz and a communicating antenna for data transmission/reception within the UWB band were reported in another design in [15]. In this configuration, the height and angle of the ground plane are changed for frequency tuning. A compound tunable antenna with sensing capability was reported in [16], and it was able to detect the band from 3 to 12 GHz and communicate in five narrowbands from 1.73 to 5.23 GHz. For cognitive radio applications, [17] describes a multi-port antenna system. The antenna has four narrowband communication modes and can sense between 3.1 to 10.6 GHz. Separate antennas for sensing the channel and transmission/reception of data within the unoccupied channel are used in the literature [13-17]. These designs have complex structures, large physical size, and occupies more space. As a consequence, for cognitive radio applications, a single antenna with both sensing and communicating capabilities is highly desirable.

For CR applications, the designs reported in [18-23] use a single antenna capable of switching between wideband and narrowband. For reconfigurability, the proposed antenna in [18] uses filtering structures in the ground plane. In UWB mode, the antenna had a sensing range of 3.1 to 10.6 GHz and in narrowband mode, it covered five switchable states. The paper [19] introduced a frequency reconfigurable antenna with an integrated reconfigurable filter for CR application. The antenna can detect the spectrum from 2.63 to 3.7 GHz and function in four sub-bands. The reconfigurable antenna mentioned in the article [20] incorporates sensing and communication functionality. In sensing mode, the antenna covered 2.8-11.4 GHz and nearly covered three communication bands in the entire UWB spectrum for cognitive radio applications. A wideband radiator and a multifunctional tunable filter are included in the design in [21]. The antenna in this proposed design has all-pass, bandpass, and bandstop characteristics in the frequency range 2.5-4.2 GHz. For the sub-6 GHz band (2.5-4.2GHz), a single antenna system incorporating both MIMO and CR technologies has been reported [22]. It has sensing frequency range of 2.0-5.7GHz, with a bandpass response of 3.2-4.0GHz, and a band-notch response of 3.37-4.0GHz. The design [23] included a frequency reconfigurable MIMO mode as well as UWB MIMO mode. For reconfigurability, this design uses PIN and varactor diode. The antenna has broad sensing capability in the sensing band, ranging from 1 to 4.5 GHz, and continuous frequency tunability in the communication band, ranging from 0.9-2.6 GHz.

In the literature [13-17], two different antennas were used to provide sensing and communication functionality, resulting in increased physical size, increased space requirements, and complex structures. The designs in [18,21-23] use a single antenna for sensing and communication, but they do not cover the nR79 (4.4-5.0GHz) and sub-6 GHz range above 4.8-6.0 GHz. This paper describes a sub-6 GHz antenna system for cognitive radio. A square patch monopole antenna, slotted ground plane, and PIN diodes are included in the structure for frequency reconfigurability. In this structure, diodes are located in the ground plane and do not require vias and biasing lines for activation. The presented antenna has two modes: wideband mode and communication mode. The unique features and contribution of the proposed antenna are:

1) A single antenna performs both sensing and communication operation, so port isolation is not necessary.
2) Wide band sensing from 3.31 to 6.03 GHz, with data transmission/reception in three bands in communication mode, covering the bands 3.31-4.32, 3.78-4.98, 4.98-5.96 GHz.
3) A multifunctional filter is used to achieve sensing and communication functionality through a single antenna system.
4) Compact size of $25\times25\times1mm^3$ reduces the space requirement for CR applications.

The design aspects, parametric analyis and optimization is discussed in the Section II. Section III presents the simulated and experimental outcomes. Section IV is devoted to the conclusion.

II. METHODOLOGY

This section covers the basic antenna configuration as well as the evolution stages of quad-band frequency tunable antennas. The electromagnetic solver (HFSS) is used to model the proposed antenna. A lumped model of the diode is used as a switching device in the simulation environment to obtain frequency reconfigurability. The design has been optimized to get good impedance characteristics and far-field patterns.

A. Antenna geometry

Figure 1 presents the basic antenna configuration. It is etched on the top surface of FR-4 substrate with a thickness of 1mm, a relative permittivity of 4.4, and a dissipation factor of 0.02. The substrate has a total volume of 25×25×1mm^3. The microstrip feed line is simple and offers a planar structure. Therefore, the microstrip feed line is chosen to excite the radiator. A switchable C-shaped slot on the ground plane and a reconfigurable stub attached to the feed are used to make the antenna frequency reconfigurable. The slot length and shape determine the desired operating band and bandwidth.

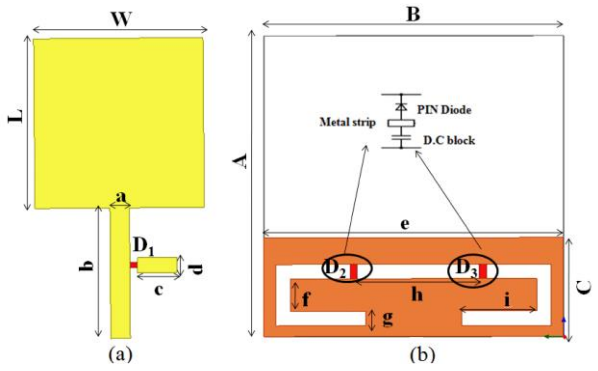


Fig. 1. The basic geometry: (a) top side and (b) back side.

B. Design and analysis

Figures 2 and 3 depict the narrow and wideband antenna design stages and their corresponding frequency responses. At first, the conventional square patch monopole UWB radiator is created, and the radiator length has selected to produce a lower resonance of 3.3 GHz using the following design equation [24,25].

The lower resonance f_L is calculated by:

$$f_L = \frac{7.2}{L + r + p} GHz \,. \tag{1}$$

The L and r are estimated as follows for the square-shaped monopole radiator. If W is the width of radiator, then,

$$L = W \,, \quad r = \frac{W}{2\pi} \,, \tag{2}$$

and p is the distance between the monopole radiator and the ground plane.

Figure 2 depicts the frequency response of the UWB radiator without a slot, which has an impedance bandwidth (S_{11} < -10 dB) of 3.3-6.57 GHz.

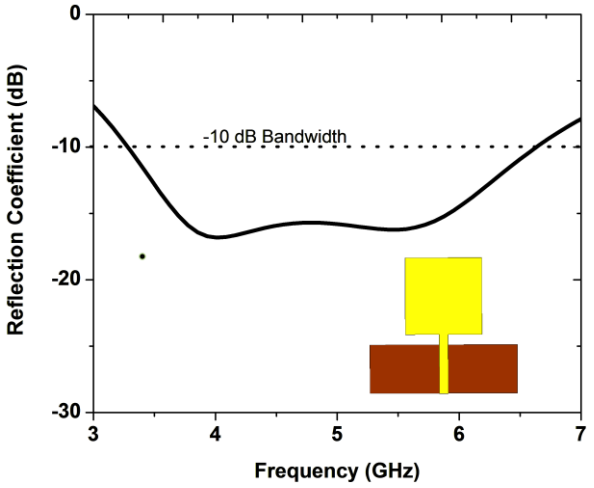
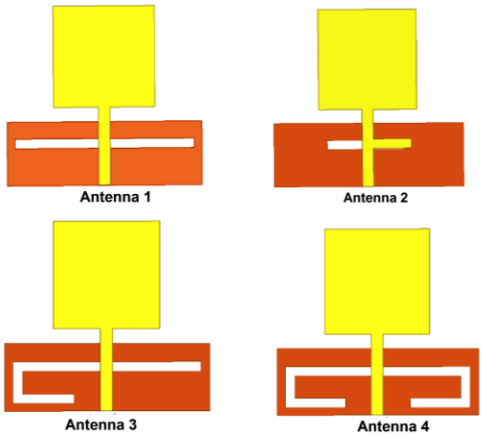


Fig. 2. UWB antenna and reflection coefficient.

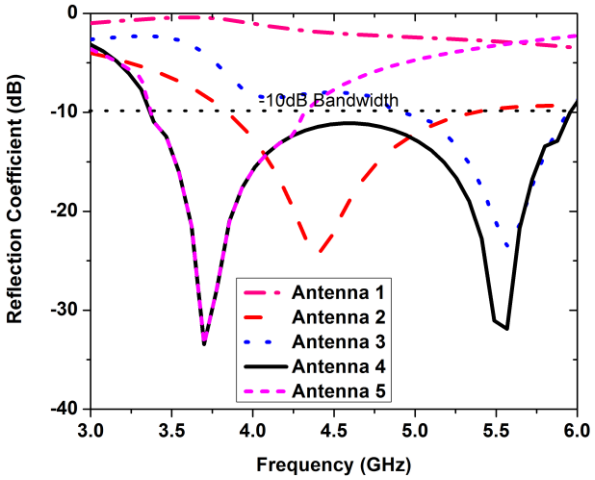
Next, a horizontal slot is placed on the ground plane of the UWB radiator. The embedded slot below the feed line function as a bandstop filter that attenuates the desired frequency band (3.3-6.57 GHz) and leaves a passband between them [Antenna 1]. The effective length of the horizontal slot is determined using the following relation given in [24]:

$$L_{slot} = \frac{c}{2 f_r \sqrt{\varepsilon_e}} \,, \tag{3}$$

where f_r is the lower resonance, c is the velocity of light and ε_e is an effective dielectric constant of the substrate material. The width of the stopband is controlled by the width of the slot. It has been set to 1.2mm, which is approximately the physical dimension of the PIN diode. The next step is shortening the horizontal slot length in Antenna 1 [Antenna 2], which results in a narrow passband of 3.85-4.86 GHz, but the matching is poor. The rectangular stub is added, provides additional reactive loading, and dramatically improves the feed line and radiator matching. Hence, Antenna 2 has strong matching and passband characteristics of 3.85-4.86 GHz. Following that, the introduction of a U-shaped slot in Antenna 2 without stub [Antenna 3] resulted in bandpass characteristics of 4.86-5.87 GHz. Antenna 3 was then changed by inserting another U-shaped slot, resulting in Antenna 4, which has a higher bandwidth due to additional capacitive effects. As a result, Antenna 4 acts as a UWB radiator with a bandwidth of 3.38-5.96 GHz. The proposed antenna is then constructed by combining narrowband (Antenna 2) and wideband (Antenna 4) antennas. It has bandpass response from 3.38-4.24GHz.



(a) Evaluation of narrow and wideband antenna



(b) Reflection coefficient

Fig. 3. Various design structures and their corresponding reflection coefficient.

The evolution of the proposed design providing capability of narrow and wideband responses is shown in Fig. 4. The combination of the horizontal slot, two U-shaped slots, and rectangular stub forms the multi-purpose filter, controlling the frequency between narrow and wideband responses.

Fig. 4. Evolution of the proposed antenna.

The parametric investigation is conducted to understand the functioning of the antenna, i.e., in Antenna 2 and Antenna 4, and to avoid measurement errors. There are over 12 design variables to study. The height of the ground (C), length (e,f, i) and width of the C-shaped slot (g) and stub (c and d), have a significant influence on frequency response.

The S-parameter curves in Fig. 5 (a) visualize the changes in reflection coefficient (S_{11} in dB) as the height of the ground plane in Antenna 4 is varied while all other parameters remain constant. The gap between the radiator and the ground plane changes as the height grows from 7.8mm to 8.4mm, causing the bandwidth to shift. When C is 8.2mm, and all other parameters are held constant, a good characteristic for S_{11} and bandwidth is obtained. Fig. 5 (b) indicates the effect in reflection coefficient for various lengths of the long horizontal slot. The graph illustrates that shortening the slot decreases antenna bandwidth while also affecting matching. The value of e=23mm was found to be the optimum compromise for improved reflection coefficient and bandwidth. The impact of changing the length of the vertical slot (f), the small horizontal slot (i), and the slot width (g) is shown in Figs. 5 (c)-(e). The antenna loses matching at the mid-frequency region as the vertical slot length decreases. Similarly, changing the length of the horizontal slot has the same effect. The antenna has an excellent S_{11} and bandwidth for parameter values of e=23mm, f=2.7mm, i=6.3mm, and g=1.2mm.

The impact of the rectangular stub on antenna performance has also been studied in Antenna 2. Figure 5 (d) depicts the reflection coefficient of the antenna for various stub lengths. When the stub length is trimmed from 5mm to 1mm, the antenna deviates significantly from the target application band, and antenna impedance matching is particularly poor for small length.The value of c equal to 4.5mm yields the best simulation results for S_{11} characteristics and bandwidth. The stub width was also varied to test the antenna's performance, as illustrated in Fig. 5 (e). As the width drops from 1.5mm to 0.8mm, the reflection coefficient declines from its maximum value, causing the antenna to deviate significantly from the desired application band. The values of d equal to 1mm, 1.1mm, and 1.2mm deliver the best simulation results. However, d equal to 1.2mm produces the best results not only in Antenna 2 but also in Antenna 5. The proposed antenna,according to the study, has the best frequency response for the optimized parameters, which are as follows: A = 25mm, B = 25mm, C = 8.2mm, L = 13mm, W = 13mm, a = 1.5mm, b = 10mm, c = 4.4mm, d = 1.2mm, e = 23mm, f = 2.7mm, g = 1.2mm, h = 12.2mm, and i = 6.3mm.

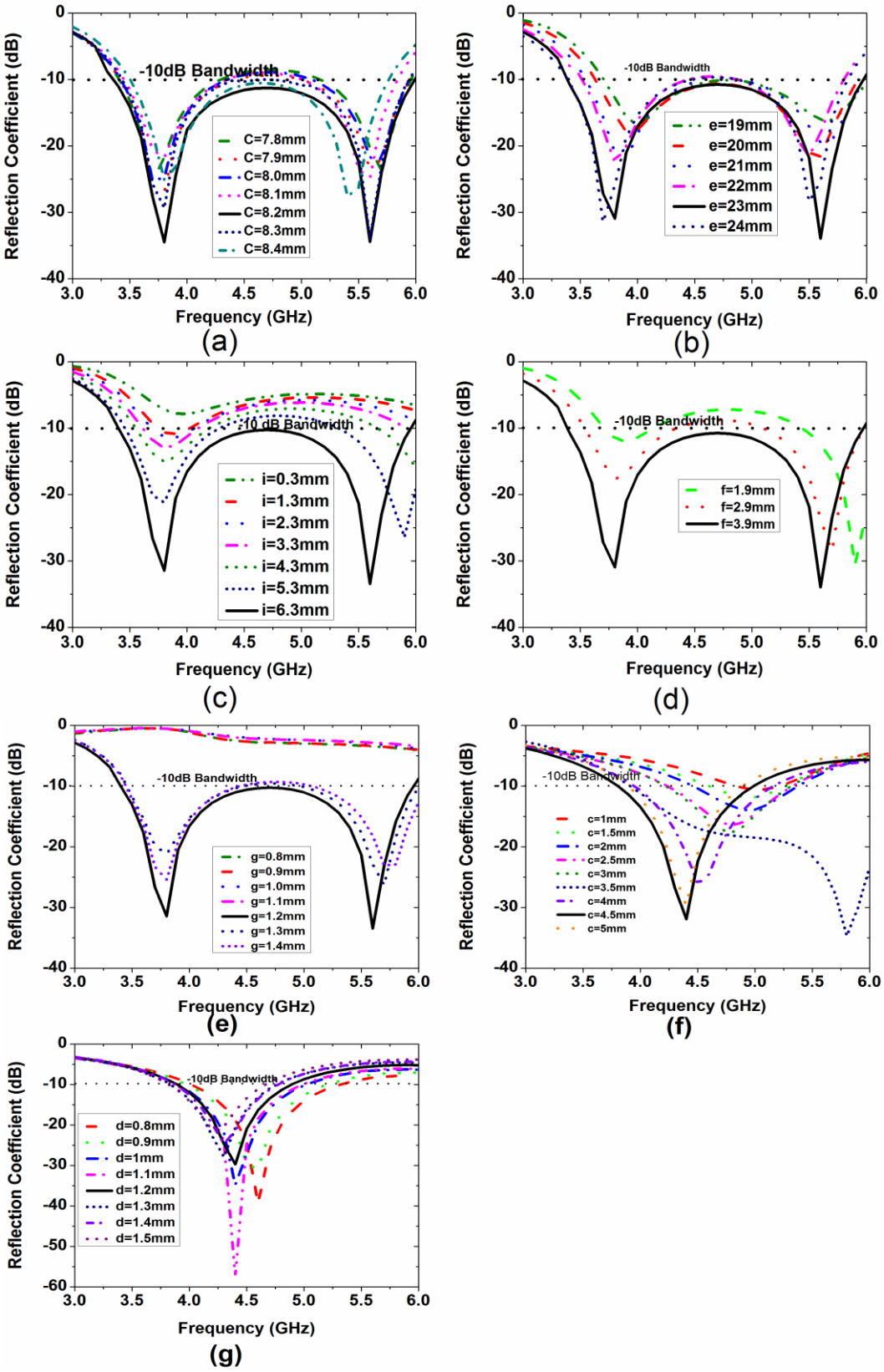


Fig. 5. Simulated reflection coefficients for (a) variations in ground plane height, (b)-(e) variations in slot parameters, and (f)-(g) variations in stub parameters.

C. Optimization

After the parametric study is performed, the genetic algorithm is used for further optimization. The height of the ground (C), length (e,f, i) and width of the C-shaped slot (g) and stub (c and d) is chosen to be the optimized parameters. Based on the optimetric study, the minimum and maximum values are taken for each parameters are 50% and 150% of the nominal values of initial design. The reflection coefficient is defined as cost function (≤-10dB) in the frequency range (3-6GHz) for Antenna 4 and (3.8-5.3GHz) for Antenna 2.

In the GA optimizer, the following setup is kept for the optimization process, maximum number of generations is 200, number of parents are 30, number of mating pools is 30, number of children is 30, number of survivors is 10, and selection pressure for next generation is 10. A uniform distribution mutation type is used in the reproduction setup with 0.1 uniform mutation probability, 0.5 individual mutation probability, 0.2 variable mutation probability, and 0.05 standard deviation [26]. The optimized values of the parameters are presented in Table 1.

Table 1: Initial and optimized values obtained from the genetic algorithm

Antenna 2			Antenna 4		
Parameters	**Initial Design Values (mm)**	**Optimized Values (mm)**	**Parameters**	**Initial Design Values (mm)**	**Optimized Values (mm)**
C	8.2	8.25	C	8.2	8.27
c	4.4	4.47	c	23	23
d	1.2	1.17	d	2.7	2.7
e	12.2	12.26	e	6.3	6.3
g	12	1.19	g	1.2	1.2

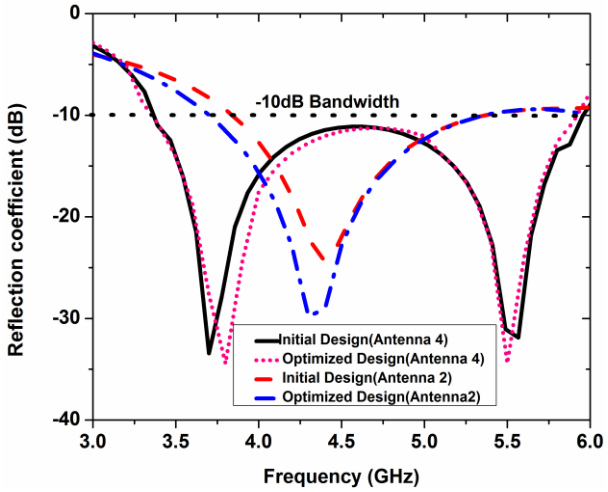


Fig. 6. Reflection coefficient curves for initial and optimized design.

Figure 6 depicts the reflection coefficient findings for the initial and optimized designs. It can be observed that the optimized Antenna 2 has a better bandwidth than the original design and a low reflection coefficient. However, in Antenna 4, a minor shift in bandwidth and reflection coefficient is seen.

D. Reconfigurable antenna

A quad-band reconfigurable antenna is designed by controlling the slot length and connecting/disconnecting the stub to the feed line using Antenna 5 [proposed design] in Fig. 4. To obtain multi-mode characteristics, two switches D_2 and D_3 are connected in the slots, and D_1 is attached between the feed line and the stub (Proposed deisgn in Fig. 1). The various structures [Antenna 1-5] depicted in Fig. 3 are determined by the placement of the switches D_2 and D_3 in the slot at the appropriate positions. This design uses RF PIN diodes as switching element. Figures 7 (a)-(c) depicts the analogous lumped RLC boundary conditionof the PIN diode describing the ON and OFF states. In the ON condition, the diode has an ohmic resistance (R_s) of 2.1Ω and an inductance (L_s) of 0.6nH. In the OFF state, it has a series inductance (L_s) of 0.6nH and is connected to a parallel combination of resistance (R_p) and capacitance (C_t) of 3 KΩ and 0.17pF. However, since the lumped model provides short and open circuit action in the ON and OFF states, we did not consider the inductance and capacitance values in simulation. As a result, in simulation, the RF PIN diode is merely represented as lumped resistor. In the ON state, a lumped model with low resistance (2.1 Ω) exhibits short circuit behavior, allowing current to flow on the antenna structure. The lumped model with high resistance (3 KΩ), on the other hand, exhibits an open circuit behavior in the OFF state, preventing current flow in the antenna structure.

Figure 8 depicts the simulated frequency response of the proposed frequency-reconfigurable antenna. The antenna works in four different operating bands depending upon the configuration of the switches. If all of the switches are OFF, the antenna works in UWB mode; otherwise, it works in communication mode. In-state 1 (all diodes are ON), the antenna works at a single narrowband with an impedance bandwidth of

3.85-5.32 GHz. The antenna also operates at a single narrowband in state 2 (D1 ON, D2 OFF, D3 OFF) and state 3 (D1 OFF, D2 ON, D3 OFF). It has an impedance bandwidth of 3.38-4.24 GHz and 4.86-5.87 GHz, respectively. Finally, in state 4 (all diodes are OFF), the antenna works at a wideband covering 3.38-5.96 GHz.

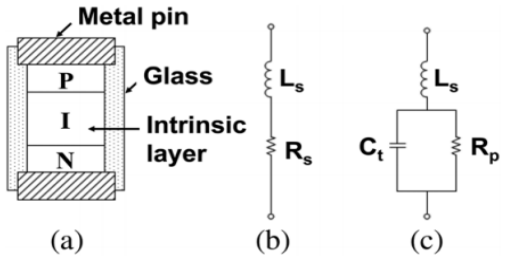


Fig. 7. PIN diode and its lumped model: (a) cross-section of diode, (b) forward bias, and (c) reverse bias.

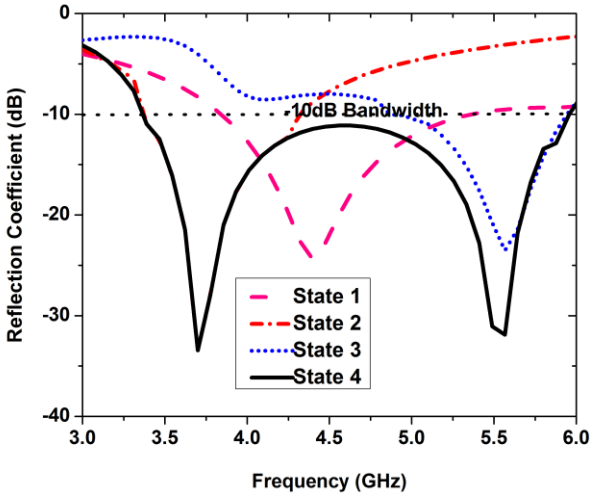


Fig. 8. Simulated reflection coefficient for different operating states.

Figure 9 depicts the surface current distribution in the slot for different switching states in order to better understand multi-band operation. The maximum current distribution aids in determining the slot length/part that contributes to each frequency band. Figure 9 (a) shows the maximum current density in state 1 at the middle of the slot and rectangular stub, As a consequence, the antenna opreates at 4.3 GHz, which covers the operating frequency range of 3.85 to 5.32 GHz.

The physical length of the slot, or the effective electrical length of the current distribution along the slot, is determined to be 0.25 λg, which corresponds to the resonant frequency 4.3 GHz. Where λg is the guide wavelength. Figures 9 (b)-(c) depicts the current distribution for states 2 and 3. In state 2, as visualized in Fig. 9 (b), the current is distributed throughout the slot and dominates in the rectangular stub. The antenna resonates at 3.7 GHz covers the frequency band 3.38-4.24 GHz with electrical slot lengths of 1.13 λg. In state 3, however, the current distribution only dominates in half of the total slot length, as seen in Fig. 9 (c). As a result, the antenna resonates at 5.5 GHz with a slot length of 0.5λg, covering the frequency range of 4.86-5.87 GHz. Figures 9 (d)-(e) shows the current prevalent maximum across the entire slot and feed line, with no distribution in the rectangular stub. With a slot length of 1.13 λg, the antenna resonates at 3.7 and 5.5 GHz, covering the frequency range of 3.38-5.98 GHz. According to the analysis, a variation in the electrical length of the current distribution causes frequency reconfiguration/multiband operation in the antenna structure.

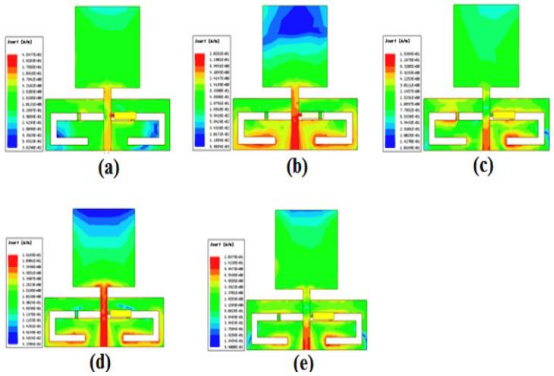


Fig. 9. The surface current distribution in various operational conditions: (a) state1 at 4.3 GHz, (b) state 2 at 3.7 GHz, (c) state 3 at 5.5 GHz, and (d)-(e) state 4 at 3.7 and 5.5 GHz.

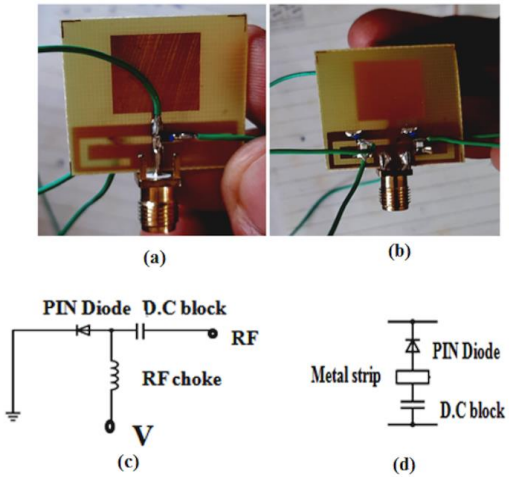


Fig. 10. Prototype and biasing structure (a) top side, (b) bottom side, (c) biasing circuit, and (d) biasing structure.

Table 2: Various operating states and their corresponding frequency band

Diodes	D1	D2	D3	Mode	Freq. Bands (GHz) Simulated/Measured
State 1	ON	ON	ON	Communication	3.85-5.32 / 3.78-4.98
State 2	ON	OFF	OFF	Communication	3.38-4.24 / 3.31-4.32
State 3	OFF	ON	ON	Communication	4.86-5.87 / 4.98-5.96
State 4	OFF	OFF	OFF	Sensing	3.38-5.96 / 3.31-6.03

III. FABRICATION AND RESULTS

The antenna is simulated and modeled using EM solver-High frequency structure simulator. Figures 10 (a)-(d) illustrates the antenna prototype with biasing circuits. After placing the PIN diodes in the proper locations, the proposed antenna is tested, and measured. The antenna is powered by a 50Ω SMA connector, and the frequency is tuned by controlling the slot length with an Infineon BAR 64-02 diode. The antenna structure has been modified to realize the proposed design with real diodes, as shown in Fig. 10 (d). A metal strip measuring 2.0×0.6 mm^2 is inserted into the slot to supply a dc voltage to the PIN diodes D_2 and D_3. A 100pF dc blocking capacitor is attached to the ground plane to provide an RF link to the diode while retaining dc isolation from the RF signal. To bias the diode, a battery applies a 1.5 V dc voltage to the metal strip. The d.c supply is directly given to diode D_1. The matching stub serves as an RF choke, preventing RF signals from entering the d.c supply.

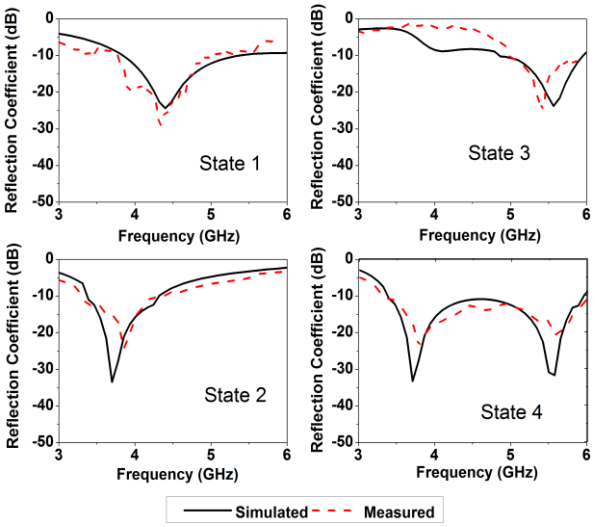


Fig. 11. Reflection coefficient for different operating states.

A. Reflection coefficient

The frequency response of the presented antenna is determined using a vector network analyzer. Figure 11 shows the input response of the presented design in simulation and measurements for various operating conditions of the diode. The designed antenna works in either the ultra wide-band or one of the three communication modes, based on the biasing conditions of the diodes. Table 2 summarizes the antenna switching conditions and their operating modes. In states 1, 2, and 3, the proposed antenna has narrow passband characteristics, while state 4 has wide-band characteristics. The simulated impedance bandwidth in state 1 is 3.85-5.32 GHz, while the measured value is 3.78-4.98 GHz. The antenna impedance bandwidth is 3.38-4.24 GHz in simulation, and 3.31-4.32 GHz in measurements for state 2. Meanwhile, the antenna operating bandwidth in state 3 is 4.86-5.87 GHz in simulation and 4.98-5.96 GHz in measurements. In simulations and measurements, the antenna in state 4 covered the operating bands 3.38-5.96 and 3.31-6.03 GHz, respectively. Across all operational bands, the proposed antenna exhibits excellent impedance matching. The simulated and observed values varied marginally due to diode mapping, prototyping tolerance levels, and connection losses.

B. Far field analysis

The far-field analysis is crucial and investigated for evaluating the proposed antenna performance using an electromagnetic simulator and anechoic chamber. Figure 12 depicts the prototyped antenna radiation pattern in all states. In the yz-plane, the antenna produces omni-directional patterns, while in the xz-plane, it produces bi-directional patterns. Figures 12 (a)-(b) shows the antenna pattern in state 1 at 4.3 GHz, which has a simulated gain of 2.1 dB but a measured gain of 1.91 dB. Figures 12 (c)-(d) depict the measured radiation pattern in state 2 at 3.7 GHz.

In the simulation, the antenna has a gain of 2.3 dB, while in measurements, it has a gain of 1.86 dB. Similarly, in states 3 and 4, the pattern in Figs. 12 (e)-(h) is obtained at 5.5 and 3.7 GHz. The simulated and observed gains are 2.2/2.0 and 2.3/2.0 dB, respectively. In the simulation, the antenna has radiation efficiencies of 87.1, 86.4, 88.9, and 92.1 percent, and in measurements, it has radiation efficiencies of 80, 78, 83, and 86 percent. The antenna cross-polarization is < -15 dB in all states. Table 3 presents the simulated and observed values of the presented antenna in all four operating states.

Table 3: Operating states and performance

States	Freq. Bands (GHz) Simulated/Measured	Antenna Gain (dB) Simulated/Measured	Antenna Radiation Efficiency (%) Simulated/Measured
1	3.85-5.32 / 3.78-4.98	2.1 / 1.91	87.1 / 80
2	3.38-4.24 / 3.31-4.32	2.3 / 1.86	86.4 / 78
3	4.86-5.87 / 4.98-5.96	2.2 / 2.0	88.9 / 83
4	3.38-5.96 / 3.31-6.03	2.3 / 2.0	92.1 / 86

Table 4: Performance comparison with other works

Ref.	Size (mm^2)	Sensing Band (GHz)	Communicating Band (GHz)	Is same Antenna have both Sensing and Communication Capabilities?	Radiation Efficiency (%)
[9]	50×46	-	2.2-4.75(Six Bands)	No	-
[15]	70×70	2-6	2.4, 2.75, 4.2	No	>70
[19]	68×51	2.63-3.7	2.63-3.7	Yes	-
[21]	60×50	2-5.7	3.21-4.0	Yes	>75
[22]	37.5×37.5	2.3-4.5	3.15-4.15, 2.65-3.85	Yes	>70
[23]	27×56	1-4.5	0.9-2.6(continuous)	Yes	> 70
This work	25×25	3.31-6.03	3.78-4.98, 3.31-4.32 4.98-5.96	Yes	>78

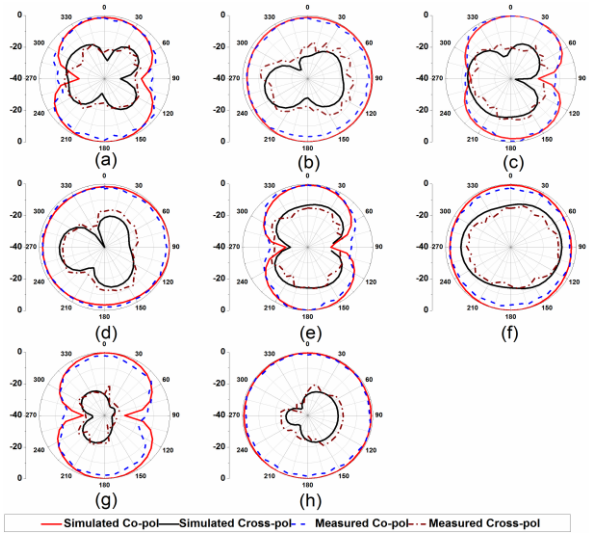


Fig. 12. Normalized gain pattern. State 1 (a) E-plane at 4.3 GHz; (b) H-plane at 4.3 GHz; State 2 (c) E-plane at 3.7 GHz; (d) H-plane at 3.7 GHz; State3 (e) E-plane at 5.5 GHz; (f) H-plane at 5.5 GHz; State 4 (g) E-plane at 3.7 GHz; (h) H-plane at 3.7 GHz.

IV. COMPARISON OF PERFORMANCE WITH PREVIOUS WORKS

Table 4 compares the proposed antenna to other antenna designs for similar applications that have been published. In contrast to the designs in [9,15,19,21,22, and 23], the proposed design is more compact. In addition, the antenna has radiation efficiency over 78 percent in all operating bands. Finally, this design operates in the sub-6 GHz band, covering almost the entire lower band (3.1-3.85 GHz), nR77 (3.3-4.2 GHz), and nR79 (4.4-5.0 GHz) in the 5G spectrum, as well as the entire higher band (4.8-6.0 GHz). In addition, the proposed antenna supports a variety of wireless communication systems, including WiMAX (3.3-3.8, 5.15-5.85 GHz), WLAN (5.1-5.3 GHz), and UWB (3.0-6.0 GHz).

V. CONCLUSION

This article demonstrated a quad-band frequency tunable antenna for 5G sub-6GHz applications. A slotted ground structure and a reconfigurable matching stub loaded to the feed line were used to achieve frequency reconfigurability with a consistent radiation pattern. Three PIN diodes were used to electronically tune the antenna. The antenna covered 3.31-6.03 GHz in sensing mode, and three bands in communication mode: 3.31-4.32, 3.78-4.98, and 4.98-5.96 GHz. The designed antenna has compact size, measuring 25×25×1 mm^3. The antenna has an excellent cross polar level, gain, and efficiency. The prototyped antenna is capable of detecting and communicating in the sub-6 GHz band. As a result, the design shown here could be used in cognitive radio and future communication systems.

ACKNOWLEDGMENT

We sincerely thanks CADD LINE systems for the fabrication of antenna and SSN engineering college for testing.

REFERENCES

[1] Spectrum for 4G and 5G, *Qualcomm Technol. Inc*., San Diego, CA, USA, Dec. 2017.

[2] IEEE: IEEE 5G and Beyond Technology Roadmap White Paper, October, 2017.

[3] International Telecommunication Union, ITU-R: Technical feasibility of IMT in bands above 6 GHz, July 2015.

[4] X. Liu, X. Yang, and F. Kong, "A frequency-reconfigurable monopole antenna with switchable stubbed ground structure," *Radio Engineering*, vol. 24, pp. 449-454, 2015.

[5] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206-1232, 2015.

[6] P. Gardner, M. R. Hamid, P. S. Hall, J. Kelly, F. Ghanem, and E. Ebrahimi, "Reconfigurable antennas for cognitive radio: Requirements and potential design approaches," *Institution of Engineering and Technology Seminar on Wideband, Multiband Antennas and Arrays for Defense or Civil Applications*, pp. 89-94, Mar. 2008.

[7] L. Ge and K. Luk, "Frequency-reconfigurable low-profile circular monopolar patch antenna," *IEEE Trans. Antennas Propag*., vol. 62, no. 7, pp. 3443-3449,2014.

[8] G. Chen, X. Yang, and Y. Wang, "Dual-band frequency-reconfigurable folded slot antenna for wireless communications," *IEEE Antennas Wirel. Propagation Lett.*, vol. 11, pp. 1386-1389, 2012.

[9] H. A. Majid, M. K. A. Rahim, M. R. Hamid, and M. F. Ismail, "A compact frequency-reconfigurable narrowband microstrip slot antenna," *IEEE Antennas Wirel. Propag. Lett.*, vol. 11, pp. 616-619, 2012.

[10] V. Rajeshkumar and S. Raghavan, "A compact frequency reconfigurable split ring monopole antenna for WLAN/WAVE applications," *Applied Computational Electromagnetics Society Journal*, vol. 30, no. 3, pp. 338-344, 2015.

[11] S. Chilukuri, K. Dahal, A. Lokam, and W. Chen "A CPW fed T-shaped frequency reconfigurable antenna for multi radio applications," *Applied Computational Electromagnetics Society Journal*, vol. 33, pp. 1276-1285, 2018.

[12] H. Rajagopalan, J. M. Kovitz, and Y. Rahmat-Samii, "MEMS reconfigurable optimized E-shaped patch antenna design for cognitive radio," *IEEE Trans. Antennas Propag.*, vol. 62, no. 3, pp. 1056-1064, 2014.

[13] E. Erfani, J. Nourinia, C. Ghobadi, M. Niroo-Jazi, and T. A. Denidni, "Design and implementation of an integrated UWB/Reconfigurable-slot antenna for cognitive radio applications," *IEEE Antennas Wirel. Propag. Lett.*, vol. 11, pp. 77-80, 2011.

[14] Y. Tawk and C. G. Christodoulou, "A new reconfigurable antenna design for cognitive radio," *IEEE Antennas Wirel. Propag. Lett.*, vol. 8, pp. 1378-1381, 2009.

[15] J. Costantine, Y. Tawk, J. Woodland, N. Flaum, and C. Christodoulou, "Reconfigurable antenna system with a movable ground plane for cognitive radio," *Microw. Antennas Propag. IET*, vol. 8, pp. 858-863, 2013.

[16] S. S. Tripathi and C. C, "A novel reconfigurable antenna with separate sensing mechanism for CR system," *Prog. Electromagn. Res. C*, vol. 72, pp. 187-196, 2017.

[17] A. Nella and A. S. Gandhi, "A five-port integrated UWB and narrowband antennas system design for CR applications," *IEEE Trans. Antennas Propag.*, vol. 66, no. 4, pp. 1669-1676, 2018.

[18] H. Boudaghi, M. Azarmanesh, and M. Mehranpour, "A frequency-reconfigurable monopole antenna using switchable slotted ground structure," *IEEE Antennas Wirel. Propag. Lett.*, vol. 11, pp. 655-658, 2012.

[19] A. Mansoul, F. Ghanem, M. R. Hamid, and M. Trabelsi, "A selective frequency-reconfigurable antenna for cognitive radio applications," *IEEE Antennas Wirel. Propag. Lett.*, vol. 13, pp. 515-518, 2014.

[20] G. Srivastava, A. Mohan, and A. Chakrabarty, "A compact reconfigurable UWB slot antenna for cognitive radio applications," *IEEE Antennas Wirel. Propag. Lett.*, vol. 16, pp. 1139-1142, 2016.

[21] S. R. Thummaluru, M. Ameen, and R. K. Chaudhary, "Four-port MIMO cognitive radio system for midband 5G applications," *IEEE Trans. Antennas Propag.*, vol .67, no. 8, pp. 5634-5645, 2019.

[22] T. Alam, S. R. Thummaluru, and R. K. Chaudhary, "Integration of MIMO and cognitive radio for sub-6 GHz 5G applications," *IEEE Antennas Wirel. Propag. Lett.*, vol. 18, no. 10, pp. 2021-2025, 2019.

[23] X. Zhao, S. Riaz, and S. Geng, "A reconfigurable MIMO/UWB MIMO antenna for cognitive radio applications," *IEEE Access*, vol. 7, pp. 46739-46747, 2019.

[24] K. G. Thomas and M. Sreenivasan, "A simple ultrawideband planar rectangular printed antenna with band dispensation," *IEEE Trans. Antennas Propag.*, vol. 58, no. 1, pp. 27-34, 2009.

[25] K. P. Ray, "Design aspects of printed monopole antennas for ultra-wide band applications," *International Journal of Antennas and Propagation*, vol. 2008, 8 pages, 2008.

[26] D. Kapoor, V. Sangwan, C. M. Tan, V. Paliwal, and N. Tanwar, "Optimization of a T-shaped MIMO antenna for reduction of EMI," *Applied Sciences*, vol. 10, no. 9, p. 3117, 2020.

E. Sivakumar received his M.Tech in Applied Electronics in 2006 from Dr M.G.R Educational Research Institute of Science and Technology. He is pursuing Ph.D. in SRM Institute of Science and Technology, India. His research interest includes antenna design, reconfigurable antennas, RF and microwave circuit design.

B. Ramachandran received his bachelor's degree in Electronics and Communication Engineering from Thiagarajar College of Engineering, Madurai, in 1990 and master's degree in Satellite Communications from National Institute of Technology, Trichy, in 1992. He obtained his Ph.D. degree in the area of Wireless Mobile Networks from Anna University, Chennai, in 2009. At present, he is working as Professor at SRM University in Chennai. He authored a textbook on digital signal processing. His teaching and research interests include digital communication, wireless networks, UWB antenna design, network security, and mobile computing. He has published 37 research papers in national and international conferences and journals. He was awarded IETE-SK Mitra Memorial Best Research Paper Award in 2009. He is a member of ISTE, and fellow of IE(I) and IETE.

Size-Reduced Equilateral Triangular Metamaterial Patch Antenna Designed for Mobile Communications

Guoxiang Dai [1], Xiaofei Xu [1,2,*], and Xiao Deng [1]

[1] School of Communication and Information Engineering
Shanghai University, Shanghai, 200444, China

[2] Key Laboratory of Specialty Fiber Optics and Optical Access Networks
Shanghai Institute for Advanced Communication and Data Science
Shanghai University, Shanghai 200444, China
*xfxu@shu.edu.cn

Abstract ─ A size-reduced equilateral triangular metamaterial patch antenna (ETMPA) is proposed for the 5G mobile communications. The new ETMPA is formed from a conventional equilateral triangular patch antenna (ETPA) by additionally loading triangular-shaped mushroom metamaterials. One ETMPA is experimentally demonstrated. It is shown to resonate at 3.488GHz. The side length is only $0.483\lambda_g$, which is much smaller than that for a conventional ETPA with a length of $0.66\lambda_g$. Despite its compactness, the ETMPA has an acceptable bandwidth of 2.1% and antenna gain of 6.3dBi in measurement. These performances make the compact ETMPA proposing to be used in the wireless communications at 3.5GHz.

Index Terms ─ Equilateral triangular patch antenna, metamaterials, size-reduced.

I. INTRODUCTION

With the rapid development of mobile wireless communications, there have been increasing demands for compact patch antennas. Among various choices of different patch shapes, it is interesting to note that an equilateral triangle patch antenna (ETPA) [1-6] has a relatively smaller patch than its square and circular counterparts, although the latter two kinds of antennas are better known [1]. The side length of ETPA is about $0.66\lambda_g$ (λ_g is the guided wavelength of substrate at the resonant frequency) [1], assuming it works at the lowest fundamental mode. The associated patch area is nearly $0.19\lambda_g^2$. However for a square patch antenna, the side length is nearly $0.5\lambda_g$ [1] while patch area is about $0.25\lambda_g^2$. The case for the circular patch antenna is similar. The diameter for the circular patch is about $0.58\lambda_g$. The patch area is as large as $0.26\lambda_g^2$ [1]. Both of them are assumed to operate at their fundamental modes respectively.

To further compact the ETPA for the potential 5G mobile communications, a new approach is proposed in this work by employing the metamaterial technology [7-14]. Metamaterials are man-made subwavelength structures that exhibit an effective permittivity or permeability, and thus equivalent to some natural materials. Inspired by the square-shaped mushroom metamaterials designed for a compact rectangular patch antenna [14], triangle-shaped mushroom metamaterials (TSMMs) are used in this contribution to adapt for an ETPA. A new equilateral triangle metamaterial patch antenna (ETMPA) is hence formed by filling the TSMMs into a conventional ETPA. The TSMMs are physically made from low permittivity substrates, but work as some effective materials with an enhanced high dielectric constant, similar with their square-shaped equivalents [14].

A practical ETMPA is designed to operate at 3.5GHz. The antenna characteristics are studied in both numerical simulations and experiments. Results are observed in good agreement. The practical ETMPA has a side length of 24mm. However it resonates at 3.488GHz in experiments. Considering the TSMMs are made from substrates with low dielectric constant of 3, the electrical length for the ETMPA is only $0.483\lambda_g$, by normalizing 24mm to the λ_g (49.67mm) at 3.488GHz. The patch area is about $0.1\lambda_g^2$, which is much smaller than that for a conventional ETPA. The new ETMPA has a -10dB bandwidth of 73MHz (2.1%), ranging from 3.451 to 3.524GHz in measurement. The peak measured antenna gain is 6.3dBi. Within the working bandwidth, the antenna gain is observed not less than 5.8dBi. Considering the compact patch size, the bandwidth and antenna gain for the ETMPA are however not much sacrificed. Both are acceptable that make the new antenna promising in the applications of 5G mobile communication at the 3.5GHz band.

Submitted On: June 20, 2021
Accepted On: July 21, 2021

https://doi.org/10.47037/2021.ACES.J.360811

II. ANTENNA DESIGN

The configuration of the proposed compact ETMPA is shown in Fig. 1. It is shown with numerous metallic triangle-shaped mushroom structures embedded in the substrate. These triangle-shaped structures are designed to fit for the configuration of an ETPA. The substrate of ETMPA is divided into two sub-layers by the mushroom structures. The upper layer is assumed with thickness h_1, dielectric constant ε_{r1}, and loss tangent $tan\delta_1$, while the lower layer is with h_2, ε_{r2}, and $tan\delta_2$. Both can be made from low index laminates that are commonly found. The mushroom structures are designed in the lower substrate layer. Each mushroom element has a small triangle cap mounted on the top of lower substrate, which is connected to the ground plane by a conducting via located in the center of element. A total number of m^2 (m is an integer) triangle-shaped mushroom elements are included beneath the patch. In Fig. 1, an example of $m = 4$ is shown. Considering that the ETMPA is with a side length of L, and the side length for the mushroom element is p, we have $L = m \times p$. The triangle mushroom cap is assumed with side length of a, which is very close to (but still less than) p. The outer diameter for the perforated vias is d. The length for the ground plane is G. The antenna is fed by a 50Ω coaxial probe. In order to prevent the inner pin of probe from overlapping those metallic mushrooms, two mushroom elements around the feeding position are designed defected by removing the conducting vias.

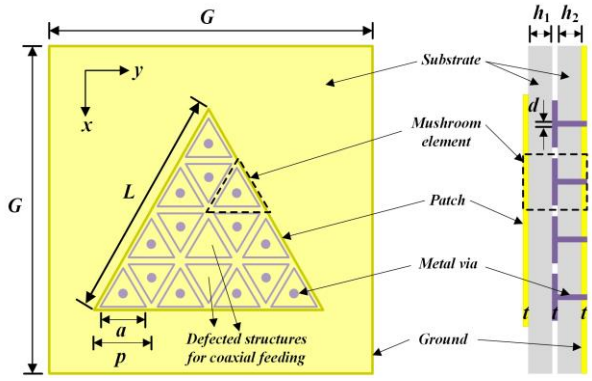


Fig. 1. Configuration of the ETMPA.

These mushroom structures can be regarded as effective metamaterials if p is much smaller than wavelength [7-8]. The effective permittivity ε_{reff} is irrelevant to the shapes of the structure since it originates from the enhanced capacitive effect in the substrate [14]. The underlying mechanism is much different from the electromagnetic bandgap or high impedance surface applications in [15-18], in which the period p is generally at the order of 0.25 to 0.5 wavelength. For the composite TSMMs, the ε_{reff} is approximate to that for those rectangular-shaped mushroom metamaterials [14] as:

$$\varepsilon_{reff} \approx \varepsilon_{r1} \frac{h_1 + h_2}{h_1}. \quad (1)$$

The resonant frequency f_0 of fundamental TM_{10} mode can be hence estimated [1-6] using:

$$f_0 \approx \frac{2c}{3L\sqrt{\varepsilon_{reff}}} \approx \frac{2c}{3L\sqrt{\varepsilon_{r1}}} \sqrt{\frac{h_1}{h_1 + h_2}}, \quad (2)$$

where c is the light speed in free space.

The side length L for the ETMPA normalized to λ_g is:

$$\frac{L}{\lambda_g} \approx \frac{2}{3} \sqrt{\frac{h_1}{h_1 + h_2}}, \quad (3)$$

in which an additional reduction ratio $\sqrt{h_1/(h_1+h_2)}$ is included. It makes the electrical length of the ETMPA much reduced than the conventional. The principles introduced above provide good starting points to design a compact ETMPA.

III. RESULTS

One practical ETMPA is designed operating at 3.5GHz for the 5G mobile communications in the Sub-6GHz band [19-20]. It is first optimized in the full wave solver Ansys HFSS. An optimal set of parameters are obtained as following: L = 24 mm, G = 50 mm, p = 6 mm, a = 5.4 mm, $h_1 = h_2$ = 1.52mm, and d = 1mm.

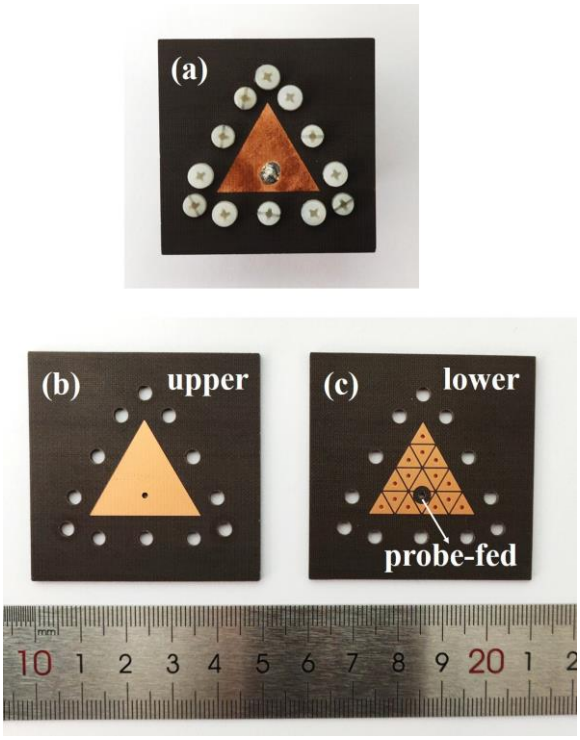


Fig. 2. The photograph for (a) the assembled ETMPA, (b) the upper layer, and (c) the lower layer.

The dielectrics for the two sub-layers to support the TSMMs are both chosen to be the SCGA-500 GF300 provided by Shengyi Technology Co. Ltd. (SYTECH), with $\varepsilon_{r1} = \varepsilon_{r2} = 3$, and $tan\delta_1 = tan\delta_2 = 0.0023$.

An antenna sample is further manufactured for the experimental demonstration using these optimal parameters. It is manually assembled from two substrate layers which are fabricated independently, with the help of nylon screws. The assembled ETMPA is shown in Fig. 2 (a). The upper layer is given in Fig. 2 (b) showing the equilateral triangle patch. The lower layer is shown in Fig. 2 (c). The triangle mushroom caps of the TSMMs are seen on the substrate. The ETMPA is fed by a coaxial probe. Two mushroom elements near the feeding point are defected without vias.

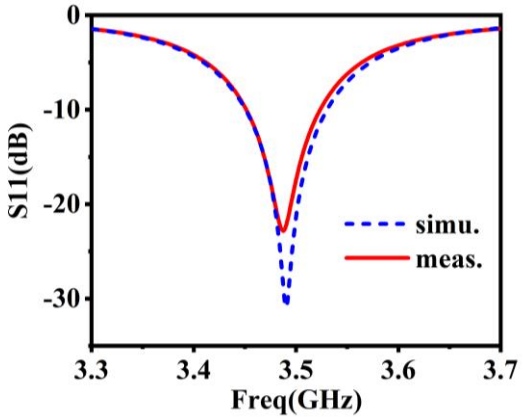


Fig. 3. The reflection coefficients of the ETMPA.

The simulated and measured reflection coefficients (S11s) for the proposed ETMPA is given in Fig. 3. They are seen in good agreement. It is revealed from the measured results that the antenna resonates at 3.488GHz. This value is close to 3.4GHz that is predicted from equation (2). The measured -10dB bandwidth for the ETMPA is 73MHz (2.1%) ranging from 3.451 to 3.524GHz. The simulated bandwidth is slightly wider as 79MHz from 3.451 to 3.53GHz.

It is interesting to notice that the electrical length for the practical ETMPA is now $0.483\lambda_g$, by normalizing the physical length 24mm to the λ_g (49.67mm) at 3.488GHz. This result agrees well with the theoretical prediction of about $0.47\lambda_g$ according to equation (3). The patch area is calculated to be only about $0.1\lambda_g^2$. However according to [1-6], the side length for a conventional ETPA is at the order of $0.66\lambda_g$, while the patch area is nearly $0.19\lambda_g^2$. These results convincingly prove that the new ETMPA has a much reduced patch area than the conventional ETPA. The size reduction effect is mainly influenced by the thickness ratio of h_1 and h_2. By further increasing the ratio of h_2 over h_1, the ETMPA can be designed even smaller in future. If we alternatively use the free space wavelength λ_0 (86mm) at 3.488GHz to measure the ETMPA, the new antenna now occupies a patch area of $0.034\lambda_0^2$. The total thickness is $0.035\lambda_0$ by normalizing 3.04mm to 86mm.

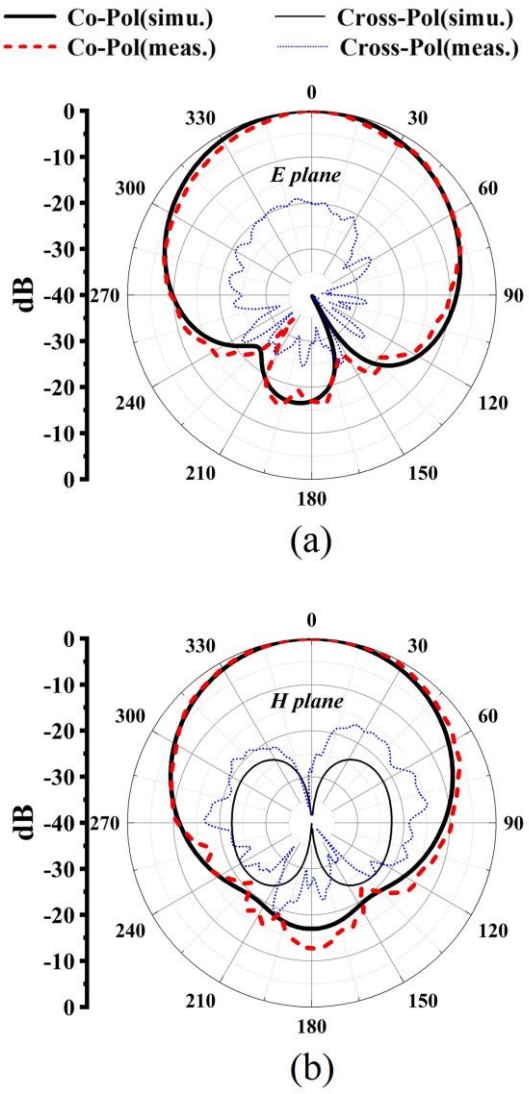


Fig. 4. The radiation patterns for the ETMPA at 3.5GHz on the (a) E plane and (b) H plane.

The radiation patterns for the ETMPA at 3.5GHz are also studied and given in Fig. 4. The patterns on the E plane (*xoz* plane) and H plane (*yoz* plane) are shown in Fig. 4 (a) and Fig. 4 (b), respectively. It is observed that the simulated and measured co-polarized patterns are almost the same. The front-to-back ratio is at the order of 15dB. The cross-polarization levels are however different on the E and H planes. The simulated cross-polarized levels are below -40dB on the E plane that are almost unseen, and about -22dB on the H plane. The measured cross-polarizations are nevertheless much

higher, which are seen to be -18.8dB on the E plane, and -14.6dB on the latter H plane.

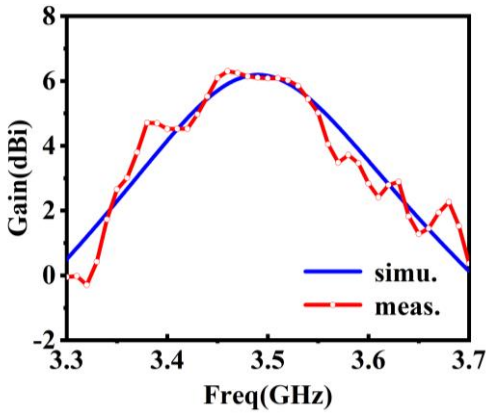


Fig. 5. The antenna gain.

The antenna gain curves are given in Fig. 5 obtained from numerical predictions and experiments. Their trends versus frequency are similar. The peak (realized) antenna gain predicted from HFSS is 6.2dBi. For the measured (realized) gain, the peak value is 6.3dBi. Within the operating frequency band, the measured antenna gain is all over 5.8dBi. The total (realized) efficiency of the fabricated ETMPA is not measured here. However, it is predicted to be about 91% from HFSS. Considering the similarities between the simulated and measured radiation patterns as shown in Fig. 4, the directivity for the fabricated ETMPA should be very close to that of the antenna model used in simulations. Hence the total efficiency for the practical ETMPA can be reasonably estimated at the level of 90%, according to the simulated value.

Table 1: Performances of several compact metamaterial antennas operating near 3.5GHz

Ref.	f_0 /GHz	Patch area	Thick-ness	BW /%	Gain /dBi
[11]	3.3	$0.036\lambda_0^2$	$0.034\lambda_0$	3.1	0.79
[12]	3.494	$0.063\lambda_0^2$	$0.087\lambda_0$	2.2	7
[13]	3.5	$0.053\lambda_0^2$	$0.018\lambda_0$	3.7	4.6
This	3.488	$0.034\lambda_0^2$	$0.035\lambda_0$	2.1	6.3

The antenna performances are shown in Table 1, in contrast to several other compact metamaterial antennas [11-13] operating near 3.5GHz. In Table 1, all of the patch areas and thicknesses are normalized to the free space wavelength λ_0 at the resonant frequency f_0. In the BW column, results are summarized from the -10dB relative bandwidth of these antennas. From Table 1, the ETMPA is observed to occupy a very compact antenna volume but maintain an acceptable bandwidth and gain. To further improve the bandwidth, several broadband techniques [1], e.g., adding a stacked patch, or using a U-slot patch etc., might be used in future.

IV. CONCLUSION

A new compact ETMPA is designed and realized operating at 3.5GHz. By loading periodic TSMMs into a conventional ETPA, the new ETMPA is demonstrated with a reduced patch area than conventional. Moreover, the bandwidth and radiation performances for the size-reduced ETMPA are also evaluated. The measured bandwidth for the practical compact ETMPA is 2.1% while the peak antenna gain is about 6.3dBi. They are both sufficient to meet the demand of 5G mobile communications.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (No. 11904222) and the Natural Science Foundation of Shanghai (No. 16ZR1446100).

REFERENCES

[1] R. Garg, P. Bhartia, I. Bahl, and A. Ittipiboon, *Microstrip Antenna Design Handbook*. Artech House, USA, 2001.

[2] J. Helszajn and D. S. James, "Planar triangular resonators with magnetic walls," *IEEE Trans. Microw. Theory Techn.*, vol. 26, pp. 95-100, 1978.

[3] E. F. Kuester and D. C. Chang, "A geometrical theory for the resonant frequencies and Q-factors of some triangular microstrip patch antennas," *IEEE Trans. Antennas Propag.*, vol. 31, pp. 27-34, 1983.

[4] K.-F. Lee, K. Luk, and J. S. Dahele, "Characteristics of the equilateral triangular patch antenna," *IEEE Trans. Antennas Propag.*, vol. 36, pp. 1510-1518, 1988.

[5] N. Kumprasert and K. W. Kiranon, "Simple and accurate formula for the resonant frequency of the equilateral triangular microstrip patch antenna," *IEEE Trans. Antennas Propagat.*, vol. 42, pp. 1178-1179, 1994.

[6] M. Biswas and M. Dam, "Fast and accurate model of equilateral triangular patch antennas with and without suspended substrates," *Microw. Opt. Techn. Lett.*, vol. 54, pp. 2663-2668, 2012.

[7] J. B. Pendry, A. J. Holden, D. J. Robbins, and W. J. Stewart, "Magnetism from conductors and enhanced nonlinear phenomena," *IEEE Trans. Microw. Theory Techn.*, vol. 47, pp. 2075-2084, 1999.

[8] C. Caloz and T. Itoh, *Electromagnetic Metamaterials: Transmission Line Theory and Microwave Applications*. Wiley, USA, 2005.

[9] Y. Dong and T. Itoh, "Metamaterial-based antennas," *Proc. IEEE*, vol. 100, pp. 2271-2285,

2012.

[10] J. Wang, Y. Li, Z. H. Jiang, T. Shi, M.-C. Tang, Z. Zhou, Z. N. Chen, and C.-W. Qiu, "Metantenna: when metasurface meets antenna again," *IEEE Trans. Antennas Propag*., vol. 68, pp. 1332-1347, 2020.

[11] J. Zhu and G. V. Eleftheriades, "A compact transmission-line metamaterial antenna with extended bandwidth," *IEEE Antennas Wireless Propag. Lett*., vol. 8, pp. 295-298, 2009.

[12] X. M. Yang, Q. H. Su, Y. Jing, Q. Cheng, X. Y. Zhou, H. W. Kong, and T. J. Cui, "Increasing the bandwidth of microstrip patch antenna by loading compact artificial magneto-dielectrics," *IEEE Trans. Antennas Propag*., vol. 50, pp. 373-378, 2011.

[13] T. Cai, G. M. Wang, X. F. Zhang, Y. W. Wang, B. F. Zong, and H. X. Xu, "Compact microstrip antenna with enhanced bandwidth by loading magneto-electro-dielectric planar waveguided metamaterials," *IEEE Trans. Antennas Propag.*, vol. 63, pp. 2306-2311, 2015.

[14] X. Xu and J. Wei, "Miniaturisation design of patch antenna using a low-profile mushroom type meta-substrate tailored with high permittivity," *IET Microw. Antennas Propag.*, vol. 12, pp. 1216-1221, 2018.

[15] D. Sevenpiper, L. Zhang, R. F. Broas, and N. G. Alexopolous, "High-Impedance electromagnetic surfaces with a forbidden frequency band," *IEEE Trans. Microw. Theory Techn*., vol. 47, pp. 2059-2074, 1999.

[16] L. Li, Q. Chen, Q. Yuan, C. Liang, and K. Sawaya, "Surface-wave suppression band gap and plane-wave reflection phase band of mushroomlike photonic band gap structures," *J. Appl. Phys*., vol. 103, p. 023513, 2008.

[17] T. Jiang, T. Jiao, and Y. Li, "Array mutual coupling reduction using L-loading E-shaped electromagnetic band gap structures," *Int. J. Antennas Propag.*, 2016, 6731014, 2016.

[18] T. Jiang, T. Jiao, and Y. Li, "A low mutual coupling MIMO antenna using periodic multi-layered electromagnetic band gap structures," *Applied Computational Electromagnetics Society Journal*, vol. 33, pp. 305-311, 2018.

[19] WRC-15 Press Release. (Nov. 27, 2015). World Radio communication Conference Allocates Spectrum for Future Innovation. Online website: http://www.itu.int/net/pressoffice/press_releases/2015/56.aspx

[20] Z. Li, J. Han, Y. Mu, X. Gao, and L. Li, "Dual-band dual-polarized base station antenna with a notch band for 2/3/4/5G communication systems," *IEEE Antennas Wireless Propag. Lett*., vol. 19, pp. 2462-2466, 2020.

Analysis of a Serrated Ground Plane for a Low-Loss Reflectarray Antenna

Jiawei Ren [1,2], Hongjian Wang [1,2], Weichun Shi[1,2], and Minzheng Ma [1,2]

[1] Department of Electrical Engineering
CAS Key Laboratory of Microwave Remote Sensing, National Space Science Center
Chinese Academy of Sciences, Beijing 100190, China
18810203041@163.com, wanghongjian@mirslab.cn, 232050977@qq.com, 1144515770@qq.com

[2] Department of Electrical and Computer Engineering
University of Chinese Academy of Sciences, Beijing 100049, China

Abstract ─ A novel serrated ground plane (NSGP) for na low-loss reflectarray antenna is presented in this paper. Compared with a conventional smooth ground plane (CGP), the NSGP consists of a series of serrated elements, which can reflect the incident waves in the main beam direction, so the losses of the reflectarray can be effectively reduced. The principle and losses of reflectarray antennas are studied and analyzed. Then, a low-loss NSGP is proposed, and two design methods for the NSGP are given in this paper. Finally, a reflectarray antenna with elements arranged in a 15×15 grid is designed, simulated and measured with the NSGP and CGP respectively within the frequency from 12.88 to 13.88GHz. The results show that the reflectarray antenna with the NSGP can effectively utilize the reflected waves and has a maximum higher gain of 0.681 dB compared with the gain of the reflectarray with the CGP. This NSGP has a potential to be used in the high accurate design of the reflectarray which requires to realize beam forming, low-loss, high-efficiency, etc.

Index Terms ─ High-accuracy, high gain, low-loss, reflectarray antenna, reflector antenna, serrated ground plane.

I. INTRODUCTION

A reflectarray antenna is a high gain antenna that is easier to manufacture and takes up less space than a parabolic reflector antenna [1]. Reflectarray antennas can realize beam forming [2] and beam scanning [3] and have the advantages of low loss and ease of fabrication compared with phase array antennas. Reflectarray antennas were first proposed in 1963 [4], and reflectarray antennas have attracted a great deal of interest from researchers and have rapidly developed since the 1970s with the development of microstrip antennas.

The reflectarray antenna has the advantages of high gain, a low profile and low loss, while the losses of the reflectarray antenna cannot be ignored, especially at high frequencies [5]. In general, the losses of reflectarray antennas include dielectric losses, conductor losses, surface wave excitation [6], phase dispersion compensation and the losses caused by the reflection of the ground plane. To date, much work has been done to analyze the losses of reflectarray antennas. The losses of microstrip-patch antennas have been studied in [7]. Conductor and dielectric losses are more severe for thinner substrates, and the model of the dielectric losses and the role of the substrate loss tangent were investigated in [8]. The effects of the substrate thickness and permittivity on the antenna's performance were studied in [9]. A detailed study of losses for mm-wave microstrip arrays was presented in [10]. The modeling of printed reflectarray elements through a lumped-element circuit, which includes the effect of metal and dielectric losses, is presented in [11].

The reflectarray antenna is designed with the assumption that the reflected waves travel along the main beam direction, while part of the waves reflected by the ground plane deviates from the main beam direction most of the time [12], and this part of the waves can increase the sidelobes and cause losses of the reflectarray antenna. To date, the losses of reflectarray antennas caused by the mirror reflection of the ground plane have rarely been studied.

In this paper, a novel serrated ground plane (NSGP) for a low-loss reflectarray antenna is proposed. The principle and losses of the reflectarray antenna are studied. Then, a low-loss NSGP that can effectively use the energy of the reflected waves is designed and analyzed, and a reflectarray with conventional ground plane (CGP) and the NSGP is fabricated and tested respectively. Measured results show the reflectarray with the NSGP has an average higher gain and aperture efficiency of 0.54dB and 6.1% compared with the reflectarray with CGP within the working frequency of 12.88GHz to 13.88GHz. The geometry and analysis of the NSGP is presented in Section III and Section IV, and

Submitted On: March 20, 2021
Accepted On: July 7, 2021

https://doi.org/10.47037/2021.ACES.J.360812

a reflectarray antenna operating at 13.38 GHz with the NSGP and CGP respectively is designed, simulated and tested in Section V. Finally, comparisons are given, and the conclusions are summarized in Section VI.

II. PRINCIPLE AND LOSSES OF THE REFLECTARRAY ANTENNA

A. Phase compensation of the reflectarray antenna

A reflectarray antenna is usually designed by the following steps: the first step is obtaining the amplitude and phase distributions that are needed to be compensated for the reflectarray surface, and there are some studies to synthesize reflectarray [13-15]. After the required phase compensation is obtained, an appropriate element must be designed to provide the required phase compensation, and finally, there are some methods to calculate the far field radiation pattern of the reflectarray, such as the conventional array theory [16], [17] and the equivalent aperture field method based on the physical optic (PO).

From the analysis above, we can know that by properly tailoring the elements phase of the reflectarray antenna [18], the reflectarray can reflect the incident waves in a desired direction so that we can obtain a very directional pencil beam in the far field, which is the principle of reflectarray antenna. The coordinate system of the reflectarray antenna is shown in Fig. 1.

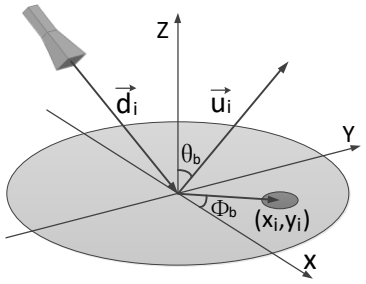


Fig. 1. Coordinate system of the reflectarray antenna.

The phase compensation required by the reflectarray element (x_i, y_i) to generate the required beam direction (θ_b, φ_b) as shown in Fig. 1 is:

$$\varphi(x_i, y_i) = -k_0 sin\theta_b cos\varphi_b x_i - k_0 sin\theta_b sin\varphi_b y_i, \quad (1)$$

where $k_0 = \frac{2\pi}{\lambda} = \frac{2\pi f}{c}$ is the wave number in free space. The phase of the reflected wave is the sum of the phase of the incident wave and the phase shift introduced by the reflectarray unit cell after the incident wave is reflected by the reflectarray, as shown in the following formula:

$$\varphi(x_i, y_i) = -k_0 d_i + \varphi_R(x_i, y_i), \quad (2)$$

where $\varphi_R(x_i, y_i)$ is the phase shift introduced by the reflectarray element and d_i is the distance between the reflectarray unit cell and the feed source. Based (1) and (2), the required phase compensation for the reflectarray element is:

$$\varphi_R = k_0 \left(d_i - \left(x_i cos\varphi_b + y_i sin\varphi_b \right) sin\theta_b \right). \quad (3)$$

B. Losses of the reflectarray antenna

Based on the analysis of the reflectarray antenna, the losses of the reflectarray antenna can be summarized into the following parts [19]: the losses caused by the dielectric substrate, the losses caused by the resonance of the conductor, the losses caused by surface waves and the losses caused by the mirror reflection of the ground plane. Much work has been done to reduce the losses of the reflectarray, such as a substrate with a relatively small dielectric loss tangent has been used, and a reflectarray antenna without a substrate has been designed to reduce the losses caused by the substrate [20]. For conductor losses of the reflectarray antenna, materials with lower losses, such as light-controlled plasma materials, graphene materials and quartz crystals, have been used to make reflectarray patches [21], [22]. For losses caused by surface waves, some special structures, such as metamaterials, have been used to suppress surface waves [23]. However, there are no studies that analyze the losses caused by the mirror reflection of the ground plane, and it is necessary to study the losses caused by the reflection of the ground plane to reduce the losses and improve the gain of the reflectarray antenna.

When electromagnetic waves are incident on the reflectarray antenna, all the incident waves are reflected by the ground plane, which mainly includes three parts: the first part is the re-radiated waves caused by the resonance of reflectarray element patches, the second part is the waves reflected by the ground plane, and the third part is the non-resonant radiated waves of the reflectarray elements. Since the thickness of the substrate of the reflectarray is usually small $(< 0.1\lambda_0)$, the reflected waves mainly include the first two parts [23], as shown in Fig. 2 (a).

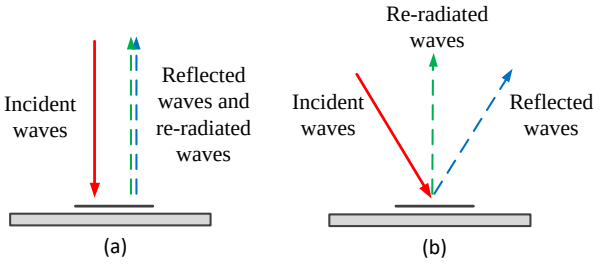


Fig. 2. Re-radiated and reflected waves of the reflectarray element when (a) the incident waves with a small incident angle, and (b) the incident waves with a large incident angle.

When the angle of incident waves is small or not very large as shown in Fig. 2 (a), the direction of re-radiated waves caused by the resonance of the reflectarray patches and the reflected waves caused by reflection of the ground plane are both along the main beam direction and this follows the assumption of our design of the reflectarray. But when the electromagnetic waves are incident on the elements at the edge of the reflectarray as shown in Fig. 2 (b), since the angle of reflected waves is equal to the angle of incidence, the direction of the re-radiated waves and the reflected waves are not the same based on the law of reflection, only part of the reflected waves contributes to the gain of the reflectarray antenna, and others will cause the losses and increase the sidelobes of the reflectarray antenna.

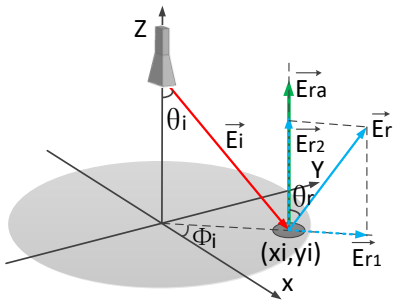


Fig. 3. Re-radiated and reflected waves of the reflectarray element at the edge of the reflectarray.

As shown in Fig. 3, when electromagnetic wave is incident on the element at the edge of the reflectarray, the total reflected wave include the re-radiated wave caused by the element patche and the wave reflected by the ground plane. According to the law of reflection, the angle of the reflected wave and incident wave is equal $\theta_r = \theta_i$, and the reflected wave $\vec{E}_r$ can be divided into $\vec{E}_{r1}$ and $\vec{E}_{r2}$ in the Cartesian coordinate system. $\left|\vec{E}_{r2}\right| = \left|\vec{E}_r\right| cos\theta_r$ is along the main beam direction and contributes to the gain of the reflectarray antenna, and $\left|\vec{E}_{r1}\right| = \left|\vec{E}_r\right| sin\theta_r$ is vertical to the main beam direction and will increase the sidelobes of the reflectarray. For example, when the focal diameter ratio (f / D) is 0.8, the maximum angle of the incident wave can be calculated as $\theta_i = arctan\left(D/2f\right) = arctan(0.625)$, so the reflected wave along the main beam direction is:

$$\begin{aligned}\left|\vec{E}_{r2}\right| &= \left|\vec{E}_r\right| cos\theta_r \\ &= \left|\vec{E}_r\right| cos\left[arctan(0.625)\right] \\ &\approx 0.85\left|\vec{E}_r\right|. \end{aligned} \quad (4)$$

The losses of the reflected wave along the main beam direction compared with the incident wave are:

$$\begin{aligned}\left|\vec{E}_{losses}\right| &= \left|\vec{E}_r\right| - \left|\vec{E}_{r2}\right| \\ &= \left|\vec{E}_r\right| - \left|\vec{E}_r\right| cos\left[arctan(0.625)\right] \\ &\approx 0.15\left|\vec{E}_r\right|. \end{aligned} \quad (5)$$

From the analysis above, the waves reflected by the ground plane and the re-radiated waves caused by the element patches are not in the same direction when the incident angle is large, which will cause losses and increase the sidelobes of the reflectarray antenna, using a large f / D ratio can reduce the losses [12], but it does not fundamentally solve the problem. With many studies and simulations, a novel serrated ground plane (NSGP) for a low-loss reflectarray is proposed, which can reflect the incident waves to the main beam direction, thereby reducing the losses of the reflectarray antenna.

III. DESIGN OF THE NSGP

The reflector antenna can be regarded as a two-dimensional parabola rotating around a fixed axis, and when the electromagnetic waves emitted at the focal point are reflected by the reflector antenna, the reflected waves are parallel to each other. The reflector antenna is shown in Fig. 4.

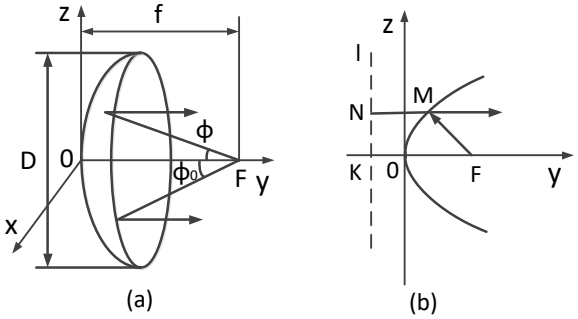


Fig. 4. The reflector antenna.

The Fresnel-Zone Plate Reflector (FZPR) was developed many years ago as an inexpensive alternative to a parabolic reflector [24], [25]. As shown in Fig. 5, there are many concentric zones, which are separated by two heights or thicknesses, and all the reflected waves are more or less in phase when the heights or thicknesses are properly designed [26].

The FZPR has many advantages, such as simplicity in design and manufacturing. However, there is a principle drawback of FZPR: the FZPR has relatively low efficiency due to the phase dispersion of the reflected signals in each zone [27], the distance from one ring to the next ring varies from zone to zone and due to

the lack of individual elements, and the FZPR cannot achieve beam scanning [12].

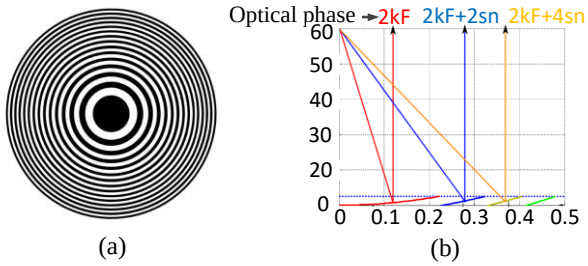


Fig. 5. The FZPR.

A. A NSGP for the low-loss reflectarray antenna

Here, we introduce a NSGP for a reflectarray antenna that combines the principle of reflector antenna, FZPR and reflectarray antenna, and the elements of the reflectarray can be designed individually. According to the analysis of the reflectarray antenna, the reflected waves mainly include two parts: the re-radiated waves caused by the element patches and the waves reflected by the ground plane. When a reflectarray is designed, the direction of the re-radiated waves is determined, so we can just adjust the direction of the reflected waves. A NSGP is proposed that can change the direction of the reflected waves to the main beam direction, thereby reducing the losses of the reflectarray, and there are two methods for designing it.

B. Design of the NSGP based on the law of reflection

As shown in Fig. 6, when electromagnetic waves are obliquely incident on a ground plane, according to Maxwell's equations and the boundary conditions, the incident waves will be reflected by the ground plane, and the reflected angle a_r is equal to the incident angle a_i. If the ground plane rotates by an angle of β, then the reflected waves will rotate by an angle of $2*\beta$, as shown in Fig. 6 (b). If a series of small conductive planes on the same horizontal rotate by a given angle, which is half of its incident angle, it is evident that the incident waves will be reflected back to the same direction, as shown in Fig. 6 (c). If we use these small planes as the reflectarray ground plan, the directions of the reflected waves and re-radiated waves will be the same, and both along the main beam direction, as shown in Fig. 6 (d), so the losses of the reflectarray caused by the reflection of the ground plane can be reduced.

C. Design of the NSGP based on the reflector antenna

As discussed above, the waves that are reflected by the reflector antenna are parallel to each other when the waves are emitted from the focal point as shown in Fig. 7 (a).

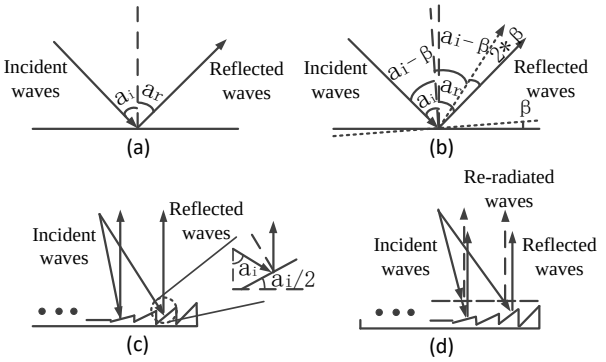


Fig. 6. Design of the NSGP based on the law of reflection.

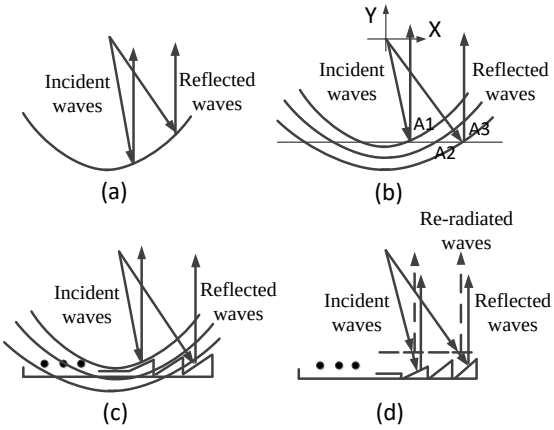


Fig. 7. Design of the NSGP based on the reflector antenna.

If a series of electromagnetic waves emitted from the focal point are reflected by a series of parabolas with formulas $y=\frac{x^2}{2f_i}-\frac{f_i}{2}$, which have the same focal point $(0,0)$ but different focal lengths $\frac{f_i}{2}$, then the waves will be reflected back to the vertical direction and parallel to each other as shown in Fig. 7 (b). If a series of small planes is used to approach the series of parabolas at the same horizontal positions, as shown in Fig. 7 (c), it is evident that the waves reflected by the small planes will be parallel to each other and all along the vertical direction. This structure can be used for the reflectarray antenna so that the directions of the reflected waves and re-radiated waves will be the same for all the elements and the losses caused by the reflection of the ground plane can be effectively reduced as shown in Fig. 7 (d).

The slopes of the element ground planes at different positions can be calculated as follows: Assuming that the small element ground planes are on the same horizontal plane $y=-y_0$, and the coordinates of element planes

are $(np, -y_0)$, where p is the period of each unit cell, $n = 1, 2, 3.....$, then take the coordinates of element planes $(np, -y_0)$ into the parabolic formula and a series of focal lengths $\frac{f_n}{2}$ can thus be obtained, finally, the slopes of element ground planes can be calculated by the formula $y_n^{'} = \frac{nf}{f_n}$.

From the discussion above, it is evident that the NSGP uses the serrated structure to adjust the directions of the reflected waves at each unit cell, so the directions of each elements' reflected waves are all along the main beam direction, the energy of the reflected waves can be used effectively and the gain of the reflectarray antenna can be improved. To learn more about the NSGP, the models of the NSGP and CGP are analyzed in the following section respectively.

IV. ANALYSIS OF THE LOW-LOSS NSGP

As discussed above, the reflected waves of the reflectarray antenna mainly include the re-radiated waves caused by the resonance of the element patches and the waves reflected by the ground plane as shown in the following formulas:

$$\vec{E}_I = E_0 e^{jk_0(xu_i + yv_i + z\cos\theta_i)}, \quad (6)$$

$$\vec{E}_T = \vec{E}_R + \vec{E}_S = [R(\theta_i, \varphi_i) + S(\theta_i, \varphi_i)] \times E_0 e^{jk_0(xu_i + yv_i + z\cos\theta_i)}, \quad (7)$$

$$\begin{bmatrix} E_\theta^T \\ E_\varphi^T \end{bmatrix} = \left(\begin{bmatrix} R_{\theta\theta} & 0 \\ 0 & R_{\varphi\varphi} \end{bmatrix} + \begin{bmatrix} S_{\theta\theta} & S_{\theta\varphi} \\ S_{\varphi\theta} & S_{\varphi\varphi} \end{bmatrix} \right) \times \left(\begin{bmatrix} E_{0\theta} \\ E_{0\varphi} \end{bmatrix} e^{jk_0(xu_i + yv_i + z\cos\theta_i)} \right), \quad (8)$$

where $\begin{bmatrix} R_{\theta\theta} & 0 \\ 0 & R_{\phi\phi} \end{bmatrix}$ is the reflected factor of the reflectarray antenna, $\begin{bmatrix} S_{\theta\theta} & S_{\theta\phi} \\ S_{\phi\theta} & S_{\phi\phi} \end{bmatrix}$ is the re-radiated factor of the reflectarray antenna, and $\begin{bmatrix} E_{0\theta} \\ E_{0\varphi} \end{bmatrix} e^{jk_0(xu_i + yv_i + z\cos\theta_i)}$ is the incident waves.

When a reflectarray antenna is designed, the structures of the element patches and the substrate of the reflectarray are determined, so the re-radiated waves of the reflectarray are determined. At this time, the gain of the reflectarray antenna mainly depends on the waves reflected by the ground plane, which will be studied and analyzed in detail as follows.

A. Analysis of the CGP

As shown in Fig. 8, when the electromagnetic wave is incident on the CGP at an incident angle of (θ_i, φ_i), according to the law of reflection, the magnitude of the reflected wave and the incident wave is equal $|\vec{E}_r| = |\vec{E}_0|$, and the angle of the reflected wave is equal to the angle of incident wave $\theta_r = \theta_i$. It can be obtained that the magnitude of the reflected wave along the main beam direction is:

$$|\vec{E}_{er1}| = |\vec{E}_0| cos\theta_i \quad . \quad (9)$$

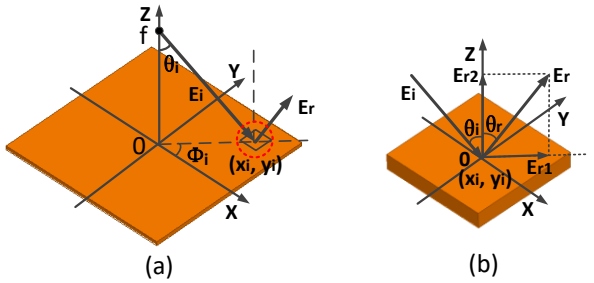


Fig. 8. Model of the CGP.

B. Analysis of an idea novel serrated ground plane (INSGP)

The model of the INSGP for the reflectarray is shown in Fig. 9, and the slope of each element ground plane is designed by the analysis in Section III. According to the analysis above, when the electromagnetic wave $\vec{E}_i$ at the focal point $(0, 0, z_0)$ is incident on the element ground plan $(x_i, y_i, 0)$ with an angle of (θ_i, φ_i), the wave $\vec{E}_i$ will be reflected back by the ground plane along the main beam direction. So the magnitude of the reflected wave along the main beam direction is:

$$|\vec{E}_{er2}| = |\vec{E}_r| = |\vec{E}_0| \quad . \quad (10)$$

In this case, both the re-radiated wave and the reflected wave are along the main beam direction, and the reflectarray will fully utilize the energy re-radiated by the element patches and reflected by the ground plane.

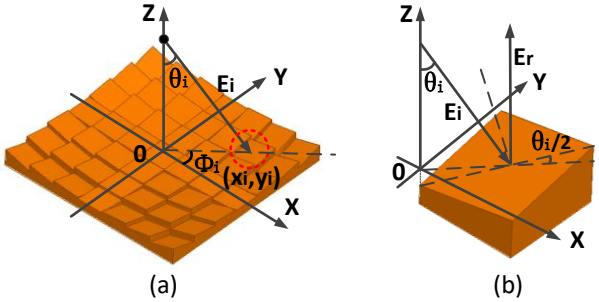


Fig. 9. Model of the INSGP.

C. Analysis of a NSGP I

Considering that the number of reflectarray unit cells is usually large and the difficulty of manufacturing,

the INSGP for the reflectarray is simplified. Assume that the electromagnetic wave $\vec{E}_i$ at the focal point $(0, 0, z_0)$ is incident on the element ground plane $(x_i, y_i, 0)$ with an angle of (θ_i, φ_i) as shown in Fig. 10.

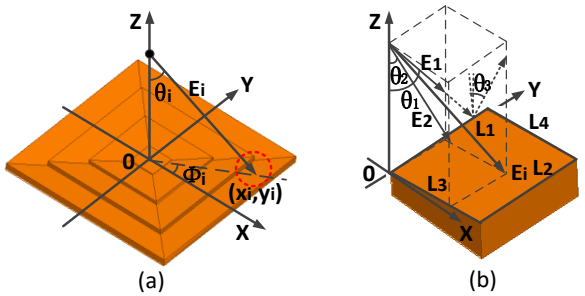


Fig. 10. Model of the NSGP I.

As shown in Fig. 10 (b), the incident wave $\vec{E}_i$ is decomposed into $\vec{E}_1$ and $\vec{E}_2$ in the Cartesian coordinate system and $\vec{E}_i = \vec{E}_1 + \vec{E}_2$, where $\vec{E}_1 = |\vec{E}_i| sin\theta_i sin\varphi_i * \vec{e}_y - \frac{1}{2}|\vec{E}_i| \cos\theta_i * \vec{e}_z$ is in the yoz plane, the angle between $\vec{E}_1$ and the negative z-axis is $\theta_1 = arctan\left(|y_i| / (\frac{z_0}{2})\right)$, and the magnitude of $\vec{E}_1$ is:

$$|\vec{E}_1| = |\vec{E}_i| \sqrt{sin^2\theta_i sin^2\varphi_i + \frac{1}{4}\cos^2\theta_i} \,. \quad (11)$$

$\vec{E}_2 = |\vec{E}_i| sin\theta_i cos\varphi_i * \vec{e}_x - \frac{1}{2}|\vec{E}_i| \cos\theta_i * \vec{e}_z$ is in the xoz plane, the angle between $\vec{E}_2$ and the negative z-axis is $\theta_2 = arctan\left(|x_i| / (\frac{z_0}{2})\right)$, and the magnitude of $\vec{E}_2$ is:

$$|\vec{E}_2| = |\vec{E}_i| \sqrt{sin^2\theta_i cos^2\varphi_i + \frac{1}{4}\cos^2\theta_i} \,. \quad (12)$$

The element ground plane can thus be designed according to the analysis above, the two sides of the element ground plane are parallel $L_1 // L_2 // \vec{e}_y$ and the other two sides of the element ground plane are $L_3 // L_4$, besides, the angle between L_3 and the x-axis is $\frac{\theta_2}{2}$. According to the law of reflection, the incident wave $\vec{E}_2$ will be reflected back to the vertical direction, and the magnitude of $\vec{E}_{2re}$ that contributes to the gain of the reflectarray is:

$$|\vec{E}_{2re}| = |\vec{E}_2| = |\vec{E}_i| \sqrt{sin^2\theta_i cos^2\varphi_i + \frac{1}{4}\cos^2\theta_i} \,. \quad (13)$$

The sides L_1 and L_2 are parallel to the y-axis, while the ground plane rotates along the y-axis by an angle of $\frac{\theta_2}{2}$, the reflected wave $\vec{E}_{1r}$ of the incident wave $\vec{E}_1$ is complicated. The magnitudes of the reflected wave $\vec{E}_{1re}$ along the main beam direction is given directly as:

$$|\vec{E}_{1re}| = |\vec{E}_{1r}| cos\theta_1 cos\theta_2$$
$$= |\vec{E}_i| cos\theta_1 cos\theta_2 \sqrt{sin^2\theta_i sin^2\varphi_i + \frac{1}{4}\cos^2\theta_i} \,. \quad (14)$$

Therefore, the sum of the magnitudes of the reflected waves along the main beam direction is:

$$|\vec{E}_{er3}| = |\vec{E}_{1re}| + |\vec{E}_{2re}|$$
$$= |\vec{E}_i| (cos\theta_1 cos\theta_2 \sqrt{sin^2\theta_i sin^2\varphi_i + \frac{1}{4}\cos^2\theta_i} \,. \quad (15)$$
$$+ \sqrt{sin^2\theta_i cos^2\varphi_i + \frac{1}{4}\cos^2\theta_i}\,).$$

D. Analysis of a NSGP II

Similar to the analysis in Section. IV. C, the incident wave $\vec{E}_i$ is decomposed into $\vec{E}_1$ and $\vec{E}_2$ as shown in Fig. 11, where $\vec{E}_1$ is in the yoz plane and is parallel to the y-axis, $\vec{E}_1 = |\vec{E}_i| \sin\theta_i sin\varphi_i * \vec{e}_y$. $\vec{E}_2$ is in the xoz plane, $\vec{E}_2 = |\vec{E}_i| sin\theta_i cos\varphi_i * \vec{e}_x - |\vec{E}_i| cos\theta_i * \vec{e}_z$, the angle between $\vec{E}_2$ and the negative z-axis is $\theta_2 = arctan\left(\frac{x_i}{z_0}\right)$, and the magnitude of $\vec{E}_2$ is:

$$|\vec{E}_2| = |\vec{E}_i| \sqrt{sin^2\theta_i cos^2\varphi_i + cos^2\theta_i} \,. \quad (16)$$

The element ground plane can thus be designed based on the analysis above, the sides of the reflectarray element ground plane are $L_1 // L_2 // \vec{e}_y$ and $L_3 // L_4$, and the angle between L_3 and the x-axis is $\frac{\theta_2}{2}$. According to the law of reflection, $\vec{E}_1$ is parallel to the y-axis, the ground plane does not reflect it, and $\vec{E}_2$ will be reflected back to the main beam direction, and the magnitude of the reflected wave $\vec{E}_{2re}$ is:

$$|\vec{E}_{2re}| = |\vec{E}_2| = |\vec{E}_i| \sqrt{sin^2\theta_i cos^2\varphi_i + cos^2\theta_i} \,. \quad (17)$$

Therefore, the magnitude of the reflected wave along the main beam direction is:

$$|\vec{E}_{er4}| = |\vec{E}_{2re}| = |\vec{E}_i| \sqrt{sin^2\theta_i cos^2\varphi_i + cos^2\theta_i} \,. \quad (18)$$

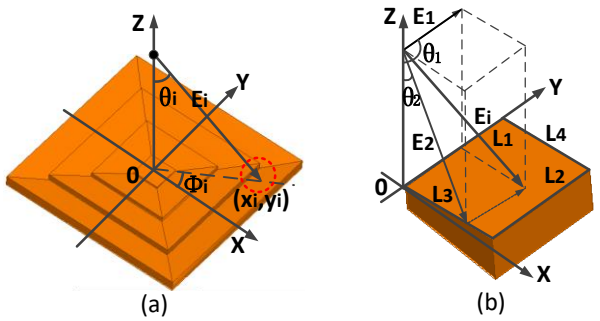


Fig. 11. Model of the NSGP II.

E. Comparisons

From the analysis above, the magnitudes of the reflected waves in Section IV. A, Section IV. B, Section IV. C and Section IV. D along the main beam direction are:

$$\left|\vec{E}_{er1}\right| = \left|\vec{E}_i\right| cos\theta_i , \tag{19}$$

$$\left|\vec{E}_{er2}\right| = \left|\vec{E}_r\right| = \left|\vec{E}_i\right| , \tag{20}$$

$$\begin{aligned}\left|\vec{E}_{er3}\right| &= \left|\vec{E}_{1re}\right| + \left|\vec{E}_{2re}\right| \\ &= \left|\vec{E}_i\right| \left(cos\theta_1 cos\theta_2 \sqrt{sin^2\theta_i sin^2\varphi_i + \frac{1}{4}\cos^2\theta_i} + \sqrt{sin^2\theta_i cos^2\varphi_i + \frac{1}{4}\cos^2\theta_i} \right)\end{aligned} , \tag{21}$$

$$\left|\vec{E}_{er4}\right| = \left|\vec{E}_i\right| \sqrt{sin^2\theta_i cos^2\varphi_i + cos^2\theta_i} . \tag{22}$$

Usually, $0.8 \le f/D \le 1.5$ and $-45° \le \varphi_i \le 45°$, then it is easy to calculate that $\left|\vec{E}_{er2}\right| \ge \left|\vec{E}_{er4}\right| \ge \left|\vec{E}_{er3}\right| \ge \left|\vec{E}_{er1}\right|$. Compared to the INSGP as shown in Section. IV. B, the NSGP II simplifies the manufacturing process, and the gain of NSGP II is lower:

$$\begin{aligned}\left|\vec{E}_{low}\right| &= \left|\vec{E}_{er2}\right| - \left|\vec{E}_{er4}\right| \\ &= \left|\vec{E}_i\right| \left(1 - \sqrt{sin^2\theta_i cos^2\varphi_i + cos^2\theta_i}\right).\end{aligned} \tag{23}$$

Compared with the CGP, the NSGP II effectively increases the gain of the reflectarray, and the gain of the NSGP II is higher:

$$\begin{aligned}\left|\vec{E}_{high}\right| &= \left|\vec{E}_{er4}\right| - \left|\vec{E}_{er1}\right| \\ &= \left|\vec{E}_i\right| \left(\sqrt{sin^2\theta_i cos^2\varphi_i + cos^2\theta_i} - cos\theta_i\right).\end{aligned} \tag{24}$$

Considering the difficulties associated with manufacturing, the NSGP II has the advantages of reduced manufacture and can also reduce the losses caused by the reflection of ground plane.

V. SIMULATIONS AND COMPARISONS

A. Unit cell element

A simple reflectarray element is designed, and the reflectarray element patch consists of three square rings and a vertical strip that combines the three square rings to form a unit cell. After a series of simulations and optimizations, the structure and parameters of the element are shown in Fig. 12. The length of the unit cell L is 10 mm, and the slot length $L1$ is 9 mm. The length of middle square ring $L2$ is $3.5\times d$ mm, and the slot length $L3$ is $3\times d$ mm. The length of the inner square ring $L4$ is $2.5\times d$ mm, and the slot length $L5$ is $2\times d$ mm, and the reflection phase of the unit cell is controlled by adjusting the variable length of d. The width of the vertical strip w is 0.5 mm, the thickness of the substrate (RT/duroid5880) is h and the height of the air substrate is $h_1 = H - h$.

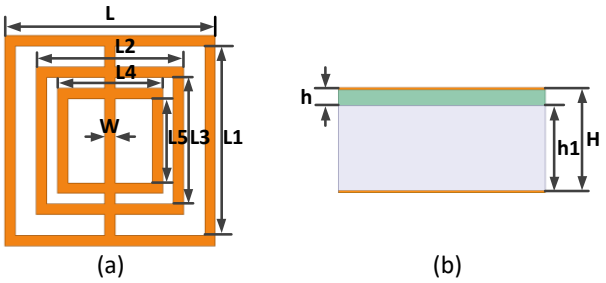


Fig. 12. Configuration and geometry of the reflectarray element.

A parametric study is carried out to optimize element structure in order to obtain a linear reflection phase curve. The influence of the thickness of substrate h and H are shown in Fig. 13 respectively, it can be observed that the substrate thickness h has little influence on the reflection phase curve. Considering the manufacture factor, the thickness of the substrate h is 0.762mm. It can be concluded that the thickness of H has great influence on the reflection phase curve as shown in Fig. 13 (b), so the thickness of $H = 5mm$ is chosen for better linear reflect phase and large phase variation range.

When the element patch rotates a certain angle relative to the ground plane, the reflected phase of element with variable length d operating at 13.38 GHz is plotted in Fig. 14. It is observed that the range of the reflection phase is over 450°, which satisfies the required phase compensation for the reflectarray antenna. The rotation between the element patch and the ground plane has little impact on the reflection phase, and the deviation between the curves when the element patch and ground plane are rotated is less than 18°.

B. Design and simulation of reflectarray combined with different ground planes

To verify the design concept of the proposed NSGP, simulations and comparisons will be made for the reflectarray antenna with different ground planes.

A reflectarray antenna with the proposed square ring elements working at 13.38 GHz is designed. The reflectarray surface is composed of 15×15 elements. The

period of reflectarray elements p is 10 mm ($0.45 \times \lambda$), the thickness h of substrate RT5880 is 0.762 mm and the thickness ($H - h$) of the air substrate is 4.238mm, the focal length f is 120 mm, and the focal diameter radio (f/D) is 0.8, which is determined by the radiation pattern of the horn feed. According to the analysis in Section IV, the CGP and the NSGP are designed, and the front views of the ground planes are shown in Fig. 15 below.

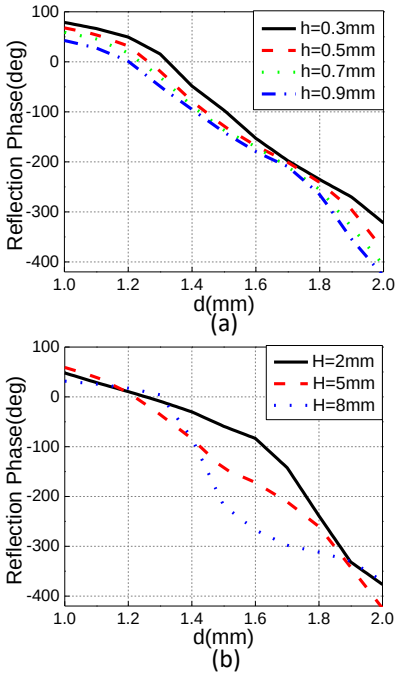


Fig. 13. (a) Reflection phase of the element for different h. (b) Reflection phase of the element for different H.

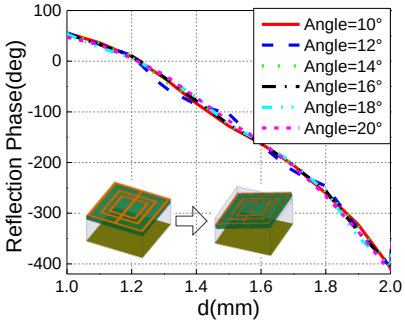


Fig. 14. Reflection phase of the element for different rotation angles.

The CGP of the reflectarray antenna is shown in Fig. 15 (a), which has a smooth surface and is used for most reflectarray antennas. Fig. 15 (b) shows the INSGP, which is designed by the method of "Design of the NSGP based on the law of reflection" in Section III. B. The NSGP II analyzed in Section IV is designed as shown in Fig. 15 (c). Compared with the INSGP, the NSGP is easier to manufacture and can also effectively improve the gain of reflectarray antennas.

The reflectarray antenna with different ground planes discussed above is designed and shown in Fig. 16 respectively. It is worth noting that the reflectarray is exactly the same except for the ground planes.

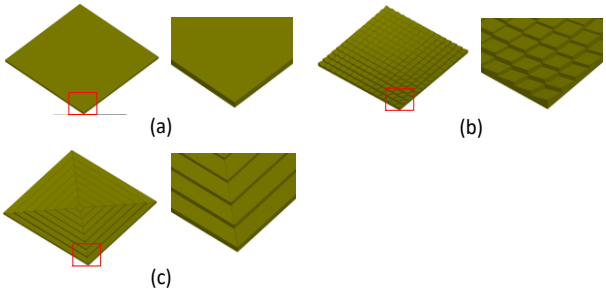


Fig. 15. Geometry of the ground planes. (a) The CGP, (b) the INSGP, and (c) the NSGP II.

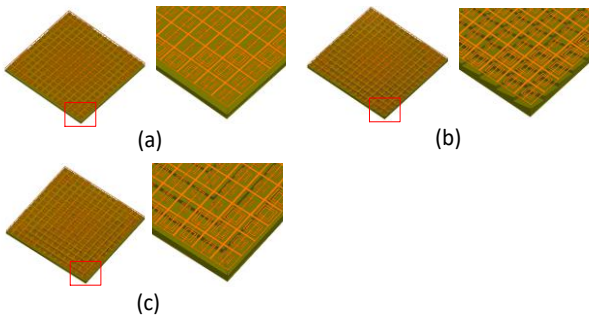


Fig. 16. Geometry of the reflectarray with different ground planes. (a) The reflectarray with the CGP. (b) The reflectarray with the INSGP. (c) The reflectarray with the NSGP II.

The radiation patterns of the reflectarray antenna with different ground planes above are simulated and the results are shown in Fig. 17 respectively.

It can be observed that the gain of the reflectarray with the CGP is 24.610 dB, and the gain of the reflectarray with the INSGP is 25.279 dB. Compared with the CGP, the INSGP can reflect the incident waves in the main beam direction, so the losses of the reflectarray antenna caused by the reflection of the ground plane are reduced, and the gain of the reflectarray is increased by 0.669 dB. It can be seen that the gain of the reflectarray with the NSGP is 25.015 dB, which is 0.405 dB higher than the gain of the reflectarray with the CGP and 0.264 dB lower than the gain of the reflectarray with the INSGP.

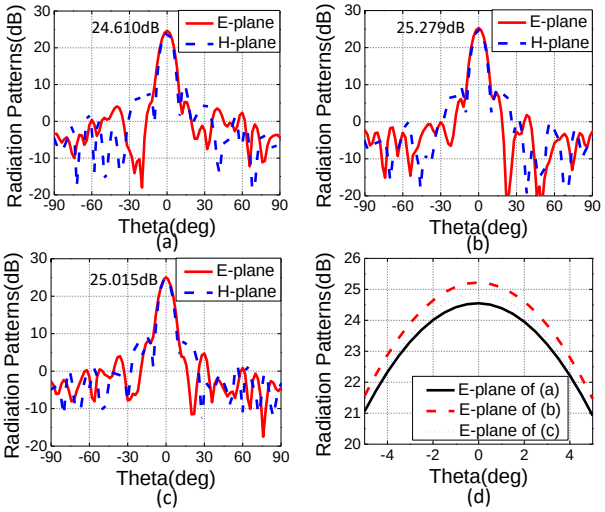


Fig. 17. Radiation patterns of the reflectarray. (a) The radiation patterns of the reflectarray with the CGP. (b) The radiation patterns of the reflectarray with the INSGP. (c) The radiation patterns of the reflectarray with the NSGP. (d) The comparison of the radiation patterns.

From the comparisons of the radiation patterns, the gain of the reflectarray from high to low is the reflectarray with the INSGP, the reflectarray with the NSGP and the reflectarray with the CGP. Considering the difficulty associated with manufacturing in practice, the NSGP greatly reduces the difficulty of manufacturing compared with the INSGP, and the losses caused by the reflection of the ground plane can be effectively reduced.

C. Design of the reflectarray with offset feed.

Considering the problem of feed blockage, a 15×15 reflectarray antenna fed by a linearly polarized horn with an offset angle of $\theta_i = 15°$ with respect to the broadside direction is designed as shown in Fig. 18.

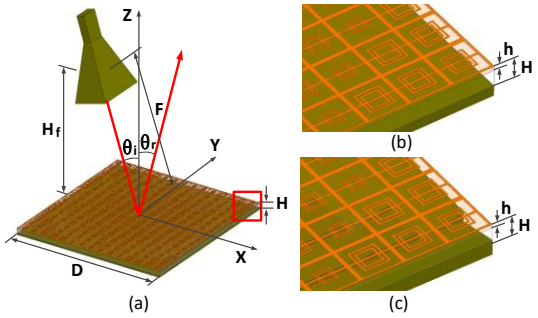


Fig. 18. Geometry of the reflectarray with offset feed. (a) The reflectarray with the CGP. (b) The reflectarray with the NSGP.

The element discussed in Section V is adopted to make the reflectarray antenna. The square aperture of the reflectarray D is 150 mm, and the focal length F = 120 mm, which indicates an F / D ratio of 0.8. The position of offset feed horn is $x_{feed} = -31.06mm$, $y_{feed} = 0mm$, $z_{feed} = 115.91mm$ based on the coordinate system in Fig. 18. The NSGP is designed as shown in Fig. 19 based on the analysis in Section IV, and the slopes of the element ground planes are shown in Table 1.

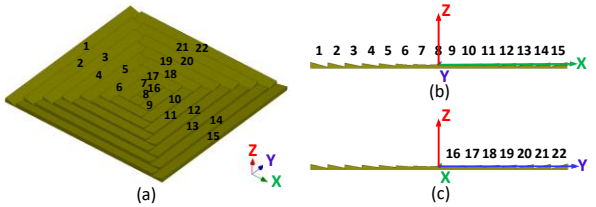


Fig. 19. Geometry and configuration of the NSGP.

Table 1: Slopes of the element ground planes of the NSGP

Number	Slope	Number	Slope
1	0.2320	12	0.1710
2	0.2050	13	0.2150
3	0.1760	14	0.2588
4	0.1450	15	0.3016
5	0.1120	16	0.0431
6	0.0767	17	0.0856
7	0.0393	18	0.1273
8	0	19	0.1677
9	0.0411	20	0.2065
10	0.0835	21	0.2435
11	0.1270	22	0.2785

The reflectarray is measured by NSI planner near-field system and the photographs of the reflectarray antenna are shown in Fig. 20. It is worth noting that two measurements of the reflectarray are exactly the same except for the ground planes.

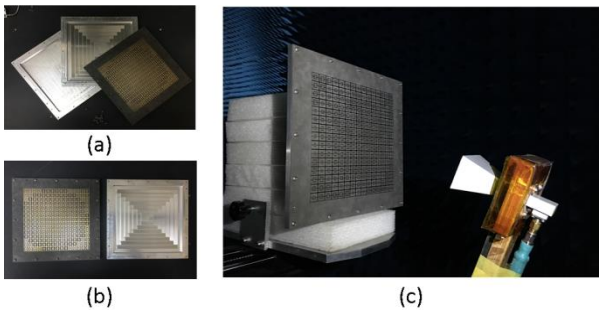


Fig. 20. (a), (b) Photographs of the reflectarray, and (c) near-field measurement.

The simulated and measured radiation patterns in both E and H plane at the center frequency 13.38GHz

are shown in Fig. 21. Good agreement between the simulated and measured results is observed. The highest measured side lobe level (SLL) is about -18.3dB, and the cross polarization discrimination is less than -20dB at the center frequency 13.38GHz. The comparisons of the measured gain between the reflectarray with the NSGP and CGP are shown in detail in Fig. 22.

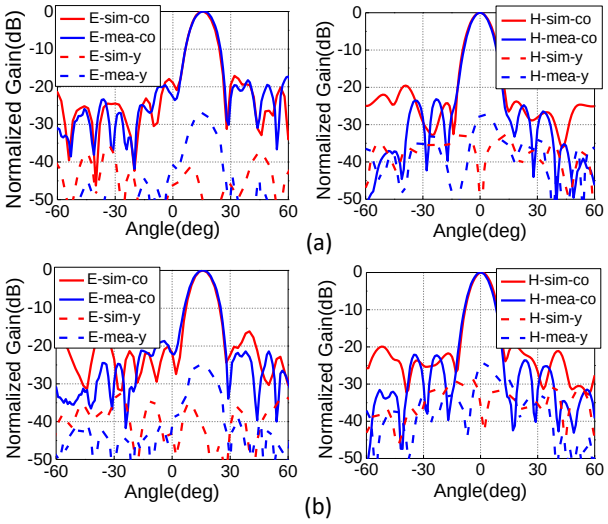


Fig. 21. Radiation patterns of the reflectarray. (a) Radiation patterns of the reflectarray with the NSGP. (b) Radiation patterns of the reflectarray with the CGP.

From Fig. 22 (a), it can be seen that, for the normalized gain of the reflectarray with the NSGP in E-plane, the left SLL, right SLL and cross polarization are -19.531dB, -19.575dB, and -26.908dB respectively, the average SLL is -19.553dB, and the right SLL is 0.024dB higher than the left SLL. For the reflectarray with the CGP, the left SLL, right SLL and cross polarization discrimination are -18.648dB, -21.433dB, and -24.735dB respectively, the average SLL is -20.0405dB, and the left SLL is 2.785dB higher than the right SLL. From the comparisons between the SLLs of the reflectarray with the NSGP and CGP, it can be observed that although the right SLL and average SLL of the reflectarray with CGP is lower than the reflectarray with the NSGP, the difference between the left SLL and right SLL, and the cross polarization are higher than the reflectarray with the NSGP, which means the reflectarray with the NSGP has a more balanced SLL and a higher gain than the reflectarray with the CGP.

From the comparison between the normalized gain of the reflectarray with the NSGP and CGP in H-plane as shown in Fig. 22 (b), it can be observed that the reflectarray with the NSGP has a lower left SLL, right SLL and cross polarization, which will contribute to improve the gain and efficiency of the reflectarray.

The measured gain of the reflectarray with the NSGP and CGP are given in Fig. 23. The gain of the reflectarray with the NSGP varies from 24.413dB to 24.967dB and is average 0.54dB higher than the CGP reflectarray within the working frequency of 12.88GHz to 13.88GHz. The aperture efficiencies are also shown in Fig. 23, the efficiency of the reflectarray with the NSGP varies from 50.1% to 54.5%, peaking at 12.88GHz, and is increased from 4.2% to 7.9% compared with the reflectarray with CGP.

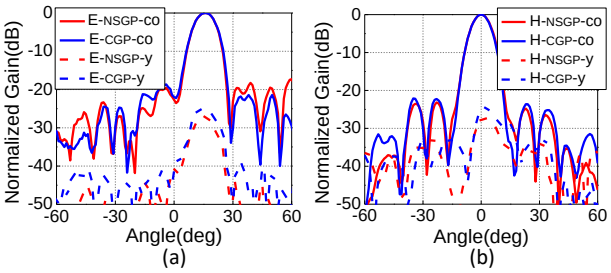


Fig. 22. Comparisons of the measured gain. (a) The comparison of normalized gain with the NSGP and CGP in E-plane. (b) The comparison of normalized gain with the NSGP and CGP in H-plane.

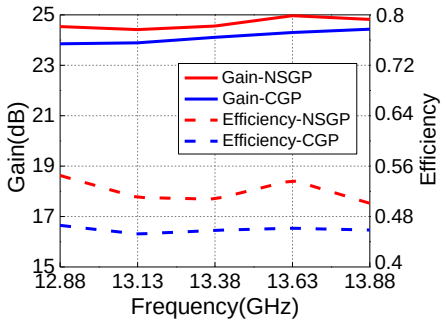


Fig. 23. Measured gain and aperture efficiency of the reflectarray with the NSGP and CGP.

It is worth to mention that although the improved gain and efficiency of reflectarray is not very large, the NSGP provides a high accurate design of the reflectarray and another method to increase the gain, that is to say, we can combine the sub-wavelength elements, metal only elements, some special materials as well as the NSGP to improve the gain and efficiency of the reflectarray. Based on the principle of the NSGP, the gain and efficiency will be improved significantly for a large reflectarray and the NSGP can be used in the place where requires the low-loss, high gain and high accurate design of the reflectarray.

VI. CONCLUSION

A NSGP for a low-loss reflectarray is proposed based on the principle of the reflector antenna and the law of reflection. The principle and losses of the

reflectarray are studied first. Then, a NSGP is proposed that can reflect the incident waves to the main beam direction, therefore reducing the losses of the reflectarray. To verify the design concept of the proposed NSGP in reducing the losses of the reflectarray antenna, a reflectarray with different ground planes is simulated, fabricated and tested respectively, the results show that the NSGP can increase the accurate design of the reflectarray and has an average higher gain and efficiency of 0.54dB and 6.1% compared with the CGP reflectarray within the working frequency of 12.88GHz to 13.88GHz. With the development of the 3D print technology, the NSGP can be fabricated easily, and the NSGP is promising for the applications where a high accurate design such as the beam forming, low-loss and high-efficiency reflectarray is needed.

REFERENCES

[1] J. Shaker, M. R. Chaharmir, M. Cuhaci, and A. Ittipiboon, "Reflectarray research at the communications research center canada," *IEEE Antennas Propag. Mag.*, vol. 50, no. 4, pp. 31-52, Aug. 2008.

[2] M. Karimipour, and A. Pirhadi, "A novel approach to synthesis of non-uniform conformal reflectarray antennas," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 28, no. 11, pp. 1040-1048, Nov. 2013.

[3] M. Arrebola, E. Carrasco, and J. A. Encinar, "Beam scanning antenna using a reflectarray as sub-reflector," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 26, no. 6, pp. 473-483, June 2011.

[4] D. G. Berry, R. G. Malech, and W. A. Kennedy, "The reflectarray antenna," *IEEE Trans. Antennas Propag.*, vol. 11, pp. 645-651, 1963.

[5] R. Deng, F. Yang, S. Xu, and M. Li, "A 100-GHz metal-only reflectarray for high-gain antenna applications," *Antennas Wireless Propag. Lett*., vol. 15, pp. 178-181, 2016.

[6] H. Rajagopalan and Y. Rahmat-Samii, "Dielectric and conductor loss quantification for microstrip reflectarray: Simulations and measurements," *IEEE Trans. Antennas Propag*., vol. 56, no. 4, pp. 1192-1196, 2008.

[7] D. R. Jackson, S. A. Long, J. T. Williams, and V. B. Davis, "Computer-aided design of rectangular microstrip antennas," in K. F. Lee and W. Chen (eds.), *Advances in Microstrip and Printed Antennas*, New York, John Wiley & Sons, Chapter 5, 1997.

[8] D. M. Pozar, S. D. Targonski, and H. D. Syrigos, "Design of millimeter waves microstrip reflectarray," *IEEE Trans. Antennas Propag*., vol. 45, no. 2, pp. 286-295, Feb. 1997.

[9] F. S. Farida, P. M. Hadalgi, P. V. Hunagumd, and S. R. Ara, "Effect of substrate thickness and permittivity on the characterisitics of rectangular microstrip antenna," *Precision Electromagnetic Measurements Digest*, pp. 598-599, July 1998.

[10] A. Sabban, "A comprehensive study of losses in mm-waves microstrip antenna arrays," *27th European Microwaves Conference*, vol. 1, pp. 163-167, Oct. 1997.

[11] M. Bozzi, S. Germani, and L. Perregrini, "A figure of merit for losses in printed reflectarray elements," *Antennas Wireless Propag. Lett*., vol. 1, no. 3, pp. 257-260, 2004.

[12] J. Huang and J. A. Encinar, *Reflectarray Antennas.* New York, N Y, USA: Wiley, 2007.

[13] P. Nayeri, A. Elsherbeni, and F. Yang, "Design, full-wave simulation, and near-field diagnostics of reflectarray antennas using FEKO electromagnetic softeave," *28th Annual Review of Progress in Applied Computational Electromagnetics (ACES)*, Columbus, Ohio, pp. 503-508, Apr. 2012.

[14] O. Bucci, G. Franceschetti, G. Mazzarella, and G. Panariello, "Intersection approach to array pattern synthesis," *IEEE Proceedings*, vol. 137, no. 6, PP. 349-357, Dec. 1990.

[15] J. Zornoza, R. Leberer, J. Encinar, and W. Menzel, "Folded multilayer microstrip reflectarray with shaped pattern," *IEEE Trans. Antennas Propag.*, vol. 54, no. 2, pp. 510-518, Feb. 2006.

[16] G. Zhao and Y. Jiao, "Design of broadband dual-polarization contoured-beam reflectarray for space applications," *28th Annual Review of progress in Applied Computational Electromagnetics (ACES)*, Columbus, Ohio, pp. 790-794, Apr. 2012.

[17] A. Trastoy, F. Ares, and E. Moreno, "Phase-only synthesis of non-φ-symmetric patterns for reflectarray antennas with circular boundary," *Antennas Wireless Propag. Lett*., vol. 3, pp. 246-248, 2004.

[18] D. M. Pozar and T. A. Metzler, "Analysis of a reflectarray antenna using microstrip patches of variable size," *Electronics Letters*, vol. 29, no. 8, PP. 657-658, 1993.

[19] X. Fei, "Study on broadband dual-band reflectarray and broadband transmitarray antennas," Ph.D. dissertation, *Electromagnetic Field and Microwaves Technology*, Chinese Academy of Sciences, China, pp. 33-35, 2017.

[20] W. An, S. Xu, and F. Yang, "A metal-only reflectarray antenna using slot-type elements," *Antennas Wireless Propag. Lett*., vo. 13, pp. 1553-1556, 2014.

[21] G. Perez-Palomino, J. A. Encinar, M. Barba, and E. Carrasco, "Design and evaluation of multi-resonant unit cells based on liquid crystals for reconfigurable reflectarrays," *IET Microwaves, Antennas & Propagation*, vol. 6, no. 3, pp. 348-354, 2012.

[22] E. Carrasco and J. Perruisseau-Carrier, "Reflectarray antenna at terahertz using graphene," *Antennas Wireless Propag. Lett*., vol. 12, pp. 253-256, 2013.

[23] H. Li, B. Z. Wang, G. Zheng, W. Shao, and L. Guo, "A reflectarray antenna based on FSS for low RCS and high radiation performances," *Progress In Electromagnetics Research C*, vol. 15, pp. 145-155, 2010.

[24] L. F. Van Buskirk and C. E. Hendrix, "The zone plate as a radio frequency focusing element," *IEEE Trans. Antennas Propag.*, vol. 9, pp. 319-320, May 1961.

[25] V. Manohar, J. M. Kovitz, and Y. Rahmat-Aamii, "Synthesis and analysis of low profile, meta-only stepped parabolic reflector antenna," *IEEE Trans. Antennas Propag.*, vol. 66, no. 6, pp. 2788-2797, June 2018.

[26] B. Khayatian and Y. Rahmat-Sam, "A novel antenna concept for future solar sails: application of fresnel antennas," *IEEE Antennas Propag. Mag.*, vol. 46, no. 2, pp. 50-63, Apr. 2004.

[27] J. Gutierrez-Rios and J. Vassal'lo, "Technology aspects of Fresnel Zone reflectoers," *Advances on Antennas, Reflectors and Beam Control*, edited by A. Tazon, Research Signpost, Kerala, India, 2005.

Jiawei Ren was born in Hebei, China, in 1992. He received the B.S. degree from the Hebei University, Hebei, China, in 2015, and the M.S. degree from the University of Chinese Academy of Sciences, Beijing, China, and Nation Space Science Center, Chinese Academy of Sciences, Beijing, China in 2019. He is currently pursuing the Ph.D. degree in the University of Chinese Academy of Sciences, Beijing, China, and Nation Space Science Center, Chinese Academy of Sciences, Beijing, China.

His current research interests include reflectarray, microstrip antennas, low-loss antenna, wideband antenna.

Hongjian Wang received the B.S. and M.S. degrees from Wuhan University, Ph.D. degree from the Beijing Institute of Technology, Beijing, China.

He began his career with the radar group, China Airborne Missile Academy as a Radar Antenna Engineer. From 2002 to 2003, he was a Research Assistant in City University of Hong Kong, Hong Kong. He is currently a Professor of the National Space Science Center, Chinese Academy of Science, Beijing, Course Professor of University of CAS, Beijing. His current research interests include Space-borne Antenna theory and technology, inflatable/deployable antenna, mm-wave/ Terahertz antenna, mm-wave components, antenna measurement.

Weichun Shi was born in Gansu, China, in 1996. He received the B.S. degree from the Beijing Jiaotong University, Beijing, China, in 2018, and he is currently pursuing the M.S. degree with the University of Chinese Academy of Sciences, Beijing, China.

Minzheng Ma was born in Shandong, China, in 1997. He received the B.S. degree from the Southwest Jiaotong University, Chengdu, China, in 2019, and he is currently pursuing the M.S. degree with the University of Chinese Academy of Sciences, Beijing, China.

SDR Based Modulation Performance of RF Signal under Different Communication Channel

S. Habib

Department of Information Technology, College of Computer
Qassim University, 6633, Buraidah, 51452, Saudi Arabia
s.habibullah@qu.edu.sa

Abstract ─ Hardware components are an integral part of Hardware Define Radio (HDR) for seamless operations and optimal performance. On the other hand, Software Define Radio (SDR) is a program that does not rely on any hardware components for its performance. Both of the latter radio programmers utilize modulation functions to make their core components from signal processing viewpoint. The following paper concentrates on SDR based modulation and their performance under different modulations. The bit error rate (BER) of modulations such as PSK, QAM, and PSAM were used as indicators to test channel quality estimation in planar Rayleigh fading. Though it is not commonly used for channel fading, the method of the adder determines the regionally segmented channel fading. Thus, the estimation error of the channel change substantially reduces the performance of the signal, hence, proving to be an effective option. Moreover, this paper also elaborates that BER is calculated as a function of the sample size (signal length) with an average of 20 decibels. Consequently, the size of the results for different modulation schemes has been explored. The analytical results through derivations have been verified through computer simulation. The results focused on parameters of amplitude estimation error for 1dB reduction in the average signal-to-noise ratio, while the combined amplitude deviation estimation error results are obtained for a 3.5 dB reduction.

Index Terms ─ Bit error rate, receiver operation, RF signal, signal noise ratio, transmitter, wireless channel.

I. INTRODUCTION

The domain of mobile communication continues to have a greater impact on our daily life as compared to other technological alterations. In fact, this field has been witnessing the fastest pace of changes since the dawn of the 21st century with regard to design. At the same time, the services offered by mobile communication has revolutionized the current practices and systems in areas of health, finances, and education, across the world by accelerating transactions that would otherwise takes months to complete and completing them within a matter of a few hours. In more ways than one, mobile communication has extended the scope of business in unimaginable ways. One such transformative change has emerged as a significant reduction in costs made possible in the wake of modern tools used to design a wireless system [1]. Simulation is advantageous in that it can lower the cost associated with design testing, despite potentially necessitating investments in computing resources. Cognitive radio plays an important role in the reconfiguration of HDR in wireless communication. Based on the reconfiguration requirements antennas are designed. The designed antennas are used for cognition operation such as reconfiguration [2].

Channel modeling prepares any wireless communication system's core component to help determine whether packets are unable to reach the supposed destination, a phenomenon referred to as packet-loss. The wireless system's simulation could encompass channel coding, speech coding, as well as other issues pertaining to interleaving in modulation. It is possible to use different methods to estimate the wireless system's the overall performance by simulating the channel under different conditions [3]. Certain models leverage the simulation of bit-error-rate (BER) or signal-to-noise ratio (SNR), while others may concentrate on the as alterations occurring over a longer duration, including packet error rate or segmentation [4], [5].

The objective of this article is to provide radio frequency signals for various modulation modes in Rayleigh fading channels on the basis of data and pilot symbols. To that end, a number of modulation techniques are utilized for performance assessment by using the RF signal's BER. These simulation findings are found to be in alignment with our analytical results.

The remainder of this paper is organized in the following manner: Section II presents related work while the technique of the proposed work is explained in Section III. The channels using for SDR are discussed in Sections IV, V, VI and Section VII. Experimental results with associated discussions are given in Section VIII.

Submitted On: March 30, 2021
Accepted On: August 12, 2021

https://doi.org/10.47037/2021.ACES.J.360813

Conclusion and future research directions are drawn in Section IX.

II. RELATED WORK

Wireless channel models are commonly used to study the performance of transmission or link-layer protocols using simulation. For example, the performance of transmission protocol (TCP) over the wireless link is studied by Chaskar et al. [6]. Analysis of TCP/ IP over wireless connections is presented by Cheyenne et al. [7]. Leibniz [8] studied the performance of error correction code on wireless connections. In all these cases, a frame loss model has been applied to the surface of the link layer. In addition, the performance of other communication protocols over wireless links (such as ATMs [9], [10]) is also studied by simulation. Channel detection has proven an effective technique in M-QAM demodulation as it accurately compensates for channel dimension and phase distortion [11]. Many authors have studied channels in the voice of PSAM [12], [13] and these relays have proven to be useful for fading channels. Previous studies on the performance of PSAM M-QAM were primarily based on computer simulation and experimental implementation. The only result of the analysis is that the upper limit of the 16-QAM [14] symbol error rate is strict. These results provide an efficient method to evaluate the performance of various system design parameters.

III. PROPOSED FRAMEWORK

In an attempt to transmit information from one point to another, the signal has to be sent through the medium to reach the recipient. The path from the transmitter to the receiver is referred to as a channel. Some examples of channels include copper cable, fiber optic cable, or space. The channel characterizing features may include.

A. Additive white Gaussian noise channel

The channel model commonly used in communication system analysis is that of the Additive White Gaussian Noise (AWGN) channel, and it is longer and easier than the Gaussian Noise Channel.

The term "additive" refers to the superposition or addition of noise to a signal, thus limiting the receiver's ability to make decisions about the correct signal and the rate of information. Therefore, AWGN is the effect of thermal noise generated through the movement of electrons in all electronic components (Resistors, wires, etc.) featuring dissipation property [15].

For wireless channels, their properties are usually determined by specific locations, atmospheric effects, transmission objects, multi-pathing effects, and so on. Transmitter and receiver in this study are assumed to be fixed as a state of default and for the line of sight (LOS). In other words, the transmitter and receiver are not moving, and the theory between each other is very intuitive.

The rationale for a fixed LOS wireless channel is the AWGN channel, which has frequency selectivity in the case of matte channels, rather than frequency selectivity. Given that, both, the frequency converter and the receiver are default, this study does not take into account signal delay and the use of AWGN to terminate the signal in such a mobile communication channel.

Assume that the AWGN channel bandwidth has a constant power spectral density (PSD) and Gaussian amplitude probability density. As depicted in Fig. 1, this Gaussian noise is incorporated into the transmitted signal before it is received by the receiver. Mathematically, thermal noise is expressed by the zero-mean Gaussian random process, where the signal is a random variable of Gaussian noise, and the DC signal as (1):

$$\boldsymbol{r(t) = s(t) + n(t)}. \quad (1)$$

In the formula defined in equation (1), r(t) is the received signal, s(t) is the transmitted signal and n(t) is the noise signal between the transmitted input and the received output signal.

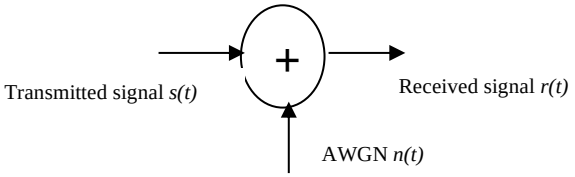


Fig. 1. The Gaussian channel diagram

In this model, the noise power (white noise) of uniform spectral density is incorporated into the actual signal. The result of the distribution of sound is the Gaussian process at zero average. Although not always realistic, this assumption simplifies the mathematical process associated with estimating the performance of a given communication system. In fact, most of the BER curves are generated by analyzing Gaussian noise channels.

B. Rayleigh fading channel

Since the signal is propagated in the air and near the ground, in addition to the influence of the freeway loss Ls, the most important effect of signal attenuation is the multipath propagation effect. This effect will cause the amplitude, phase and angle to fluctuate in the received signal due to multipath blur.

In general, mobile communication has two blurring effects: large-scale blurring and small-scale blurring. Large-scale blurring, when moving over a large area, represents an increase in average signal strength or damage along the way due to the effects of shadows. On the other hand, small-scale fading refers to sharp changes in signal amplitude and phase, caused by minute changes

in the spatial separation between the receiver and the transmitter (as small as a half-wavelength). Small-scale fading is also called Rician fading because the signal envelope received can be expressed by the Rician Probability Density function (PDF).

IV. MULTIPATH CHANNEL

The most destructive feature of the mobile radio in the communication system is its gradual disappearance of path. The author focused on multipath characteristics of the channel with high and low waves travelling in different directions before reaching to the receiver antenna. The experimental test were taking in the daytime with fixed transmitter antenna and moved the receiver antenna. This radio signal can reach the receiver after various delays, in amplitudes and phases due to multiple interferences of multi-fading in the channel. The receiver antenna is installed in a vehicle to measure the received signal with different disappearances shown in Fig. 2.

Therefore, in a multipath fading channel, the pulse transmitted signal at the receiver end receives multiple pulses due to multiple fading in the channel. These fading can blur the amplitude and phases of the received signal, can lead to careful fluctuations, which effect the transmitting information and reliability. This problem causes very small destructive interventions in recipients [16] where many textbooks have extensively introduced multidimensional minerals for optimal results. The main features of the multipath fading channel with frequency non-selective fading are introduced in the following subsections.

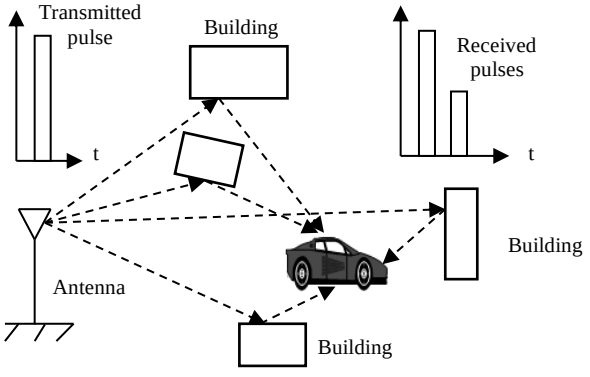


Fig. 2. The Gaussian channel diagram.

V. FADING CHANNEL CHARACTERIZATION

As mentioned earlier, the delay and destructive feature will reflect to the transmitting behaviour of the signal. Therefore, multiple pulses received by recipients cannot solve the problem arise by multiple path. The problem can be reduced, if the transmitting signal can transmit a very short signal (pulses in an ideal case). These transmitting signals can be transmitted to different fading channels at different times and on the receiver side, it can be received in the form of a series of pulses as shown in Fig. 3.

For the use of multipath channel, the signal received at different time slot in order to receive the transmitting signal seems random and unpredicted. Therefore, it is important to characterize the channel from a statistical point of view. Thus, first check the channel influence on the transmission signal, usually as (2):

$$\boldsymbol{S_P(t) = Re\{s(t)e^{j2\pi f_c t}\}}, \quad (2)$$

Where;

Re - the real part;

$S_P(t)$ - band-pass transmission pulse;

s(t) - baseband input signal whose bandwidth is limited by the filter in the transmitter with a carrier frequency f_c.

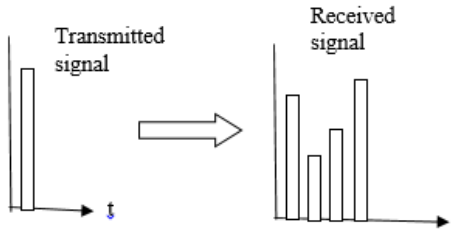


Fig. 3. The Gaussian channel diagram.

The multiple propagation paths associated with each path has propagation delay τ(t) and attenuation factor $a_n(t)$. Due to the change of the medium structure in the wireless communication system, the propagation delay and the amplitude attenuation factor have all shown to be time-varying. Therefore, the bandpass signal received after multipath propagation can be derived as (3):

$$\boldsymbol{x(t) = \sum_n a_n(t) s_p\{t - \tau_n(t)\}}, \quad (3)$$

Where;

$\boldsymbol{a_n(t)}$ - amplitude attenuation factor of the received signal in the nth path;

$\boldsymbol{\tau_n(t)}$ - propagation time delay of the nth path.

Substituting, $\boldsymbol{S_P(t)}$ from equation (4) into equation (5) gives the result:

$$\boldsymbol{x(t) = Re\left[\sum_n a_n\, s[t - \tau_n(t)] e^{j2\pi f_c[t-\tau_n]}\right]}, \quad (4)$$

and the baseband filter received the signal at the receiver side as (5):

$$\boldsymbol{x(t) = Re\left[\sum_n a_n\,(t) e^{-j2\pi f_c \tau_n(t)} s[t - \tau_n(t)]\right]}. \quad (5)$$

Since $\boldsymbol{x(t)}$ is the filter response to the input signal, $\boldsymbol{s(t)}$, it can be expressed as h(τ,t) by using the time-varying impulse response as (6):

$$\boldsymbol{h(\tau, t) = \sum_n a_n\,(t) e^{-j2\pi f_c \tau_n(t)} s[t - \tau_n(t)]}. \quad (6)$$

Given that $\boldsymbol{f_c}$ is the transmitted unmodulated carrier frequency and $\boldsymbol{s(t) = 1}$ for all $\boldsymbol{t}$, then the received signal reduces as depicted in (7):

$$\boldsymbol{r(t) = \sum_n a_n\,(t) e^{-j2\pi f_c \tau_n(t)} = \sum_n a_n\,(t) e^{-j\theta_n(t)}}, \quad (7)$$

where $\boldsymbol{\theta_n(t) = j2\pi f_c \tau_n(t)}$.

Thus, it is concluded that the received signal is a summary of the amplitude and phase with varying vectors of different times. Since the changes of $\boldsymbol{a_n(t)}$ and $\boldsymbol{\theta_n(t)}$ occur on different scales, when the randomly changing vector is added catastrophically, the multipath propagation model in equation (9) disappears with a strong signal. When this happens, the input signal received is too small or almost zero. Due to the effect of $\boldsymbol{a_n(t)}$ and $\boldsymbol{\theta_n(t)}$, the received signal r (t) can also be changed as a random operation. Therefore, rewriting the fading response as in equation (8):

$$\boldsymbol{h(\tau, t) = \sum_n a_n(t) e^{-j\theta_n(t)} s[t - \tau_n(t)]}, \quad (8)$$

where $\boldsymbol{h(\tau, t)}$ the modulated process with varying time ($\boldsymbol{t}$). Central limit theorem in [16] explains the complex value of Gaussian filters with a different number of paths. When there are a large number of paths, r(t) can be simplified to a complex-valued Gaussian stochastic process. Therefore, $\boldsymbol{r(t)}$ can be simplified as channel fading response by Gaussian filter process based with variable time ($\boldsymbol{t}$).

VI. DOPPLER SHIFT

The Doppler Effect leads to having different signals shift at one time which when combined at some other time cancel the effect of fading due to multiple paths staying as time-dependent parameters. Therefore, fast and slow fading normalize the maximum effect of the fading rate as described in [17] and are reproduced as in (9):

$$\boldsymbol{f_d = \frac{f_m}{B} = \frac{f_c\, v cos\propto}{B}}, \quad (9)$$

Where;

$\boldsymbol{f_m}$- maximum Doppler shift frequency;

$\boldsymbol{B}$ - bandwidth of the baseband signal;

$\boldsymbol{v}$- speed of the wireless signal;

$\boldsymbol{f_c}$ - carrier frequency;

$\propto$ - arrival angle of the path with the maximum Doppler shift frequency;

c - speed of light constant.

VII. FREQUENCY OF SLOW RALEIGH FADING CHANNEL

Channel fading can also be classified as frequency selective fading or frequency non-selective repetitive fading as given in [18] by $\Delta \boldsymbol{f_{coh}}$ as:

$$\Delta \boldsymbol{f_{coh} = \frac{1}{T_m}}. \quad (10)$$

If the bandwidth of the blurred channel is less than the bandwidth (Δfcoh) of the transmission signal, the channel is said to be the frequency selection attenuation. In this case, the channel will severely distort the signal and may cause inter-signal interference. In frequency non-selective fading channels, all frequency components present in the transmission signal experience almost the same focus and phase shift.

According to [16] and [18], in the frequency non-selective fading channel, the received signal reaches the receiver through the fading path. Therefore, the signal can be simplified as the product of the transmitted signal and α, showing time-varying features of fading multipath channels.

To make the analysis simpler, it has been omitted. Also, unlike signals, noise does not necessarily attenuate the number of multipath channels that come into the channel. In this communication system, it is assumed that AWGN has a connected power spectrum, and that the deviation of its centre frequency will not change its statistical properties. It is generally assumed that the band Gaussian has got a circular symmetry in complex Gaussian noise.

By itself, the real and imaginary parts of random variables are free, and so is the Gaussian distribution. Carrier frequency and phase offset will not change its statistical characteristics. Therefore, noise can be expressed as an additional term in the received signal expression.

Considering the above discussion, the frequency non-selective and dim Rayleigh fading channel can be approximated as a multiplication factor of the transmission signal. Therefore, for noise, the received signal can be expressed in the same way:

$$\boldsymbol{r(t) = c(t) \times s(t) + n(t)}, \quad (11)$$

Where;

$\boldsymbol{r(t)}$ - received signal;

$\boldsymbol{c(t)}$- Fading distortion, its envelope has a Rayleigh distribution;

$\boldsymbol{s(t)}$- Transmitted signal;

n(t) - average of zero additive white Gaussian noise and the power spectral density N_0

The systematic convolution code (SCC) model is used for transmitted signal where Fading and AWGN signal appears as the part of received signal. The SCC model is shown in Fig. 4 encoded the representation of signal. The SCC decision is used for soft decision with the configurable frequency factor.

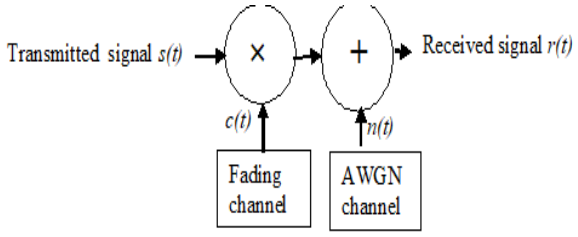


Fig. 4. Transmission channel model.

VIII. EXPERIMENTAL RESULTS AND DISCUSSION

Antenna plays an important role in the reconfigurable procedure of wireless communication. The

adaptation reconfigurable antennas is used to maximize the performance by operation during communication. The common factor that is used for SDR algorithms are used as RF reconfiguration. The Matlab program has been used for SDR simulation to evaluate the efficiency of RF signal transmission. Transmitted and received signals of Fig. 5 show the BER and SNR, and the acquisition performance of the proposed system to indicate the transmitted and received signals using the proposed SDR system. It can be concluded that the nature and shape of the sent and received signals are the same. However, due to the noise and filtering effect of the synchronization in the transmission, the signal size will vary but the overall transmission remains accurate.

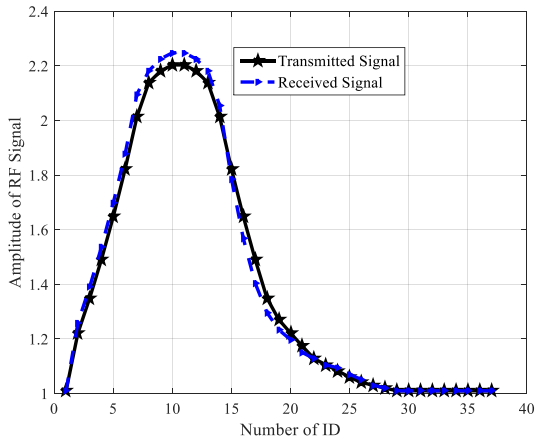


Fig. 5. Transmitted and received signal using proposed SDR system.

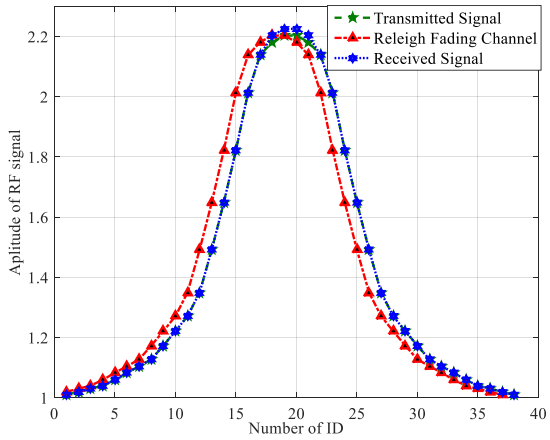


Fig. 6. BER degradation using SDR system.

Figure 6 shows a typical SNR value of 4dB. Iterative decoding algorithms have been used to obtain transmitted RF signals using Rayleigh fading channels. The number of repetitions affects the bit error rate. In Fig. 8, it can be seen that the performance of the system is well estimated in a large number of numbers. Further, the decline in the performance of the SDR-based BER modulation scheme provides an approximate value with slight fluctuations compared to the input value. The performance index can be estimated by plotting the relationship between BER and SNR as EB / No. It can be seen that this fluctuation will cause BER to drop and reduce efficiency due to the influence of probability error on the detection process as compared to the input probability error. It has also been observed that BER decreases with increasing SNR. In the PSAM modulation scheme, the BER degradation is almost parallel to the input degradation curve.

Figure 7 shows the BER performance of RF data communication via the SDR-type digital modulation scheme on the fading channel. In all cases, the performance of the proposed system in PSAM and QAM will be low, and the performance in PSK modulation will be satisfactory. For the typical 4 dB SNR value, the BER values of the PSK, QAM and PSAM modules are 0.002035, 0.5086 and 0.7586, respectively. The system performance is better from 10 to 18 dB. From ambient SNR (4-16 dB) values, the system shows almost flat degradation performance. For PSK, when SNR is greater than 12 dB, BER is close to zero. It can be seen from Figure 7 that the performance of QAM and PSAM system is poor compared to PSK using the receiver.

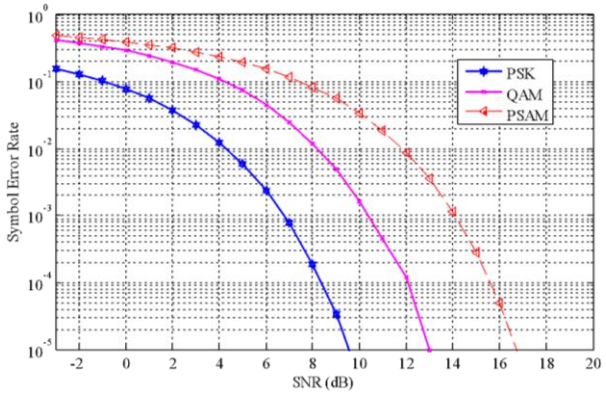


Fig. 7. SDR performance under different modulation schemes.

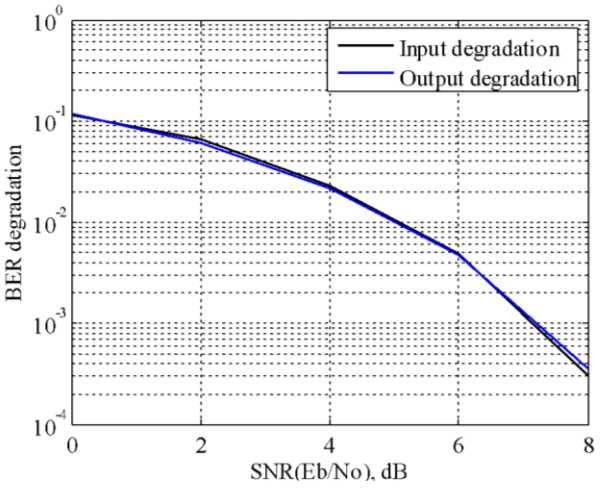


Fig. 8. BER performance of RF transceiver in Rayleigh fading channel.

As illustrated in Fig. 8, the performance of the system is thoroughly examined in a large number of iterations. It shows that the decline in performance of the SDR-based BER modulation scheme provides an estimate with a slight change compared to the input value. Measure performance indicators can be found due to the detection process and the probability error impact on the input probability error index. These changes tamper with the efficiency of the BER and, hence, the performance. In the PSAM modulation scheme, the BER degradation curve is almost parallel to the input degradation curve.

IX. CONCLUSION

This paper designs the SDR system based on RF simulation. The SNR of 10 dB has been used to test the degradation of the performance of the synthetic softening channel. The results obtained indicate PSAM with the optimal performance in terms of BER under various roll-off coefficients, thus, reflecting the performance of the proposed SDR. It can be seen that a fixed SNR with a higher roll-off factor will reflect lower results. Consequently, the RF transmission capacity is improved. Therefore, it can be concluded that in case of high bit transmission rate requirements, the PSAM used as a modulation scheme stands a chance of achieving better transmission efficiency.

ACKNOWLEDGMENT

We acknowledge the overall paper editing support by Dr. Sheroz Khan and Dr. Muhammad Islam, Department of Electrical Engineering and Renewable Engineering, Onaizah College of Engineering & Information Technology, Al-Qassim, Saudi Arabia; 2053, Saudi Arabia. Zarak S. Khan is acknowledged for proof reading.

REFERENCES

[1] O. Popescu, S. El-Tawab, S. Abraham, and S. Abraham, "A mobile platform using software defined radios for wireless communication systems experimentation," *ASEE Annual Conference & Exposition*, Columbus, Ohio, pp. 1-12, 2017.

[2] Z. Zhang, L. Xiao, X. Su, J. Zeng, and X. Xu, "A channel estimation method based on the improved LMS algorithm for MIMO-OFDM systems," *12th International Symposium on Medical Information and Communication Technology (ISMICT)*, 2018. doi:10.1109/ismict.2018.8573728.

[3] M. F. Flanagan and A. D. Fagan, "Iterative channel estimation, equalization, and decoding for pilot-symbol assisted modulation over frequency selective fast fading channels," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 4, pp. 1661-1670, 2007.

[4] H. Ning, H. Liu, and Y. Zhang, "Scalable and distributed key array authentication protocol in radio frequency identification-based sensor systems," *IET Communication*, vol. 5, no. 12, pp. 1755-1768, 2011.

[5] M.-D. Kim, J. Lee, J. Liang, and J. Kim, "Multipath channel characteristics for propagation between mobile terminals in urban street canyon environments," *17th International Conference on Advanced Communication Technology (ICACT)*, 2015. doi: 10.1109/icact.2015.72249179/eusipco.2016.7760675.

[6] P. Zetterberg and R. Fardi, "Open source SDR front end and measurements for 60-GHz wireless experimentation," *IEEE Access*, vol. 3, pp. 445-456, 2015.

[7] M. Chiani, E. Milani, and R. Verdone, "Optimization for weighed cooperative spectrum sensing in cognitive radio network," *Applied Computational Electromagnetic Society (ACES) Journal*, vol. 26, no. 10, pp. 800-914, 2011.

[8] J. Dai, "Bit-error-rate analysis of raptor codes over rician fading channels," *Journal of Electrical and Computer Engineering*, vol. 2020, 2020. doi.org/10.1155/2020/2685075.

[9] M. Abirami and A. Vimala, "A review of various antenna design methods for cognitive radio application," *4th International Conference on Electronics and Communication Systems, (ICECS)*, 2017. doi:10.1109/ecs.2017.8067850.

[10] C. Schuler, "Research on correction algorithm of propagation error in wireless sensor network coding," *EURASIP Journal on Wireless Communications and Networking*, vol. 1, 2020.

[11] H. Katiyar and R. Bhattacharjee, "Average capacity and signal-to-noise ratio analysis of multi-antenna regenerative cooperative relay in Rayleigh fading channel," *IET Communication*, vol. 5, no.

14, pp. 1971-1977, 2011.

[12] S. Ohno and G. B. Giannakis, "Average-rate optimal PSAM transmissions over time-selective fading channels," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 712-720, 2002.

[13] M. C. Valenti and B. D. Woerner, "Iterative channel estimation and decoding of pilot symbol assisted turbo codes over flat-fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 9, pp. 1697-1705, 2001.

[14] S. J. Lee, W. Kang, and J. Seo, "Performance enhancement of OFDM-SQ2AM in distorted channel environments," *IEICE Electronics Express*, vol. 7, no. 14, pp. 1020-1026, 2010.

[15] S. Bernard, *Digital Communications: Fundamentals and Applications*. Prentice-Hall, 2nd Edition, pp. 30-33, 2001.

[16] Y. Li, Y. Wang, and T. Jiang, "Norm-adaption penalized least mean square/fourth algorithm for sparse channel estimation," *Signal Processing*, vol. 128, pp. 243-251, 2016. doi:10.1016/j.sigpro.2016.04.003.

[17] G. A. Ellis, "Wireless propagation in non line of sight urban areas using uniform theory of diffraction," *Applied Computational Electromagnetic Society (ACES) Journal*, vol. 18, no. 3, pp. 162-171, Nov. 2003.

[18] Y. Zhang, S. B. Gelfand, and M. P. Fitz, "Soft-output demodulation on frequency selective Rayleigh fading channels using AR channel models," *IEEE Transactions on Communications*, vol. 55, no. 10, pp. 1929-1939, 2007.

Shabana Habib received the M.Sc. degree in Computer Science from the Postgraduate College Mandian Abbott Abad, Pakistan, in 2003, the M.S. degree in IT from the Institute of Management Science, in 2008, and the Ph.D. degree in Electrical, Electronics and Systems Engineering from the Faculty of Engineering and Alam Bina (FKAB), Universiti Kebangsaan Malaysia (UKM), Malaysia, in 2017. From January 2004 to June 2009, she was working as a Lecturer with Frontier Education Foundation (FEF) Degree College Peshawar, Pakistan. From February 2010 to February 2011, she worked with the Legenda Education Group, Mantin, Negeri Sembilan, Malaysia. Since September 2018, she has been working as an Assistant Professor with the Department of Computer Science, Al Qassim University, Saudi Arabia. She has published more than ten journal and international proceeding articles in the field of image and signal processing. Her research interests include image processing, signal processing, networking, and wireless communications.

Light Wave Propagation Model for Indoor Visible Light Communication Systems Employing Small LED Sources

Marwa M. A. Elsaaty [1], Adel Zaghloul [2], and Khalid. F. A. Hussein [3]

[1] Institute of Aviation Engineering and Technology (IAET), Egypt
eng.marwamohamed1991@gmail.com

[2] Zagazig University, Faculty of Engineering, Egypt
adel6730101@yahoo.com

[3] Electronics Research Institute (ERI), Cairo, 11843, Egypt
fkhalid@eri.sci.eg, khalid_elgabaly@yahoo.com

Abstract ─ The aim of the present work is to build a robust and computationally efficient model for the light wave propagation in indoor visible light communication (VLC) systems. It is assumed that a small (point) LED source is used inside a room of relatively small dimensions (room area $\leq 5\text{m} \times 5\text{m}$). The light wave is treated as an electromagnetic wave suffering multiple reflections on the walls of the room. The Geometrical Theory of Diffraction (GTD) is applied for evaluation of the light wave reflection on the rough walls of the room. Also, the present work is concerned with developing a new computational method for the assessment of inter-symbol interference (ISI) encountered in such indoor VLC systems. The signal strength, the power of ISI, and hence, the signal-to-ISI ratio (SISIR) are evaluated over the horizontal plane of the mobile units (at a height of about 1m above the room floor). The effects of the room dimensions and some structural parameters such as the reflectance of the side walls on the SISIR are numerically investigated.

Index Terms ─ Inter-Symbol Interference (ISI), Visible Light Communications (VLC).

I. INTRODUCTION

Most of the currently operational wireless communication systems are employing the radio frequency (RF) and microwave spectra. As a consequence of the rapidly increasing number of mobile communication users, the wireless communications face a serious shortage regarding spectrum allocation. The visible light communication (VLC) systems utilize the unlicensed wide spectrum of the visible light for wireless communications [1-5]. The visible light spectrum is capable of offering much higher transmission rates than those offered by RF and microwave spectra. Moreover, the VLC systems are more immune to electromagnetic interference than the microwave wireless communication systems. Also, the VLC systems have the advantage of using the general lighting sources such as the light-emitting diode (LED) through high-speed on/off switching to transmit data at high rates. Photosensitive components such as photodetectors can be used to receive the signals sent by the lighting sources. Thus, the VLC enables the efficient use of the simple on/off keying (OOK) modulation as a digital intensity modulation technique for high data rate transmission with high noise immunity.

In indoor VLC system, the light source should meet the requirement for room lighting in addition to signal coverage. This requires the installation of LED source with appropriate shape and distribution. For a room of, relatively, small area ($\leq 5\text{m} \times 5\text{m}$) the lighting may be performed by employing a single point LED source mounted at the center of the room ceil; see Fig. 1.

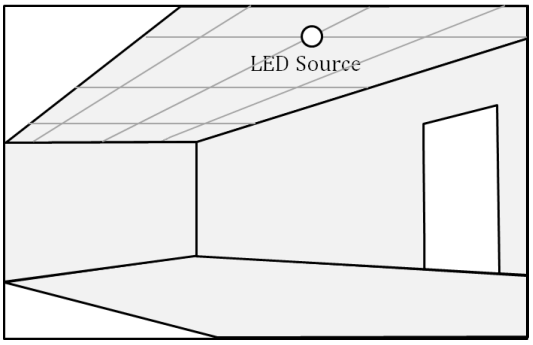


Fig. 1. The VLC system considered in the present work employs a single point LED source for signal coverage inside the room.

In the present study, it is assumed that the responsivity of the photodetector used in the mobile receivers is unity, and its FOV is 90°. The sidewalls of the room are assumed to be Lambertian reflective

Submitted On: April 24, 2021
Accepted On: July 7, 2021

https://doi.org/10.47037/2021.ACES.J.360814

surfaces with unity Lambert index and arbitrary reflectance.

The present work provides a complete description of the method used to evaluate the channel impulse response and inter-symbol interference (ISI) in the plane of the mobile receivers for both types of LED sources. It is shown that when a single point LED source is employed, the ISI results only from the reflection of light on the side walls of the room, i.e., from the non-line-of-sight (non-LOS) propagation.

II. INDOOR VLC CHANNEL MODEL

It is assumed that the room is empty and has its interior of a cuboidal shape whose dimensions are W_x, W_y, and W_z in the x-, y-, and z-directions, respectively. The LED source is assumed to be Lambertian radiator with unity Lambert index and is mounted on the room ceil with the maximum luminous intensity directed normal to the room ceil in the downward direction. On the other hand, the mobile, or even stationary, receivers are assumed as photodetectors that can freely move in the horizontal plane at a height of z_r above the room floor

In the present study, the walls of the room are assumed to be optically rough and, hence, the reflection of the light on these walls depends on the wall reflectance and roughness and can be accurately evaluated using the Geometrical Theory of Diffraction in conjunction with the Ray-Tracing (GTD-RT) method described in [6-8]. For electromagnetic simulation, the room walls are modeled as rough surfaces using the Savitzky-Golay method as described in [9]. The resulting models of the wall surface are found to be Lambertian surfaces with unity Lambert index. As the present study focuses only on the light wave propagation model, the receiving photodetector responsivity is assumed to be unity; $\eta = 1$. Also, the FOV of the receiving photodetector is assumed to be 90°.

Under these conditions, a point in the plane of the receiver units can be illuminated from the following sources:

1. Direct illumination (LOS) from the LED source.
2. Indirect illumination (non-LOS) due to the reflection of the light on the sidewalls of the room.
3. Noise resulting from ambient light of the sun, fluorescent lambs, or any other unplanned source.

A. Impulse response for point light source

It is assumed that a small LED source is used for VLC in the room as shown in Fig. 1. This light source is located at the center of the room ceil as shown in Fig. 2 and can be considered a point source. The luminous flux intensity $I(\theta)$ due to a point LED source mounted in the room ceil is the direction making angle θ with the normal to the ceil (see Fig. 2) is defined as the luminous flux within the unit solid angle at this direction:

$$I_{\mathrm{a}}(\theta) = \frac{d\Phi(\theta)}{d\Omega}, \tag{1}$$

where $d\Phi(\theta)$ is the luminous flux within the solid angle $d\Omega$ in the direction θ.

The radiation from this LED source can be described by the spatial angular distribution of the luminous flux intensity $I_a(\theta)$, which is assumed to be Lambertian with unity Lambert index:

$$I_{\mathrm{a}}(\theta) = I_{\mathrm{a}_0} \cos\theta, \tag{2}$$

where I_{a_0} is maximum luminous flux intensity (in the direction $\theta = 0$ i.e. normal to the room ceiling).

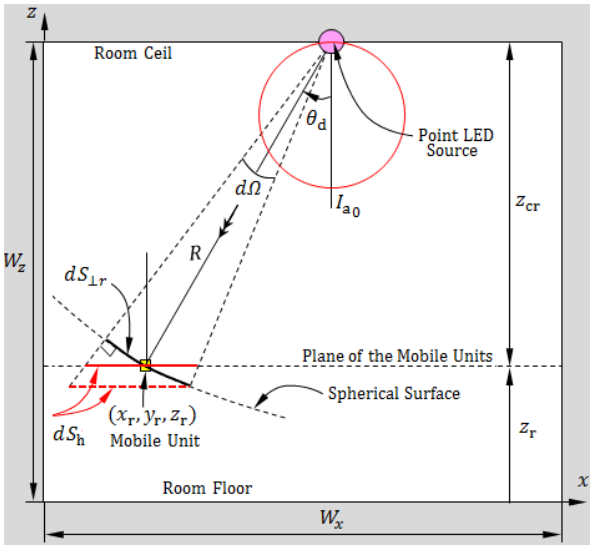


Fig. 2. Light wave propagation model for calculation of the direct (LOS) horizontal illuminance at the location of the mobile unit.

A.1. Direct (LOS) illuminance

The illuminance of a surface segment $\mathrm{dS}_{\perp \mathrm{r}}$ at the location of a receiver at $\mathrm{r_r} \equiv (\mathrm{x_r}, \mathrm{y_r}, \mathrm{z_r})$, (shown in Fig. 2) of a sphere whose center is at the point LED source can be expressed as follows:

$$E_{S_{\perp r}} = \frac{d\Phi}{dS_{\perp \mathrm{r}}} = \frac{I_{\mathrm{a}}(\theta_d)d\Omega}{dS_{\perp \mathrm{r}}}, \tag{3}$$

where $d\Omega$ is the solid angle subtending the area $dS_{\perp \mathrm{r}}$ that is given as:

$$dS_{\perp \mathrm{r}} = R^2 d\Omega, \tag{4}$$

where R is the distance between the LED source and the mobile receiver.

By substitution from (4) into (3), one gets the following expression:

$$E_{S_{\perp r}} = \frac{I_{\mathrm{a}}(\theta_d)}{R^2} = \frac{I_{\mathrm{a}_0} \cos\theta_d}{R^2}. \tag{5}$$

The illuminance of the horizontal surface at $\mathbf{r}_\mathrm{r}$ can be expressed as follows:

$$E_\mathrm{d} = E_{S_{\perp r}} \cos\theta_\mathrm{d} = \frac{I_{\mathrm{a}_0} \cos^2\theta_\mathrm{d}}{R^2}. \tag{6}$$

The distance, R, between the LED source and the mobile receiver can be expressed in terms of the ceil height, z_cr, relative to the plane of the mobile receivers as follows:

$$R = |\mathbf{r}_\mathrm{r} - \mathbf{r}_\mathrm{c}| = \frac{z_\mathrm{cr}}{\cos\theta_\mathrm{d}}, \tag{7}$$

where $\mathbf{r}_\mathrm{c} = (W_x/2, W_y/2, W_z)$ is the position vector of the center point of the room ceil.

The substitution from (7) into (6) results in the following expression for the horizontal illuminance in the plane of the mobile receivers:

$$E_\mathrm{d} \equiv E(\theta_\mathrm{d}) = \frac{I_{\mathrm{a}_0} \cos^4\theta_\mathrm{d}}{z_\mathrm{cr}^2}. \tag{8}$$

The total luminous flux is equal to the total light power, P_{Tx}, radiated from the from the LED source. Hence,

$$P_{Tx} = \int_{\text{half space}} d\Phi = \int_{\text{half space}} I_\mathrm{a}(\theta)\, d\Omega. \tag{9}$$

Considering that $d\Omega = \sin\theta\, d\theta\, d\phi$, the last integral can be expressed as follows:

$$P_{Tx} = \int_0^{2\pi} \int_0^{\pi/2} I_{\mathrm{a}_0} \cos\theta\ \sin\theta\, d\theta\ d\phi. \tag{10}$$

Thus,

$$P_{Tx} = I_{\mathrm{a}_0} \int_0^{2\pi} d\phi \int_0^{\pi/2} \tfrac{1}{2}\sin(2\theta) d\theta. \tag{11}$$

Thus, the maximum luminous flux intensity, I_{a_0}, of the LED source can be expressed in terms of the transmitted optical power, P_{Tx}, as follows:

$$I_{\mathrm{a}_0} = \frac{P_{Tx}}{\pi}. \tag{12}$$

The impulse response due to a point LED source can be expressed as follows:

$$h_\mathrm{d}(t) = \frac{P_{Tx} \cos^4\theta_\mathrm{d}}{\pi z_\mathrm{cr}^2}\ \delta(t - t_\mathrm{a}), \tag{13}$$

where t_a is the time taken by an optical impulse to travel from the LED source at $\mathbf{r}_\mathrm{c}$ to the receiver at $\mathbf{r}_\mathrm{r}$:

$$t_\mathrm{a} = \frac{1}{c_0}|\mathbf{r}_\mathrm{r} - \mathbf{r}_\mathrm{c}| = \frac{z_\mathrm{cr}}{c_0 \cos\theta_\mathrm{d}}, \tag{14}$$

where c_0 is the speed of light in free space.

A.2. Indirect (non-LOS) illuminance

Referring to Fig. 3, the illuminance of the side wall $x = 0$, at a point $\mathbf{r}_\mathrm{w} \equiv (0, y_\mathrm{w}, z_\mathrm{w})$ can be given as,

$$[E_\mathrm{w}]_{\mathbf{r}_\mathrm{w}} = \left[\frac{I_\mathrm{a}(\theta_\mathrm{a})}{R_\mathrm{wa}^2} \cos\theta_\mathrm{wa}\right]_{\mathbf{r}_\mathrm{w}} = I_{\mathrm{a}_0}\left[\frac{\cos\theta_\mathrm{a} \cos\theta_\mathrm{wa}}{R_\mathrm{wa}^2}\right]_{\mathbf{r}_\mathrm{w}}. \tag{15}$$

The power (luminous flux) of the light emitted from the wall segment $dy_\mathrm{w}\, dz_\mathrm{w}$ at the point $\mathbf{r}_\mathrm{w}$ can be expressed as follows.

$$[\Phi_\mathrm{w}]_{\mathbf{r}_\mathrm{w}} = [G_\mathrm{w}\, E_\mathrm{w}]_{\mathbf{r}_\mathrm{w}}\, dx_\mathrm{w}\, dz_\mathrm{w}, \tag{16}$$

where G_w is the wall reflectance at the point $\mathbf{r}_\mathrm{w}$.

The wall segment $dy_\mathrm{w}\, dz_\mathrm{w}$ at the point $\mathbf{r}_\mathrm{w}$ can be considered as a secondary radiator. As the wall is assumed as a Lambertian rough surface, the luminous flux intensity of the light emitted from this wall segment can be expressed as follows:

$$[I_\mathrm{w}(\theta_\mathrm{w})]_{\mathbf{r}_\mathrm{w}} = \left[I_{\mathrm{w}_0} \cos\theta_\mathrm{w}\right]_{\mathbf{r}_\mathrm{w}}, \tag{17}$$

where $\left[I_{\mathrm{w}_0}\right]_{\mathbf{r}_\mathrm{w}}$ is the luminous flux intensity in the direction normal to the wall due to the emission from the wall segment $dy_\mathrm{w}\, dz_\mathrm{w}$ at the point $\mathbf{r}_\mathrm{w}$.

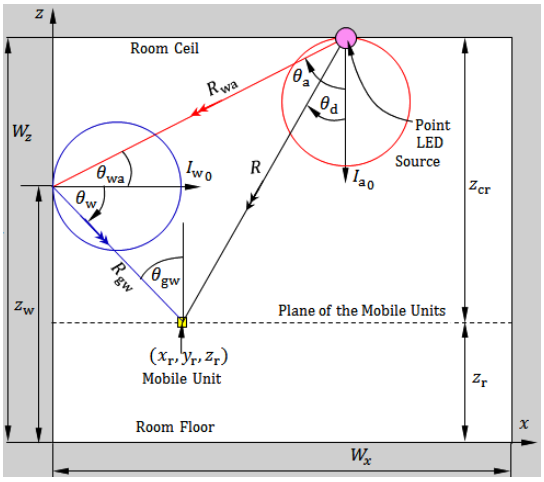


Fig. 3. Light wave propagation model for calculation of indirect (non-LOS) horizontal illuminance at the location of the mobile unit.

Also, the luminance intensity of the light reflected from the wall segment $dy_\mathrm{w}\, dz_\mathrm{w}$ in the direction of the mobile unit at $\mathbf{r}_\mathrm{r}$ can be expressed as follows:

$$[I_\mathrm{w}(\theta_\mathrm{w})]_{\mathbf{r}_\mathrm{w}} = \left[\frac{d\Phi_\mathrm{w}}{d\Omega}\right]_{\mathbf{r}_\mathrm{w}}. \tag{18}$$

Thus, the total luminous flux Φ_w emitted from the wall segment $dy_\mathrm{w}\, dz_\mathrm{w}$ at $\mathbf{r}_\mathrm{w}$ can be obtained as follows:

$$[\Phi_\mathrm{w}]_{\mathbf{r}_\mathrm{w}} = \int_{\text{half space}} [I_\mathrm{w}(\theta)]_{\mathbf{r}_\mathrm{w}}\, d\Omega. \tag{19}$$

Considering that $d\Omega = \sin\theta\, d\theta\, d\phi$, the last integral can be expressed as follows:

$$[\Phi_{\mathrm{w}}]_{\mathbf{r}_{\mathrm{w}}} = \int_0^{2\pi}\int_0^{\pi/2} \left[I_{\mathrm{w}0}\right]_{\mathbf{r}_{\mathrm{w}}} \cos\theta \sin\theta\, d\theta\; d\phi. \quad (20)$$

Thus,

$$[\Phi_{\mathrm{w}}]_{\mathbf{r}_{\mathrm{w}}} = \left[I_{\mathrm{w}0}\right]_{\mathbf{r}_{\mathrm{w}}} \int_0^{2\pi} d\phi \int_0^{\pi/2} \tfrac{1}{2}\sin(2\theta) d\theta. \quad (21)$$

Thus, the total luminous flux $[\Phi_{\mathrm{w}}]_{\mathbf{r}_{\mathrm{w}}}$ emitted from the wall segment $dy_{\mathrm{w}}\, dz_{\mathrm{w}}$ at the point $\mathbf{r}_{\mathrm{w}}$ can be expressed in terms of the maximum luminous flux intensity, $\left[I_{\mathrm{w}0}\right]_{\mathbf{r}_{\mathrm{w}}}$, as follows:

$$[\Phi_{\mathrm{w}}]_{\mathbf{r}_{\mathrm{w}}} = \pi\left[I_{\mathrm{w}0}\right]_{\mathbf{r}_{\mathrm{w}}}. \quad (22)$$

Making use of (15), (16), and (22), the maximum luminous flux intensity, $\left[I_{\mathrm{w}0}\right]_{\mathbf{r}_{\mathrm{w}}}$, can be expressed as follows:

$$\left[I_{\mathrm{w}0}\right]_{\mathbf{r}_{\mathrm{w}}} = \frac{I_{\mathrm{a}0}}{\pi}\left[G_{\mathrm{w}} \frac{\cos\theta_{\mathrm{a}} \cos\theta_{\mathrm{wa}}}{R_{\mathrm{wa}}^2}\right]_{\mathbf{r}_{\mathrm{w}}} dx_{\mathrm{w}}\, dz_{\mathrm{w}}. \quad (23)$$

The horizontal illuminance at the photodetector of the mobile receiver due to the indirect radiation from the wall segment $dy_{\mathrm{w}}\, dz_{\mathrm{w}}$ at $\mathbf{r}_{\mathrm{w}}$ can be expressed as follows:

$$[dE_{\mathrm{i}}]_{\mathbf{r}_{\mathrm{w}}} = \left[\frac{I_{\mathrm{w}0} \cos\theta_{\mathrm{w}}\; \cos\theta_{\mathrm{gw}}}{R_{\mathrm{gw}}^2}\right]_{\mathbf{r}_{\mathrm{w}}}. \quad (24)$$

By substitution from (23) into (24), the following expression is obtained:

$$[dE_{\mathrm{i}}]_{\mathbf{r}_{\mathrm{w}}} = [\xi_i]_{\mathbf{r}_{\mathrm{w}}}\, dxdz, \quad (25)$$

where,

$$[\xi_i]_{\mathbf{r}_{\mathrm{w}}} = \frac{P_{Tx}}{\pi^2}\left[\frac{G_{\mathrm{w}} \cos\theta_{\mathrm{a}} \cos\theta_{\mathrm{wa}} \cos\theta_{\mathrm{w}}\; \cos\theta_{\mathrm{gw}}}{R_{\mathrm{wa}}^2\, R_{\mathrm{gw}}^2}\right]_{\mathbf{r}_{\mathrm{w}}}. \quad (26)$$

Let the indices 1, 2, 3, and 4 be assigned to the walls $x = 0$, $y = 0$, $x = W_x$, and $y = W_y$, respectively. The component of the non-LOS impulse response due to the reflection of light on the wall #1 ($x = 0$) can be expressed as follows:

$$[h_i(\mathrm{t})]_1 = \int_{z_{\mathrm{r}}}^{W_z}\int_0^{W_y} [\xi_i]_{\mathbf{r}_{\mathrm{w}}}\, \delta(t - t_{\mathrm{w}})\, dy_{\mathrm{w}}\; dz_{\mathrm{w}}, \quad (27)$$

where t_{w} is the total time taken by an optical impulse when it is transmitted by the point LED source, then reflected on the surface segment at $\mathbf{r}_{\mathrm{w}}$ (on the wall $x = 0$) and, finally, arrives at the location of the receiver $\mathbf{r}_{\mathrm{r}}$:

$$t_{\mathrm{w}} = \frac{1}{c_0}\left(|\mathbf{r}_{\mathrm{w}} - \mathbf{r}_{\mathrm{c}}| + |\mathbf{r}_{\mathrm{r}} - \mathbf{r}_{\mathrm{w}}|\right). \quad (28)$$

The total non-LOS impulse response at the mobile receiver can be obtained by performing the integration expressed in (27) on each of the four side walls of the room and then summing the integral values as follows:

$$h_i(\mathrm{t}) = \sum_{q=1}^{4} [h_i(\mathrm{t})]_q. \quad (29)$$

Considering that the mobile receiver has its FOV $\leq 90°$, the area of the side wall #1($x = 0$) subtended between the planes $z = z_r$ and $z = W_z$ can be discretized with fine enough resolution to a number of rectangular segments of equal area $\Delta s = \Delta y \Delta z$. The illuminance at the location of the receiver due to the light reflected from the wall segment number m, n can be expressed as follows:

$$\left[\Delta E_{\mathrm{i}_{m,n}}\right]_1 = \left[\Delta s \xi_{i_{m,n}}\right]_1, \quad m = 1,2,..,M_1, \quad n = 1,2,..,N_1 \quad (30)$$

where M_1 and N_1 are the number of segments which the side wall #1 is discretized in the vertical and horizontal directions, respectively, and $\xi_{i_{m,n}}$ is given as follows:

$$\left[\xi_{i_{m,n}}\right]_1 = [\xi_i]_{\mathbf{r}_{\mathrm{w}}=[0,(m-1)\Delta y,(n-1)\Delta z]}. \quad (31)$$

The total (indirect) impulse response due to the light reflection on the four side walls is the horizontal illuminance due to the light arriving at the receiver location through non-LOS propagation and can be expressed as follows:

$$h_{\mathrm{i}}(t) = \sum_{q=1}^{4}\sum_{m=1}^{M_q}\sum_{n=1}^{N_q}\left[\Delta E_{\mathrm{i}_{m,n}}\right]_q \delta\left(t - t_{q,m,n}\right), \quad (32)$$

where M_q and N_q are the number of segments to which the q^{th} side wall is discretized in the vertical and horizontal directions, respectively; $t_{q,m,n}$ is the total time taken by an optical impulse when it is transmitted by the point LED source at $t = 0$, then reflected on the surface segment number m, n on the q^{th} side wall, and finally arrives at the location of the receiver $\mathbf{r}_{\mathrm{r}}$:

$$t_{q,m,n} = \frac{1}{c_0}\left[R_{\mathrm{wa}} + R_{\mathrm{gw}}\right]_{q,m,n}. \quad (33)$$

It is assumed that the OOK modulation technique is employed in the indoor VLC system where the transmitted pulse can be considered as a rectangular pulse of duration equal to the bit period, $T_b = 1/R_b$.

The impulse response due to a point LED source on the ceiling of a cuboidal room can be expressed as the sum of the LOS and non-LOS impulse responses as follows:

$$h(t) = h_{\mathrm{d}}(t) + h_{\mathrm{i}}(t). \quad (34)$$

Upon the knowledge of the total impulse response, $h(t)$, the received pulse can be expressed as follows:

$$s_r(t) = \eta\, s_t(t) \otimes h(t), \quad (35)$$

where η is the responsivity of the receiving photodetector. In the present study, the photodetector responsivity is assumed to be unity; $\eta = 1$.

B. Impulse response due to light source of arbitrary shape

The impulse response due to a point LED source is expressed in (34). A practical LED source that can be commercially available for room lightening may take arbitrary shape extending of a larger area than that of a point LED source. The radiating surface of such a LED source can be planar or three-dimensional. Let the impulse response of an arbitrarily-shaped LED source be designated as $h_{\mathrm{a}}(t)$. For such a LED source, the impulse response can be obtained by integrating $h(t)$ over the surface of the LED source:

$$h_{\mathrm{a}}(t) = \iint_S h(t)\, ds, \tag{36}$$

where S is surface of the LED source that can be of arbitrary shape. The numerical evaluation of (36) requires the discretization (segmentation) of the source surface with enough resolution. Each of the discrete surface elements resulting from the segmentation of LED source surface can be considered as point LED source whose impulse response can be obtained by (34). During the evaluation of the integral in (36) one should take care that the time of arrival parameters t_{a} appearing in (13) and $t_{q,m,n}$ appearing in (32) should be calculated for each discrete surface segment on the LED source surface.

C. Assessment of root-mean-squared propagation delay

In indoor VLC systems, the received optical pulse is subjected to delay spread due to multipath propagation. The transmitted optical pulse representing a bit in the OOK modulation system travels in different paths with variable lengths and, hence, arrives at the position of the receiver at varying times, which causes the received pulse representing a bit to be spread and overlapped with the subsequent bits. The RMS delay spread is a measure of the ISI caused by multipath channel of the indoor propagation.

The average propagation delay, T_{av}, can be calculated once the impulse response $h(t)$ is obtained using the following integral:

$$T_{\mathrm{av}} = \frac{\int_{-\infty}^{\infty} t\, h^2(t)\, dt}{\int_{-\infty}^{\infty} h^2(t)\, dt}. \tag{37}$$

The RMS propagation delay, T_{rms}, can be evaluated as follows:

$$T_{\mathrm{rms}} = \frac{\int_{-\infty}^{\infty} (t - T_{\mathrm{av}})^2\, h^2(t)\, dt}{\int_{-\infty}^{\infty} h^2(t)\, dt}. \tag{38}$$

The integrals (37) and (38) can be numerically evaluated once the impulse response $h(t)$ is determined.

D. Efficient computational method for assessment of power of intersymbol interference and SISIR

The SISIR is the ratio between the light energy of the received pulse to the energy interfering the other pulses due to the delay spread of the received pulse. Thus, the SISIR can be expressed as follows:

$$\mathrm{SISIR} = \frac{E_S}{E_{ISI}} = \frac{\int_{t_{\mathrm{a}}}^{t_{\mathrm{a}}+T_{\mathrm{b}}} s_r(t)\, dt}{\int_{t_{\mathrm{a}}+T_{\mathrm{b}}}^{\infty} s_r(t)\, dt}, \tag{39}$$

where t_a is the time of arrival of the received pulse. The assessment of the SISIR only can be evaluated by numerical evaluation of the integrals in (39).

D.1. Computational procedure for time-domain ray-tracing

In this section, a novel time-domain ray-tracing procedure is described for the calculation of E_S and E_{ISI} appearing in (37). The model developed in the proposed work is based on the GTD-RT method for the assessment of light reflections on the room walls. This method has been computationally applied and experimentally verified in [7] and [8].

The proposed ray-tracing method accounts for the light reflection on the four side walls. It is considered that the LED source is mounted on the room ceiling and the light propagates in the downward directions to illuminate the room and, hence, light reflection on the ceiling surface is neglected. On the other hand, the room floor is assumed to be covered with textures like carpets that are commonly light-absorptive and, hence, the reflections from the room floor are very weak and can be negligible.

According to the theoretical and experimental studies achieved in [7] and [8] (see the reference list of the manuscript), the light reflected from the wall due to second- and higher-order bounces can be negligible when compared to that reflected due to the first-order bounce. Consequently, a first-order ray-tracing technique is applied as it is enough to get accurate assessment of the impulse response and ISI in a closed room.

The following is a concise list of the steps of the proposed procedure for the application of the proposed time-domain ray-tracing method.

1. The area of each of the four sidewalls of the room is divided into a number of square segments with equal area ΔS_W that is small enough to get high resolution.
2. The time is discretized to small divisions Δt with enough resolution. The time origin is that time at which the rays start emitting from the point LED source.
3. Originating at the location of the point LED source, the range of the solid angle $\Omega = 2\pi$ covering the room space (the lower half-space: $0 < \theta < \pi$ and

$0 < \phi < 2\pi$) is equally divided into equal solid angle divisions each $d\Omega$.

4. The light emitted from the point LED source in the direction θ_d is represented by a number of rays within the solid angle division $d\Omega$ around the direction θ_d. The number of rays is proportional to the luminous flux intensity in the direction θ_d.
5. The power associated with each ray δP is equal to luminous flux within the solid angle division $d\Omega$ divided by the number of rays.
6. Each ray starts traveling from the position of the point LED source with the speed of light and is traced alone.
7. If the ray arrives at the location of the mobile receiver without hitting any of the side walls its associated power is added to the direct signal power (LOS path).
8. If the ray hits the same surface segment ΔS_W within the same time division Δt on one of the side walls its associated power is added to the power of ISI (non-LOS path) and the location and time at which this ray hits the side wall is stored.
9. The light illuminating each discrete segment ΔS_W of the side wall surface is calculated by summing power associated with each ray hitting this segment within the same time division Δt.
10. The wall segment ΔS_W is considered as a secondary radiator where the emitted light has the luminous flux can be calculated using (16).
11. The steps from 3 through 7 of the procedure listed above are repeated replacing LED source by the secondary radiator represented by the side wall segment ΔS_W.
12. The list of power and time of arrival calculated and stored as described above are now sorted to calculate the time record of both the LOS and the non-LOS light power arriving at the mobile receiver.
13. If multiple point sources are employed in the VLC, the impulse response, $h_d(t)$, due to the LOS propagation is calculated using (13) and the time record obtained in the step 12 above.
14. The impulse response, $h_i(t)$, due to the non-LOS propagation is calculated by numerical integration of the integral in (27) using the time record obtained in the step 12 above.
15. The total impulse response,$h(t)$, is calculated using (34).
16. If an arbitrarily shaped LED source is employed, the integral in (36) is numerically evaluated to get the total impulse response, $h_a(t)$.

III. RESULTS AND DISCUSSIONS

This section is concerned with the presentation and discussion of the numerical results obtained when the proposed computational model proposed in the previous sections for indoor light wave propagation is applied. It should be noted that the sidewalls of the room are assumed to be Lambertian with reflectance $G = 0.5$.

A. ISI at arbitrary point in indoor VLC system

Consider the case of a mobile unit (receiver) at the point $\mathbf{r} = (1\text{m}, 2\text{m}, 1\text{m})$. The present section is concerned with the investigation of the impulse response at the location of the indicated receiver due to a point LED source mounted at the center of the room ceil.

A.1. Impulse response at arbitrary point

The impulse response at the location $\mathbf{r} = (1\text{m}, 2\text{m}, 1\text{m})$ is presented in Fig. 4. The LOS impulse response is a Dirac-delta function $\delta(t - t_a)$, where $t_a = 8.6$ ns, is the time of arrival of the light impulse transmitted by a point LED source at the center of the room ceil. The non-LOS impulse response is shown in Fig. 4 (b). Each of the four peaks of the impulse response is caused by the light reflection on one of the four side walls of the room. The first peak is the highest one and is caused by light impulse reflection on the nearest wall to the receiver ($x = 0$). The fourth peak is the lowest one and is caused by light reflection on the farthest wall ($x = W_x$).

The total impulse response at the location $\mathbf{r} = (1\text{m}, 2\text{m}, 1\text{m})$ due to a point LED source is shown in Fig. 5. As the level of magnitude of the non-LOS response is very low relative to that of the LOS response, log scale is used for the vertical axis to show both responses together. As shown in this figure, the total impulse response has two distinct components; one is a Dirac-delta function, whereas the other component has a duration over the interval $(13 - 30$ ns) of the delay time.

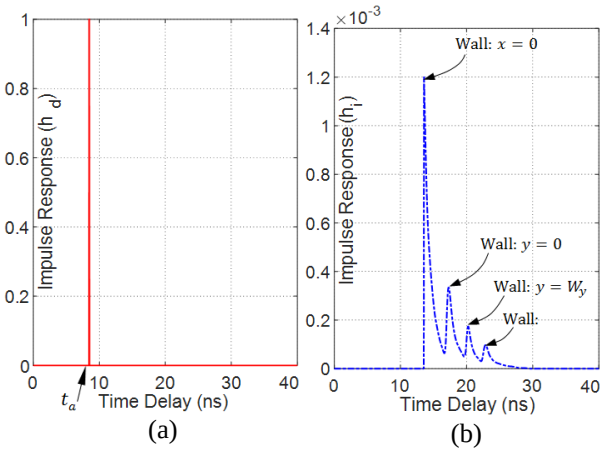


Fig. 4. Normalized impulse response at a point $\mathrm{r} = (1\text{m}, 2\text{m}, 1\text{m})$ inside a room of dimensions $5\text{m} \times 5\text{m} \times 3\text{m}$ due to a point LED source mounted at the center point of the room ceil, (a) LOS impulse response and (b) non-LOS impulse response.

A.2. Received signal at arbitrary point in the horizontal plane

It is assumed that the VLC employs intensity modulation using OOK. Also, it is assumed that the LED source transmits the data at a bit rate $R_b = 50$ Mbps. When a point LED source is used to transmit a pulse that represents a data bit, the pulses that arrive at $\mathbf{r} = (1\text{m}, 2\text{m}, 1\text{m})$ through the LOS and non-LOS are presented in Fig. 6 (a). The total received pulse is presented in Fig. 6 (b) in comparison to the ideal (reference) pulse. The reference pulse is the pulse that would be received if the impulse response were a pure Dirac-delta function, i.e., the pulse received through the LOS from a point LED source. It is shown that the received pulse has two types of distortion; the first type is amplitude distortion due to the addition of the light received through the non-LOS path and the other distortion is caused by the delay-spread due to the delayed light pulse received through the non-LOS path. The power of the ISI is calculated as described in Section II.C resulting in SISIR = 15.75 dB.

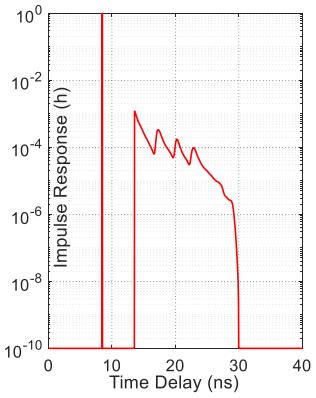


Fig. 5. Total (normalized) impulse response at a point $\mathrm{r} = (1\text{m}, 2\text{m}, 1\text{m})$ inside a room of dimensions $5\text{m} \times 5\text{m} \times 3\text{m}$ due to a point LED source mounted at the center point of the room ceil.

B. ISI at the central point in indoor VLC system

It is assumed that the VLC employs intensity modulation using on-off keying (OOK). Also, it is assumed that the LED source transmits the data at a bit rate $R_b = 50$ Mbps. It transmits a rectangular pulse that represents a data bit. The impulse response, the received pulse, and, hence the ISI are assessed at the center of the horizontal plane of the mobile units ($z = 1\text{m}$).

B.1. Impulse response at the central point in the horizontal plane

Assuming that the mobile unit is at $\mathbf{r} = (2.5\text{m}, 2.5\text{m}, 1\text{m})$, i.e., the center point of the plane $z = z_r$, the impulse response is plotted against the time delay as shown in Fig. 7. For a single LED point source at the center of the room ceil, the LOS impulse response is a Dirac-delta function $\delta(t - t_a)$ as shown in Fig. 10 (a). Thus, the LOS impulse response of a VLC channel employing a point LED source is also an impulse at the time of arrival, t_a, that can be calculated using (14) which gives $t_a = 6.7$ ns. On the other hand, the non-LOS impulse response is spread over the time interval $(8 - 26\text{ ns})$. However, as shown in Fig. 10 (a), the magnitude of the latter is very low relative to that of the former, the latter is very low relative to the former ($\max(h_i) \sim 10^{-4} \max(h_d)$). Thus, for a point LED source, no spread delay is caused by the LOS illuminance; the spread delay and, hence, the ISI are caused only by the non-LOS illuminance which is very small in comparison to the LOS illuminance.

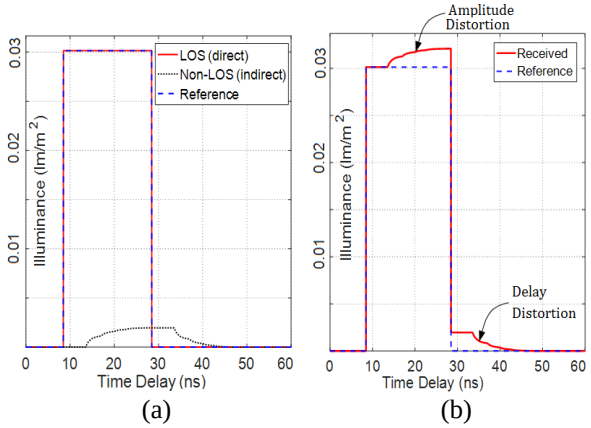


Fig. 6. The received optical pulse due to a single point LED source at the ceil center in indoor VLC at the point $\mathrm{r} = (1\text{m}, 2\text{m}, 1\text{m})$ in a room of dimensions $5\text{m} \times 5\text{m} \times 3\text{m}$. (a) Components of the received pulse: LOS and non-LOS, and (b) total received pulse.

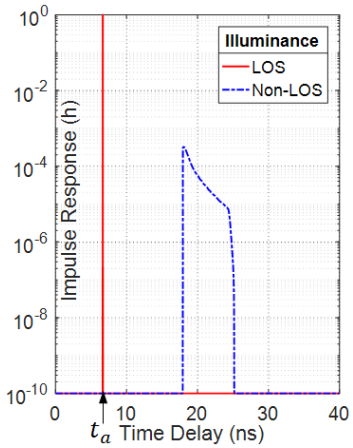


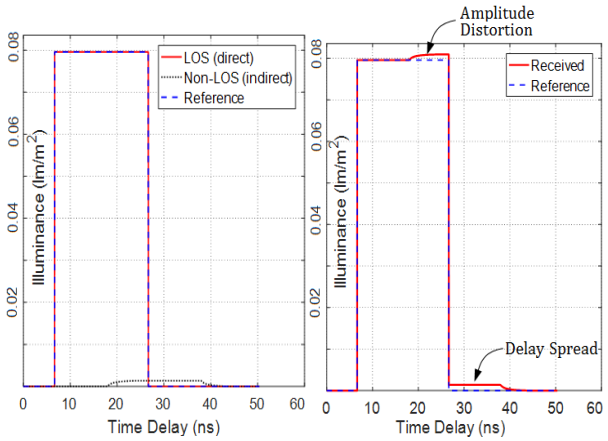
Fig. 7. Normalized impulse response due to direct (LOS) propagation and indirect (non-LOS) propagation in indoor VLC at the point (2.5m, 2.5m, 1.0m) in a room of dimensions $5\text{m} \times 5\text{m} \times 3\text{m}$ when a point LED source is mounted at the ceil center.

B.2. Received signal at the central point in the horizontal plane

When a point LED source is used to transmit a rectangular pulse of duration $T_b = 1/R_b$, the pulse arriving through the LOS and that arriving through the non-LOS at $\mathbf{r}_r = (2.5\text{m}, 2.5\text{m}, 1\text{m})$ are shown in Fig. 8. The total received pulse is shown in Fig. 1 (b) in comparison to the reference pulse. It shown that the received pulse has two types of distortion; the first type is amplitude distortion due to the addition of the light received through the non-LOS path and the other distortion is caused by the delay-spread due to the delayed light pulse received through the non-LOS path. However, the received pulse is weakly distorted and, hence, the resulting SINR is 19.4 dB.

C. Distribution of illumination on the horizontal plane

The distribution of the received signal (without ISI) in the plane of the mobile units ($z = 1\text{m}$) is presented in Fig. 9 inside a room of dimensions $5\text{m} \times 5\text{m} \times 3\text{m}$. The point LED source is mounted on the room ceil at its center point and the mobile receiving photodetectors are assumed to have FOV $= 90°$ and responsivity $\eta = 1$.



(a) (b)

Fig. 8. The received optical pulse due to a single LED point source at the ceil center in indoor VLC at the point $(2.5\text{m}, 2.5\text{m}, 1.0\text{m})$ in a room of dimensions $5\text{m} \times 5\text{m} \times 3\text{m}$. (a) Components of the received pulse: LOS and non-LOS, and (b) total received pulse.

D. Distribution of the SISIR in the horizontal plane

The distribution of the SISIR in the plane of the mobile units ($z = 1\text{m}$) is presented in Fig. 10 for a VLC system employing a point LED source inside a room of dimensions $5\text{m} \times 5\text{m} \times 3\text{m}$. The light source is mounted on the room ceil at its center point and the mobile receiving photodetectors are assumed to have FOV $= 90°$ and responsivity $\eta = 1$. It is clear that the SISIR in the horizontal plane is maximum near the room center and minimum near the room corners.

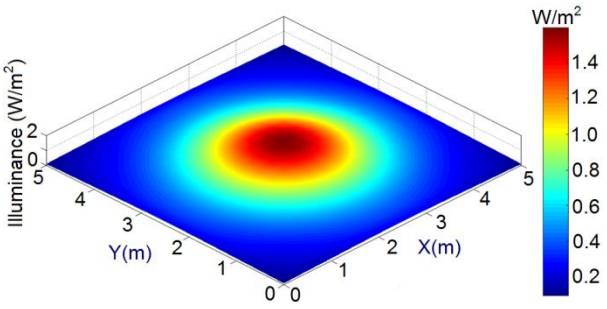


Fig. 9. Distribution of the received signal (without ISI) in the plane of the mobile receivers ($z = 1\text{m}$) in a VLC system inside a room of dimensions $5\text{m} \times 5\text{m} \times 3\text{m}$; a point LED source is mounted on the room ceil at its center point; the mobile receivers have FOV $= 90°$.

E. Effect of the room dimensions structural parameters on the SISIR

The room dimensions and the structural parameters including the optical properties of the wall material, color and suface roughness have significant effects on the perfromance of the indoor VLC system. The dependence of the SISIR on the reflectance of the side walls in a VLC system inside a room of height 3m and different areas is presented in Fig. 11 when the mobile receiver is at the center of the room area and near one of the room corners. As the ISI is mainly caused by the reflections of the side walls of the room, increasing the wall reflectance leads to significant decrease of the SISIR irrespective of the location of the mobile receviver.

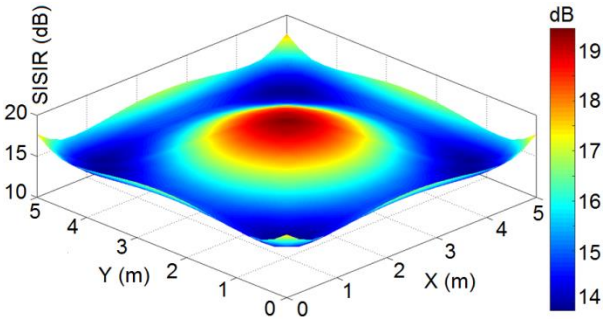


Fig. 10. Distribution of the SISIR in the plane of the mobile units ($z = 1\text{m}$) in VLC system inside a room of dimensions $5\text{m} \times 5\text{m} \times 3\text{m}$; a point LED source is mounted on the room ceil at its center point; the mobile receivers have FOV $= 90°$.

As shown in Fig. 11 (a), increasing the room dimensions has a great effect on the indoor VLC

perfromance and leads to increase the SISIR at the center point of the room because the center point is farther from the side walls when the room is wider. Howerver, increasing the room dimensions has a weak effect on the SISIR in the regions near the room corners. As shown in Fig. 11 (b), the SISIR at the point (0.5m, 0.5m, 1.0m) is slightly decreased due to increasing the room diemnsions because this point is closer to the LED source when the room is smaller.

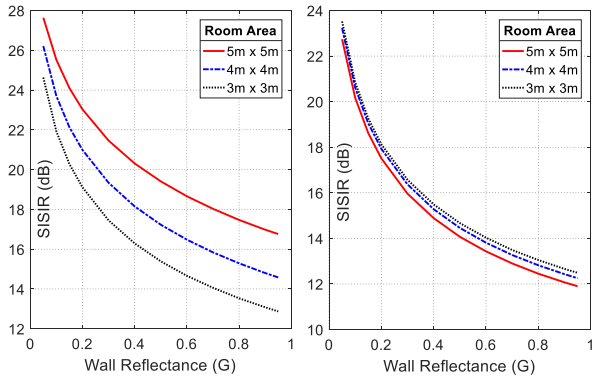


Fig. 11. Dependence of the SISIR on the reflectance of the side walls in VLC system inside a room of height 3m and different areas. The mobile receiver is at (a) the center point of the room area, and (b) the point (0.5m, 0.5m) near the room corner.

IV. CONCLUSION

A robust computational model has been introduced for the light wave propagation in indoor VLC systems employing small LED source inside a room of area $\leq 5\text{m} \times 5\text{m}$. The light wave is treated as an electromagnetic wave suffering multiple reflections on the walls of the room. The GTD-RT technique is applied for evaluation of the light wave reflection of the rough surfaces of the room walls. Moreover, an efficient computational method has been described for the assessment of the ISI encountered in indoor VLC system. The distribution of the SISIR is numerically evaluated over the horizontal plane of the mobile units (at a height of about 1m above the room floor). The effects of the room dimensions and some structural parameters such as the reflectance of the side walls on the SISIR have been investigated numerically. It has been shown that a narrower room has worse effect on the SISIR.

REFERENCES

[1] Y. Chen, S. Li, and H. Liu, "Dynamic frequency reuse based on improved tabu search in multi-user visible light communication networks," *IEEE Access*, vol. 7, pp. 35173-35183, 2019.

[2] I. Abdalla, M. Rahaim, and T. Little, "Interference in multi-user optical wireless communications systems," *Philosophical Transactions of the Royal Society A*, 378.2169, 20190190, 2020.

[3] K. Lee, H. Park, and J. R. Barry, "Indoor channel characteristics for visible light communications," *IEEE Communications Letters*, vol. 15, no. 2, pp. 217-219, 2011.

[4] K. Lee and H. Park, "Channel model and modulation schemes for visible light communications," In *2011 IEEE 54th International Midwest Symposium on Circuits and Systems (MWSCAS)*, pp. 1-4, Aug., 2011.

[5] T. Komine and M. Nakagawa, "Fundamental analysis for visible-light communication system using LED lights," *IEEE Transactions on Consumer Electronics*, vol. 50, no. 1, pp. 100-107, 2004.

[6] H. H. Qamar, K. F. A. Hussein, and M. B. El-Mashade, "Assessment of signal strength in indoor optical wireless communications using diffuse infrared radiation," In *2019 36th National Radio Science Conference (NRSC)*, pp. 108-117, 2019.

[7] H. H. Qamar, A. E. Farahat, K. F. A. Hussein, and M. B. El Mashade, "Assessment of scattering of plane waves on optically illuminated area of rough surface," *Progress In Electromagnetics Research*, vol. 86, pp. 77-102, 2020.

[8] H. H. Qamar, M. B. El-Mashade, A. E. Farahat, and K. F. A Hussein, "Convergence of ensemble averaging for optical scattering on rough surfaces using GTD-RT," In *6th International Conference on Advanced Control Circuits and Systems (ACCS) & 2019 5th International Conference on New Paradigms in Electronics & Information Technology (PEIT), IEEE*, pp. 167-175, 2019.

[9] S. A. M, Soliman, A. E. Farahat, K. F. A. Hussein, and A.E. A. Ammar, "Spatial domain generation of random surface using Savitzky-Golay filter for simulation of electromagnetic polarimetric systems," *Applied Computational Electromagnetics Society Journal*, vol. 34, no. 1, 2019.

Signal Propagation Modeling Based on Weighting Coefficients Method in Underground Tunnels

Yusuf Karaca [1] and Özgür Tamer [2]

[1] Graduate School of Natural and Applied Sciences
Dokuz Eylül University, İzmir, Turkey
yusuf.karaca@deu.edu.tr

[2] Department of Electrical and Electronics Engineering
Dokuz Eylül University, İzmir, Turkey
ozgur.tamer@deu.edu.tr

Abstract ─ The propagation of electromagnetic waves guided in tunnels and mines is an area of scientific study which is hard to model due to multiple reflections on walls and surrounding obstacles. A novel propagation model for underground tunnels based on the weighting sum of the log-distance propagation model, the modified waveguide model and, the far zone propagation model for the ultra-high frequency (UHF) band is proposed in this paper. The propagation model is divided into five regions based on the distance between the transmitter and the receiver. Each region shows a different propagation characteristic and modeled with weighting sum of the base propagation models. Our model was tested in a tunnel with 2 m x 1.5 m cross-section and 250 m length. Measurement results are consistent with the proposed propagation model.

Index Terms ─ Propagation modeling, underground tunnel, underground communications.

I. INTRODUCTION

Modelling of electromagnetic wave propagation in underground environments has been an active area of research for many years. Despite several theoretical analyses and experimental studies, electromagnetic wave propagation in underground tunnels is hard to model with high precision and low complexity due to the imperfect structure of the surrounding obstacles and walls, such as multiple reflections from the surrounding walls [1]. A large number of models were presented and discussed for underground tunnels [2]. Zhang and Hong proposed a model based on the ray optical propagation model to show signal propagation at the UHF frequency band for rectangular underground tunnels [3]. Zhang, Zheng, and Sheng modeled electromagnetic wave propagation at 900 MHz using free space and mod propagation models for coal mines [4]. Two propagation regions with different characteristics were defined, where the breakpoint between the regions in the passageways was found to be 45 m from the transmitter. The waveguide propagation model for coal mines is proposed by Emslie, Lagace, and Strong [5]. The waveguide model is investigated in rectangular coal tunnels at frequencies 1 GHz and 415 MHz for both vertical polarization and horizontal polarization. Boutin et al. also present breakpoints for different propagation models [6]. Guan and friends showed that the free-space propagation model can be used at short ranges, while the multimode propagation model can be used at medium range distances and the fundamental mode propagation model can be used for longer ranges [7]. This model proposes a solution to unify and extend existing propagation mechanism models in tunnels. Although many researchers prefer a waveguide-based model to interpret the electromagnetic wave propagation in the far region [5], the results are not satisfactory for long tunnels. If the transmitter and receiver are far enough from each other, propagation is more similar to the far zone propagation model [7]. In fact, far zone propagation was not sufficiently investigated by researchers due to a lack of experimental studies in long tunnels in underground environments. Rak and Pechac searched for experimental measurements of subterranean galleries for two frequencies: 446 MHz and 860 MHz [8]. They proposed a simple linear attenuation model using recorded experimental data.

In this paper, a novel propagation model based on a combination of three different propagation mechanisms for near, mid, and far regions is introduced. A gradual transition with the weighting coefficients is applied between propagation regions. In the second part of the paper, a theoretical background about the employed propagation models is presented and the proposed propagation model is introduced in section three. A comparison of the model with measurement results is presented in the fourth section while the results are

Submitted On: March 16, 2021
Accepted On: July 18, 2021

https://doi.org/10.47037/2021.ACES.J.360815

concluded in the fifth and the final section.

II. THEORETICAL BACKGROUND

The signal propagation model proposed in this work employs three different propagation models introduced in below subsections.

A. Log-distance path loss model

In the near zone of tunnel-like structures, the slope of the signal attenuation curve is steep [6]. In our study we adopt the log-distance signal propagation model in the near zone given in equation (1) [9]:

$$P_1 = \overline{PL}(d_0) + 10nlog_{10}\left(\frac{d}{d_0}\right), \quad (1)$$

where n is the path loss exponent, d is the transmitter-receiver separation in meters, d_0 is the distance of the reference measurement and $\overline{PL}(d_0)$ is the path loss at d_0. Path loss exponent value is used as 1.75 since there is line-of-sight between the transmitter and the receiver and the tunnel is a rectangular cross-section one [9], while the value of the d_0 is 1 m. Received power for the log distance model (P^{Log}) is given by:

$$P^{Log} = \sum_{i=1}^{+\infty} A_i P_1, \quad (2)$$

where A_i are the weighting coefficients and P^{Log} is the representation of weighting signal attenuation for the near zone.

In the proposed model, log-distance is the dominant propagation mode for the first 50λ of the distance from the transmitter where λ is the wavelength as presented in Fig. 6. The attenuation of the signal based on this propagation model is presented in Fig. 2, Fig. 3 and Fig. 4.

B. Waveguide model

We consider that the tunnel is similar to a waveguide for the preferred frequency band as presented in the literature [5], [10]. Walls are lossy media built of concrete and iron bars with the corresponding electrical parameters defined in Fig. 1 as ε_ℓ, μ_ℓ, σ_ℓ, and ε_b, μ_b, σ_b. E^i, H^i, E^r, H^r in Fig. 1 are defined as the fields in the space of the waveguide-like structure. while E_b^{tr}, H_b^{tr} are the fields on top and bottom of the tunnel walls E_ℓ^{tr}, H_ℓ^{tr} are fields at the left and right of tunnel walls. Electrical space inside the rectangular waveguide is ε_r=1, μ_r = 1 since the tunnel is filled with air. The horizontal and vertical dimensions of the waveguide are given as a and b respectively, while the horizontal lines on axis x for the model and the vertical lines on axis y and the wave propagates in the z-direction.

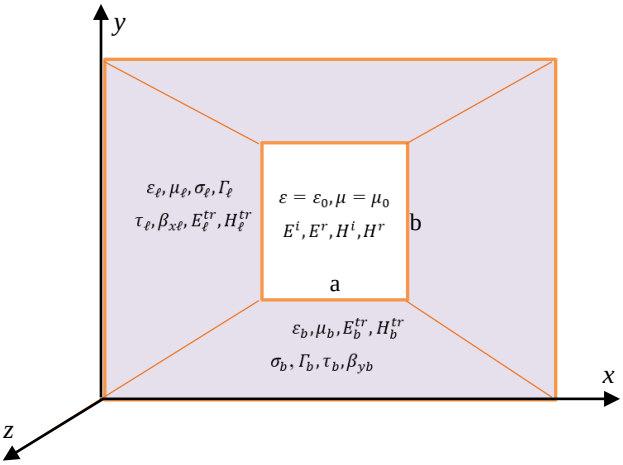


Fig. 1. Cross-section of the tunnel as a waveguide.

The complex dielectric constant for the sidewalls is given by the following equation:

$$\varepsilon_\ell = \varepsilon_\ell' - j\frac{\sigma_\ell}{\omega}, \quad (3)$$

where σ_ℓ is the conductivity of the sidewalls. Experimental studies show that ε_ℓ' is a real dielectric constant [14] while ω is working frequency [11]. Similarly, the dielectric constant for the top and bottom walls is defined as:

$$\varepsilon_b = \varepsilon_b' - j\frac{\sigma_b}{\omega}, \quad (4)$$

where σ_b is the conductivity of the top and the bottom where ε_b' is the real dielectric constant parameter.

The reflection coefficient on the sidewalls (Γ_ℓ) and the top and bottom walls (Γ_b) is given by:

$$\Gamma_\ell = \frac{\sqrt{\varepsilon}-\sqrt{\varepsilon_\ell}}{\sqrt{\varepsilon}+\sqrt{\varepsilon_\ell}}, \quad (5)$$

$$\Gamma_b = \frac{\sqrt{\varepsilon}-\sqrt{\varepsilon_b}}{\sqrt{\varepsilon}+\sqrt{\varepsilon_b}}. \quad (6)$$

The transmission coefficients for the sidewalls (τ_ℓ) and the top and bottom walls (τ_b) are given as:

$$\tau_\ell = \left(\frac{1}{\varepsilon}+\frac{\Gamma_\ell}{\varepsilon}\right)\frac{\beta_x \varepsilon_\ell}{\beta_{x\ell}}, \quad (7)$$

$$\tau_b = \left(\frac{1}{\varepsilon}+\frac{\Gamma_b}{\varepsilon}\right)\frac{\varepsilon_b \beta_y}{\beta_{yb}}, \quad (8)$$

where β_{xb} is the phase constant for the sidewalls, β_{yb} is the phase constant for the top and bottom walls and β_y is the phase constant for the tunnel cavity.

The direction of the power flow is defined in the +z direction as presented in Fig. 1. Power flow, P_{mn}, of an electromagnetic wave in TE_{mn}^z mode along the axis of +z is presented in equation (9) [13].

Table 1: Regions of the weighting signal propagation model

Log-distance Path Loss Model (P^{Log})	Transient Region I ($P^{Log} + P^{Waveg}$)	Waveguide Model (P^{Waveg})	Transient Region II ($P^{Waveg} + P^{Far}$)	Far Zone Propagation Model (P^{Far})
<50λ	50λ-120λ	120λ-700λ	700λ-800λ	>800λ

$$P_{mn} = \frac{A_{mn}{}^2 \beta_z}{2\omega\mu\varepsilon^2}\left[\beta_x{}^2\left(\frac{b\sin(2n\pi)}{4n\pi}+\frac{b}{2}\right)\left(\frac{a}{2}-\frac{a\sin(2m\pi)}{4m\pi}\right) + \beta_y{}^2\left(\frac{a\sin(2m\pi)}{4m\pi}+\frac{a}{2}\right)\left(\frac{b}{2}-\frac{b\sin(2n\pi)}{4n\pi}\right)\right], \quad (9)$$

where β_z is the phase constant in the +z direction.

The real part of surface impedance for the sidewalls is given in equation (10) and the power losses because of the conductivity is given in equation (11). I is assumed that the dissipated power for both the left wall and the right walls are the same:

$$R_{s\ell} = \sqrt{\frac{\omega\mu_\ell}{2\sigma_\ell}}, \quad (10)$$

$$\frac{P_{c\ell}}{z_0} = \frac{R_{s\ell}A_{mn}{}^2}{2(2\omega\mu_0)^2}\left|\frac{1}{\varepsilon}+\frac{\Gamma_\ell}{\varepsilon}-\frac{\tau_\ell}{\varepsilon_\ell}\right|^2\left\{\left(\frac{b\sin(2n\pi)}{4n\pi}+\frac{b}{2}\right)\beta_c{}^4 + \left(\frac{b}{2}-\frac{b\sin(2n\pi)}{4n\pi}\right)\beta_y{}^2\beta_z{}^2\right\}. \quad (11)$$

Here $P_{c\ell}$ is the dissipated power on the sidewalls due to conductivity, β_c is the cutoff phase constant, μ_0 is the permeability of the air medium in the tunnel. The conduction losses on the top and bottom surface, P_{cb} , is given by:

$$\frac{P_{cb}}{z_0} = \frac{R_{sb}A_{mn}{}^2}{2(2\omega\mu_0)^2}\left|\frac{1}{\varepsilon}+\frac{\Gamma_b}{\varepsilon}-\frac{\tau_b}{\varepsilon_b}\right|^2\left\{(\beta_x\beta_z)^2\left(\frac{a}{2}-\frac{a\sin(2m\pi)}{4m\pi}\right) + \beta_c{}^4\left(\frac{a\sin(2m\pi)}{4m\pi}+\frac{a}{2}\right)\right\}, \quad (12)$$

where P_c is the total loss in equation (13) and perturbational method is preferred as presented in equation (13):

$$\alpha_c = \frac{P_c/z_0}{2P_{mn}} = \frac{\left(\frac{2P_{cb}}{z_0}+\frac{2P_{c\ell}}{z_0}\right)}{2P_{mn}} = \frac{P_{c\ell}+P_{cb}}{z_0 P_{mn}}. \quad (13)$$

In order to find the total conduction losses associated with a rectangular waveguide, surface currents are obtained by using cross product of the magnetic field and normal unity vector along the *x-z* plane at the *y=0*, *y=b* and *y-z* plane at the *x=0*, *x=a* respectively. Total linear surface currents are obtained by means of magnetic fields on the surfaces and the total conduction losses are evaluated by integrating scalar product of the linear surface currents along the all four surfaces and the losses on the four walls are evaluated as total conduction losses (P_c) accordingly. Since the top and bottom losses are assumed as equal, total losses on these walls is equal to $2 \cdot P_{cb}$ and sidewalls losses are also assumed equal and total losses on the sidewalls is equal to $2 \cdot P_{c\ell}$. The received power is therefore given by:

$$P_2 = P_0 \exp(-2\alpha_c z), \quad (14)$$

where P_0 is the reference power. The weighting signal attenuation for the waveguide model is given by:

$$P^{Waveg} = \sum_{i=1}^{+\infty} B_i P_2, \quad (15)$$

where B_i are the weighting coefficients.

C. Far region propagation

Far region signal propagation model is given by the following equation:

$$P_3 \cong \frac{A_0^2 ab}{2\eta_0 d^2}, \quad (16)$$

where A_0 is the constant of the integration from a differential equation. a and b are the tunnel cross-section dimensions, d is the distance from the transmitter to the receiver antennas, and η_0 is the intrinsic impedance of the free space (the propagation medium) [12]. Received power for the far zone (P^{Far}) is given by:

$$P^{Far} = \sum_{i=1}^{+\infty} C_i P_3, \quad (17)$$

where C_i are the corresponding weighting coefficients. Far region propagation model comparison with other models and signal measurements is presented in Fig. 2, Fig. 3 and Fig. 4.

III. COMBINED PROPAGATION MODEL

As can be observed from Fig. 2, 3 and 4 each propagation model presents a different propagation characteristic, and it is not possible to use just a single one to model the propagation in a tunnel since signal level at different distances match different propagation models through the tunnel. This case is also verified by the usage of the measurements presented in Fig. 7. Therefore, a new model which is a combination of the three models is proposed in this study. The tunnel is divided into five zones, three of them have dominant propagation models, and two are transient regions between the dominant models as presented in Table 1. The propagation models presented in Section II are used with different weighting coefficients in these zones. The propagation model for the first and the closest distance region is the log-distance path loss model. For the second part propagation model gradually transforms from the log-distance path loss model to the waveguide-based propagation model. The propagation model for the third region is defined by the waveguide model is also proposed in this paper which is calculated via the perturbation method [13]. The fourth region is a transient region transforming gradually to the far zone propagation model from the proposed waveguide model. The

weighting coefficients of both the waveguide model and the far zone signal model are the same in the fourth region. For the fifth and the farthest region, the far zone propagation model is preferred.

IV. MEASUREMENT RESULTS

Signal measurements were taken at the heating infrastructure tunnels of the Karamanoğlu Mehmetbey University, which was built as a rectangular-shaped tunnel. The underground tunnel has a cross-section of 2 m x 1.5 m. The overall width of the tunnel is 1.5 m including the pipes. Pipes are not taken into account on signal propagation since they are thick plastic pipes carrying water. The total length of the tunnel is 250 m. including concrete walls supported with iron bars. Measurements were taken with an Anritsu MG3694C as the transmitter and an Anritsu MS2830A as the receiver. A photograph taken during measurements is presented in Fig. 5. Measurements are carried out at 980 MHz, 1000 MHz and 1150 MHz and both of the antennas are placed at a height of 1.5 m to imitate an average person's communication equipment height. Measurement results and their comparison with the propagation models are presented in Fig. 2, Fig. 3 and Fig. 4. It can be observed that none of the pre-mentioned propagation models had compatibility with the measurement results, but they do comply with the measurements at specific distances. For the closest part, the log-distance propagation model shows a good match while for most of the graph, the waveguide model shows a good match between 120λ and 700λ for the measurement. After 700λ the model is closer to the far region propagation model. The distances comply with the proposed model, then we need to determine the weights for the transient region I and II in Table 1.

As presented in Table 1, the first region is between 0λ and 50λ. For the first region, the propagation model is dominant as mentioned before while for region three in Table 1 waveguide propagation model is dominant from 120λ until 700λ. Therefore, the second region starts at 50λ and ends at 120λ. For the second region, a transition between these two models is necessary, as presented in Table 1 weighting coefficients (in Fig. 6) provide a gradual transition between the two models. After the third region, we again need a gradual transition between the waveguide model and the far region model.

This transition is different from the first transition in most of the second transition region in Table 1 waveguide model and the far region model are equally weighted. In our measurements, it was not possible to take any measurements in region 5 in Table 1 due to the physical limitations of the tunnel, but measurement results presented in Fig. 2, Fig. 3 and Fig. 4 show a better match with the far region propagation model at the end of the tunnel. A comparison between the measurement results and the proposed model is presented in Fig. 7. As can be observed the model fits the measurement results. The error of the measurements is calculated and compared for different frequencies along the z-axis in Fig. 8.

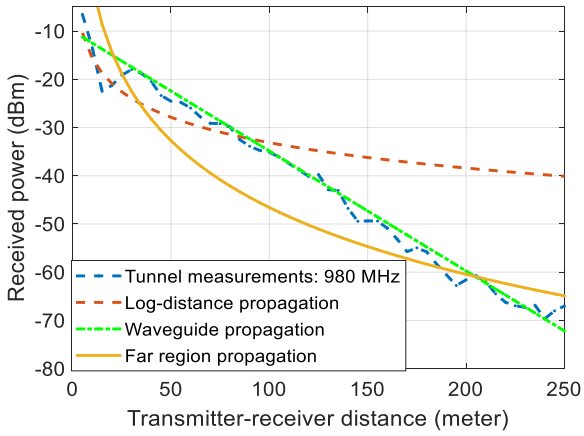


Fig. 2 Path loss of the propagation models and measurements for 980 MHz.

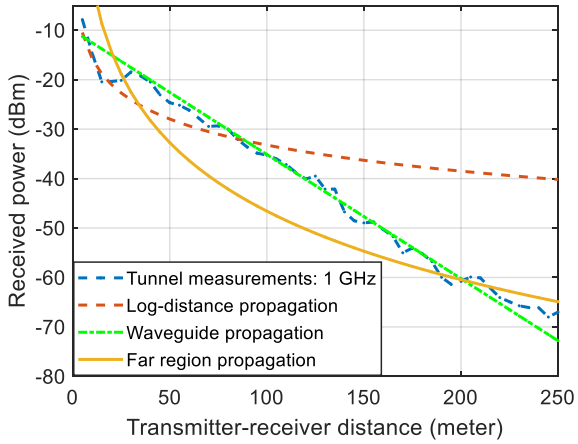


Fig. 3 Path loss of the propagation models and measurements for 1000 MHz.

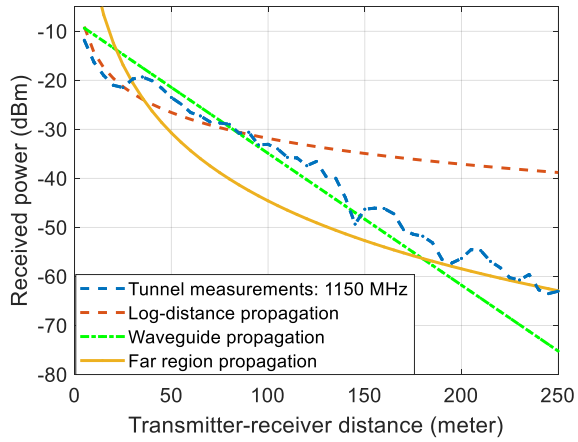


Fig. 4. Path loss of the propagation models and measurements for 1150 MHz.

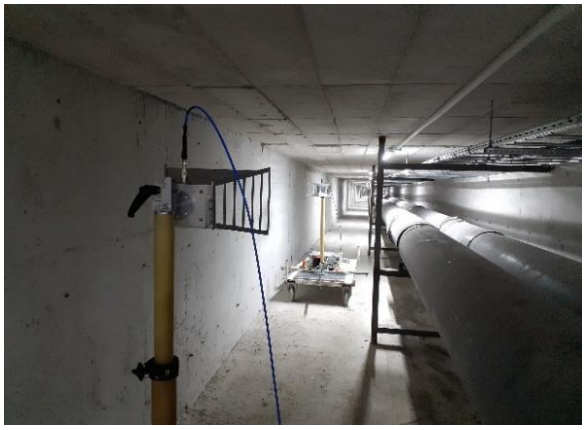


Fig. 5. Signal measurements and underground tunnel.

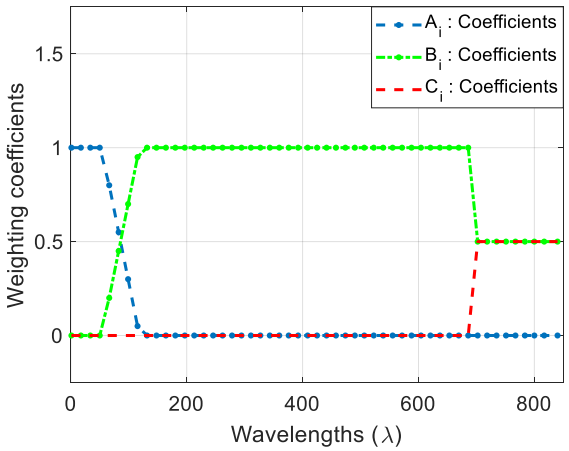


Fig. 6. Applied weighting coefficients for the regions.

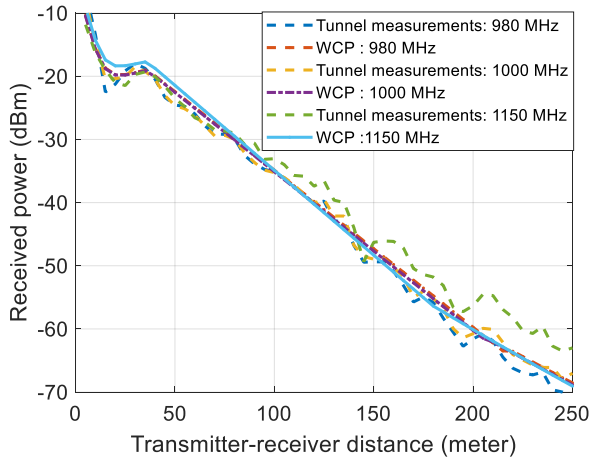


Fig. 7. Comparison of the signal measurements and weighting coefficient propagation models.

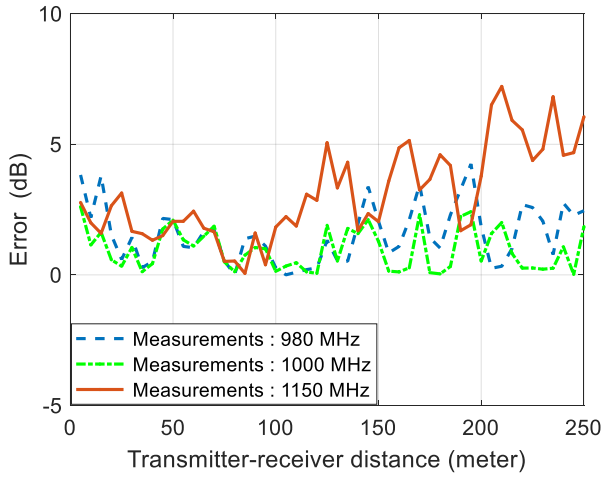


Fig. 8. Error comparison for different frequencies.

V. CONCLUSION

A novel radio wave propagation model is introduced for rectangular cross-section tunnels in this paper. The model is based on the idea of separating the propagation medium into five regions with different propagation characteristics based on the distance in terms of wavelength between the transmitter and the receiver. Three areas of the propagation model have dominant propagation characteristics when two regions are transition regions among the three regions. The transient region propagation models are weighted, which improves the consistency of the proposed model. A waveguide signal model used in the second, third, and fourth regions a novel propagation model proposed for tunnel propagation at UHF frequency band. Measurements are conducted at 980 MHz, 1000 MHz and 1150 MHz in a tunnel of the heating system. According to the results, the proposed signal propagation model complies with the underground signal measurement values. It was possible to validate the model in the fifth region at 1150 MHz since the wavelength is smaller and the length of the tunnel is barely enough for this frequency.

REFERENCES

[1] A. E. Forooshani, S. Bashir, D. G. Michelson, and S. Noghanian, "A survey of wireless communications and propagation modeling in underground mines," *IEEE Commun. Surv. Tutorials*, vol. 15, no. 4, pp. 1524-1545, 2013.

[2] A. Hrovat, G. Kandus, and T. Javornik, "A Survey of radio propagation modeling for tunnels," *IEEE Commun. Surv. Tutorials*, vol. 16, no. 2, pp. 658-669, 2014.

[3] Y. P. Zhang and H. J. Hong, "Ray-optical modeling

of simulcast radio propagation channels in tunnels," *IEEE Trans. Veh. Technol.*, vol. 53, no. 6, pp. 1800-1808, 2004.

[4] Y. P. Zhang, G. X. Zheng, and J. H. Sheng, "Radio propagation at 900 MHz in underground coal mines," *IEEE Trans. Antennas Propag.*, vol. 49, no. 5, pp. 757-762, 2001.

[5] A. G. Emslie, R. L. Lagace, and P. F. Strong, "Theory of the propagation of UHF radio waves in coal mine tunnels," *IEEE Trans. Antennas Propag.*, vol. 23, no. 2, pp. 192-205, 1975.

[6] M. Boutin, A. Benzakour, C. L. Despins, and S. Affes, "Radio wave characterization and modeling in underground mine tunnels," *IEEE Trans. Antennas Propag.*, vol. 56, no. 2, pp. 540-549, 2008.

[7] K. Guan and Z. Zhong, "Complete propagation model structure in-side tunnels," *Prog. Electromagn. Res.*, vol. 141, pp. 711-726, 2013.

[8] M. Rak and P. Pechac, "UHF propagation in caves and subterranean galleries," *IEEE Trans. Antennas Propag.*, vol. 55, no. 4, pp. 1134-1138, 2007.

[9] H. I. Volos, C. R. Anderson, W. C. Headley, R. M. Buehrer, C. R. C. M. Da Silva, and A. Nieto, "Preliminary UWB propagation measurements in an underground limestone mine," *GLOBECOM - IEEE Glob. Telecommun. Conf.*, pp. 3770-3774, 2007.

[10] O. M. Abo Seida, "Propagation of electromagnetic waves in a rectangular tunnel," *Appl. Math. Comput.*, vol. 136, no. 2-3, pp. 405-413, 2003.

[11] N. R. Peplinski, F. T. Ulaby, and M. C. Dobson, "Dielectric properties of soils in the 0.3–1.3-GHz range," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 3, pp. 803-807, 1995.

[12] J. Molina-Garcia-Pardo, J. Rodriguez, and L. Juan-Llacer, "Wide-band measurements and characterization at 2.1 GHz while entering in a small tunnel," in *IEEE Transactions on Vehicular Technology*, vol. 53, no. 6, pp. 1794-1799, 2004.

[13] D. Pozar, *Microwave Engineering.* 3rd ed., John Wiley and Sons, 2005.

[14] M. Ozturk, U. K. Sevim, O. Akgol, E. Unal, and M. Karaaslan, "Determination of physical properties of concrete by using microwave nondestructive techniques," *Applied Computational Electromagnetics Society Journal*, vol. 33, no. 3, pp. 265-272, 2018.

Yusuf Karaca received his B.S. degree from the Department of Electronic Engineering, Uludağ University, in 2010, Bursa, Istanbul, Turkey and M.S. degree from Department of Electrical-Electronics Engineering, Istanbul University, in 2014, Istanbul, Turkey. Currently, he is a graduate Research Assistant School of Department of Electrical-Electronics Engineering, Dokuz Eylül University, İzmir, Turkey. He is pursuing a Ph.D. degree under the supervision of Asst. Prof. Özgür Tamer. His current research interests are in Electromagnetic Wave Propagation, Wireless Underground Communication Networks, and Wireless Sensor Networks.

Özgür TAMER received his BSc., MSc. in and Ph.D. at Dokuz Eylul University Electrical and Electronics Eng. Dept in 1997, 2001, and 2008 respectively. He worked at NTNU Telematics Department at Trondheim/Norway as a postdoc researcher from 2010-2011 He is currently working as an Assistant Professor at Dokuz Eylul University Electrical and Electronics Eng. Dept. İzmir/Turkey. His main research interests are focused on smart antenna systems, embedded systems, and wireless sensor networks.

Characteristic Analysis and Control of a Rotary Electromagnetic Eddy Current Brake

Qiao Ren, Jimin Zhang, Hechao Zhou*, and Jinnan Luo

Institute of Railway Transit
Tongji University, Shanghai, 201800, China
R138150187Q@163.com, zjm2011@tongji.edu.cn, *zhouhechao@tongji.edu.cn, tjljn@tongji.edu.cn

Abstract ─ This article designs an electromagnetic rotating eddy current brake (ECB), which has the advantages of no wear and low noise compared with traditional friction brake. First, using the magnetic circuit analysis model, a theoretical calculation formula of the ECB's braking characteristics is given. The results show that the braking torque is negatively correlated with the thickness of the air gap as well as the electrical conductivity and the relative magnetic permeability of the brake disc material, and positively correlated with the number of ampere turns and the number of electromagnet poles. Secondly, a three-dimensional finite element (FE) model of the brake is established. The results of braking torque-speed characteristics between finite element calculation and theoretical analysis are compared, and the reasons for the differences between the two are explained. Using the FE model, the influence of the design parameters on torque characteristics is studied. Combined with the theoretical analysis model, the results are explained accordingly, providing a reference for the optimal design of the brake. Finally, a controller for the electromagnetic rotating eddy current brake is designed to control the amplitude of the desired braking torque.

Index Terms ─ Braking characteristics, braking torque control, eddy current brake, finite element analysis.

I. INTRODUCTION

With the development of high-speed railways, the running speed of trains has increased significantly [1-4]. Braking systems are the guarantee of operation safety, especially for trains running at high speeds. Because the kinetic energy of high-speed trains is much larger than that of ordinary trains, both the adhesion coefficient between wheels and rails and the friction coefficient between brake shoes and moving wheels are greatly reduced at high speeds, therefore traditional mechanical braking methods cannot meet the emergency braking requirements of high-speed trains [5-7]. And friction brake has obvious defects such as fast wear and loud noise under high-speed braking conditions. The ECB is one of the most popular braking systems because of significantly lower maintenance, no wear, and excellent braking performance at high speeds [10-11] as compared to other braking devices. The main advantage of an ECB, making it superior to the other brakes [12], is its fast dynamic response.

In this paper, an electromagnetic rotating ECB was designed, as shown in Fig. 1, mainly consisting of two parts: stator and rotor. The rotor, which is termed as the brake disc and is rotating with the axle. The stator, which is the stationary part and fixed to the train, includes several electromagnets and magnetic yokes. Each electromagnet is made up of an iron core and its surrounding coils, corresponding to the red and yellow parts in Fig. 1. When passing a certain amount of direct current (DC) into the coil, the electromagnet will generate an excitation magnetic field, meanwhile the brake disc rotating with the axle cuts magnetic induction lines in the electromagnetic field. According to the principle of electromagnetic induction, the kinetic energy of the train is converted into the eddy current in the brake disc, dissipating into the surroundings in the form of heat to achieve the purpose of braking [8-9].

Braking characteristics are defined as the variation law of braking torque with rotational speed of brake disc. While the study of braking characteristics provides the basis for making full use of ECB, current research work mostly performed their analyses of the braking characteristics of electromagnetic ECB at the theoretical and simulation level. Based on the basic macro-electromagnetic field theory, the variable separation method (VSM) was used to solve the analytical formula of braking torque [1]. In [13], a calculation method using the Laplace and Helmholtz equations was presented to analyze the eddy current generated by the transverse alternating magnetic field in two parallel conductors. In [14], Lee used the mirror image method to calculate the electric field intensity of a conductor plate with a finite radius, then introduced the magnetic flux Reynolds number to consider the armature reaction, and finally obtained a formula for calculating the braking torque through using the Lorentz formula. In [15], Liu used

Submitted On: March 12, 2021
Accepted On: July 7, 2021
https://doi.org/10.47037/2021.ACES.J.360816

the Maxwell's equations and appropriate boundary conditions to solve magnetic vector potential and magnetic flux density, and obtained their braking torque formula based on Maxwell stress tensor approach. Most recently, Reza [16] used subdomain analysis to compute the braking force for a magnetic pole moving above the induction plate.

Studies mentioned above have shown that using theoretical analysis methods to find solutions often requires a large number of premise assumptions and problem simplifications, rendering the accuracy of their results limited. For this reason, substantial efforts have been put into finite element analysis (FEA) to improve braking characteristics analysis results. For example, a 2D FE model for linear eddy current brakes was established by [17]. Compared with the simulation results, their calculation results have a higher degree of agreement in the low-speed area but greater deviation in the high-speed area. Using three-dimensional FEA, Mehmet [18] verified the effectiveness of braking torque theoretical modeling, and simulated the resistivity and permeability of the best brake disc material with finite element to obtain the maximum braking torque. Sohel [19] established a parameter model based on the FE model calculation results of braking torque characteristics, fitting the relationship between braking torque and current and speed into a polynomial function. In [20], eddy current losses of the transformer were calculated by a 2D FE model. Their use of the Galerkin method to apply surface impedance boundary conditions to finite element calculation effectively shortened the calculation time.

As for the research on braking torque control methods, literature [21] designed a constant torque control algorithm based on the principle of controlling a constant ratio between the reference current and the speed, and verified the effectiveness of this algorithm through proportional experiments. In order to keep the braking torque constant, a dynamical compensation scheme for speed control was proposed by [22].

In this paper, an electromagnetic rotating ECB of high-speed trains is designed. We first use the magnetic equivalent circuit (MEC) to study the dynamic characteristics of the ECB system, and offer a brief mathematical model of the braking characteristics to clarify the main parametric factors that affect the braking torque. Secondly, a three-dimensional FE model of the brake is established to specifically quantify the influence of each parameter factor on the braking characteristics. Then the differences between the braking torque-speed-current characteristics curve calculated by the theoretical model and the FE model are compared and analyzed to offer sensible explanations. Finally, a fuzzy controller for the electromagnetic rotating ECB is designed to accurately track the desired braking torque.

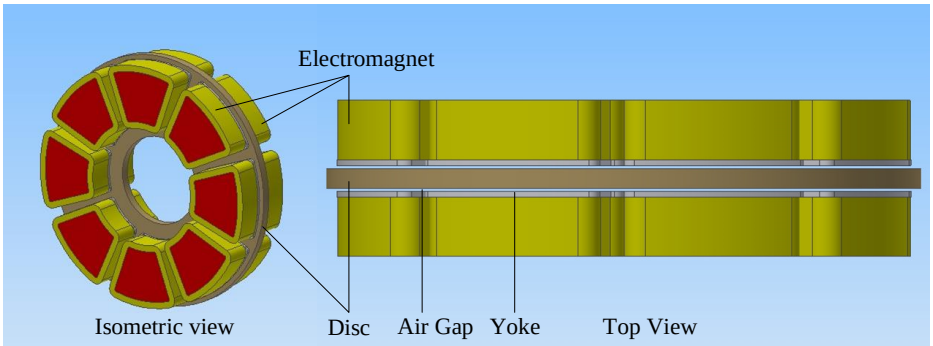


Fig. 1. Configuration of the eddy current brake.

II. ANALYTICAL MODEL

This article derives the brake disc eddy current, air gap magnetic field and eddy current power according to the differential principle. Starting from the law of energy conservation, the eddy current power is regarded equal to braking power, and the braking torque is obtained according to the relationship between power and torque.

A. Model assumptions

To facilitate the derivation of braking torque, the following assumptions are made for the model:

(1) The influence of temperature on the electromagnetic properties are neglected.

(2) The magnetic flux lines are ideally distributed only within the projection areas on the disc of yokes.

(3) The magnetic saturation and nonlinearity of magnetic properties are neglected.

B. Eddy current analysis

The projection area of a single yoke on the brake disc is a sector, as shown in Fig. 2, whose inner and outer radii are r_1 and r_2, respectively, with a 40° angle. From the perspective of differentiation principle, this sector is regarded as composed of countless similar sector-shaped rings (shadowed) with a width of dr, and the position of these sector-shaped rings can be represented by the

distance r from the inner and outer arcs to the middle radius of the large sector, where $r \in [0, \Delta r]$, $\Delta r = r_c - r_1 = (r_2 - r_1)/2$, $r_c = (r_1 + r_2)/2$.

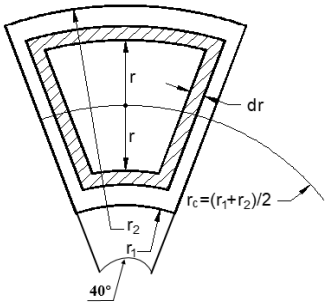


Fig. 2. Projection area of a single yoke on the brake disc.

The magnetic flux passing through the sector area can be expressed as:

$$\phi_r = BS_r \cos \omega_n t = \frac{4\pi r_c r B \cos \omega_n t}{9}, \quad (1)$$

where, S_r is the area of the inner sector surrounded by the sector ring, which can be calculated according to Eq. (2) below. B is the magnetic induction density value on the surface of the brake disc, and ω_n is the electrical angular velocity of the brake disc:

$$S_r = \frac{40°}{360°} \times \pi[(r_c + r)^2 - (r_c - r)^2] = \frac{4\pi r_c r}{9}, \quad (2)$$

$$\omega_n = N_p \cdot \omega = N_p \cdot \frac{2\pi n}{60} = \frac{\pi N_p n}{30}, \quad (3)$$

where N_p is the number of pole pairs on one side of the disc, ω is the angular velocity of the disc, and n is the rotation speed of the disc.

The induced electromotive force is:

$$\varepsilon_r = -\frac{d\phi_r}{dt} = \frac{4\pi r_c r B \omega_n \sin \omega_n t}{9}. \quad (4)$$

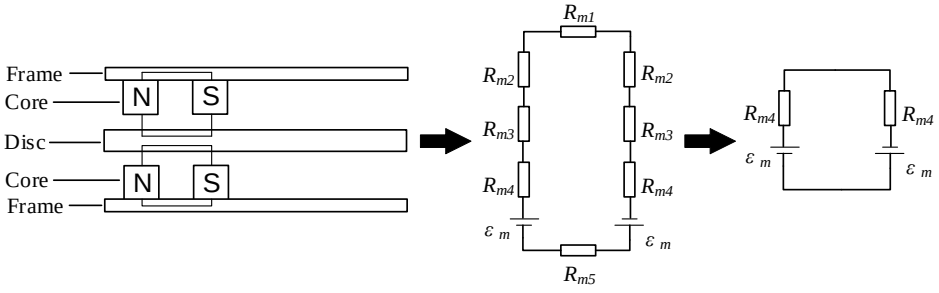


Fig. 3. Magnetic circuits.

According to Fig. 3, the magnetic circuit can be written as:

$$\varepsilon_m = R_{m4} \cdot \phi, \quad (9)$$

where the magnetoresistance of the air gap is:

$$R_{m4} = \frac{\delta}{\mu_0 \mu_{r4} S_p} \approx \frac{\delta}{\mu_0 S_p}, \quad (10)$$

where, δ and μ_{r4} are the length of air gap and air relative permeability, respectively.

The resistance of the sector ring is:

$$R_r = \rho \frac{L_r}{d' \cdot dr}, \quad (5)$$

where, d' and L_r are the skin depth of the vortex and the length of the sector ring, respectively. And d' satisfies the following formula:

$$d' = \sqrt{\frac{2\rho}{\mu_0 \mu_r \omega_n}}, \quad (6)$$

where, ρ is the resistivity of the brake disc, μ_0 is the vacuum permeability, μ_r is the relative permeability of the brake disc.

Thus, according to the Ohm's law, the current value on the sector ring, that is, the eddy current value can be expressed as:

$$i_r = \frac{\varepsilon_r}{R_r} = \frac{\sqrt{2} r_c r B \omega_n \sin \omega_n t dr}{(r + \pi r_c)\sqrt{\rho \mu_0 \mu_r \omega_n}}. \quad (7)$$

C. Magnetic field analysis

The Ohm's law for a magnetic circuit can be derived from Maxwell equations:

$$\varepsilon_m = \sum R_{mi} \cdot \phi_i, \quad (8)$$

where ε_m is the magnetomotive force, ϕ is the magnetic flux, $\phi = BS_p$, S_p is the area of the total sector shown in Fig. 2, R_{mi} represents every magnetic resistance in the magnetic circuit.

The simplified magnetic circuit and the definition of each part reluctances are shown in Fig. 3. R_{m1}, R_{m2}, R_{m3}, R_{m4}, R_{m5} are the magnetic resistances of the frame, core, yoke, air gap, and disc, respectively. Since the relative permeability of the frame, core and disc is much higher than that of air which approximately equals 1, only R_{m4} is taken into consideration in this paper.

The total magnetic motive force ε_m is given as:

$$\varepsilon_m = \varepsilon_0 - \varepsilon_e , \quad (11)$$

where ε_0 and ε_e are the magnetic motive forces generated by electromagnets and eddy current, respectively, and they are given as:

$$\varepsilon_0 = NI , \quad (12)$$

$$\varepsilon_e = k_e \cdot I_e , \quad (13)$$

where N is the number of turns per coil, I is the excitation current, k_e is the computation coefficient, I_e is the RMS eddy current which can be obtained by integration of the eddy current i_r:

$$I_e = \frac{1}{\sqrt{2}} \int_0^{\Delta r} i_r = \frac{r_c B \omega_n}{\sqrt{\rho \mu_0 \mu_r \omega_n}} \cdot (\Delta r - \pi r_c \ln \frac{\Delta r + \pi r_c}{\pi r_c}) . \quad (14)$$

From Eqs. (9)-(14), we can derive that the magnetic flux density is:

$$B = \frac{\mu_0 NI \sqrt{\rho \mu_0 \mu_r}}{\delta \sqrt{\rho \mu_0 \mu_r} + k_e r_c \mu_0 \sqrt{\omega_n} (\Delta r - \pi r_c \ln \frac{\Delta r + \pi r_c}{\pi r_c})} . \quad (15)$$

D. Power analysis

The RMS power of eddy current of the whole disc P_e can be obtained by integration of the product of eddy current i_r and the induced electromotive force ε_r:

$$\begin{aligned} P_e &= 2 \times \frac{360°}{40°} \times \frac{1}{\sqrt{2}} \int_0^{\Delta r} i_r \cdot \varepsilon_r \\ &= \frac{4\sqrt{2} \pi r_c^2 B^2 \omega_n^2}{\sqrt{\rho \mu_0 \mu_r \omega_n}} \cdot [\frac{1}{2} \Delta r^2 - \pi r_c \Delta r \\ &\quad + (\pi r_c)^2 \cdot \ln \frac{\Delta r + \pi r_c}{\pi r_c}] . \end{aligned} \quad (16)$$

E. Torque analysis

The brake power P_b equals the RMS power of eddy current P_e in accordance with the law of conservation of energy, therefore the following expression can be obtained:

$$P_b = P_e = T_b \cdot \omega , \quad (17)$$

where T_b is the torque generated by the eddy current brake. From (16)-(17) and the relation between the angular velocity ω and the train speed v, we can derive that the brake torque is:

$$T = \frac{4\sqrt{10} \pi k_1 r_c^2 \mu_0^{\frac{5}{2}} (\mu_r \rho)^{\frac{1}{2}} (NI)^2 N_p^{\frac{3}{2}} D_w^{-\frac{1}{2}} v^{\frac{1}{2}}}{3(\delta \sqrt{\rho \mu_0 \mu_r} + k_2 k_e r_c \mu_0 \sqrt{\frac{5 N_p v}{9 D_w}})^2} , \quad (18)$$

where k_1, k_2 are both coefficients that are determined by the geometric parameters of the brake device, D_w is the diameter of the wheel.

III. VALIDATION OF THE FE MODEL

Brake torque is influenced by the geometric and electromagnetic parameters according to the above theoretical model. More specifically, there is a negative correlation between brake torque and the length of air gap, the inner radius of yoke, the product of the resistivity and permeability of disc material, and a positive correlation between brake torque and the number of poles, the outer radius of yoke. In addition, brake torque is proportional to the square of ampere turns of electromagnets. In order to analyze the quantitative effects of each factor on the braking performance and verify the accuracy of the analytical model, a 3D FE model was established for further numerical calculations, as shown in Fig. 4. Only the brake disc, yokes, and electromagnets were considered in the FE model because of the assumption that the other components only act as fixed support and no magnetic lines pass through them. So, their effects on the magnetic field can be ignored. In the FE model, the material of electromagnet core and brake disc is mild steel steel_1008. The material of magnetic yoke is mild steel steel_1010, and the material of electromagnet coil is copper. A 1/4Np simplified model is adopted to reduce simulation time. For this reason, a symmetry boundary condition is added in the middle section of the brake disc. Master and slave boundary conditions are implemented in the two sides considering the electromagnets' periodic distribution. Main parameters of the FE model are given in Table 1.

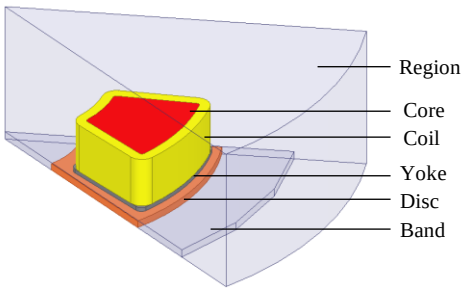


Fig. 4. FE model.

The magnetic flux density map, eddy current losses and the distribution of induced eddy currents are shown in Figs. 5 (a), (b) and (c) respectively. It can be seen that 2.2T magnetic flux density is obtained by the FE model. And at the same parameters, the air gap magnetic flux density of 2.34T is calculated based on the analytical model. The value of relative error is about 6.3%. Hence, the results of the analytical model are in good agreement with the FE model results. In addition, eddy current losses are calculated by the FE model and Eq. (16), as shown in Fig. 6. It can be seen that both the analytical model and FE model have the same linear growth trend. The eddy current losses increase with the speed increases.

Table 1: ECB Parameters

Parameter	Value
Inner radius of the disc	210mm
Outer radius of the disc	600mm
Disc thickness	20mm
Air gap width	3mm
Inner radius of the yoke	280mm
Outer radius of the yoke	580mm
Yoke thickness	6mm
Core thickness	60mm
Angle covered by a single electromagnet	40deg
Angle between two adjacent electromagnets	45deg
Number of pole pairs on one single side of the disc	4
Number of turns per coil	400
Excitation current	40A

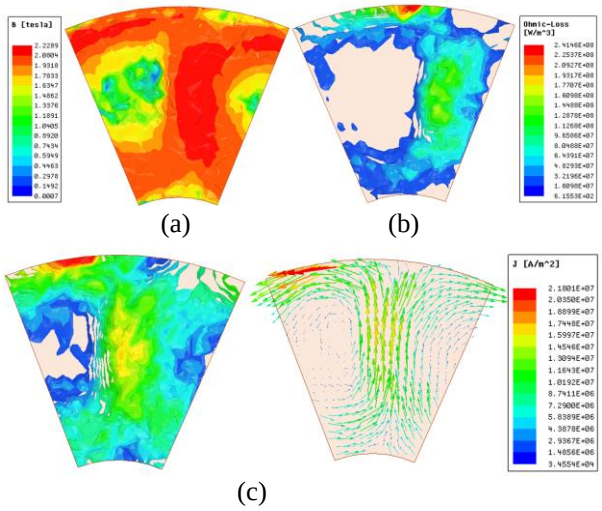


Fig. 5. (a) The magnetic flux density map, (b) eddy current loss, and (c) distribution of induced eddy currents.

At the initial simulation time, a large amplitude fluctuation of the braking torque appeared and then converged to a steady value. This phenomenon is inevitable due to the initial value of calculation changes to the target value suddenly. It was only regarded steady-state values as the effective data and calculated the final torque-speed characteristic curve. Figure 7 shows a comparison between the FE model and analytical model of braking torque-speed characteristics. It is observed that the braking characteristics tendency of the analytical model is consistent with that of the FE model. Still, the results from analytical method in Eq. (18) are slightly larger than those from FE at the whole speed range. When the vehicle speed is higher than the critical speed, the decrease of braking torque calculated by FE is more evident than that calculated by Eq. (18). This is due to the assumptions made to simplify the theoretical analysis, especially assumption (3), in which we assumed that the magnetic saturation and nonlinearity of magnetic properties are neglected. Additionally, in the analytical model, the effects of flux leakage and uneven distribution of magnetic field lines are ignored, which also cause the theoretical value to be too large.

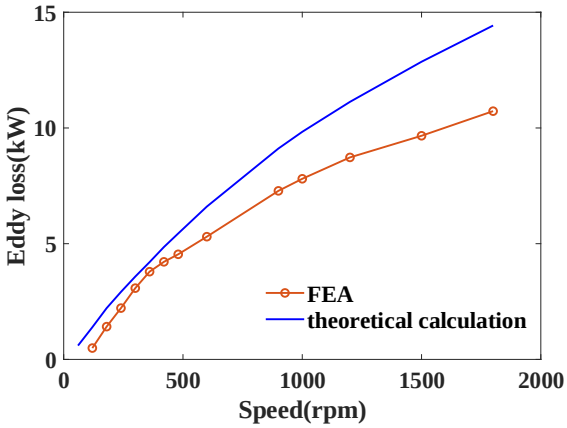


Fig. 6. Eddy current losses.

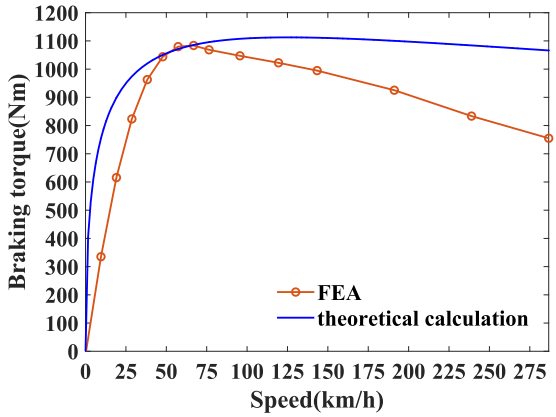


Fig. 7. Braking torque-speed characteristic.

IV. PARAMETER ANALYSIS

It is found by FEA that the excitation current, air gap thickness, brake disc thickness and electromagnet core shape have more significant influence on the braking characteristic, so the influence of these four factors on braking characteristics is specifically analyzed in this paper.

A. Influence of the excitation current

Figure 8 shows the variation of braking torque characteristics for I=40,60,80,100A while keeping other parameters unchanged. It is obvious that the braking torque increases gradually as the excitation current increases. And the peak torque increases in a trend of

linear function. However, from Eq. (18), the braking torque should be proportional to the square of the excitation current. The main reason for this difference is that electromagnets have nonlinear magnetization characteristics. The magnetic permeability decreases as the excitation current increases so that the magnetic induction intensity shows a nonlinear growth trend with a declining growth rate. Furthermore, the critical speed increases with the increase of excitation current. According to Eq. (6), the skin depth increases as the magnetic permeability decreases, which weakens the eddy current skin effect. Therefore, if the eddy current magnetic field is to reach its original strength, it needs to increase the speed [17].

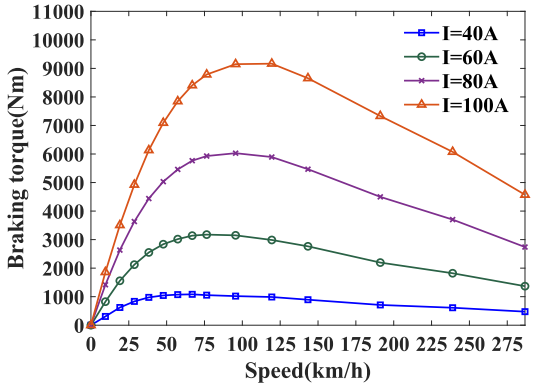


Fig. 8. Braking torque-speed characteristics for different excitation currents.

B. Influence of the air gap length

Figure 9 shows the variation of braking torque characteristics for air gap=3,4,5,6mm, while other parameters are kept to values introduced earlier. It is observed that a shorter air gap length leads to higher braking torque values. This is due to a large air gap length will result in a larger magnetoresistance, which will cause a lower magnetic density. In addition, for a large air gap, the torque-speed characteristic curves of the high-speed region are flatter than that of a small air gap. This is due to the former's eddy current being far away from the yoke compared to the latter, which causes a lesser demagnetization effect on the original magnetic field.

C. Influence of the conductor plate thickness

The effect of the conductor plate thickness on the braking torque-speed characteristic has been studied by changing the thickness from 15 to 30 mm in a step of 5mm. From Fig. 10, the thinner the conductor plate is, the higher the peak torque and the critical speed would be for the ECB. The reason is that eddy currents in the conductor plate increase as the thickness of the conductor plate increases, which intensifies the influence of the eddy current magnetic field on the excitation magnetic field. Therefore, the braking torque reaches the peak value at a low speed, and the peak torque becomes small due to the eddy current demagnetization effect.

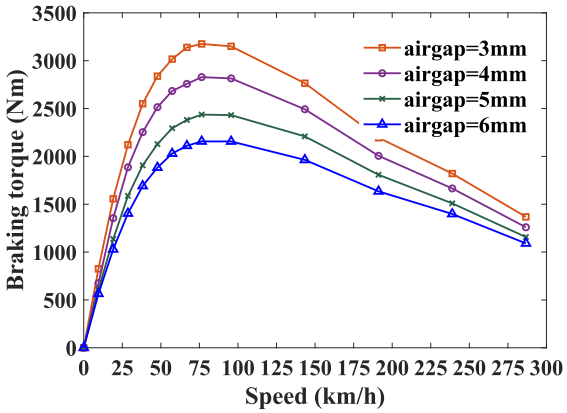


Fig. 9. Braking torque-speed characteristics for different air gap lengths.

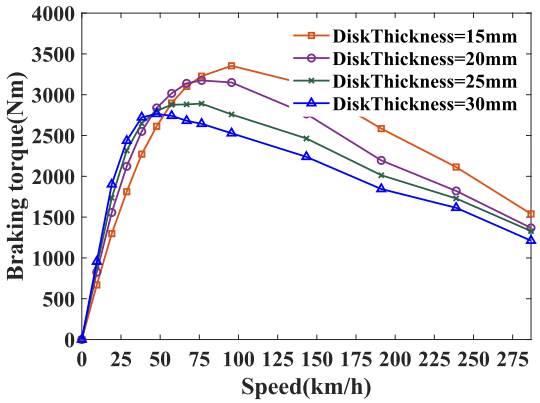


Fig. 10. Braking torque-speed characteristics for different disc thicknesses.

D. Influence of the pole shape

The electromagnet core is designed to be round, square, and sector in turn while keeping the cross-sectional area of the electromagnet core unchanged. The torque-speed characteristics for three kinds of electromagnet core shapes are shown in Fig. 11. We can observe that an optimum torque generation capacity exists for the sector-shaped section electromagnet. Because of the fact that only the tangential electromagnetic force generated by the radial eddy current can contribute to the braking torque. In contrast, the tangential eddy current is ineffective for braking torque generation. Electromagnets with circular and square cross-sections have a lesser radial coverage of the projection area on the brake disc than sector-shaped section electromagnets, resulting in a small effective

braking torque. In order to enlarge the path of the radial eddy current, a sector-shaped section electromagnet is designed. The adequate electromagnetic torque can be enhanced by increasing the eddy current flowing along the radius of the disc with the sector pole [18].

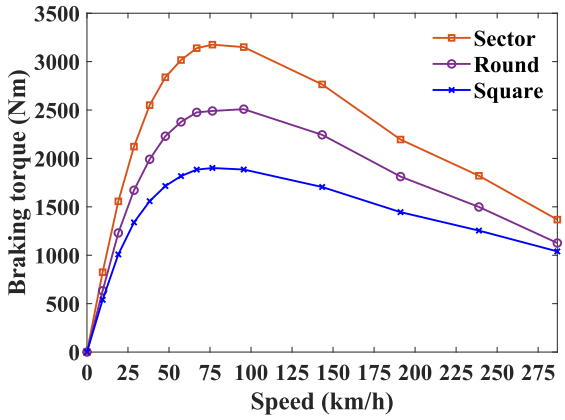


Fig. 11. Braking torque-speed characteristics for different pole shapes.

V. EDDY CURRENT BRAKE CONTROL SYSTEM

A. Braking torque controller design

Based on the above theoretical calculation of braking torque and the FEA of braking characteristics, the braking torque decreases with the increase of speed in the high-speed region. In order to track the desired braking torque accurately, this paper presents a controller based on the fuzzy method to control the amplitude of the braking torque. Figure 12 shows the control scheme of the ECB. It consists of a braking torque controller and a current controller. The braking torque controller outputs the reference of the excitation current change ΔI_{ref} according to the error between the desired braking torque T_{ref} and the actual braking torque T_b. The current controller is composed of a current hysteresis controller and a DC chopper circuit, which outputs a reference voltage u_{ref} according to the reference value of the excitation current change. One can adjust the excitation current to regulate the value of the braking torque by controlling the reference voltage.

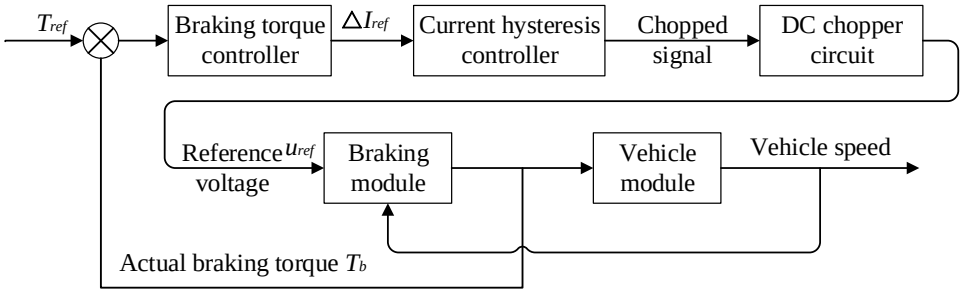


Fig. 12. The overall scheme of the torque control system.

B. Control algorithm description

In the fuzzy controller, the system input variables are the braking torque error e and its rate of change ec. The output variable is the change of reference current ΔI_{ref}. Design variables use the same set of fuzzy state words {NB, NS, PS, PB}, in which each fuzzy state corresponds to {'negative big', 'negative small', 'positive small', 'positive big'}, denoting the variables sign and absolute value size. The membership function for the fuzzy sets is set as a triangular function, as shown in Fig. 13. According to the qualitative relationship between the braking torque error and the change of reference current, the fuzzy rule table is formulated as shown in Table 2. The center of gravity method, the most commonly used defuzzification method, is used in this paper.

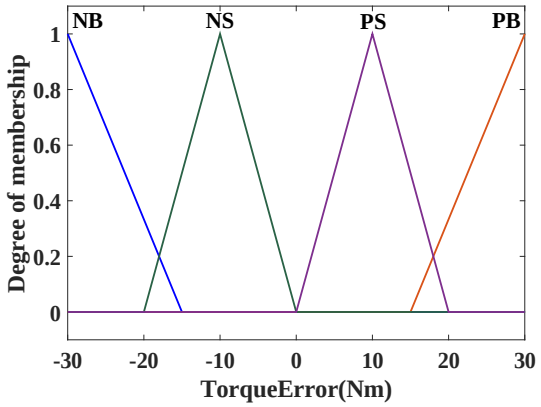


Fig. 13 Membership function for fuzzy sets.

Table 2: Fuzzy rules

Error e	Change of ec			
	NB	NS	PS	PB
NB	NB	NB	NB	NB
NS	NB	NS	NS	NS
PS	NS	PS	PS	PB
PB	PB	PB	PB	PB

VI. ANALYSIS OF THE SIMULATION RESULTS

In Fig. 14, the variation of the braking torque is plotted with respect to varying speeds of the rotating disk under open-loop control and closed-loop control algorithms. In this paper, PID control and fuzzy control are carried out. It is evident that the braking torque increases firstly and then decreases as the speed increases under open-loop control, which causes the fluctuation of the train braking deceleration. Note that the braking torque keeps constant as the speed varies under closed-loop control. This is because the control system adjusts the excitation current in real-time according to the braking torque feedback. In the speed range of 20-160 *km* / *h*, the actual braking torque is controlled near the desired braking torque. As the speed continues to decrease, the maximum braking torque output is limited by the coil current in the low-speed region, the braking torque starts to drop. It can be concluded that the two control algorithms mentioned can effectively control the braking torque to remain constant in an extensive speed range.

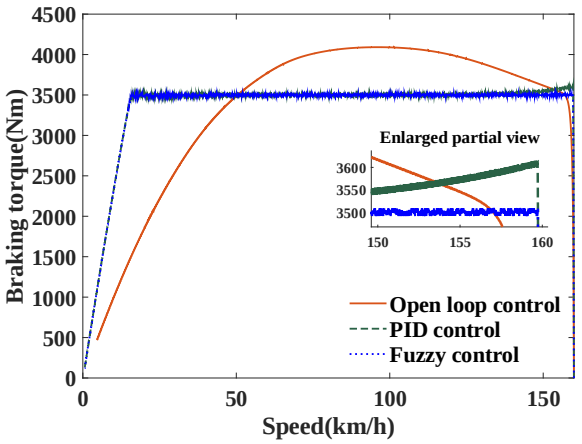


Fig. 14. Brake torque versus train speed under different control algorithms.

In Fig. 15, the time response curve of the braking torque is shown. In order to quantitatively illustrate the control effect of the two algorithms, the transient and steady-state performance indicators of the system are calculated in Table 3. It can be seen that from transient performance indicators, the maximum overshoot of the system under PID control is 111.7N·m, which is far greater than the overshoot of 26.2N·m under fuzzy control. In addition, the adjustment time of the system under PID control is 3.210s, which is longer than 0.318s under fuzzy control. In terms of steady-state performance indicators, the steady-state value of the system under the fuzzy control is closest to the desired braking torque of 3500N·m. However, the result under the PID control deviates from the expected value significantly. It can be concluded that fuzzy control is better than PID control in both transient and steady-state indicators.

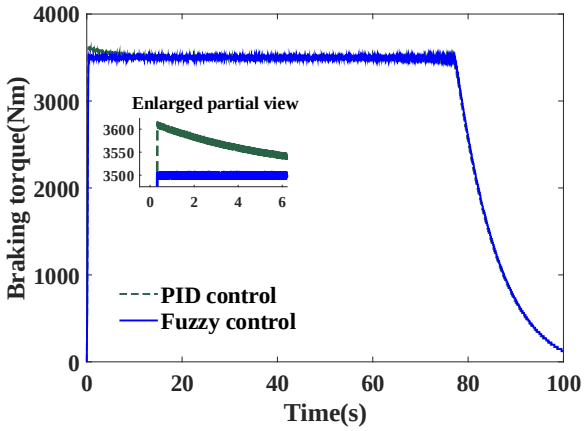


Fig. 15. Brake torque under different control algorithms.

Table 3: System performance indicators

	Performance	PID	Fuzzy Control
Transient indicator	Adjustment time t_s /(s)	3.210	0.318
	Peak time t_p /(s)	0.333	0.323
	Maximum overshoot σ /(Nm)	111.7	26.2
Steady-state indicator	Steady state average /(Nm)	3502.3	3499.3
	Steady state maximum /(Nm)	3524.4	3504.6
	Steady state minimum /(Nm)	3489.1	3497.3

VII. CONCLUSIONS

In this paper, an electromagnetic rotating eddy current brake is regarded as the research object. Studying the braking characteristic from two aspects of theoretical analysis and FEA, and designing a braking torque controller. The main conclusions can be drawn as follows:

(1) The theoretical model shows that the braking

torque is negatively correlated to the air gap length, the electrical conductivity and relative permeability of the brake disc material, and positively correlated to the number of ampere turns and the number of pole pairs.

(2) There are differences between the results of FEA and theoretical analysis. The main reason for the difference is that the theoretical model neglects nonlinear characteristics of the brake disc material, the magnetic saturation phenomenon of electromagnets, and the non-ideal distribution of the magnetic flux.

(3) The relationship between the braking torque and the design parameters such as excitation current, air gap length, brake disc thickness, and electromagnet shape on the braking characteristics is analyzed using the FEA, which provides a reference for the optimal design of the ECB.

(4) An ECB controller based on fuzzy control theory is designed. Compared with open-loop control and PID control, the proposed controller can keep the braking force stable during the train braking process.

ACKNOWLEDGMENT

This work was supported by the Shanghai Collaborative Innovation Center for Multi-network and Multi-mode Rail Transit under Grant 28002360012.

REFERENCES

[1] X. Zhu and X. Zhang, "Analysis and calculation of braking force on rail eddy current braking of high speed trains," [J]. *Journal of Tongji University (Natural Science)*, vol. 17, no. 4, pp. 1-8, Dec. 1996.

[2] C. Yin and K. Zhang, "Electromagnetic force calculation of conductor plate double Halbach permanent magnet electrodynamic suspension," *Applied Computational Electromagnetics Society Journal*, vol. 29, no. 11, pp. 916-922, Nov. 2014.

[3] M. Fujita, T. Tokumasu, T. Yamada, T. Hirose, Y. Tanaka, N. Kumagai, and S. Uchida, "3-dimensional electromagnetic analysis and design of an eddy-current rail brake system," *IEEE Transactions on Magnetics*, vol. 34, no. 5, pp. 3548-3551, Sept.1998.

[4] S. Surenkhorloo and J. K. Byun, "Analysis and case study of permanent magnet arrays for eddy current brake systems with a new performance index," *Journal of Magnetics*, vol. 18, no. 3, pp. 276-282, Sept. 2013.

[5] C. Aldo and B. Vusini, "Design of axial eddy-current couplers," *IEEE Transactions on Industry Applications*, vol. 39, no. 3, pp. 725-733, May 2003.

[6] L. Thierry and A. Rezzoug, "3-D analytical model for axial-flux eddy-current couplings and brakes under steady-state conditions," *IEEE Transactions on Magnetics*, vol. 51, no. 10, pp. 1-12, July 2015.

[7] Z. Ali and A. Mirabadi, "Railway wheel detector in the presence of eddy current brakes," *Applied Computational Electromagnetics Society Journal*, vol. 28, no. 1, pp. 77-84, Jan. 2013.

[8] B. Kou, Y. Jin, H. Zhang, L. Zhang, and H. Zhang, "Analysis and design of hybrid excitation linear eddy current brake," *IEEE Transactions on Energy Conversion*, vol. 29, no. 2, pp. 496-506, Mar. 2014.

[9] K. Kerem, A. Suleman, and E. J. Park, "Analytical modeling of eddy current brakes with the application of time varying magnetic fields," *Applied Mathematical Modelling*, vol. 40, no. 2, pp. 1168-1179, Jan. 2016.

[10] J. R. Tibola, R. L. Sari, T. D. M. Lanzanova, M. E. S. Martins, and H. Pinheiro, "Modeling and control of a low-cost driver for an eddy current dynamometer," *Journal of Control, Automation and Electrical Systems*, vol. 27, no. 4, pp. 368-378, Apr. 2016.

[11] B. Kou, Y. Jin, L. Zhang, and H. Zhang, "Characteristic analysis and control of a hybrid excitation linear eddy current brake," *Energies*, vol. 8, no. 7, pp. 7441-7464, July 2015.

[12] C. Tian, M. Wu, L. Zhu, and J. Qian, "An intelligent method for controlling the ECP braking system of a heavy-haul train," *Transportation Safety and Environment*, vol. 2, no. 2, pp. 133-147, June 2020.

[13] T. Szczegielniak, P. Jabłoński, D. Kusiak, and Z. Piątek, "Eddy currents induced in two parallel round conductors," *Applied Computational Electromagnetics Society Journal*, vol. 34, no. 12, pp. 1922-1930, Dec. 2019.

[14] L. Kapjin and K. Park, "Modeling eddy currents with boundary conditions by using Coulomb's law and the method of images," *IEEE Transactions on Magnetics*, vol. 38, no. 2, pp. 1333-1340, Aug. 2002.

[15] Y. Reza and M. Mirsalim, "Axial-flux wound-excitation eddy-current brakes: Analytical study and parametric modeling," *IEEE Transactions on Magnetics*, vol. 50, no. 6, pp. 1-10, Jan. 2014.

[16] Z. J. Liu, A. Vourdas, and K. J. Binns, "Magnetic field and eddy current losses in linear and rotating permanent magnet machines with a large number of poles," *IEE Proceedings A-Science, Measurement and Technology*, vol. 138, no. 6, pp. 289-294, Dec. 1991.

[17] M. Hecquet, P. Brochet, L. Jin, and P. Delsalle, "A linear eddy current braking system defined by finite element method," *IEEE Transactions on Magnetics*, vol. 35, no. 3, pp. 1841-1844, May 1999.

[18] G. Mehmet, E. Yolacan, and M. Aydin, "Design, analysis and real time dynamic torque control of single-rotor–single-stator axial flux eddy current

brake," *IET Electric Power Applications*, vol. 10, no. 9, pp. 869-876, Nov. 2016.
[19] A. Sohel, "A parametric model of an eddy current electric machine for automotive braking applications," *IEEE Transactions on Control Systems Technology*, vol. 12, no. 3, pp. 422-427, May 2004.
[20] J. M. Díaz-Chacón, C. Hernandez, and M. A. Arjona, "A comprehensive 2D FE-SIBC model for calculating the eddy current losses in a transformer tank-wall," *Applied Computational Electromagnetics Society Journal*, vol. 27, no. 8, pp. 646-653, Aug. 2012.
[21] H. J. Ryoo, J. S. Kim, D. H. Kang, G. H. Rim, Y. J. Kim, and C. Y. Won, "Design and analysis of an eddy current brake for a high-speed railway train with constant torque control," *Conference Record of the 2000 IEEE Industry Applications Conference*, Rome, Italy, vol. 1. pp. 277-281, Oct. 2000.
[22] S. Emmanuel and D. Georges, "Modeling and control of eddy current brake," *IFAC Proceedings Volumes*, vol. 28, no. 8, pp. 109-114, July 1995.

Qiao Ren graduated from Changan University in 2018 with a bachelor's degree. She is currently working toward the Ph.D. degree with the Institute of Railway Transit, Tongji University, China. Her current research interests include eddy current brake, high-speed magnetic levitation control.

Jimin Zhang received his M.Sc. and Ph.D. degrees from Southwest Jiaotong University, Chegdu, China, in 1999, and 2004, respectively. Since 2004, he has been with Tongji University, where he is currently a Professor. His research interests include magnetic levitation control, electromechanical coupling and active vehicle safety control, etc.

Hechao Zhou received his M.Sc. and Ph.D. degrees from Tongji University and Technische Universität Berlin, in 2010, and 2014, respectively. Since 2014, he has been with Tongji University, where he is currently an Assistant Professor. His research interests include high-speed magnetic levitation and vehicle system dynamics, etc.

Jinnan Luo received his B.S. degree in 2018 and M.Sc. degree in 2021 from Tongji University, Shanghai, China. His interest is high-speed magnetic field analysis and eddy current brake.

Modelling of Electromagnetic Fields for Shielding Purposes Applied in Instrumentation Systems

Ahmad M. Dagamseh [1*], Qasem M. Al-Zoubi [1], Qasem M. Qananwah [2], and Hamzeh M. Jaradat [3]

[1] Department of Electronics Engineering, Hijjawi Faculty for Engineering Technology
Yarmouk University, Irbid, P.O. Box 21163, Jordan
*a.m.k.dagamseh@yu.edu.jo, qzabi50@yu.edu.jo

[2] Department of Biomedical Systems and Informatics Engineering, Hijjawi Faculty for Engineering Technology
Yarmouk University, Irbid, P.O. Box 21163, Jordan
qasem.qananwah@yu.edu.jo

[3] Department of Telecommunications Engineering, Hijjawi Faculty for Engineering Technology
Yarmouk University, Irbid, P.O. Box 21163, Jordan
hamzehjaradat@yu.edu.jo

Abstract ─ In any sensory system, the Electromagnetic (EM) shielding of the channel-carrying signal is a fundamental technique to provide a noise-immune measurement system. Severe failures and uncertainty may occur if the external EM fields interfered with the measurements. Typically, the shielding is realized by enclosing the channel-carrying signal with thin-conductive hollow structures. However, with such structures, it is required to provide access to the interior components from the outside, for wires' connections, or better heat dissipation. This can be considered as a weakness in such the external magnetic fields can penetrate through the shielding structure. In this paper, the EM shielding effectiveness is considered for long hollow-cylinder structures with slots. The induced eddy current in thin-conductive shielding systems with slots together with the magnetic fields at different conditions are modeled. The objective is to determine the impact of the integrated slots along with the structure. The influence of the slots' sizes ($\boldsymbol{\alpha}$) and position relative to the excitation magnetic field (i.e., the declination angle ($\boldsymbol{\beta}$)) are investigated to evaluate the shielding effectiveness by means of the determination of the shielding factor. The results reveal the inherent relationship between the shield parameters and shielding effectiveness. The shielding effectiveness deteriorates by the slots' integration within the shielding surface. However, decreasing the size of the slots improves the shielding, significantly, towards the shielding effectiveness of the continuous cylindrical structure. Additionally, utilizing the symmetry in the structure positioning the slots in the direction perpendicular to the magnetic field flux improves the shielding effectiveness, drastically. Such a model can be considered to evaluate the degree of effectiveness or success of integrating opening slots within the shielding structure, which can be applied to different types of instrumentation systems specifically at the sensor-electronics interface.

Index Terms ─ Hollow cylinder, instrumentation system shielding, magnetic field, shielding effectiveness.

I. INTRODUCTION

Recently, electronic devices and technologies have been developing rapidly. The notable progress and systems' integrity, specifically in devices' structure, PCB designs, and packaging, led to an increase the complexity. The system requirements to overcome the error, reduce Electromagnetic Interference (EMI), and generate or transmit free-of-noise signals with proper timing have been considered intensively. Specifically, the EMI, as the major source of measurement errors in various disciplines, has been treated to make the electronic system immune. One of these disciplines is sensor design and instrumentation, where Electromagnetic Shielding (EMS) is frequently used to block or reduce either the emitted or intruded noise components. Metal sheets (e.g. aluminum, copper, etc…) formed in different structure designs can be used to fit the electronics enclosures [1,2].

EMS can be considered to design reliable systems, especially in critical applications. Examples of such systems include industry, military, medical, electric vehicle, sensors, and aerospace electronic devices, where the results of failure can extend from data loss to death [3]. However, there are no comprehensive structure designs to perform shielding. Therefore, in instrumentation

Submitted On: January 7, 2021
Accepted On: July 8, 2021
https://doi.org/10.47037/2021.ACES.J.360817

systems with EMS, the shield encloses the signal path in the entire instrumentation channel. Particularly, at the interfacing side between the sensor and the electronics, the role of EMS can be significant as the sensor's signal is still weak at this interface.

The integration of EMS in sensors has been studied intensively in the literature. There will be great advantages in combining the sensation techniques and EMS to perform proper measurements and improve the robustness and accuracy of the instrumentation system. Rienzo in his paper modeled the magnetic sensors to measure high AC and DC currents based on magnetic field sensing [4]. The application of EMS in instrumentation was used to reduce the measurement error using a ferromagnetic or conductive material. Shielding can be utilized to improve the measurements and the detection limit of the capacitive sensors. This can be attained by reducing the effects of both external noise and parasitic capacitances. With such a scheme, Dagamseh et al. modified the design of the capacitive-based artificial hair sensor using wafer-level and printed-circuit-board shielding [5]. The modified design improved the resolution of the measurement (i.e., performing localized measurement) and allowed measuring the capacitance changes originated from a single-hair sensor. Consequently, the detection limit of the new sensor design is improved down to 1 mm s–1 airflow amplitude with significant improvement in the directivity. Yang et al. presented a current sensor based on a giant magnetoresistance system with magnetic shielding [6,7]. The sensitivity and linearity of the sensors were considered by reducing the error due to the magnetic field. The results show that the sensor has a low nonlinearity error of less than 0.8% for the range of 10 mA to 20A for frequencies up to 200 kHz.

Screening by thin conductors with simple geometrical forms like planes, hollow cylinders, and hollow spheres has been investigated in the literature. Kaden investigated the screening effect of thin layers with angles and seats [8]. Lopez et al. provided a comparison between two conductive textiles with a wire mesh screen or with compact material to assess the shielding effectiveness [9]. The results showed good agreement between the modeling and measurement results with a 3 dB deviation at 1 GHz for the wire mesh model compared with a 2 dB deviation at 1.5 GHz for the compact material.

Another fundamental issue related to shielding is the shape of the shield. Azizi et al. have considered modeling the shielding effectiveness of aperture in a rectangular enclosure using circuit modeling and finite-difference time-domain method [10]. The results showed that the square and circular apertures are better to use than the rectangular shapes in electromagnetic shielding. Park et al. considered the integration of periodic metal strips within a conventional ferrite plate [11]. The period of the metal strip and the source position concerning the metal strips were considered. However, the analysis was provided for very thin strips. Fagnard et al. modeled the effect of introducing slits within a cylinder shield structure [12]. With his work, the analysis considered the magnetic properties of a hollow cylinder specifically with two axial slits that cut the cylinder in equal halves.

Several studies have analyzed and modeled the eddy currents in magnetic conductors and investigated the shielding effectiveness for different shielding structures [13-21]. H. El-Maghrabi investigated the shielding effectiveness and determined the electromagnetic shielding effectiveness of two cascaded wire-mesh sheets. The model results were compared with experiments and good agreement was obtained [13]. R. Araneo studied the effect of the shield parameters and position of the source on the effectiveness of the shield at low-frequency near0field magnetic sources [14]. Mayergoyz et al. analyzed the eddy currents for elliptical polarization of the magnetic field [17]. Delinger modeled the magnetic field inside a long cylindrical hole in a long cylindrical conductor [18]. The solution involved the summation of the magnetic fields of the current centered at the conductor and the opposite direction centered at the hole. Babic et al. modeled the magnetic field of a hollow cylinder with finite thickness in three dimensions with a longitudinal current component [19]. Sailing et al. have utilized the magnetic field of the direct current to identify the cracks in conductors [20]. A model to identify surface cracks with an elliptic shape was provided. Kvitkovic et al. investigated the shielding effectiveness within the inhomogeneous magnetic field [21].

For various instrumentation systems, the shielding at the sensor-electronics interface is a critical node; as the signal is weak. Electromagnetic shielding can be performed to protect the measured signal and thereby the accuracy of the measurement system. In this paper, the magnetic field has been modeled using a cylindrical geometrical structure with slots, taking into consideration the induced eddy currents. The target is to evaluate the impact of integrating slots within the shield in terms of shielding effectiveness utilizing studying the characteristics of the shield. These slots can be, thereby, controlled for the optimal performance of the entire instrumentation system to provide an access to the sensor side or to the interfacing electronics.

II. SYSTEM MODELING

The magnetic shielding depends on the material properties, the shield geometry, and the amplitude of the magnetic field. In this work, we investigate the effect of shielding structures on the field shape and the characteristics of the shield. A Long-hollow cylinder structure with integrated lateral slots for different designs is considered. The models will be based on determining the shape of the eddy currents in the

shielding structure and the magnetic flux through the slot to investigate the magnetic flux penetration. It is assumed that the skin-depth (δ) is much more than the thickness of the conductive layer with $\delta = \frac{1}{\sqrt{\pi \mu f \sigma}}$ with μ, f and σ are the magnetic permeability ($\frac{N}{A^2}$), frequency (Hz), and conductivity ($\frac{S}{m}$), respectively [22].

The solution for the eddy currents with two axial slots, taking into consideration the symmetry of the structure, consists of the following procedure and assumptions:

- The general solution for the homogeneous problem to determine the vector potential A (inside and outside the shielding surface) can be obtained using the solution of the Laplace equation. The excited magnetic alternative field B can then be obtained using this solution.
- The coefficients of the solution can be defined using the boundary conditions. At the boundaries of both regions (i.e., above and below the shielding structure) the following conditions can be applied:

$$A_1 = A_{2.}, \tag{1}$$

and

$$(H_{t1} - H_{t2}) = K, \tag{2}$$

where H_{t1} and H_{t2} are the magnetic field intensity outside and inside the shielding surface, respectively.

- The surface current density K can be determined using the law of induction along the circumference of the cylinder with thickness d according to:

$$K = \begin{cases} -j\omega x d A_{1,} \\ \text{for: } -\pi + \beta + \frac{\alpha}{2} \le \varphi \le \beta - \frac{\alpha}{2} \\ \text{and } \beta + \frac{\alpha}{2} \le \varphi \le \pi + \beta - \frac{\alpha}{2} \\ 0, \quad \text{else} \end{cases}. \tag{3}$$

- According to equ. (3), the surface current can be expanded in terms of the solutions for the coefficients of vector potentials A_1 and A_2 and substitute it in equ. (2). This results in a linear system of equations that can be used to find the coefficients of the Laplace solution.
- The vector potential is considered as the sum of the exciting vector potential A_o (outside the shielding surface) and the vector potential of the generated eddy current A_i (inside the shielding surface) with:

$$A = A_i + A_o. \tag{4}$$

A long hollow-cylinder with lateral slots the homogeneous alternating field H_o is considered. The magnetic field is oriented perpendicularly to the cylinder. To simplify the solution through the distribution of currents, the thickness of the cylinder wall (d) is assumed to be small. The width of the slot is defined by the angle (α). The position of the slot, in reference to the direction of the field H_o, is described by the angle (β). Figure 1 shows the design parameters of the hollow-cylinder shielding structure.

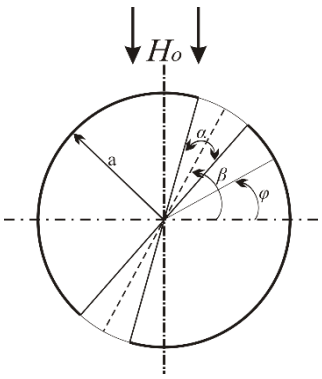


Fig. 1. A schematic representation for the hollow-cylinder shielding structure with lateral slots and the definitions of the angles (α, β, and φ).

The solution of the Laplace equation for the vector potential A in polar coordinates can be represented as shown in equ. (5):

$$\left.\begin{aligned} A_o(\rho,\varphi) &= \mu a H_0 \left[\frac{\rho}{a} cos\,\varphi + \sum_{n=1}^{\infty} \left(\frac{a}{\rho}\right)^n (C_n \cos n\varphi + D_n \sin n\varphi)\right] \\ &\text{for } \rho \ge a \\ A_i(\rho,\varphi) &= \mu a H_0 \left[\frac{\rho}{a} cos\,\varphi + \sum_{n=1}^{\infty} \left(\frac{\rho}{a}\right)^n (C_n \cos n\varphi + D_n \sin n\varphi)\right] \\ &\text{for } \rho \le a \end{aligned}\right\}, \tag{5}$$

where a is the cylinder radius, C_n and D_n are the solution coefficients.

Due to the symmetry in the design of the system, the vector potential analysis can be performed with the odd numbers n (as $A(\rho,\varphi) = -A(\rho,\varphi+\pi)$. To simplify the solution this property is used at the end of the analysis. The boundary condition in equ. (1) is satisfied by this form of solution and the surface current density becomes at $\rho = a$:

$$K(\varphi) = 2H_0 \sum_{n=1}^{\infty} n(C_n \cos n\varphi + D_n \sin n\varphi). \tag{6}$$

Because of the slots, the structure is not symmetrical along a single axis. Accordingly, equ. (3) can be represented as shown in equ. (7):

$$\left.\begin{aligned} K &= -j\omega\mu x da H_0 \times \left[\cos\varphi + \sum_{n=1}^{\infty}(C_n \cos n\varphi + D_n \sin n\varphi)\right], \\ &\qquad \text{for: } -\pi+\beta+\frac{\alpha}{2} \le \varphi \le \beta-\frac{\alpha}{2} \quad \text{and} \quad \beta+\frac{\alpha}{2} \le \varphi \le \pi+\beta-\frac{\alpha}{2} \\ K &= 0, \qquad \text{for: } \beta-\frac{\alpha}{2} < \varphi < \beta+\frac{\alpha}{2} \quad \text{and} \quad \pi+\beta-\frac{\alpha}{2} < \varphi < \pi+\beta+\frac{\alpha}{2} \end{aligned}\right\}, \tag{7}$$

with constants coefficients C_n and D_n that should be determined.

Utilizing these boundary conditions at the surface in equ. (7), the surface current density can be expanded in Fourier series form as shown in equ. (8):

$$K(\varphi) = \sum_{m=1}^{\infty}(A_m \cos m\varphi + B_m \sin m\varphi). \tag{8}$$

where:

$$\left.\begin{aligned} A_m &= c_m(\pi-\alpha) - \sum_{n=1}^{\infty}(c_n t_{mn} + d_n q_{mn}) - \sum_{\substack{n=1\\ n\neq m}}^{\infty}(c_n s_{mn} + d_n p_{mn}) \\ B_m &= d_m(\pi-\alpha) - \sum_{n=1}^{\infty}(c_n q_{mn} - d_n t_{mn}) + \sum_{\substack{n=1\\ n\neq m}}^{\infty}(c_n p_{mn} - d_n s_{mn}) \end{aligned}\right\}, \tag{9}$$

in such

$$\left.\begin{aligned} d_n &= -j\omega\mu x da H_0 . D_n \\ c_n &= -j\omega\mu x da H_0 . C_n \end{aligned}\right\}, \quad \text{for } n \neq 1. \tag{10a}$$

$$c_1 = -j\omega\mu x da H_0(1 + C_1), \quad \text{for } n = 1. \tag{10b}$$

and:

$$\left.\begin{aligned} p_{mn} &= \frac{1}{n-m}.\sin[(n-m)\beta].\sin[(n-m)\frac{\alpha}{2}] \\ q_{mn} &= \frac{1}{n+m}.\sin[(n+m)\beta].\sin[(n+m)\frac{\alpha}{2}] \\ s_{mn} &= \frac{1}{n-m}.\cos[(n-m)\beta].\sin[(n-m)\frac{\alpha}{2}] \\ t_{mn} &= \frac{1}{n+m}.\cos[(n+m)\beta].\sin[(n+m)\frac{\alpha}{2}] \end{aligned}\right\}. \tag{10c}$$

The left sides of equ. (6) and equ. (8) are identical to the right sides. This results in two linear-system of equations to determine the constants C_n and D_n for the expression of the vector- potentials, represented in equ. (5). The coefficient matrix that defines C_n and D_n can be represented as:

$$\begin{bmatrix} e & f \\ u & v \end{bmatrix}\begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} g \\ w \end{bmatrix}. \tag{11}$$

Accordingly, the elements of the coefficient matrix are defined as follows:

$$\left.\begin{aligned} e_{mn} &= s_{mn} + t_{mn} \quad , \quad f_{mn} = p_{mn} + q_{mn} \\ u_{mn} &= q_{mn} - p_{mn} \quad , \quad v_{mn} = s_{mn} - t_{mn} \end{aligned}\right\} \text{ for } m \neq n, \tag{12a}$$

and

$$\left.\begin{aligned} e_{mn} &= j\frac{\pi\delta^2 m}{ad} - (\pi - a) + t_{mn,} \\ f_{mn} &= q_{mn} \\ v_{mn} &= j\frac{\pi\delta^2 m}{ad} - (\pi - a) - t_{mn} \\ u_{mn} &= q_{mn} \end{aligned}\right\} \text{for } m = n, \tag{12b}$$

with:

$$\begin{pmatrix} g_m = (\pi - a) - t_{11} & w_m = -q_{11} & \text{for } m = 1 \\ g_m = -t_{m1}\text{-s}_{m1} & w_m = p_{m1} - q_{m1} & \text{for } m \neq 1 \end{pmatrix}. \tag{12c}$$

The system of equations was solved numerically for a limited number of equations and coefficients. The shielding factor (*S*) was used to evaluate the shielding effectiveness of the structure (i.e., shielding factor), which represents the ratio of the magnetic fields at the middle of the cylinder B_{max} to the magnetic field at the same point for the continuous cylinder without slots B_u in such:

$$S = \frac{B_{max}}{B_u}, \tag{13a}$$

with:

$$B_u = \frac{\mu H_0}{\sqrt{1 + \left(\frac{ad}{\delta^2}\right)^2}}. \tag{13b}$$

In general, the induction at the middle of the cylinder with the slots is not proportional to the excitation field and it can be represented by:

$$B\left(0, \frac{\pi}{2}\right) = -\mu H_0[(1 + C_1)\hat{\boldsymbol{e}}_\rho + D_1\hat{\boldsymbol{e}}_\varphi]. \tag{14}$$

Subsequently,

$$B\left(0, \frac{\pi}{2}\right) = \mathrm{B}_\rho\hat{\boldsymbol{e}}_\rho + B_\varphi\hat{\boldsymbol{e}}_\varphi. \tag{15}$$

III. RESULTS AND DISCUSSIONS

The shielding effectiveness while utilizing a shield with a geometry of hollow-cylinder with slots is expressed through the shielding factor (S). Once *S* approaches unity, the shield performance of the hollow cylinder with slots can be approximated to the continuous cylinder. The effect of various design parameters has been investigated to analyze the magnetic flux penetration through the slot such as; the position of the slot relative

to the magnetic field lines (i.e. field orientation), slot size, and excitation frequency. Figure 2 and Fig. 3 show examples for a plot of the magnetic field flux applied to the hollow-cylindrical structure in the presence of slots at different slots' conditions.

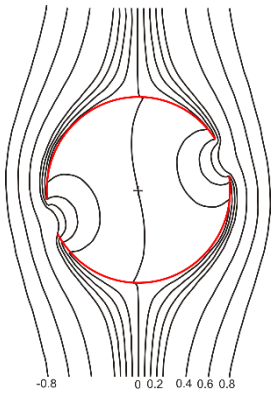


Fig. 2. Modeled magnetic field flux lines for the shielding structure with ($\frac{a.d}{\delta^2} = 5$).

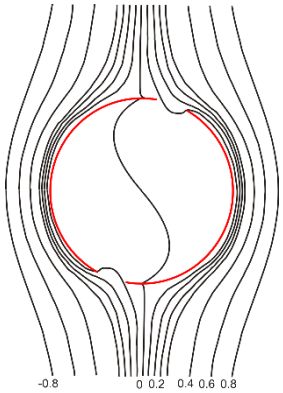


Fig. 3. Modeled magnetic field flux lines for the shielding structure with ($\frac{a.d}{\delta^2} = 10$).

As expected and away from the slot positions, the magnetic field profile exhibits a shape similar to the bulk cylinder [23]. With the presence of the slots, this profile is altered and the behavior becomes dependent on the slots' positions relative to the direction of the magnetic field flux. The effect of the geometry can be revealed by altering the shape and the behavior of the magnetic field lines by means of the presence of the induced currents.

The magnetic field lines, which are the lines of the constant vector potential in reference to $\mu a H_o$, have shown an interval of 0.8 away from the shield surface while at the interface between the conductive side and the slot have 0.2 intervals. These intervals decrease at the proximity of the slot position to a quarter of this interval value. This is due to the effect of the induced currents which provide an additional field source superimposed to the main magnetic field. This modifies the effect of the main magnetic field lines at the surface of the cylinder and causes compression of these lines. Therefore, the magnetic field penetration in the vicinity and beyond the surface of the cylinder vanishes, significantly. It can be observed from Fig. 2 that the field flux lines are coupled through the slots considerably into the interior region of the cylinder. The induced magnetic field vector can be decomposed into two components. One component is in the vertical direction and the other is in the parallel direction. The vertical field component inside the cylinder is always directed against the excitation field. Whereas the region outside the cylinder, the field adds up to the excitation field. Both slot size and its direction have a great impact on the magnetic field strength that couples through the slot. As the slot position gets closer to the horizontal line, the strength of the induced field's parallel component is greater than the vertical component. Consequently, the combined fields in the vicinity of the slot opening have a dominant parallel component compared with the vertical component, which is responsible for the fields' lines to connect through the openings. In Fig. 3, the slot position is closer to the vertical line. The vertical component of the induced field is more dominant and hence, less field couples through the slot and thereby more shielding effectiveness.

Figure 4 reveals the effect of the slot size and orientation relative to the magnetic field flux on the shielding factor of long-hollow cylinder structure with slots. The slot opening is represented by the angle α and the relative orientation of the slot to the magnetic field flux by the declination angle β. The effect of the slots' size and position relative to the magnetic field flux can be observed, as the angles α and β are modified. With the presence of the slots, the generated eddy currents got disturbed and break into two components at both ends of each slot. When increasing the slot size, the shielding factor deteriorates and the field penetration increases. This is attributed to the degradation of the counter field represented by the vertical field component generated by the shield surface at the slot position. Figure 5 shows the effect of increasing the slot size on the shielding factor of the shield at different orientations of the magnetic field lines. It is noted that for slots positioned at $\beta = 0^o$, increasing the slots' size would reduce the shielding effectiveness. This is accredited to the reduction of the induced vertical field component that is responsible to impede the excitation field.

For the angle β = 0 (i.e., when the excitation magnetic field is in parallel with the slot axis), small slot size has a negligible effect on the magnetic field shape. Under this condition, the shielding factor approaches unity and the field shape will approach the behavior of a continuous cylinder with no slots as can be observed in Fig. 4.

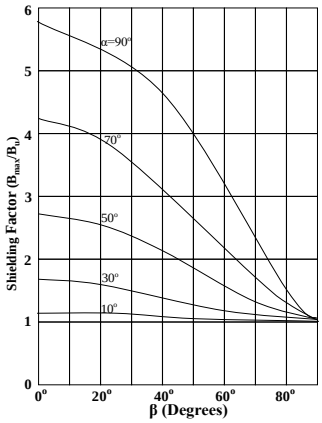


Fig. 4. The shielding factor (***S***) of the cylindrical shielding structure at different slots sizes and slots positions with ($\frac{a.d}{\delta^2} = 10$).

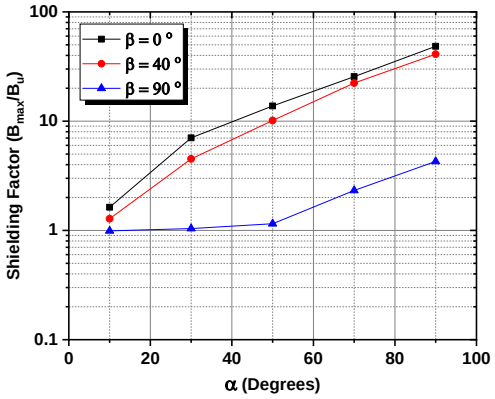


Fig. 5. The shielding factor (***S***) of a long-hollow cylinder shielding structure for different slot sizes with ($\frac{a.d}{\delta^2} = 10$).

When the slot position is modified (i.e., varying β > 0), the flux lines of the magnetic field at the center of the cylinder have shown a deflection compared with the lines without the shield. This can be due to the presence of the penetrated magnetic field through the slots. This tends to interrupt the symmetry of the structure relative to the direction of the field and thereby modifies the direction of the field lines. Similar results have been obtained by Fagnard et al. [12]. They found that the field direction at the center of the cylinder is disturbed compared with the excitation field.

Figures 2 and Fig. 3 show the shape of the magnetic field lines at the center of the cylinder while varying the position of the slots (i.e., varying β). However, a minimum variation occurs when $\beta = 90^o$, as this angle allows a direct penetration of the magnetic field lines towards the center of the cylinder. Figure 6 shows the magnetic field flux lines with $\beta = 90^o$.

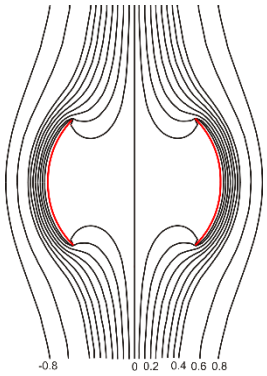


Fig. 6. Modeled magnetic field flux lines for the shielding structure at $\beta = 90^o$ with ($\frac{a.d}{\delta^2} = 10$).

The results indicate that when increasing the slot opening, the shielding factor deteriorates. However, the effect of slots' sizes is less prominent when increasing the angle β to approach the 90^o. The shielding factor for the cylinder with slots at fixed slots' sizes (i.e., fixed α) is optimal at angle $\beta = 90^o$. Figure 5 shows the behavior of the magnetic field when the pair of shielding slots is positioned in the direction of the magnetic field. At these conditions, the pair of slots is symmetrical and positioned in the direction of the excitation field. Therefore, the excitation field is directly aligned through the slots. This facilitates the direct penetration of the magnetic field through the slots.

However, the results show that regardless of the opening size of the slot, the shielding effectiveness improves in such the flux of the magnetic field at the middle of the cylinder B_{max} approaches to the magnetic field flux for the continuous cylinder B_u. Due to the highly symmetrical current distributions around the vertical line, the induced surface current densities at both conductors are equal in magnitude but opposite in direction, such that the induced magnetic field lines oppose the excitation field lines. As a result, magnetic field reduction occurs in the region inside the cylinder. Figure 6 shows the magnetic field flux lines when the slots' positions relative to the magnetic field flux is at $\beta = 90^o$ for the cylindrical shielding structure. The slot position has a significant impact on the field's vertical component strength. It has the maximum strength when $\beta = 90^o$ and it decays as β goes below 90^o. Therefore, maximum field's cancelation occurs in the inner region, while the field is strengthening in the outer region, where the field lines wrap around the conductor's exterior surface.

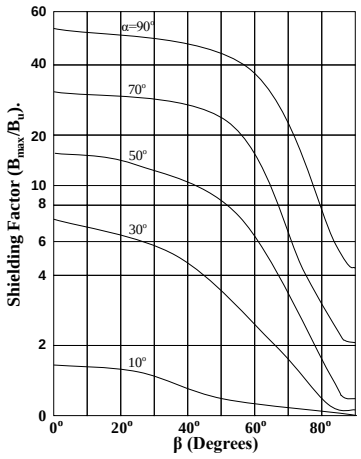


Fig. 7. The shielding factor (***S***) of the cylindrical structure at different slots sizes and positions with ($\frac{a.d}{\delta^2} = 100$).

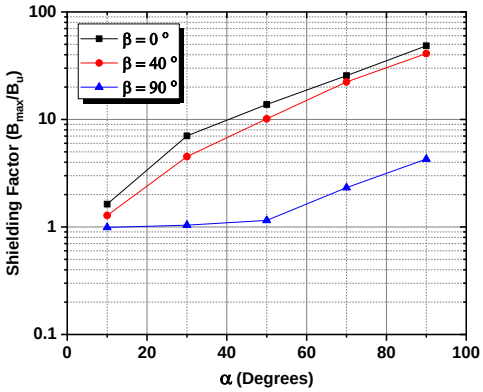


Fig. 8. The shielding factor (***S***) of a long-hollow cylinder shielding structure for different slot sizes with ($\frac{a.d}{\delta^2} = 100$).

It has been observed that when varying the excitation frequency, the behavior of the long-hollow cylindrical shielding structure with slots is the same. It is found that the shielding factor increases drastically with the excitation frequency and thereby less shielding effectiveness regardless of the slot size and positions. Figure 4 and Fig. 7 represent the effect of varying the excitation frequency (represented by the inverse proportionality relation between the excitation frequency and the skin depth while maintaining the rest of the design parameters fixed) on the shielding factor of the structure at different slots' sizes and positions. For $\alpha = 30^o$, $\beta = 40^o$ the shielding factor was 1.37 at $a.d/\delta^2 = 10$ (which indicates a low-frequency range). While at the same conditions, the shielding factor increases to about 4.51 at $a.d/\delta^2 = 100$. This indicates that at low-frequency ranges, in particular, the efficiency of the shield with slots is higher compared with the high-frequency ranges. Additionally, the effect of increasing the excitation frequency can be observed at slots' position of $\beta = 90^o$ (see Fig. 8 compared with Fig. 5).

VI. CONCLUSIONS

In this paper, modeling of the magnetic field for a long hollow cylinder with slots has been presented. The aim is to evaluate the impact of slots integration while considering the shielding effectiveness of the shielding structure and compare it with a continuous-solid conductive system. The induced eddy currents in a thin conductive system with slots have been calculated. The influence of the slot size, relative position to the excitation field, and frequency of excitation were investigated. The shielding factor was determined numerically for specific points inside the shielded space. The results show that for small slots' sizes the shielding effectiveness is comparable with the continuous cylinder structure. Additionally, positioning the slots relative to the direction of the excitation magnetic field (i.e., increasing β improves the shielding effectiveness; benefiting from the symmetry in the structure. Such a study expands on the concepts of electromagnetic shielding topology by including some considerations related to the reduction of interference utilizing the structure relative to the magnetic fields; targeting instrumentation system applications.

REFERENCES

[1] S. C. Tang, S. Y. R. Hui, and H. S. H. Chung, "Evaluation of the shielding effects on printed-circuit-board transformers using ferrite plates and copper sheets," *IEEE Trans. Power Electron.*, vol. 17, no. 6, pp. 1080-1088, 2002.

[2] F. Wen and X. Huang, "Optimal magnetic field shielding method by metallic sheets in wireless power transfer system," *Energies*, vol. 9, no. 9, pp. 733-748, 2016.

[3] Y. Kitano, H. Omori, T. Morizane, N. Kimura, and M. Nakaoka, "A new shielding method for magnetic fields of a wireless EV charger with regard to human exposure by eddy current and magnetic path," in *Proceedings - 2014 International Power Electronics and Application Conference and Exposition, IEEE PEAC 2014*, pp. 778-781,2014.

[4] L. Di Rienzo, "Modelling of magnetic current sensors for measuring high AC and DC currents," *Meas. Control*, vol. 34, no. 9, pp. 272-275, 2001.

[5] A. M. K. Dagamseh, C. M. Bruinink, R. J. Wiegerink, T. S. J. Lammerink, H. Droogendijk, and G. J. M. Krijnen, "Interfacing of differential-capacitive biomimetic hair flow-sensors for optimal sensitivity," *J. Micromechanics Microengineering*, vol. 23, no. 3, pp. 035010-035026, 2013.

[6] X. Yang, C. Xie, Y. Wang, Y. Wang, W. Yang, and G. Dong, "Optimization design of a giant magneto resistive effect based current sensor with a magnetic shielding," *IEEE Trans. Appl. Supercond.*, vol. 24,

no. 3, 2014.
[7] X. Yang, H. Liu, Y. Wang, Y. Wang, G. Dong, and Z. Zhao, "A giant magneto resistive (GMR) effect based current sensor with a toroidal magnetic core as flux concentrator and closed-loop configuration," *IEEE Trans. Appl. Supercond.*, vol. 24, no. 3, pp. 1-5, 2014.
[8] H. Kaden, *Wirbelströme und Schirmung in der Nachrichtentechnik*. Berline: Springer, 1959.
[9] A. Lopez, L. Vojtech, and M. Neruda, "Comparison among models to estimate the shielding effectiveness applied to conductive textiles," *Adv. Electr. Electron. Eng.*, vol. 11, no. 5, pp. 387-391, 2013.
[10] H. Azizi, F. Tahar Belkacem, D. Moussaoui, H. Moulai, A. Bendaoud, and M. Bensetti, "Electromagnetic interference from shielding effectiveness of a rectangular enclosure with apertures - Circuital approach, FDTD and FIT modelling," *J. Electromagn. Waves Appl.*, vol. 28, no. 4, pp. 494-514, 2014.
[11] H. H. Park, J. H. Kwon, S. Il Kwak, and S. Ahn, "Magnetic shielding analysis of a ferrite plate with a periodic metal strip," *IEEE Trans. Magn.*, vol. 51, no. 8, pp. 1-8, 2015.
[12] J.-F. Fagnard, S. Elschner, A. Hobl, J. Bock, B. Vanderheyden, and P. Vanderbemden, "Magnetic shielding properties of a superconducting hollow cylinder containing slits: Modelling and experiment," *Supercond. Sci. Technol.*, vol. 25, no. 10, p. 104006, 2012.
[13] H. M. El-Maghrabi, "Electromagnetic shielding effectiveness calculation for cascaded wire-mesh screens with glass substrate," *Applied Computational Electromagnetics Society Journal*, vol. 33, no. 6, pp. 641-647, 2018.
[14] R. Araneo, G. Lovat, S. Celozzi, and P. Burghignoli, "Shielding effectiveness of finite width shields against low-impedance magnetic near-field sources," *2018 Int. Appl. Comput. Electromagn. Soc. Symp. Denver, ACES-Denver 2018*, 2018.
[15] T. Cvetković, V. Milutinović, N. Dončov, and B. Milovanović, "Numerical investigation of monitoring antenna influence on shielding effectiveness characterization," *Applied Computational Electromagnetics Society Journal*, vol. 29, no. 11, pp. 837-846, 2014.
[16] A. M. Hussein and P. P. Biringer, "Redistribution of current in flat conductors using magnetic bars," *J. Appl. Phys.*, vol. 55, no. 4, pp. 1188-1194, 1984.
[17] I. Mayergoyz, C. Serpico, and P. McAvoy, "Analysis of eddy currents in magnetically nonlinear conductors," *J. Appl. Phys.*, vol. 109, no. 7, pp. 2011-2014, 2011.
[18] W. G. Delinger, "Magnetic field inside a hole in a conductor," *Phys. Teach.*, vol. 28, no. 4, pp. 234-235, 1990.
[19] S. Babic, Z. Andjelic, B. Krstajic, and S. Salon, "Analytical magnetostatic field calculation for a conductor with uniform current in the longitudinal direction," *J. Appl. Phys.*, vol. 67, no. 9, pp. 5827-5829, 1990.
[20] S. He and V. G. Romanov, "Explicit formulas for crack identification in conductors using boundary measurements of direct current fields," *J. Appl. Phys.*, vol. 85, no. 9, pp. 6822-6827, 1999.
[21] J. Kvitkovic, K. Burnside, M. Zhang, and S. Pamidi, "Magnetic shielding of long paraboloid structures in the inhomogeneous magnetic field," *J. Phys. Conf. Ser.*, vol. 1559, no. 1, pp. 1-9, 2020.
[22] Y. He, B. Gao, A. Sophian, and R. Yang, *Transient Electromagnetic-Thermal Nondestructive Testing: Pulsed Eddy Current and Transient Eddy Current Thermography*. Butterworth-Heinemann, 2017.
[23] W. Carr Jr, *AC Loss and Macroscopic Theory of Superconductors*. CRC Press, 2001.

Ahmad M. Dagamseh received his Ph.D. degree from the University of Twente in the Netherlands in 2011 in MEMS. His research interests include Sensors, instrumentation systems and modeling.

Qasem M. Al-Zobi received his Ph.D. degree from the Technische Universitaet Berlin/Germany in 1990. His research interests industrial electronics and external magnetic field screening.

Qasem M. Qananwah received the Ph.D. degree in Biomedical Engineering from Karlsruhe Institute of Technology, Karlsruhe, Germany. His research interest focuses in instrumentation system, design and modeling.

Hamzeh M. Jaradat received the Ph.D. degree in Electrical and Computer Engineering from the University of Massachusetts Lowell (UML), USA. His current research includes electromagnetics modeling.

Electromagnetic Acoustic Transducer for Detection and Characterization of Hidden Cracks inside Stainless Steel Material

H. Boughedda [1*], T. Hacib [1], Y. Le Bihan [2], M. Chelabi [1], and H. Acikgoz [3]

[1] L2EI Laboratory, Jijel University, BP 98 Ouled Aissa, Jijel 18000, Algeria
*houssem.boughedda@univ-jijel.dz

[2] Group of Electrical Engineering — Paris (GeePs)
CNRS - Univ. Paris-Saclay - CentraleSupelec - Sorbonne Univ, 11, rue Joliot Curie
91192 Gif sur Yvette Cedex, France

[3] Material Science Engineering Department, KTO Karatay University, Konya, Turkey

Abstract ─ Industrial structure are exposed to microstructural changes caused by fatigue cracking, corrosion and thermal aging. Generally, a hidden crack is very dangerous because it is difficult to detect by Non-Destructive Evaluation (NDE) techniques. This paper presents a new approach to estimate the hidden cracks dimensions inside a stainless steel plate based on the EMAT signal. The received signal by EMAT is simulated using the Finite Element Method (FEM). Then, the identification of the hidden crack sizes is performed via the combination of two techniques; the first one is the Time-of-Flight (ToF) technique which was applied to estimate the crack height by the evaluation of the difference between the ToF of the healthy form and the defective form. Then, the crack width is estimated by the solution of the inverse problem from the received signal based on a meta-heuristic algorithm called Teaching learning Based optimization (TLBO). The obtained results illustrate the sensitivity of the EMAT sensor to the variation of the crack sizes. Moreover, the quantitative evaluation of the cracks dimensions, show clearly the efficiency and reliability of the adopted approache.

Index Terms ─ Characterization of hidden cracks, FEM, NDE, Time-of-Flight, TLBO algorithm.

I. INTRODUCTION

The requirement of structural integrity has become an indispensable process to ensure the reliability of the system by detecting the apparition of damages. The development of non-destructive testing (NDT) techniques during the last century, provides an effective way to monitor the health of structures [1]. Electromagnetic Acoustic Transducer (EMAT) is an ultrasonic NDT technique that has offered a non-contact inspection of conductive materials, compared to traditional piezoelectric transducers [2]. One of the most important advantages of EMAT is its ability to generate several types of waves by simply changing its structure or excitation frequency, such as Rayleigh wave, longitudinal waves and shear waves. The advantage of the ultrasonic shear wave compared to other waves is its possibility to travel deep inside the materials with low energy loss, which makes it suitable for inspecting deep structures with high efficiency [3,4].

EMAT launches the magnetic forces into the conductive sample under test via three transduction mechanisms: Lorentz forces, magnetization forces and magnetostriction. However, the application of magnetizing and magnetostrictive forces is confined only in a ferromagnetic materials. Concerning the non-ferromagnetic materials, Lorentz forces have the major contribution to generate ultrasonic waves [5-7]. In this research, the transduction efficiency of a shear wave EMAT that is applied on stainless steel will be investigated, which means that only the Lorentz force transduction mechanism needs to be considered.

The examination of cracks using EMAT sensor has been widely investigated by a number of researchers. In [8], the first numerical model has been developed using the FEM, which was capable of modeling the transmitter and the receiver EMAT system. In [9], a new EMAT configuration has been proposed based on a periodic permanent-magnet (PPM) to generate and receive shear horizontal guided waves propagating in the circumferential direction. In [10], the estimation of the surface crack depth in pitch-catch mode by EMAT sensor has been proposed. It was carried out experimentally by determining the reflection and the transmission coefficients of Rayleigh waves scattered at a surface crack. However, this approach is limited only for the estimation of the cracks depth. In [11], the detection and localization of vertical cracks was

Submitted On: December 31, 2020
Accepted On: July 6, 2021

https://doi.org/10.47037/2021.ACES.J.360818

employed based on the diffraction phenomena of compressional waves by measuring the Time of Flight (ToF) of the diffracted waves. However, this approach has shown a low efficiency for assessing the crack width. In [12], a new approach for estimating the dimensions of surface cracks has been proposed. this approach is based on the combination of EMAT probe and partial least squares regression (PLSR) algorithm. This approach shows acceptable measurement accuracy in the sizing of the surface crack.

In recent years, the majority of EMAT researches have focused on cracks detection and to solve the problem of low conversion efficiency [13-15]. Whilst, the quantitative description of the defects didn't receive much concern. The main contribution of the current work is the detection and the characterization of hidden cracks inside a stainless steel plate. The cracks' height has been estimated using ToF technique in pulse-echo mode by measuring the time of flight of the transmitted and the reflected shear waves. This technique is based on the difference between the ToF of the healthy and the defective state of the inspected materials. Whereas, the cracks width has been estimated using a new stochastic optimization algorithm, that is called Teaching-Learning Based Optimization (TLBO) algorithm, which is used to solve the inverse problem from the EMAT signal. This technique was applied due to the linearity relation between the change in crack width and the received signal by EMAT. Moreover, TLBO algorithm has proved its efficiency over other stochastic algorithms [16,17].

The remainder of this paper is arranged as follows: first, developing a two dimensional (2D) numerical model of EMAT based on the FEM. This model includes an evaluation of the mechanical displacement and the output signal by EMAT. Then, the sensitivity of EMAT sensor to different hidden crack sizes has been checked. Finally, the identification of the cracks depth using ToF technique, and the characterization of the cracks width by combining TLBO algorithm and EMAT received signal.

II. EMAT SHEAR WAVE CONFIGURATION

Electromagnetic Acoustic Transducer (EMAT) configurations vary according to the type of waves that we want to produce during the test [9]. Usually, EMAT configurations consists of two main components, a magnet and a coil, which are placed above the material under test. EMAT probe can be used as a pulse-echo mode or on pitch-catch mode. In this work, the pulse-echo mode has been preferred to check the bottom of the stainless steel plate, i.e., the same transducer is used to generate and receive the shear and the longitudinal waves. The EMAT structure that is used in this work is shown in Fig. 1. It consists of a two adjacent permanent magnets above a spiral coil; each magnet produces biasing field normal to the surface but in opposite direction (with flux density of 1 T). The coil and the permanent magnet have the same central axis. The test specimen is a non-magnetic material ($\mu=\mu_0$) represented as a stainless steel plate with electrical conductivity $\sigma=3.6\times10^6$ [S/m]. The spiral coil is made by copper with six turns, it is fed by an alternative current $i(t)$ with frequency f=2 MHz in order to create a dynamic magnetic field.

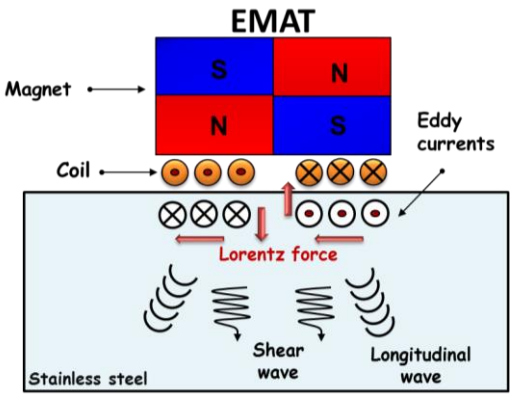


Fig. 1. Schematic representation the configuration and working principle of EMAT.

The eddy currents will be inducing at the surface of this the material due to this dynamic magnetic field [11]. The EMAT transduction mechanism in non-ferromagnetic material is the Lorentz force, which is generated by the interaction between the static magnetic field (from the permanent magnet) and the induced eddy currents. The Lorentz force launches elastodynamic waves (mechanical vibrations), which propagate inside the material under test [18]. In the receiving mode, the static magnetic field, of the receiver EMAT, interacts with the velocity of the mechanical displacement at the surface of the plate to generate spatial eddy currents underneath. These currents induce a voltage in the coil of the EMAT [12]:

$$\boldsymbol{i}(t) = \beta\, \mathrm{e}^{-\alpha.(t-\tau)^2}.\cos\left(2\pi f(t-\tau)\right), \qquad (1)$$

where, β is the current amplitude 5 (A), α is the bandwidth factor 4×10^{-12} (s^{-2}), τ is the arrival time 1(μs).

III. FINITE ELEMENT SIMULATION

A two-dimensional (2D) finite element simulation has been performed using the COMSOL Multiphysics software, to simulate the generation and the reception of the ultrasonic waves by EMAT, based on the combination of electromagnetic and mechanical models. The simulation includes an evaluation of the eddy current, the magnetic flux density, the Lorentz force, the mechanical displacement inside the stainless steel plate and the output voltage by the EMAT.

A. Governing equations

The Maxwell's equations are used to calculate the static magnetic field and the distribution of induced eddy currents in the aluminum plate [4]. After some manipulations, we get the following Partial Deferential Equation (PDE):

$$\nabla\times(\frac{1}{\mu}(\nabla\times\boldsymbol{A_z}-\boldsymbol{B_r}))+\sigma\frac{\partial\boldsymbol{A_z}}{\partial\boldsymbol{t}}-\sigma\boldsymbol{v}\times(\nabla\times\boldsymbol{A_z})+\sigma\nabla\boldsymbol{V}=\boldsymbol{J_{ex}}, \quad (2)$$

where σ, μ, v, $\boldsymbol{J_{ex}}$ and $\boldsymbol{B_r}$ are conductivity, permeability, velocity, external current density of the coil and magnetic flux density respectively.

The solving of (2) in static and transient analysis allows calculating the MVP $\boldsymbol{A_z}$, which used to calculate the static magnetic field $\boldsymbol{B_s}$ and the induced current $\boldsymbol{J_e}$ in the conductive material:

$$\boldsymbol{B_s}=\nabla\times\boldsymbol{A_z}, \quad (3)$$

$$\boldsymbol{Je}=-\boldsymbol{\sigma}\frac{\partial\boldsymbol{A_z}}{\partial\boldsymbol{t}}. \quad (4)$$

By considering that the Lorentz force as the only contribution body force and the material as satisfying the continuous elastic isotropic, we have Navier's equation [18]:

$$-\boldsymbol{\mu}\nabla\times\nabla\times\mathbf{u}-(\boldsymbol{\lambda}+\boldsymbol{\mu})\nabla(\nabla.\mathbf{u})+\boldsymbol{\rho}\frac{\partial^2\mathbf{u}}{\partial\mathrm{t}}=\mathbf{F}_L. \quad (5)$$

where λ and μ are the Lame constants.

The ultrasonic wave produced in the stainless steel plate, with length of 200 mm and thickness of 20 mm. The plate has the following properties: Young's modulus of 210 GPa, Poisson's ratio of 0.28, mass density of 7850 kg/m^3.

B. Lorentz force evaluation

The Lorentz force $\boldsymbol{F_L}$ results by the interaction of the current density $\boldsymbol{J_e}$ and the static magnetic flux density $\boldsymbol{B_S}$ that were calculated in transient and static analysis respectively, according to the following equation [18]:

$$\boldsymbol{F_L}=\boldsymbol{J_e}\times\boldsymbol{B_s}. \quad (6)$$

Figure 2 shows the eddy currents density at the surface of the plate, which is extracted under wires 1 and 6. It can be observed that the eddy currents under wires 1 and 6, have the same density but in opposite directions; this is due to the excitation current in the EMAT coils.

Figures 3 and 4 show the spatial distribution and the density of the Lorentz force in the x-direction $\boldsymbol{F_L(x)}$ at the surface of the plate. It appears that the $\boldsymbol{F_L(x)}$ density has the same density and direction under each coil segment; this is due to the non-uniformity of the $\boldsymbol{B_y}$ under the two symmetrical sides of the coil.

C. Shear wave generation

Figure 5 shows the time history of the mechanical displacement that has been extracted at the bottom of the plate (see Fig. 6 (b)). The first upcoming wave is the longitudinal wave at the moment 34 μs and the second reaching wave is the shear wave at the moment 63 μs. This time is called the Time of Flight (ToF) of the waves [16]. The longitudinal and shear wave velocity in the stainless steel material are C_l=5.8 mm/μs and C_s=3.2 mm/μs respectively [9].

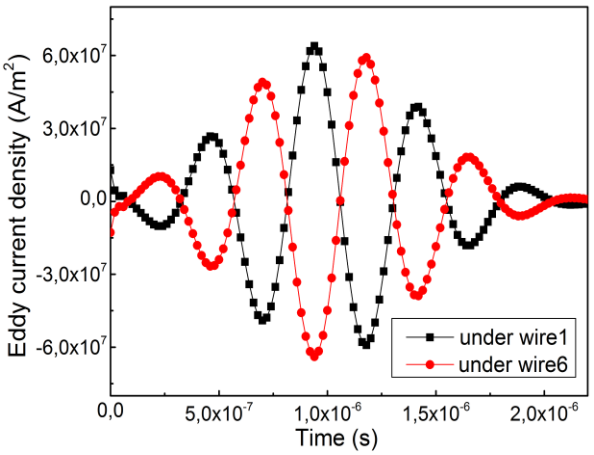


Fig. 2. Eddy current density under wire 1 and 6.

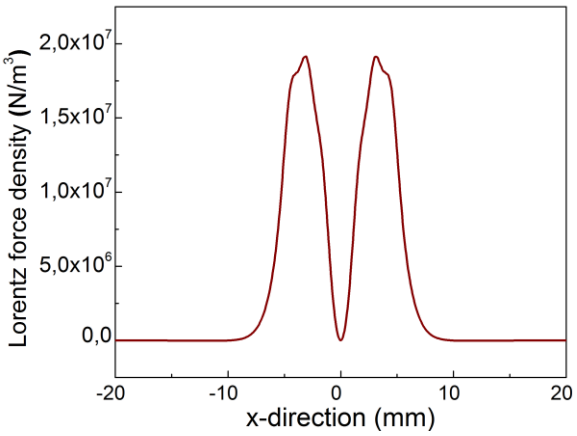


Fig. 3. Lorentz force density $\boldsymbol{F_L(x)}$ at the surface edge.

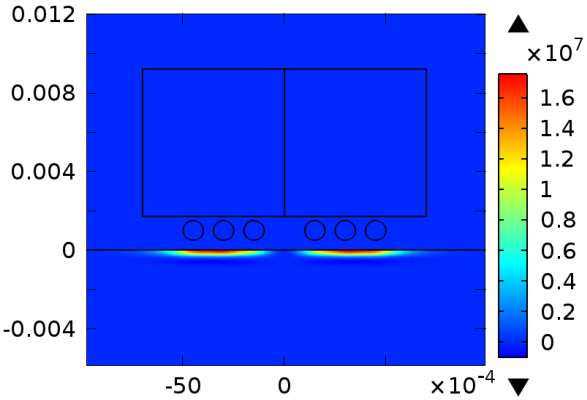


Fig. 4. Spatial distribution of the Lorentz force $\boldsymbol{F_L(x)}$.

Table 1 summarizes the comparison results between the theoretical velocity and the simulation velocity for the longitudinal and shear waves diffused in the material. From the observed error (0.81% ~ 1.41%), it can be said that there is a good agreement between theoretical and simulation velocity waves, which mean the reliability of the EMAT model to identify the hidden cracks dimensions.

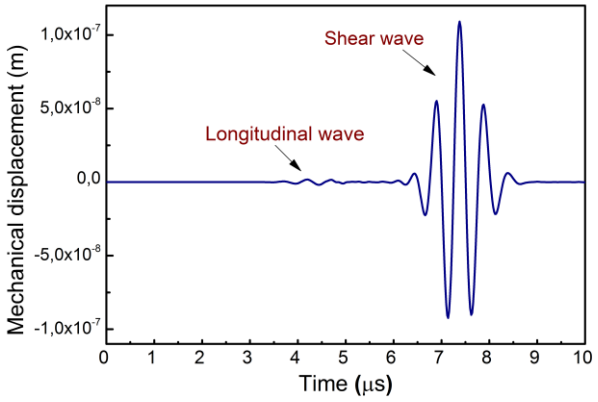


Fig. 5. Mechanical displacement at the bottom of plate.

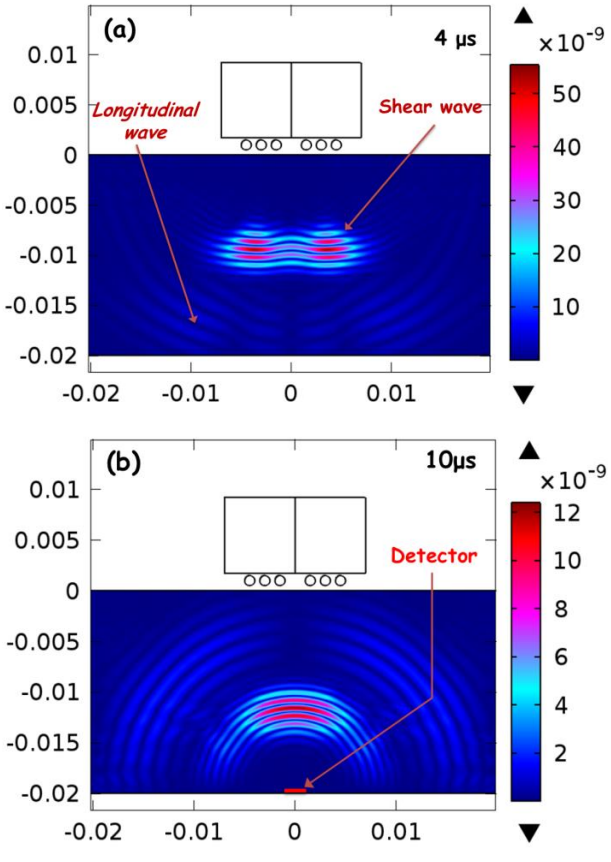


Fig. 6. Mechanical field distribution: (a) transmitted mode at (4.5 µs), and (b) reflected mode at (10.5 µs).

Table 1: Validation of the ultrasonic waves velocity

Wave Mode	Theoretical Velocity	Simulation Velocity	MAPE (%)
Shear wave	3200 (m/s)	3174 (m/s)	0.81%
Longitudinal	5800 (m/s)	5882 (m/s)	1.41%

The ultrasonic waves generated by EMAT are illustrated in Fig. 6 (a); at the instant 4 µs after the EMAT excitation, we can distinguish two types of waves. The first wave generated is the longitudinal wave, that seems faster but with low diffusion strength. The second is the shear wave; it diffuses straight towards the bottom of the plate. Moreover, it appears stronger compared to the longitudinal wave. Figure 6 (b) shows the reflected ultrasonic waves at instant 10 µs.

IV. CHARACTERIZATION OF THE HIDDEN CRACKS

A. Characterisation of cracks height by ToF

Time-of-Flight approach is a technique employed to measure the distance between the sensor and the targets for each point during the scan**,** see Fig. 7. The EMAT probe sends out an ultrasonic signal (pulse), and then picks up the returning information (echoes) from the bottom edge to the EMAT probe, i.e., measuring the round trip time of the EMAT signal.

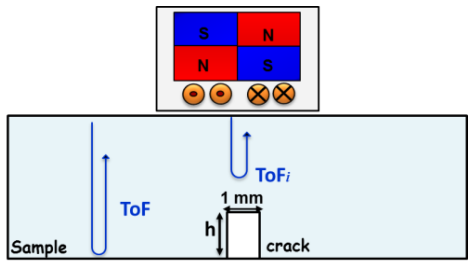


Fig. 7. Schematic represents the methodology cracks height evaluation using ToF technique.

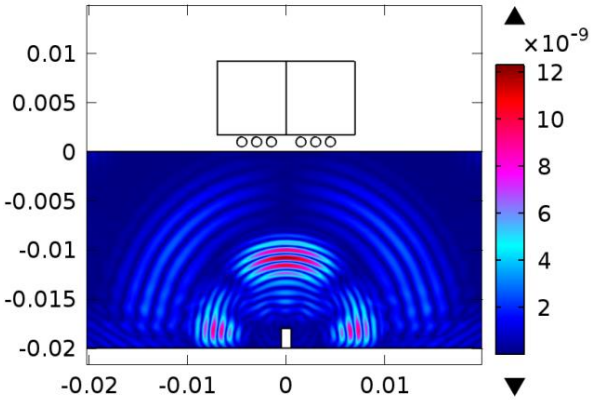


Fig. 8. Reflected ultrasonic waves from the hidden crack.

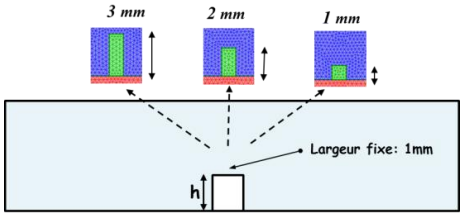


Fig. 9. Schematic represents the cracks dimensions

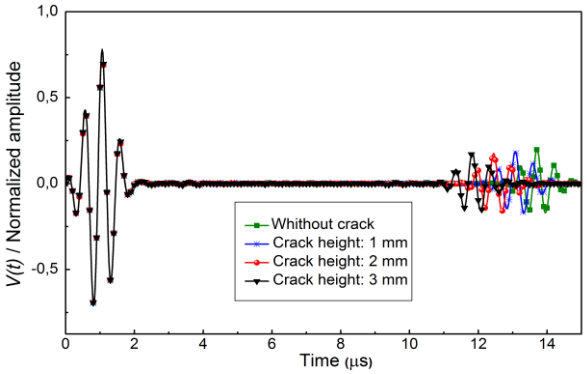


Fig. 10. Mechanical displacement for different hidden cracks height.

The crack height is calculated via multiplying the shear wave velocity C_s and the difference between the ToF of the healthy shape and that of the defective shape (ToF_i) according to the following expression:

$$h_i = \left[C_s \times \left(ToF - ToF_i \right) \right] / 2 . \quad (7)$$

In this section we introduce an artificial crack at the bottom of the stainless steel plate. Figure 8 shows the interaction of the ultrasonic waves with the hidden crack. The model was checked for five different sizes of cracks height as shown in Fig. 9.

The amplitude of the output signal ***V(t)*** by EMAT was recorded for the different sizes that are proposed. The results shown in Fig. 10, prove that the model is able to detect the existence and the changes in the crack height; this is due to the reflection caused by the presence of the hidden cracks. Table 2 presents a comparison between the desired cracks height and the estimated crack height by ToF technique. The observed Error show clearly the efficiency of the ToF approach to estimate hidden crack height.

Table 2: Estimated and desired crack height by ToF

Crack *No*:	**ToF(*i*) *µs***	**Crack Height (*m*)**		**MAPE (%)**
		Estimated	**Desired**	
1	11.9	0.96×10^{-3}	1×10^{-3}	4%
2	11.28	1.95×10^{-3}	2×10^{-3}	2.4%
3	10.9	2.57×10^{-3}	2.5×10^{-3}	2.8%
4	10.56	3.1×10^{-3}	3×10^{-3}	3.3%
5	10.2	3.6×10^{-3}	3.5×10^{-3}	2.9%

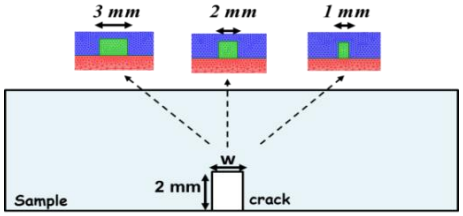


Fig. 11. Schematic represents the cracks dimensions.

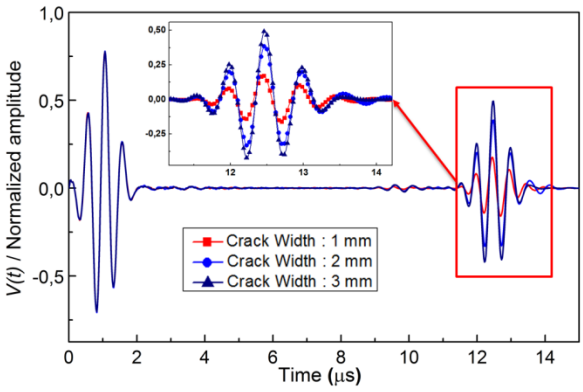


Fig. 12. Mechanical displacement for different hidden cracks width.

B. Characterisation of cracks width by TLBO

The sensitivity of EMAT to the variation of the cracks width has been checked; the crack height is fixed at 2 mm while the width is varied from 1 to 3 mm, see Fig. 11. The output signal is illustrated in Fig.12; the efficiency of the EMAT sensor to detect the crack width variation is clearly shown. Moreover, the results prove that the change of crack width involves a change in the received signal. In the next, we propose a hidden crack with unknown width, and measure the received signal by EMAT sensor. The crack width has been estimated by resolved an inverse problem. Inverse problems in NDT field are mostly stated in order to solve the optimization problems. The numerical model such as FEM is used to represent the forward problem. However, iterative methods are used to solve the inverse problem in order to deduce geometrical information about the defects [12].

In this section, the Teaching Learning Based Optimization (TLBO) algorithm is used to estimate the hidden crack width via the resolve of the inverse problem from the acquired EMAT signal. The TLBO algorithm was proposed by *Rao*; the main idea in TLBO algorithm is the philosophy of teaching and learning approach. TLBO is a meta-heuristic method for global optimization; it is easy to implement, it can be applied to unconstrained or constrained problems. *Rao and Waghmare* have been applied for many multi-objective unconstrained and constrained test functions; the results were compared with other optimization algorithms and have demonstrated

that TLBO technique is faster and precise [19]

The algorithm of this method depends on a random initialization of the population in the feasible region. The process of TLBO is divided into two phases namely: the 'Teacher Phase' and the 'Learner Phase'. In teacher phase, the teacher is considered as a highly knowledgeable person in the class and imparts his knowledge directly to learners; the best solution is considered as the teacher. Supposed that a good teacher is the one who improves his/her learners up to his/her level in terms of knowledge. In learner phase, the interaction of learners one with another is the basic idea of this phase, so random interaction between learners improves their knowledge. Several investigations have improved the performance of the TLBO algorithm and have proved that this algorithm is a powerful and a very useful tool for the optimization problems; TLBO code-algorithm and more details are available in [17]. The code of TLBO is written in MATLAB software.

The inverse problem is an evaluation of cracks width by minimizing a predefined objective function. That was formulated by the Root Mean Square Error (RMSE), which represents the differences between predicted signal $V(x_i)$ and desired output signal V_d, the objective function can be written as follows:

$$\begin{cases} \textit{minimize } f(X), \; X=(x_i) \\ f(X) = RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(V_d - V(x_i)\right)^2}, \end{cases} \quad (8)$$

where x_i represent the cracks width, the hidden crack width ranges from 250µm to 3 mm, which are distributed randomly in the investigation space N.

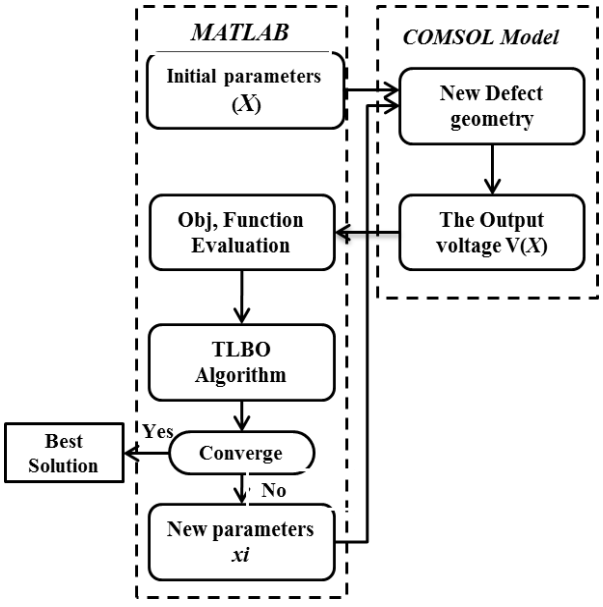


Fig. 13. Flowchart of the estimation process of the crack width by TLBO algorithm.

The resolution of the inverse problem has been carried out by the resolution of the forward problem that represented an evaluation of the output voltage in COMSOL Multiphysics. Then, the TLBO algorithm has been performed in MATLAB. Accordingly, the exhibited steps in Fig. 13 should be followed.

- Firstly, the initial parameter x_i are chosen randomly by TLBO algorithm.
- Secondly, the corresponding signal $V(X)$ of each parameter x_i is evaluated in COMSOL software. Then, the values of $V(X)$ is loaded in MATLAB in order to formulate the objective function that was represented in (8).
- Next, the TLBO algorithm is used to solve the inverse problem. The objective function has been evaluated several times for each iteration.
- Finally, if the stopping criteria is verified, the algorithm stops, otherwise, the algorithms choses new parameters from the investigation space, and repeat the evaluation of the algorithm loop until the stop criteria is verified.

Table 3: The estimated and desired width by TLBO

Defect No:	Cracks Width (*m*)		MAPE %
	Estimated	Desired	
1	$0{,}95\times10^{-3}$	1×10^{-3}	5%
2	$1{,}55\times10^{-3}$	$1{,}5\times10^{-3}$	3.3%
3	$2{,}08\times10^{-3}$	2×10^{-3}	4%
4	$2{,}45\times10^{-3}$	$2{,}5\times10^{-3}$	2%
5	$2{,}92\times10^{-3}$	3×10^{-3}	2.7%

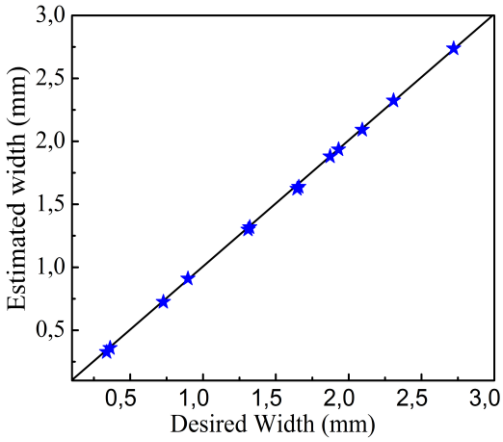


Fig. 14. Comparison of Desired and estimated of crack width by TLBO algorithm.

Table 3 shows five values of crack width desired and its corresponding estimated width using TLBO algorithm. According to the obtained results in Fig.14, it can be said that the estimation cracks width by TLBO has given results very close to the expected values, and the MAPE observed show clearly the effectiveness of

the adopted algorithm to deduce the crack width.

VI. CONCLUSION

In this work, the EMAT probe is used to generate vertical shear waves inside a stainless steel plate for the detection and characterization of the hidden cracks. ToF technique is applied to evaluate the cracks height via the comparison of the ToF between the healthy shape and the defective shape. Then, the cracks width is estimated by resolving the inverse problem from the acquired EMAT signal based on TLBO algorithm. The FEM method that was implemented in the software COMSOL Multiphysics is employed to model the generation and the reception of the ultrasonic waves by EMAT, which involved the resolve of the electromagnetic and the mechanical fields. According to the achieved results, the EMAT model illustrates a high sensitivity to the existence or the change in the size of the hidden cracks. Moreover, the results have clearly shown the efficiency and the accuracy of the adopted approach to deduce the hidden cracks dimensions.

REFERENCES

[1] P. Burrascano, S. Callegari, A. Montisci, M. Ricci, and M. Versaci, "Ultrasonic nondestructive evaluation systems: Industrial application issues," *Springer*, pp. 1-4, 2015.

[2] B. Helifa, M. Féliachi, I. Lefkaier, F. Boubenider, A. Zaoui, and N. Lagraa, "Characterization of surface cracks using Eddy current NDT simulation by 3D-FEM and inversion by neural network," *Applied Computational Electromagnetics Society Journal*, vol. 31, no. 2, pp. 187-194, 2016.

[3] J. Parra-Raad, P. Khalili, and F. Cegla, "Shear waves with orthogonal polarisations for thickness measurement and crack detection using EMATs," *NDT & E. Int.*, vol. 111, p. 102212, 2020.

[4] M. Chelabi, T. Hacib, Y. Le Bihan, N. Ikhlef, H. Boughedda, and M. Mekideche, "Eddy current characterization of small cracks using least square support vector machine," *J. Phys. D: Appl. Phys.*, vol. 49, no. 15, 155303, 2016.

[5] R. Ribichini, F. Cegla, P. Nagy and P. Cawley, "Study and comparison of different EMAT configurations for SH wave inspection," *IEEE Trans Ultrason Ferroelectr Freq Control*, vol. 58, no. 12, pp. 2571-2581, 2011.

[6] K. Mirkhani, C. Chris, M. Chris, J. Maciej, D. Tomas, S. Anthony, J. R. Shapoorabadi, K. Adalbert, and P. Marcello, "Optimal design of EMAT transmitters," *NDT & E. International*, vol. 37, no. 3, pp. 181-193, 2004.

[7] R. Dhayalan, A. Kumar, B. Rao, and T. Jayakumar, "A hybrid finite element model for spiral coil electromagnetic acoustic transducer (EMAT)," *Int. J. Appl. Electromagn. Mech.*, vol. 46, no. 3, pp. 491-500, 2014.

[8] R. Ludwig, Z. You, and R. Palanisamy, "Numerical simulations of an electromagnetic acoustic transducer-receiver system for NDT applications," *IEEE Trans. Magn.*, vol. 29, no. 3, pp. 2081-2089, 1993.

[9] M. Hirao and H. Ogi, "An SH-wave EMAT technique for gas pipeline inspection," *NDT & E. Int.*, vol. 32, no. 3, pp. 127-132, 1999.

[10] C. He, D. Peng, L. Yan, L. Xiucheng, L. Zenghua, J. Jingpin, and W. Bin, "Estimation of surface crack depth using Rayleigh waves by electromagnetic acoustic transducers," *Int. J. Acoust. Vibr.*, vol. 22, no. 4, 2017.

[11] N. Yacef, T. Bouden, and M. Grimes, "Accurate ultrasonic measurement technique for crack sizing using envelope detection and differential evolution," *NDT & E. International*, vol. 102, pp. 161-168, 2019.

[12] H. Boughedda, T. Hacib, Y. L. Bihan, and H. Acikgoz, "Cracks characterization of non-ferromagnetic material using emat probe and PLSR technique," *PIER C. J.*, vol. 103, pp. 199-209, 2020.

[13] C. Pei, S. Zhao, P. Xiao, and Z. Chen, "A modified meander-line-coil EMAT design for signal amplitude enhancement," *Sens. Actuators, A*, vol. 247, pp. 539-546, 2016.

[14] Z. Cai, Y. Yan, and G. Tian, "Enhancement of Lamb-EMAT signal using a modified one-side pitch-catch design," *IEEE Access*, vol. 7, pp. 138556-138566, 2019.

[15] C. Thring, S. Hill, S. Dixon, and R. Edwards, "The effect of EMAT coil geometry on the Rayleigh wave frequency behaviour," *Ultrasonics*, vol. 99, p. 105945, 2019.

[16] A. Habibpour-Ledari and F. Honarvar, "Three dimensional characterization of defects by ultrasonic time-of-flight diffraction (ToFD) technique," *J. Nondestr. Eval.*, vol. 37, no. 1, 2018.

[17] H. R. E. H Bouchekara and M. Nahas, "Optimization of electromagnetics problems using an improved teaching-learning-based-optimization technique," *Applied Computational Electromagnetics Society Journal*, vol. 30, no. 12, pp. 1341-1347, 2015.

[18] S. Wang, R. Su, X. Chen, L. Kang, and G. Zhai, "Numerical and experimental analysis of unidirectional meander-line coil electromagnetic acoustic transducers," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 60, no. 12, pp. 2657-2664, 2013.

[19] R. Rao, V. Savsani, and D. Vakharia, "Teaching–learning-based optimization: A novel method for constrained mechanical design optimization problems," *Computer-Aided Design*, vol. 43, no. 3, pp. 303-315, 2011.

Reduction of Cogging Torque in AFPM Machine Using Elliptical-Trapezoidal-Shaped Permanent Magnet

S. Ali [1], J. Ikram [1], C. P. Devereux [2], S. S. H. Bukhari [3,4*], S. A. Khan [1], N. Khan [1], and J.-S. Ro [3,5*]

[1] Department of Electrical and Computer Engineering, COMSATS University Islamabad, Islamabad, Pakistan

[2] Research and Development Lab. Omnidex Group Limited, Guangzhou 510050, P.R. China

[3] School of Electrical and Electronics Engineering, Chung-Ang University, Seoul, South Korea

[4] Department of Electrical Engineering, Sukkur IBA University, Sukkur, Sindh, Pakistan

[5] Department of Intelligent Energy and Industry, Chung-Ang University, Seoul, South Korea
*jongsukro@gmail.com, *sabir@iba-suk.edu.pk

Abstract ─ The reduction in cogging torque enables smooth operation and an increase in the torque density of the machine. This research aims to minimize cogging torque in dual rotor single stator axial flux permanent magnet (AFPM) machine. Reduction in cogging torque makes the back EMF sinusoidal and reduces the torque ripples in AFPM machine. In this paper, an elliptical trapezoidal-shaped permanent magnet (PM) is proposed to minimize torque ripples of the AFPM machine. The 3D finite element analysis (FEA) is used for the analysis of AFPM machine. The optimization of AFPM machine is done by employing the asymmetric magnet-overhang along with the parameters of elliptical-shaped PM using Genetic algorithm (GA).

Index Terms ─ Axial flux machine, cogging torque, elliptical-trapezoidal magnet, FEA, slot-less, torque ripples.

I. INTRODUCTION

Axial flux permanent magnet (AFPM) machines are getting popularity for electric vehicle and wind energy system, nowadays [1]–[2]. AFPM machines are beneficial as compared to radial flux machines due to higher torque and power density because of their high D/L ratio [2]–[4]. Furthermore, AFPM machine are suitable for low-speed direct drive applications since it can accommodate large number of poles. In addition, it is feasible due to its adjustable air-gap capability. Moreover, AFPM machine has high torque characteristics at low speed without utilizing mechanical gearbox. It is worth noted that slotless AFPM machine has less torque ripples as compared to the slotted AFPM machines. To further, dual rotor slotless AFPM machine has balanced magnetic force as compared to a single rotor AFPM machine. Due to higher D/L ratio, end winding in AFMs is smaller with drum type winding configuration as compared to the ring type winding configuration. In drum winding, for slotless dual rotor AFPM machine, only N-N magnet configuration is used [5]–[9].

Different techniques are used to minimize the torque ripples, *i.e.*, change in the shape of magnets, optimizing pole arc to pole pitch ratio (α), skewing the slots or magnets, and adding dummy slots [9]–[11]. Also, torque ripples are reduced by adopting different winding types. Among different shapes of magnets, trapezoidal-shaped PMs exhibit better performance as compared to the circular and rectangular-shaped PMs. An AFPM machine having arc-shaped trapezoidal PM was proposed in [12], where torque ripples were reduced as compared to the trapezoidal-shaped PM. However, using the arc-shaped trapezoidal PM, the back EMF was reduced too.

In this paper, an elliptical-trapezoidal-shaped PM AFPM machine is proposed to reduce the torque ripples and consequently enhance the output power and torque. By the use of elliptical-trapezoidal shape PM, an effective reduction of torque ripple and an improvement in back EMF is observed. To further reduce the torque ripple and increase in the back EMF, magnet overhang configuration is employed in proposed PM-shaped machine. The 3-D analysis of dual sided AFPM machine having elliptical trapezoidal-shaped magnets and drum winding on the stator is performed by using time stepped 3-D FEA. The results of AFPM machine having elliptical-trapezoidal PMs are compared to the conventional model. Furthermore, the optimization is performed by using genetic algorithm (GA). The results show a salient

Submitted On: July 16, 2019
Accepted On: July 6, 2021

https://doi.org/10.47037/2021.ACES.J.360819

reduction in the torque ripples and an increase in output torque.

II. COMPARISON BETWEEN PROPOSED AND THE CONVENTIONAL MODEL

The structure of slot-less AFPM machine consists of slotless iron stator sandwiched between the two rotor discs. Each rotor has surface mounted PMs of alternate magnetic polarity on it. Furthermore, Drum winding is used on the stator of AFPM machine in the middle of assembly. N-N type magnet configuration is used instead of N-S configuration because in drum winding the flux should flow through the stator core for maximum utilization of the windings.

Various PMs Shapes are used in AFPM machine for reducing torque ripples. However, trapezoidal-shaped PM for an AFPM machine provides better effective utilization of the rotor back iron surface as compared to the rectangular, sine wave, orthogonal and circular-shaped PMs. The increased effective utilization makes enhanced output torque. In addition, trapezoidal-shaped PM provides more uniform inter-polar separation between PMs along radial length of the disc which results in less cogging torque. AFPM machine with flat trapezoidal-shaped PMs and arc-shaped trapezoidal PMs are categorized as conventional-shaped and arc-shaped models, respectively. While an elliptical trapezoidal PM-shaped is proposed in this paper for the reduction of torque ripples.

Exploded view of elliptical trapezoidal double rotor slotless stator AFPM machine is shown in Fig. 1. Conventional, arc-shaped, and proposed models have flat top trapezoidal magnet, arc top trapezoidal magnet and elliptical trapezoidal magnet shapes, respectively. The design parameters of dual rotor single stator AFPM machine are listed in Table 1.

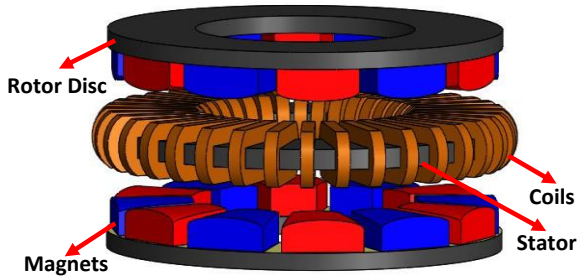


Fig. 1. Exploded view of slot-less AFPM machine.

Volume of PM in all three cases is kept constant. Height of inner and outer edge of flat top trapezoidal PM is same; however, it is different in the arc and elliptical top trapezoidal PMs as shown in Fig. 2.

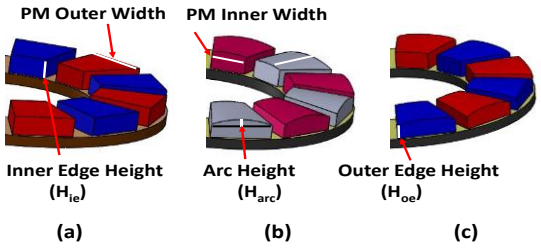


Fig. 2. Comparison of PM shapes: (a) flat trapezoidal, (b) arc trapezoidal, and (c) elliptical trapezoidal.

Table 1: Parameters of machines

Parameter	Value	Parameter	Value
Speed	1100 rpm	Stator yoke height	6.5 mm
Poles	12	Rotor yoke height	5.5 mm
Coils	36	Conductor size	0.71 mm
Air-gap	1 mm	$H_{ie\_conventional}$	10 mm
B_r	1.4 T	$H_{oe\_conventional}$	10 mm
L_m	30 mm	$H_{ie\_arc\ shape}$	9.1 mm
Coil height	16.7 mm	$H_{oe\_arc\ shape}$	7.64 mm
N_{ph}	420	$H_{ie\_proposed}$	8.7 mm
D_o/D_i	152/84.6	$H_{oe\_proposed}$	7 mm
α	0.8	$H_{arc\_arc\ shape}$	3 mm
Magnet volume	7433.1 mm^3	Coil resistance	0.1 ohm

In proposed elliptical top trapezoidal PM model, air-gap is larger than flat top trapezoidal PM model and smaller than arc top trapezoidal PM model. Comparison of air-gap lengths in flat, arc, and proposed elliptical top models is shown in the Fig. 3. Air-gap length l_g is constant and smaller in conventional flat trapezoidal PM shape due its flat top shape. The air-gap length l_g is largest in arc-shaped model and at intermediate sized in proposed elliptical model as ellipse is flatter than arc. However, there is a small difference of l_g in arc top and proposed elliptical model. Air-gap length comparison for various PM-shaped models is given in the following equation:

$$I_{g\_Arc\ shape} < I_{g\_Proposed} < I_{g\_Conventional}\ . \quad (1)$$

In conventional PM model, ϕ_g is greater than arc top and proposed elliptical model. Arc-shaped and proposed elliptical models have very minor difference of ϕ_g while arc-shaped model has the lowest value of ϕ_g. Air-gap flux comparison for the various PM shapes is given in the following equation:

$$\phi_{g\_Arc\ shape} < \phi_{g\_Proposed} < \phi_{g\_Conventional}\ . \quad (2)$$

$T_{cogging}$ is directly proportional to air-gap flux ϕ_g and the factor $dR/d\theta$ which is a change in air-gap reluctance with respect to change in the rotor position.

The relationship of cogging torque, air-gap flux, and air-gap reluctance is given as in the equation (3) [12]:

$$T_{cogging} = -\frac{1}{2}\phi_g^2 \frac{dR}{d\theta}, \quad (3)$$

where, ϕ_g is the air-gap flux, R is the air-gap reluctance, and θ is the position of rotor.

The change in the flux and change of reluctance is responsible for the cogging torque in any permanent magnet machine. Abrupt change in the reluctance is caused by the stator slots whereas the sharp edges of the magnet cause the sudden change in MMF and hence the air-gap flux to change abruptly. Air-gap flux change is smoother in elliptical magnet shape which results in less cogging torque.

The MMF is given by:

$$F_m = H_c l_m, \quad (4)$$

where F_m is the magneto-motive force, H_c is the magnetic field intensity and l_m is the height of magnet.

For flat top magnet l_m is same throughout the magnet and it has the sharp edges at both ends of magnet, which causes the sudden fall for MMF from a constant value but for arc-shaped and elliptical top magnet l_m changes gradually, *i.e.*, MMF is different at *a*, *b*, *c*, *d*, *e*, *f*, *g*, *h*, and *i*, points and MMF is reduced smoothly to the edges of the magnet.

Reduction of MMF in arc-shaped magnet causes EMF to reduce considerably because of sharp slope of l_m, from center of magnet to the edges.

The Elliptical top magnet is proposed as the solution of the problem because with the smoother surface, the l_m changes smoothly by a little value through point a, to point *i*, which causes the smaller change in MMF from point *a*, to point *i*, and hence the better distribution of air-gap flux with the less cogging torque and better EMF induction.

Figure (3) illustrates the different magnet shapes and the change in l_m for each magnet.

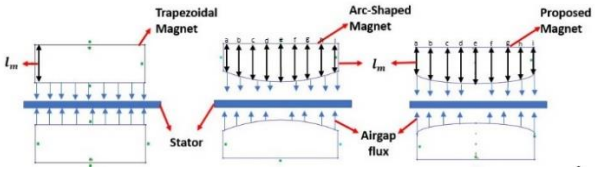


Fig. 3. Air-gap length comparison: (a) flat trapezoidal, (b) arc trapezoidal, and (c) elliptical trapezoidal.

A. Performance comparison of the conventional, arc, and proposed shape models

In this section, analysis of conventional flat trapezoidal PM model, arc trapezoidal PM model and proposed elliptical trapezoidal PM model is carried out by using 3-D finite element analysis (FEA). The results show that the proposed model has less $T_{cogging}$ as compared to the conventional and arc-shaped models. Furthermore, results show that proposed-shaped model results in increased back EMF as compared to the arc-shaped model.

The comparison of air-gap flux densities in conventional, arc-shaped, and proposed models is shown in Fig. 4. Results show that conventional model has the most and arc model has the least air-gap magnetic flux density. Air-gap flux density of proposed model lies between the air-gap flux densities of conventional and arc shape models.

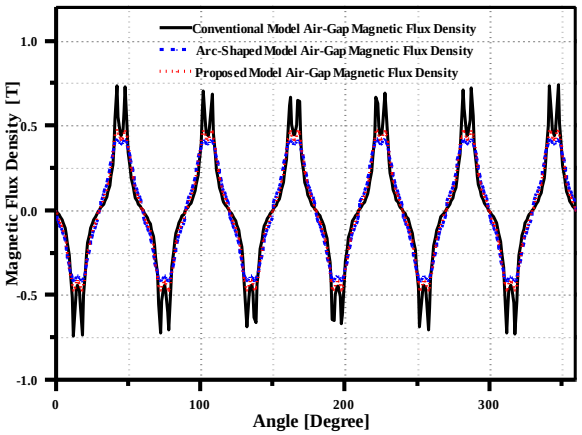


Fig. 4. Air-gap flux density comparison of conventional, arc shape and proposed model.

The comparison of cogging torque and back EMF for the conventional, arc-shaped, and proposed models a re made. The results show that cogging torque is maxim um in conventional model while minimum in proposed model. Also, the back EMF voltage is maximum in conventional and minimum in arc-shaped models. Comparison of cogging torque and back EMF for conve ntional, arc-shaped, and proposed models are shown in Figs. 5 and 6.

Conventional model has an average output power of 2208.5 W, whereas the average power the proposed and arc-shaped model is 2181W and 2108 W respectively, as shown in Fig. 7. Moreover, average output torque of conventional model is 22.82 Nm, arc-shaped model is 21.92 Nm, and the proposed model is 22.42 Nm. Furthermore, the Torque ripple in conventional, arc-shaped, and proposed model are 50.6%, 35.1% and 30.6%, respectively. Figure 8 shows the comparison of output torque of the conventional, arc-shaped, and proposed models.

The comparison of various performance parameters of conventional, arc-shaped, and proposed models is shown in Table 2.

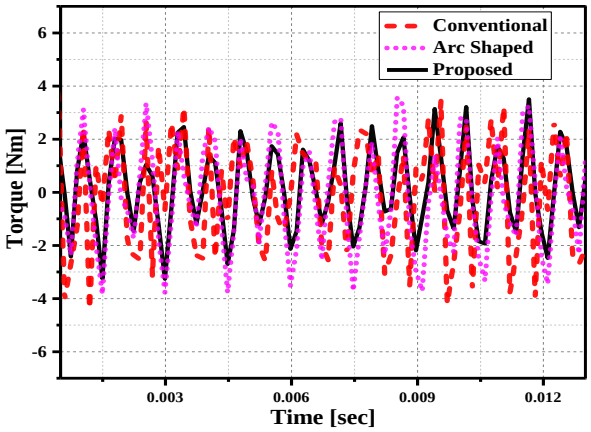


Fig. 5. Cogging torque comparison of conventional, arc shape and proposed model.

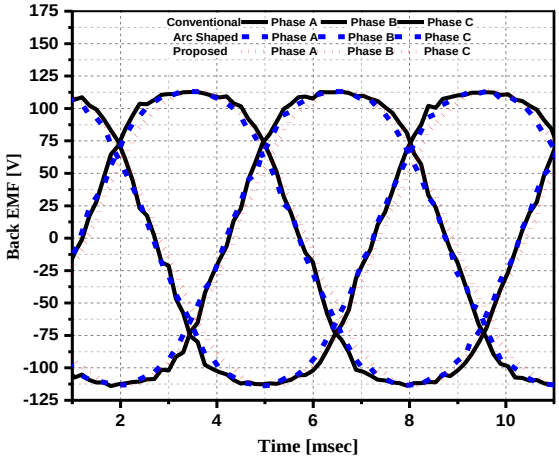


Fig. 6. Back EMF voltage comparison of conventional, arc shape and proposed model.

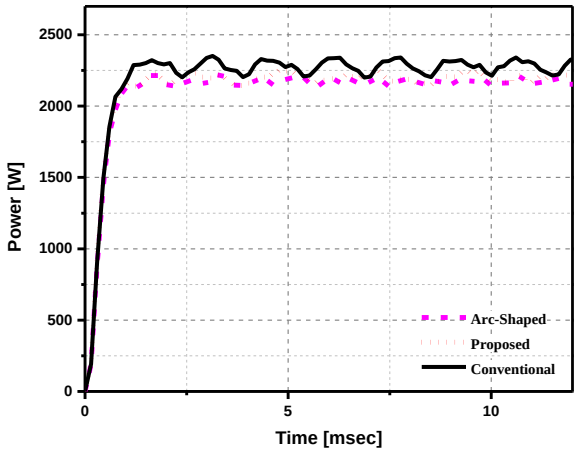


Fig. 7. Output power comparison of conventional, arc shape and proposed model.

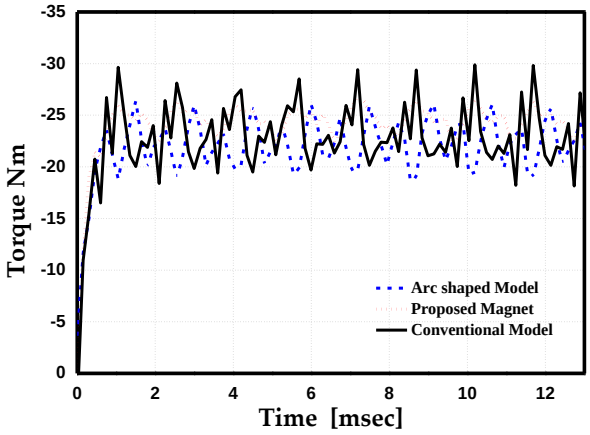


Fig. 8. Output torque comparison of conventional, arc shape and proposed model.

Table 2: Performance comparison of conventional, arc shape and proposed models

Parameters	Conventional Model	Arc-Shaped Model	Proposed Model
V_{rms}	87.6	83.1	84.4
$T_{cogging\ (pk2pk)}$	8.1	7.4	6.7
B_g (T)	0.335	0.282	0.293
V_{THD} (%)	12.2	5.6	8.8
1st & 3rd Harmonic	124.23, 12.04	118.19, 4.67	119.88, 6.7
τ_{avg} (Nm)	22.82	21.92	22.42
P_{avg} (W)	2208.5	2108	2181
$\tau_{ripples}$ (%)	50.6	35.1	30.6

III. OPTIMIZATION OF THE PROPOSED MODEL

A. Design variables

To increase the output torque as compared to the conventional model optimization of the proposed model is performed in this section.

Asymmetric magnet overhang is employed to optimize the proposed model and hence increase its back EMF and reduce the torque ripples. During optimization process, volume of elliptical trapezoidal PM is kept constant. Length of the magnet is varied along inner and outer radii by asymmetric overhang. Extending the magnet towards outer radii is termed as outer overhang and along inner radii is called inner overhang.

Latin Hypercube Sampling (LHS) was used get the samples of variables X_1, X_2 and X_3 from MATLAB. Total sixteen experiment were done with different values of parameters and the output EMF, and the cogging torque was analyzed. Genetic Algorithm (GA) was used to get the optimized value of the variables and the objectives. Values of variables and the Objectives is given in the Table 3.

Furthermore, magnet pole-arc to pole-pitch ratio and height of magnet are also varied. Height of PM is varied to make volume constant. Genetic algorithm (GA) is used to get the optimized results for average torque of proposed model. The limits of design variables are:

$$0\ mm < X_1 < 6.76mm,$$
$$0\ mm < X_2 < 8.5mm,$$
$$0.45 < X_3 < 0.8,$$

where X_1 the inner overhang of PM, X_2 is the outer overhang of PM, and X_3 the is pole-width to pole-pitch ratio (α).

According to the values of the variables X_1, X_2 and X_3, obtained by using Latin hyper cube sampling (LHS), trapezoid height of the magnet is varied to make the volume constant. Major axis of ellipse varies according to the width of the magnet and minor axis is fixed at 6 mm (3 mm half of ellipse). Pole pitch, width and height of PM are elaborated in Fig. 9.

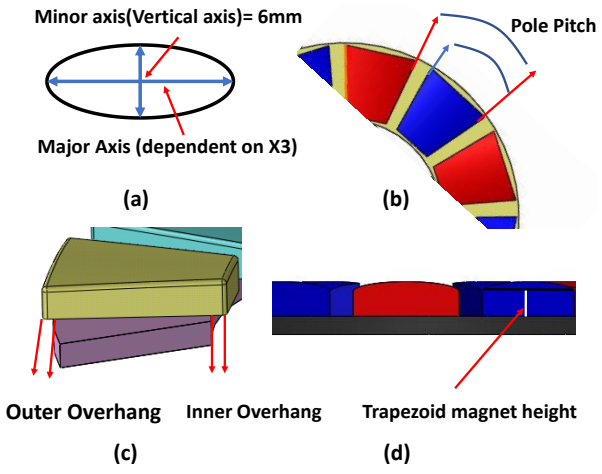


Fig. 9. Design variables of magnet.

Table 3: Experimental models for optimization and their results

No.	X_1	X_2	X_3	$T_{cogging}$	V_{rms}
1	6.03	6.16	0.54	7.34	87.51
2	3.55	5.72	0.76	6.75	89.91
3	0	1.76	0.56	6.70	80.01
4	0.35	6.6	0.63	7.46	88.77
5	5.68	4.84	0.62	5.66	88.99
6	1.42	8.36	0.47	10.01	68.19
7	6.39	7.92	0.78	5.31	90.35
8	2.84	7.48	0.74	7.79	91.04
9	3.90	3.08	0.8	7.46	89.85
10	1.06	3.52	0.58	6.49	82.46
11	5.32	1.32	0.73	3.36	87.47
12	4.97	2.64	0.65	7.21	87.01
13	2.13	2.2	0.69	8.65	87.47
14	1.77	3.96	0.51	9.33	80.74
15	6.74	7.04	0.67	7.50	89.90
16	2.48	0.88	0.45	8.25	72.17

B. Performance comparison after optimization

A comparison of various parameters for optimized elliptical-shaped permanent magnet machine with other machines is elaborated in this section. Flux density distribution of optimized elliptical-shaped permanent magnet model, conventional model and arc-shaped model is compared in Fig. 10. The optimization reduced maximum flux density from 1.8 T to 1.6 T. which shows significant improvement in the result.

The comparison of air-gap flux densities of conventional, arc-shaped, and optimized elliptical models using 3-D FEA analysis is shown in Fig. 11. The result shows a noticeable reduction in air-gap flux density.

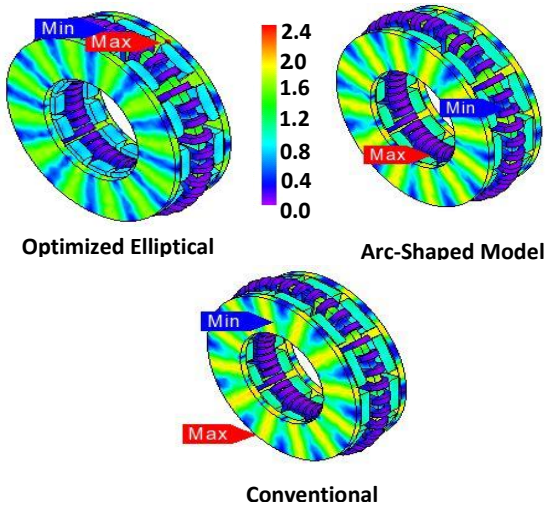


Fig. 10. Flux density distribution analysis of conventional, arc-shaped, and optimized elliptical models using 3D FEA.

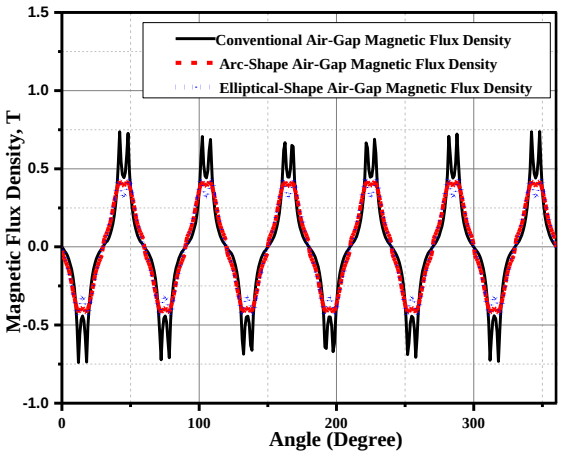


Fig. 11. Air-gap flux density comparison.

Figure 12 shows comparison of the back EMF of conventional, arc-shaped and optimized elliptical

models. The result shows that back EMF voltage is improved from 84.45 V_{rms} to 87.6 V_{rms} in elliptical-shaped model after optimization, which is an increment of 3.7 % in the back EMF of the model. Back EMF of conventional and arc-shaped model is 87.6 and 83.1, respectively.

The comparison of VTHD of conventional, arc-shaped and optimized elliptical models is made. Results show that the optimization improves the first harmonic from 119.88 V to 124.58 V as well as reduces the 3rd harmonic from 6.7 V to 2.23 V. Harmonic's comparison of conventional, arc-shaped and optimized elliptical models is shown in Fig. 13. VTHD is reduced from 8.8% to 4.3%. While VTHD in conventional and arc-shaped model is 12.2% and 5.6%, respectively.

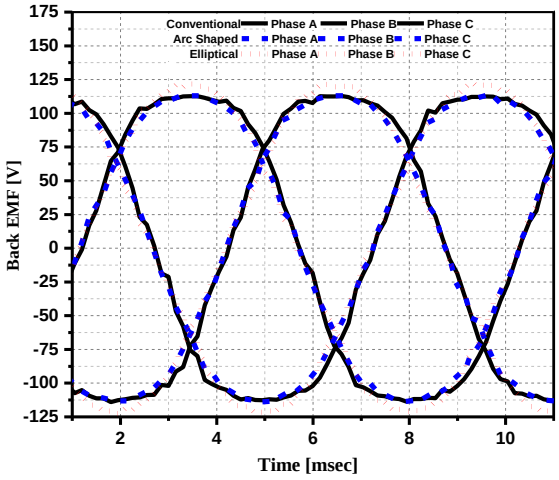


Fig. 12. Back EMF comparison.

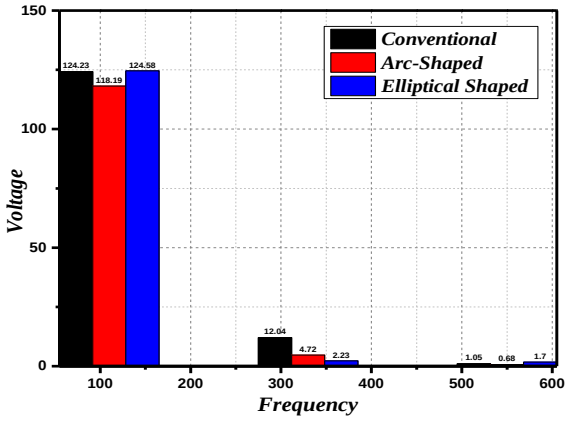


Fig. 13. Harmonics comparison.

Figure 14 shows comparison of the cogging torque for conventional, arc-shaped and optimized elliptical models. The peak-to-peak cogging torque of the proposed model is reduced from 6.77 Nm to 3.358 Nm, this shows a decrease of 50.4% through the optimization process. These results witnessed that cogging torque of the optimized model is further reduced while back EMF is also improved. So, the cogging torque of conventional, arc-shaped and optimized elliptical models is 8.1 Nm, 7.4 Nm, and 3.358 Nm, respectively.

A comparison of output power of conventional, arc-shaped and optimized elliptical models is made as shown in Fig. 15. The machine output power is improved by 7% after optimization from 2181.07 W to 2342.3 W, while output power of conventional model is 2208.5 W and that of arc-shaped model is 2108 W.

A comparison of output torque of conventional, arc-shaped and optimized elliptical models is made as shown in Fig. 16. After the optimization, output torque of the machine is improved by 7.8% and torque tipples are reduced to 17.58%. While torque ripples in arc-shaped and conventional models are 35.1% and 50.6% respectively.

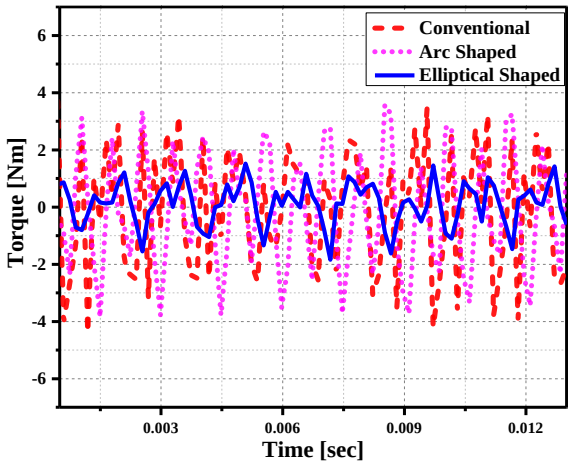


Fig. 14. Cogging torque comparison.

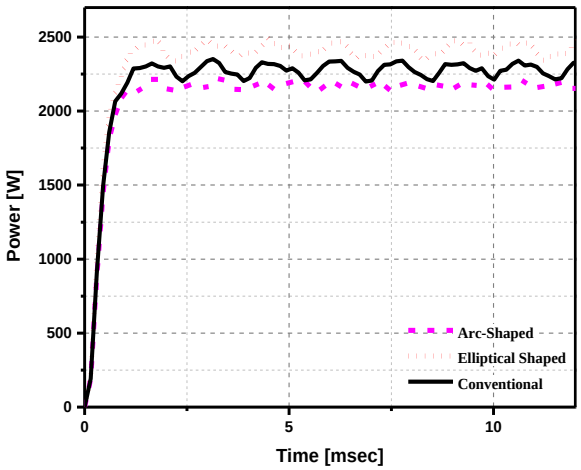


Fig. 15. Output power comparison.

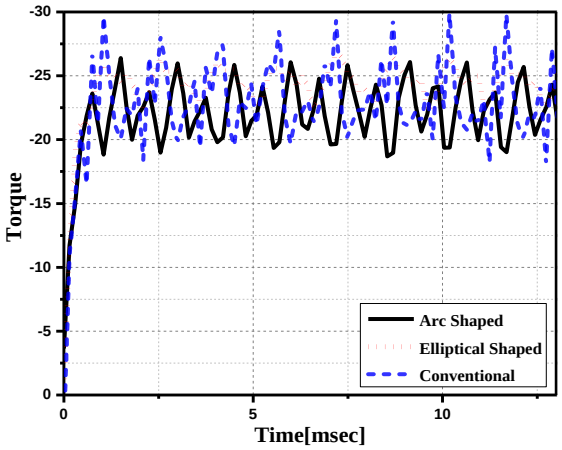


Fig. 16. Output torque comparison.

Performance comparison for proposed and optimized models is presented in Table 3. The results show that V_{rms} of optimized model is improved while $T_{cogging}$, VTHD and B_g (rms) is reduced in optimized model.

Table 3: Performance analysis of proposed and optimized model

Parameters	Elliptical Model	Conventional Model	Arc-Shaped Model
V_{rms}	87.6	87.6	83.1
$T_{cogging\ (pk2pk)}$	3.358	8.1	7.4
B_g (T)	0.282	0.335	0.282
V_{THD} (%)	4.3	12.2	5.6
1st & 3rd Harmonic	122.88, 2.23	124.23, 12.04	118.19, 4.67
τ_{avg} (Nm)	24.17	22.82	21.92
P_{avg} (W)	2342	2208.5	2108
$\tau_{ripples}$ (%)	17.58	50.6	35.1

IV. CONCLUSION

A model of the slotless AFPM machine using an elliptical trapezoidal-shaped PM is investigated in this paper. Proposed model has reduced cogging torque and torque ripples as compared to the conventional model. Furthermore, optimized elliptical-shaped magnet model has proved to be the most efficient model in terms of reduce cogging torque by an average of 58.54% as compared to the conventional model. Torque is also enhanced by 5.92% as the result of optimization of the proposed model as compared to the conventional model. Moreover, the decrease of 65.26% in torque ripples is also achieved. Additionally, power of AFPM generator is improved by 5.72%, from 2208 W to 2342 W. Hence, the optimal design demonstrates better performance as compared to the conventional and proposed models.

ACKNOWLEDGMENTS

This work was supported and funded by research and development department of Omnidex Group Limited, a multinational company with head office in 705, Building No. 7, Beijiaocheng Heng lu 1, Guangzhou 510050, P.R. China and article processing charge is supported by the Chung-Ang University Research Grant in 2020.

REFERENCES

[1] F. G. Capponi, G. De Donato, and F. Caricchi, "Recent advances in axial-flux permanent-magnet machine technology," *IEEE Trans. Ind. Appl*., vol. 48, no. 6, pp. 2190-2205, Nov./Dec. 2012.

[2] J. F. Gieras, R. Wang, and M. J. Kamper, Axial Flux Permanent Magnet Brushless Machines, *Dordrecht, The Netherlands: Kluwer*, 2004.

[3] A. Mahmoudi, S. Kahourzade, N. A. Rahim, W. P. Hew, and M. N. Uddin, "Design and prototyping of an optimised axial-flux permanent-magnet synchronous machine," *IET Electr. Power Appl*., vol. 7, no. 5, pp. 338-349, 2013.

[4] Y. P. Yang and D. S. Chuang, "Optimal design and control of a wheel motor for electric passenger cars," *IEEE Trans. Mag.*, vol. 43, no. 1, pp. 51-61, 2007.

[5] M. Andriollo, M. De Bortoli, G. Martinelli, A. Morini, and A. Tortella, "Permanent magnet axial flux disc generator for small wind turbines," *Proc. 18th Int. Conf. Electrical Machines*, pp. 1-6, 2008.

[6] P. K. Goel, B. Singh, S. S. Murthy, and S. K. Tiwari, "Three-phase four-wire autonomous wind energy conversion system using permanent magnet synchronous generator," *Electr. Power Compon. Syst.*, vol. 38, no. 4, pp. 367-386, 2010.

[7] S. Kahourzade, A. Mahmoudi, H. W. Ping, and M. N. Uddin, "A comprehensive review of axial-flux permanent-magnet machines," *Can. J. Elect. Comput. Eng*., vol. 37, no. 1, pp. 19-33, 2014.

[8] S. Wu, S. Zuo, X. Wu, F. Lin, and J. Shen, "Magnet modification to reduce pulsating torque for axial flux permanent magnet synchronous machines," *Applied Computational Electromagnets Society Journal*, vol. 31, no. 3, Mar. 2016.

[9] N. P. Gargov, A. F. Zobaa, and I. Pisica, "Investigation of multi-phase tubular permanent magnet linear generator for wave energy converters," *Electr. Power Compon. Syst.*, vol. 42, no. 2, pp. 124-131, 2014.

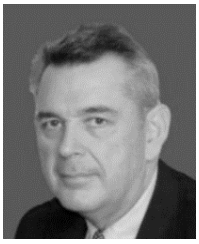
[10] E. Aycicek, N. Bekiroglu, I. Senol, and Y. Oner, "Rotor configuration for cogging torque minimization of the open-slot structured axial flux permanent magnet synchrnous motors," *Applied Computational Electromagnets Society Journal*, vol. 30, no. 4, pp. 396-408, 2015.

[11] T. M. Jahns and W. L. Soong, "Pulsating torque minimization techniques for permanent magnet ac motor drives: A review," *IEEE Trans. Ind. Electron*., vol. 43, no. 2, pp. 321-330, Apr. 1996.

[12] J. Ikram, N. Khan, Q. Junaid, S. Khaliq, and B. I. Kwon, "Analysis and optimization of the axial flux permanent magnet synchronous generator using an analytical method," *Journal of Magnetics*, vol. 22, pp. 257-265, 2017.

[13] K. Y. Hwang, H. Lin,S. H. Rhyu, and B. I. Kwon, "A study on the novel coefficient modeling for a skew permanent magnet and overhang structure for optimal design of BLDC motor," *IEEE Transactions on Magnetics*, vol. 50, pp. 1918-1923, 2012.

Salman Ali was born in Pakistan on 1st April 1992. Ali received his B.S. degree in Electrical Engineering from Federal Urdu University Islamabad in 2015 and the M.S. degree in Electrical Engineering from COMSATS University Islamabad in 2019. He was Research Associate in COMSATS University from 2017 to 2019. Currently, He is serving as Electro-Mechanical Development Engineer in R&D Department of Omnidex Group Limited since 2019. His research interests are design, analysis and prototyping different topologies of electrical machines. Until now, most of his research is in the design, analysis, and prototyping of axial flux permanent magnet machines.



Christopher P. Devereux was born in England in 1947 but spent the next 4 years of his early life in Abbottabad, Pakistan. On returning to England Christopher Devereux progressed through an uneventful British schooling until 17 when he organized and led an expedition into the Amazonas of Brazil. This inspired him to start a lifetime of travel basing himself first in Central and South America, then back to the UK and then on to China where he started and now runs an engineering company. He follows in his grandfather's footsteps (who filed over 200 engineering patents for the Indian railway system) in developing new engineering concepts and is the holder, himself, of several patents.

Syed Sabir Hussain Bukhari received his B.E degree in Electrical Engineering from Mehran University of Engineering and Technology Jamshoro, Pakistan, in 2009, and Ph.D. from the Department of Electronic Systems Engineering, Hanyang University, South Korea in 2017. He joined Sukkur IBA University as an Assistant Professor in December 2016. He is currently working as a Research Professor at Chung-Ang University, Seoul, South Korea under Korean Research Fellowship (KRF) program. His main research interests include electric machine design, power quality, and drive controls.

Junaid Ikram received his B.E. degree in Electrical Engineering from University of Engineering and Technology Lahore, Pakistan in 2005, M.S. degree from Hanyang University, South Korea in 2009 and he did his Ph.D. degree in Electrical Engineering from Comsats University Islamabad, Pakistan in 2017. Currently he is working as a Senior Engineer in COMSATS University Islamabad. His research interests include design, analysis and optimization of axial flux machines, vernier machines, wound rotor synchronous machine, hybrid flux switching machines and modeling of the machine losses.

Shahid A. Khan received the B.S. degree in Electrical Engineering from the University of Engineering and Technology, Lahore, Pakistan, and the M.S. and Ph.D. degrees in Radio Communications from the Universality of Portsmouth, U.K. In 2003, he joined COMSATS University Islamabad (CUI) as an Assistant Professor and elevated to a professor within few years. He worked as the HOD, the Chairman, and the Dean of the Faculty of Engineering. He is currently the Founder of various engineering programs in different campuses of CIIT. He has established many teaching and research labs by securing funding of about ten million USD. He has authored over 80 research articles published in international journals and conferences. On the basis of his meritorious services, he has been honored with multiple awards and memberships in the field of engineering education.

Nasrullah Khan received the B.Sc. degree in Electrical Engineering degree from UET Lahore, in 1984, the M.Sc. degree from Reading University, U.K., in 1988, and the Ph.D. degree from Essex University, U.K., in 1992. He has 21 years teaching and research and 12 years practical field experience. He has done several research projects in COMSATS and abroad. He has 34 years local and foreign teaching, research, and field experience.

Jong-Suk Ro received the B.S. degree in Mechanical Engineering from the Han-Yang University, Seoul, South Korea, in 2001, and the Ph.D. degree in Electrical Engineering from Seoul National University (SNU), Seoul, South Korea, in 2008. He conducted research at the Research and Development Center of Samsung Electronics as a Senior Engineer, from 2008 to 2012. From 2012 to 2013, he was with the Brain Korea 21 Information Technology of SNU, as a Postdoctoral Fellow. He conducted research at the Electrical Energy Conversion System Research Division, Korea Electrical Engineering and Science Research Institute, as a Researcher, in 2013. From 2013 to 2016, he worked with the Brain Korea 21 Plus, SNU, as a BK Assistant Professor. In 2014, he was with the University of Bath, Bath, U.K. He is currently an Associate Professor with the School of Electrical and Electronics Engineering, Chung-Ang University, Seoul, South Korea. His research interests include the analysis and optimal design of next-generation electrical machines using smart materials such as electromagnet, piezoelectric, and magnetic shape memory alloy.

Characteristics Analysis of Double-Sided Permanent Magnet Linear Synchronous Motor with Three-Phase Toroidal Windings

Xiaobao Chai [1], Jikai Si [1], Yihua Hu [2], Yingsheng Li [3], and Dongshu Wang [1]

[1] Department of Electrical Engineering, Zhengzhou University, Zhengzhou, 450001, China
xbcstu@126.com, sijikai@zzu.edu.cn, wangdongshu@zzu.edu.cn

[2] Department of Electronic Engineering, University of York, York, YO 10 5DD, U.K.
yihua.hu@york.ac.uk

[3] Zhengzhou Runhua Intelligent Equipment Co., Ltd., Zhengzhou, 450004, China
xdlys@vip.sina.com

Abstract ─ This paper proposes three-phase 120° phase belt toroidal windings (120°-TW) and are applied in a double-sided permanent magnet linear synchronous motor (DSPMLSM), in which the incoming ends of all coils are on the same side and have the same incoming direction. First, the structure of the proposed motor is introduced and its operation principle is analyzed by describing the variation in the armature magnet field versus time. Second, based on the similar volume, magnetic load and electrical load, the initial parameters of the DSPMLSM with different winding arrangements are presented. Then, the finite-element models (FEMs) of the DSPMLSM with 120°-TW (120°-TWDSPMLSM) and traditional toroidal windings (TTW) are established to analyze the distribution of magnetic field, back electromotive force (back-EMF), detent force, thrust, efficiency and so on. Besides, the primary optimization of the detent force is designed. Finally, the results show that the thrust density and efficiency of the 120°-TWDSPMLSM is higher than that of DSPMLSM with TTW (TTWDSPMLSM).

Index Terms ─ Character analysis, double-sided permanent magnet linear synchronous motor, operation principle, three-phase 120° phase belt toroidal windings, thrust density.

I. INTRODUCTION

Double-sided permanent magnet (PM) linear synchronous motors (DSPMLSMs) are widely used in various industrial applications and daily life due to advantages of simple structure, high thrust density, low maintenance cost and unilateral magnetic force in recent years [1-4]. However, in limited space applications, the thrust density of the DSPMLSM is expected to be higher. Therefore, many researchers worldwide are pursuing the goal of improving the thrust density of the DSPMLSL [5-7].

Researchers have done much excellent work to improve the thrust density, which can be summarized four improvement techniques. The first technique adds magnetic field modulator blocks, the thrust density is improved by adopting the method magnetic field modulation in [8-9]. The second technique improve the thrust density by changing the PM arrangement, the thrust density is improved by adopting the structure of serial magnetic circuit [10-11]. The third technique is to change the cooling mode, the force-air cooling and water-cooling method are adopted, which improves the thrust density [12-13]. The last technique changes the winding configurations, and changing winding configurations are more simple comparing with these alternatives. The concentrated and full-pitch windings are applied in [14], but the sinusoidal waveform of back electromotive force (back-EMF) is poor and has some higher-order harmonics. In [15] the DSPMLSM adopting distributed and short-pitch windings are designed and reduces the harmonic content of back-EMF, but it has great trouble to manufacture the motor. The DSPMLSM with fractional-slot concentrated windings is proposed and is suitable for limited space applications due to small size, high efficiency and easy manufacture [16-17]. However, the winding factor of the fractional-slot concentrated winding is relatively low, which has negative effects on the amplitude of back-EMF and thrust density. To improve the thrust density, the DSPMLSM with traditional toroidal windings (TTWDSPMLSM) is proposed in [18], and the thrust density of the TTWDSPMLSM is higher than the DSPMLSM with fractional-slot concentrated windings. In [19], the torque density of a direct-drive permanent magnet synchronous motor with 120° phase belt toroidal windings is higher than that with traditional toroidal windings. The 120°-TW also can be applied to the DSPMLSM to improve

Submitted On: April 6, 2021
Accepted On: July 12, 2021

https://doi.org/10.47037/2021.ACES.J.360820

the thrust density.

Summarized the above topology structures, this paper presents a double-sided permanent magnet linear synchronous motor with three-phase 120° phase belt toroidal windings (120°-TWDSPMLSM), in which the incoming ends of all coils are on the same side and have the same incoming direction. The structure and the winding connected method of the 120°-TWDSPMLSM is introduced in Section II. Besides, its operating principle is analyzed. In Section III, the design of the 120°-TWDSPMLSM is illustrated in detail. Then, the finite element models (FEMs) of the two motors are established in Section IV. Comprehensive comparisons between the 120°-TWDSPMLSM and TTWDSPMLSM from the aspect of operating characteristic are presented, and the primary optimization of the detent force is designed. Finally, a brief summary of this paper as well as the directions for future research work are proposed in Section V.

II. STRUCTRE AND OPERATION PRINCIPLE OF THE 120°-TWDSPMLSM

A. Structure of the 120°-TWDSPMLSM

The structure sketches of the 120°-TWDSPMLSM and TTWDSPMLSM are shown in Fig. 1 and Fig. 2, respectively. It can be seen from Fig. 1 and Fig. 2 that both the two motors consist of one primary and two secondary components. Besides, in the secondary, the permanent magnets of the two motors are arranged alternately along the moving direction. The prominent advantage for the double-sided secondary topology is that the toroidal windings can be adopted.

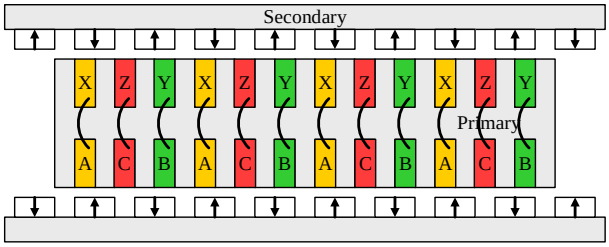


Fig. 1. Structure sketch of the 120°-TWDSPMLSM.

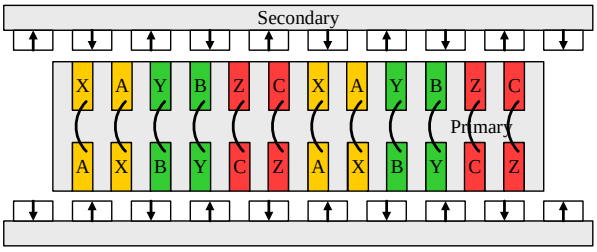


Fig. 2. Structure sketch of the TTWDSPMLSM.

Moreover, it could be also seen that the 120°-TWDSPMLSM is different with the TTWDSPMLSM, and the differences between the two motors are found as following.

(1) The unit motor of the 120°-TWDSPMLSM has three slots and two poles, as shown in Fig. 1, However, it can be found from Fig. 2 that the unit motor of the TTWDSPMLSM has six slots and four poles.

(2) It should be noted that all incoming line ends of the 120°-TW are on the same side and have the same incoming direction, however, the incoming line ends of the TTW are on both side of the motor.

B. Operation principle of the 120°-TWDSPMLSM

A three-slot unit motor is selected as an example to clearly clarify the operation principle of the 120°-TWDSPMLSM in this subsection. Three-phase current supplied to the three-phase windings are shown in Fig. 3. The current directions of the three-phase winding coils at different time are shown in Table 1, where '+' and '−' represent that the current flows into and out the windings, respectively. Based on the distributions of the three-phase winding currents at different times and the faraday law of electromagnetic induction, the armature magnetic field distribution in each moment is described in detail, as shown in Fig. 4, where '+' and '•' represent that the current flows into and out the windings, respectively.

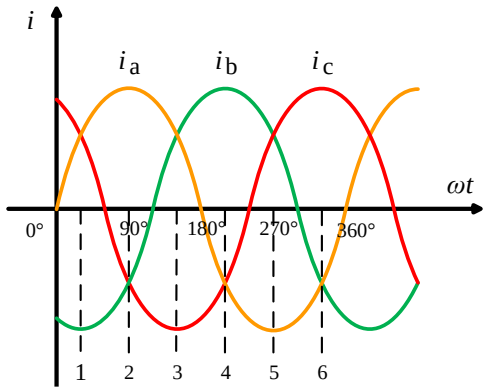


Fig. 3. Current waveforms supplied to three-phase windings.

Table 1: Current directions of three-phase winding coils versus time

Phase	Time					
	1	2	3	4	5	6
A	+	+	+	−	−	−
B	−	−	+	+	+	−
C	+	−	−	−	+	+

The variation of the magnetic poles can be found from Fig. 4 as following.

(1) Compared with the magnetic fields at Time 1,

Time 2 shows that the N pole of the armature magnetic fields has no movement, but the S pole moves 120° in the left direction.

(2) The N pole of the armature magnetic fields at Time 3 has moved by 120° in the left direction, while the S pole has no movement compared with that at Time 2.

(3) Compared with the magnetic fields at Time 3, Time 4 shows that the N pole of the armature magnetic fields does not move, but the S pole moves 120° in the left direction.

(4) The N pole of the armature magnetic fields at Time 5 has moved by 120° in the left direction, while the S pole does not move compared with that at Time 4.

(5) Compared with the magnetic fields at Time 5, Time 6 shows that the N pole of the armature magnetic fields has no movement, but the S pole moves 120° in the left direction.

As a result, it can be concluded that the armature magnetic field of the three-slot unit motor forms a pair of poles and changes periodically.

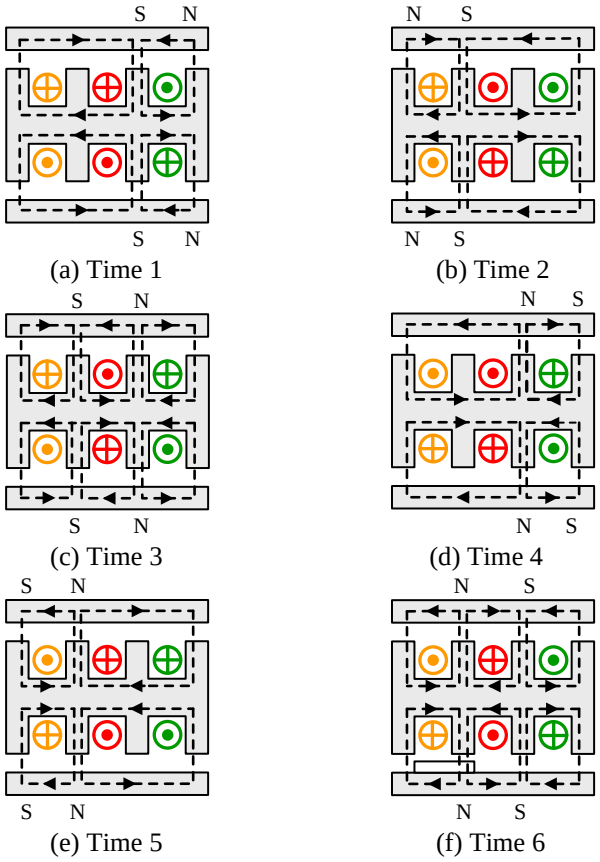


Fig. 4. Armature magnetic field versus time.

III. INITIAL PARAMETER OF THE 120°-TWDSPMLSM AND TTWDSPMLSM

To verify the feasibility of the proposed motor, the120°-TWDSPMLSM is designed and compared with the TTWDSPMLSM. In [18], the main dimensional parameters of two motors are presented in detail. The design needs to meet the following criteria to ensure a fair comparison.

(1) The two motors have the similar air-gap flux density.

(2) The wire diameter of the 120°-TWDSPMLSM must be consistent with that of TTWDSPMLSM.

(3) The two motors have the same moving speed.

(4) The two motors are excited by a current source and has the same electrical load.

(5) The same materials are adopted in the corresponding parts of the two motors.

(6) The same mesh refine is adopted.

Based on the principles above, the initial parameters of the 120°-TWDSPMLSM and TTWDSPMLSM are summarized in Table 2.

Table 2: Initial parameters of the 120°-TWDSPMLSM and TTWDSPMLSM

Items	120°-TWDSPMLSM	TTWDSPMLSM
Pole numbers	8	8
Slot numbers	12	12
Pole pitch (mm)	16.5	16.5
PM width (mm)	12.4	12.4
PM thickness (mm)	3	3
Air-gap length (mm)	1	1
Slot pitch (mm)	11	11
Slot width (mm)	6	6
Teeth width (mm)	5	5
Coil number/phase	4	4
Turns/coil	135	135
Primary width (mm)	60	60
synchronous speed (m/s)	0.33	0.33

IV. CHARACTERISTIC ANALYSIS OF THE 120°-TWDSPMLSM AND TTWDSPMLSM

With the parameters reported in Table 2, the Finite Element Models (FEMs) of the two motors are established to compare and analyze the electromagnetic characteristics under no-load and on load conditions.

A. No-load characteristic

To analyze the internal magnetic field intuitively and effectively, the magnetic flux distribution, air-gap flux density, back-EMF and detent force under no-load condition of the two motors are compared and analyzed.

The no-load magnet flux distributions of the two motors are same, because the flux generated by the permanent magnet is independent of the winding arrangements. Figure 5 shows the no-load magnetic flux

distribution and the approximate closed loops in the 1/4 model of the two motors.

The two motors can generate parallel magnetic circuits due to the corresponding magnets on the two sides that have the same polarity facing the primary. In addition, it can be seen that there are three types of paths in 120°-TWDSPMLSM and TTWDSPMSLM, which are called the long loop, middle loop, and short loop in this paper, respectively. The magnetic flux lines of the long loop and middle loop are close through the secondary yoke, the air gap, the primary teeth and the primary yoke along the magnetization direction to form a loop, which is called effective magnet flux. By contrast, the magnetic flux lines of the short loop do not pass through the primary yoke, which is called ineffective magnet flux.

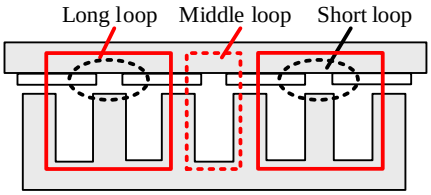


Fig. 5. The no-load magnetic flux distribution and the approximate closed loops in 120°-TWDSPMLSM and TTWDSPMSLM (1/4 model).

The no-load air-gap flux density waveforms and the corresponding harmonics distribution of the two motors within two pole pitches are presented in Fig. 6.

It can be seen that the air-gap flux density waveforms of the 120°-TWDSPMLSM is consistent with the air-gap flux density waveforms of the TTWDSPMLSM. By Fourier analysis, the amplitudes of the fundamental component of the air-gap flux density in the 120°-TWDSPMLSM and TTWDSPMLSM both are 0.78T.

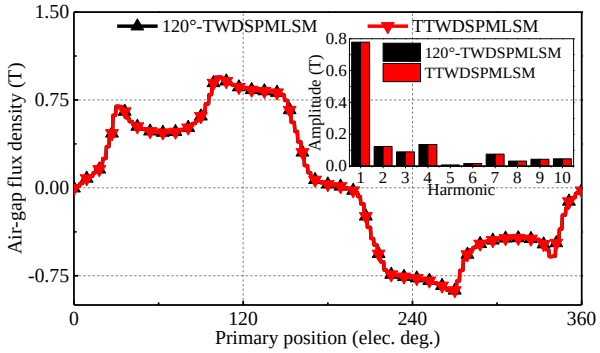


Fig. 6. No-load air-gap flux density waveforms of the 120°-TWDSPMLSM and TTWDSPMLSM.

The no-load back-EMF waveforms and the corresponding harmonics distribution of the 120°-TWDSPMLSM and TTWDSPMLSM at the same speed are presented in Fig. 7.

By the Fourier analysis, it can be seen that the no-load back-EMF of the two motors have little even harmonics. Moreover, the amplitudes of the fundamental component of the no-load back-EMF of the 120°-TWDSPMLSM (17.24V) is higher 18.98% than that of TTWDSPMLSM (14.49V). The total harmonic distortions (THDs) of back-EMF of the 120°-TWDSPMLSM and TTWDSPMLSM are 2.63% and 1.80%, respectively. And the THDs of back-EMF of the 120°-TWDSPMLSM is a little higher than that of the TTWDSPMLSM.

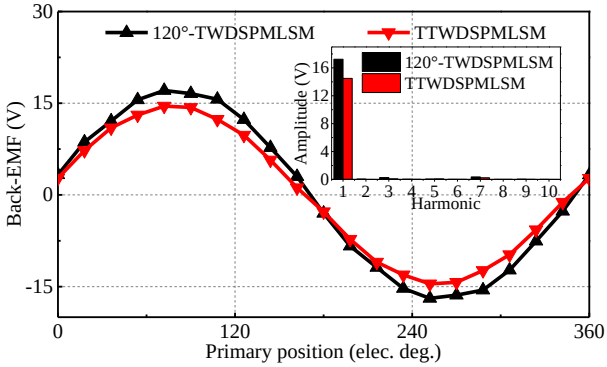


Fig. 7. No-load back-EMF waveforms of the 120°-TWDSPMLSM and TTWDSPMLSM.

The fundamental harmonic amplitude of no-load back-EMF of one side of a coil of the 120°-TWDSPMLSM can be expressed as [20]:

$$E_1 = NBlv, \tag{1}$$

where, N is the number of the conductor, B is the fundamental harmonic amplitude of the air gap flux density, l is conductor effective length, v is the speed of the motor.

Since each E_1 is superposition the same direction. Therefore, the fundamental harmonic amplitude of no-load back-EMF of the 120°-TWDSPMLSM can be further expressed as:

$$E = zE_1 = zNBlv, \tag{2}$$

where, z is the slot number.

According to formula (2) and the parameters of the 120°-TWDSPMLSM in Table 2, it is can be calculated that the E is equal to 16.68 V. The result has small deviation comparing with simulation result (17.24 V), but the error (3.35%) is less than 5%. Thus, the EFA is credible.

The detent force, which can be divided into the end force and cogging force, can cause vibration and noise of DSPMLSM. therefore, it is a key problem to be considered and solved in the design and manufacture of high performance DSPMLSM. The detent force F_{dent} is shown in formulas (3-5) [21]:

$$F_{dent} = F_{end} + F_{slot}, \tag{3}$$

$$F_{end} = \frac{2K_c\delta\Phi_m^2}{\mu_0 k_1 \tau^2 l}\frac{4}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}(\sin n\frac{2\pi}{\tau}x), \tag{4}$$

$$F_{solt} = \frac{\pi^2 zl}{2\mu_0 p\tau}(\frac{p\tau}{\pi}+\delta^2) \sum_{n=1}^{\infty} nG_n B_{nz/2p}\sin(\frac{2n\pi z}{p\tau}x), \tag{5}$$

where K_c is the carter's coefficient, δ is the length of air-gap, Φ_m is the maximum flux amplitude across the end of primary, μ_0 is the permeability of vacuum, k_1 is the flux compression coefficient, τ is the pole pitch, $B_{nz/2p}$ is the $nz/2p$-th harmonic of residual magnetism. G_n is the n-th harmonic of permeance, F_{end} is the detent force, F_{slot} is the cogging force.

Based on the finite element models and formulas (3-5), the detent force values of the two motors are calculated, as shown in Fig. 8. It can be seen that the peak-to-peak values of detent force of the two motors both are 100.84 N. According to the formulas (3-5), since all parameters are the same, the detent forces of the two motors are equal. The detent force is large, and it is important to minimize the detent force in a later subsection.

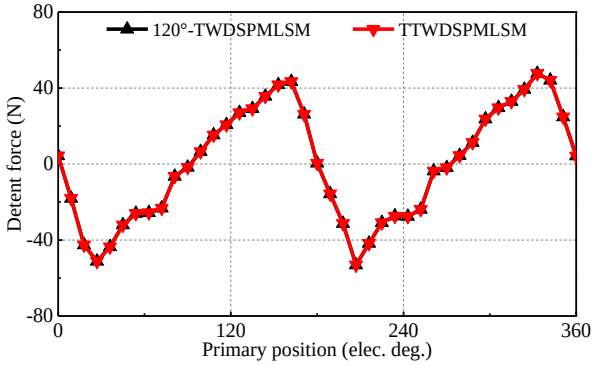


Fig. 8. Detent force waveforms of the 120°-TWDSPMLSM and TTWDSPMLSM under no load.

B. Load characteristic

The average thrust F is shown in formula (6) [22]:

$$F = \frac{3EI\cos\varphi}{\sqrt{2}v}, \tag{6}$$

where, I is the average value of phase current, cosφ is the power factor.

The thrust waveforms of the two motors at the same current density (6.0 A/mm^2) are shown in Fig. 9. It can be seen that the average thrust of the 120°-TWDSPMLSM and TTWDSPMLSM are respectively 250.40 N and 205.88 N. Besides, the thrust density of the 120°-TWDSPMLSM (4.84×10^2 kN/m^3) is 21.62% higher than that of the TTWDSPMLSM (3.98×10^2 kN/m^3).

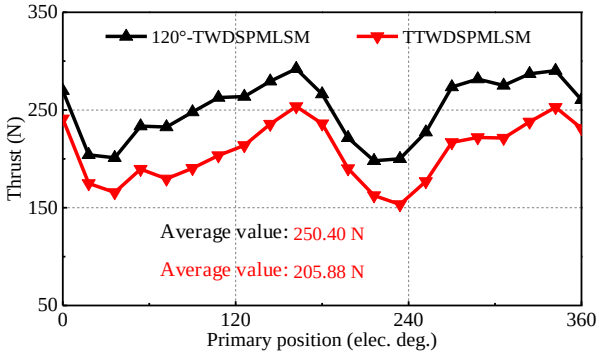


Fig. 9. Thrust waveforms of the 120°-TWDSPMLSM and TTWDSPMLSM on load.

According to the formula (6), it can be calculated that the average thrust F of the 120°-TWDSPMLSM is equal to 254.12 N. The simulation result broadly concurs with (6), with the small deviation which can be attributed to the harmonic which is ignored in (6), whereas is considered in the FEM simulation model.

Additionally, the thrust ripple F_{ripple} is shown in formal (7) [18]:

$$F_{ripple} = \frac{F_{max} - F_{min}}{F_{av}}, \tag{7}$$

where F_{max}, F_{min}, and F_{av} are the maximum, minimum and average value of the thrust, respectively.

The thrust ripples of the 120°-TWDSPMLSM and TTWDSPMLSM are 37.60% and 48.70%, respectively. Based on the above of detent force, it can be known that the main cause of thrust ripple is too large detent force. Therefore, in the next work, to reduced thrust ripple, it is important to minimize the detent force.

The Table 3 presents the losses and efficiency of the two motors.

Table 3: Losses and efficiency

Items	120°-TW DSPMLSM	TTW DSPMLSM
Output power (W)	82.63	67.94
Iron loss (W)	0.37	0.37
Copper loss (W)	81.89	81.89
Efficiency	50.11%	45.23%

It can be found that the output power of the 120°-TWDSPMLSM and TTWDSPMLSM are 82.63 W and 67.94 W, respectively. The copper losses of the two motors are same (81.89 W), because the number of conductors and the wire diameter are same in the two motors. Besides, the iron losses of the 120°-TWDSPMLSM and TTWDSPMLSM are both 0.37 W. the iron losses are lower than the copper losses of the two motors, which is negligible due to the frequency is 10 Hz. Therefore, the losses of the two motors are

mainly copper losses. The efficiency of the 120°-TWDSPMLSM and TTWDSPMLSM is 50.11% and 45.23%, respectively. Therefore, the efficiency of the 120°-TWDSPMLSM increases by 4.88%.

Figure 10 shows the thrust-current angle characteristic of the 120°-TWDSPMLSM and TWDSPMLSM.

It should be noted that the current angle indicates the relative angle between the current phasor and the d-axis. The change of current angle affects the amplitude of q-axis current (I_q). Then increasing I_q can lead to the increasement of the average thrust. If the current angle is between 0 and 90 elec. deg., I_q increases with the current angle. On the contrary, if the it is between 90 and 180 elec. deg., I_q decreases as the current angle increases. When it is 90 elec. deg., I_q is injected only, i.e., the amplitude of d-axis current (I_d) is equal to 0 A. The motors achieve maximum thrust.

Therefore, when the current angle is between 0 and 180 elec. deg., the average thrust increases first and then decreases.

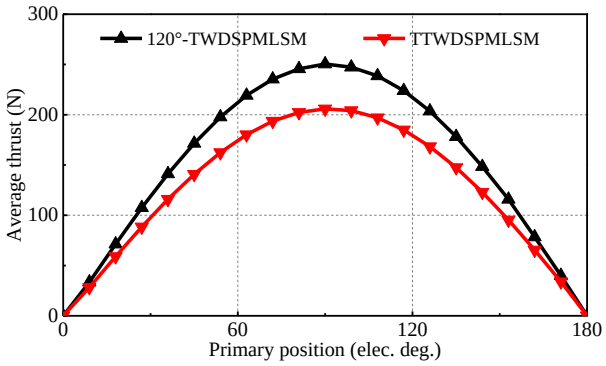


Fig. 10. Thrust-current angle characteristic of the 120°-TWDSPMLSM and TWDSPMLSM.

Figure 11 shows the waveforms of the average thrust versus current density in two motors.

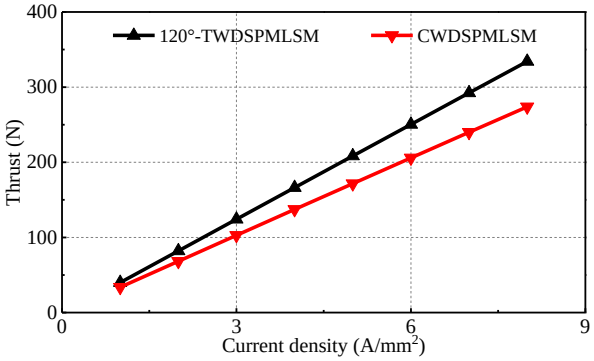


Fig. 11. Thrust versus current density.

It can be found that the average thrust of the 120°-TWDSPMLSM is always larger than that of TTWDSPMLSM. Moreover, the average thrust of the two motors both increase linearly with the armature current density.

The waveforms of the loss versus frequency in the two motors are shown in Fig. 12.

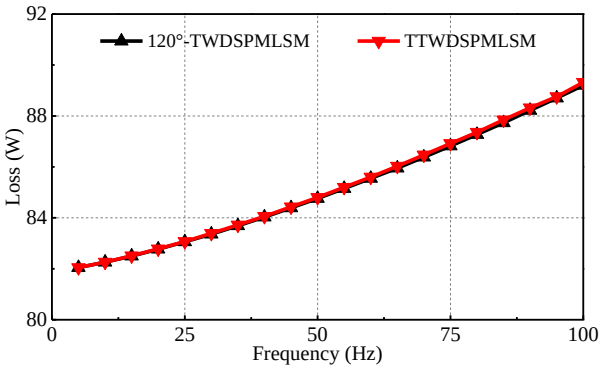


Fig. 12. Loss versus frequency.

It can be seen the losses of the two motors increase with frequency. Moreover, it can be found that the losses of the 120°-TWDSPMLSM and TTWDSPMLSM are same in various frequencies.

Figure 13 shows the waveforms of the efficiency versus frequency in the two motors.

It can be seen that the efficiencies of the two motors increase with the frequency and tend to be stable, finally. Besides, the efficiency of the 120°-TWDSPMLSM is higher than that of TTWDSPMLSM in various frequencies.

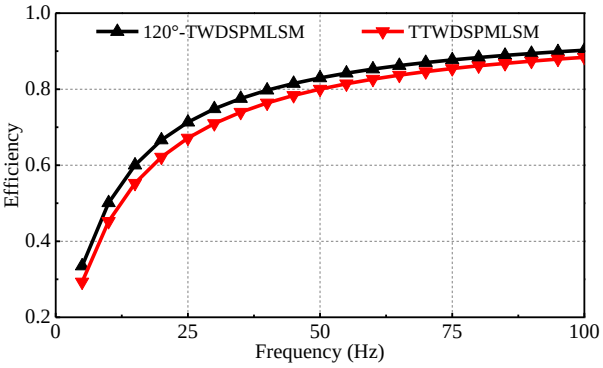


Fig. 13. Efficiency versus frequency.

C. Optimization of the detent force

The primary optimization of the detent force of the two motor is shown in Table 4.

As shown in the Table 4, the different length of the end tooth corresponds to the different value of the detent force. While the length of the end tooth is 6.0mm, the detent force is 100.84N. After optimization, it drops to

51.50N, when the length of end tooth is 8.0mm, which means that 48.92% of detent force can be reduced by changing the length of end tooth.

Table 4: Length of end tooth

Length of end tooth (mm)	5.5	6.0	6.5	7.0
Detent force (N)	103.77	100.84	92.25	79.46
Length of end tooth (mm)	7.5	8.0	8.5	9.0
Detent force (N)	67.64	51.49	69.10	83.40

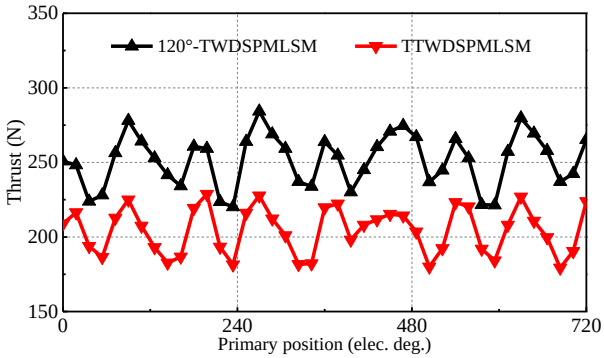


Fig. 14. Thrust waveforms of the 120°-TWDSPMLSM and TTWDSPMLSM on load.

The primary optimization waveforms of the thrust of the two motors are shown in Fig. 14. By primary optimization, it can be seen that the thrust ripples of the 120°-TWDSPMLSM and TTWDSPMLSM are 25.32% and 23.73%, respectively. Therefore, the thrust ripples of the 120°-TWDSPMLSM and TTWDSPMLSM can reduced by are 12.28% and 24.97% by changing the length of the end tooth, respectively.

D. FEA credibility and 120°-TW reasonableness validation

In order to verify the credible finite element analysis method and the reason of the proposed winding configuration (120°-TW), a permanent magnet rotation motor (PMRM)with the same 120°-TW can be used in the experiment. Figure 15 shows the prototype of the PMRM with 120°-TW (120°-TWPMRG), which includes the pedestal, rotor, stator yoke and 120°-TW. Figure 16 shows the corresponding test platform for the motor.

Figure 17 shows the comparison of the no-load back-EMF between EFA and the experiment at the constant linear velocity of 1 krpm. Figure 18 shows the comparison of the torque between FEA and the experiment at different current.

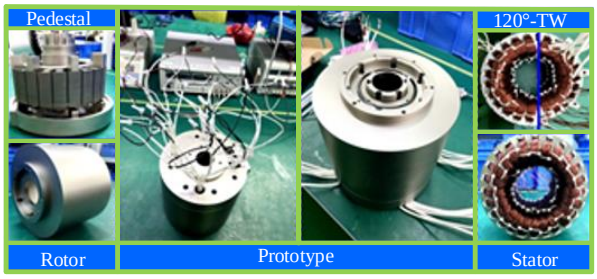


Fig. 15. Prototype of the 120°-TWPMRM.

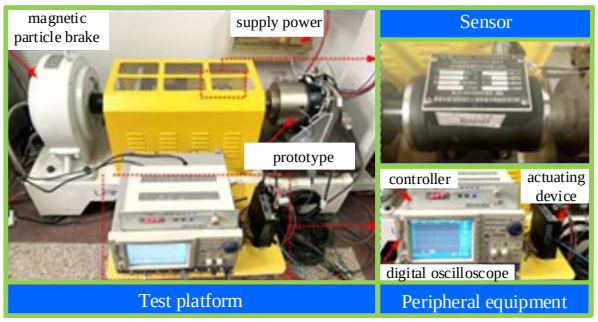


Fig. 16 Test platform.

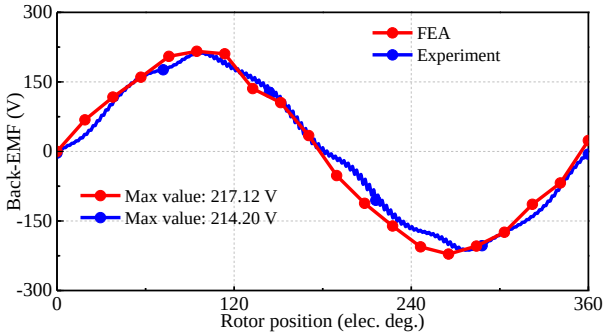


Fig. 17. Comparison of the no-load back-EMF between FEA and experiment.

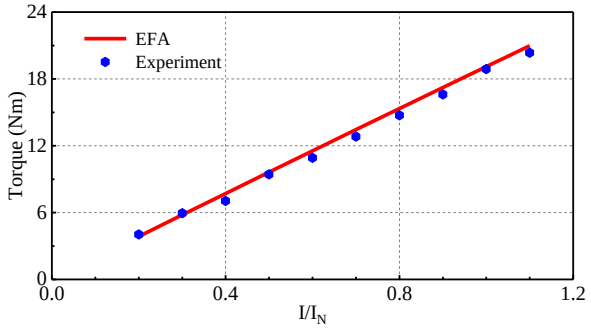


Fig. 18. Comparison of the torque between FEA and experiment at different current.

It can be seen that the no-load back-EMF obtained by FEA agrees with the experimental ones, which demonstrated the FEA method used to analyze the performance is accurate. Therefore, the analysis results obtained by the same analysis method (FEA) are credible in this paper.

V. CONCLUSION

This paper presented a 120°-TWDSPMLSM, which has simple structure, high thrust density. The main work is focused on the characteristic analysis and comparison of the 120°-TWDSPMLSM and TTWDSPMLSM. The structure of the proposed motor is introduced and operation principle is clarified. Additionally, the key operating characteristics of the two motors are analyzed and compared. Moreover, the detent forces of two motors are preliminarily optimized. The comparison results validate the reasonability of the 120°-TWDSPMLSM and exhibit the proposed motor possesses some advantages which can be concluded as below.

(1) The back-EMFs of the 120°-TWDSPMLSM is higher 18.98% than that of the TTWDSPMLSM on the same moving speed.

(2) Compared with that of the TTWDSPMLSM, the thrust density of the 120°-TWDSPMLSM is increased by 21.62% on the similar magnetic load and electrical load.

(3) The efficiency of the 120°-TWDSPMLSM is higher than that of TTWDSPMLSM in various frequencies.

The primary drawback of the 120°-TWDSPMLSM is the high detent force and thrust ripple. Therefore, the further optimization of the 120°-TWDSPMLSM will be a future investigation and shall be reported soon.

ACKNOWLEDGMENT

This work is partially supported by Natural Science Foundation of China under grant No.51777060, in part by the Major Special Project for Collaborative Innovation in Zhengzhou No. 20XTZX12023.

REFERENCES

[1] C. E. Kim, S. H. Lee, D. H. Lee, and H. J. Kim, "The analysis of permanent magnet double-sided linear synchronous motor with perpendicular arrangement," *IEEE Transactions on Magnetics*, vol. 49, no. 5, pp. 2267-2270, May 2013.

[2] Y. S. Kwon and W. J. Kim, "Electromagnetic analysis and steady-state performance of double-sided flat linear motor using soft magnetic composite," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 3, pp. 2178-2187, Mar. 2017.

[3] Z. J. Zhang, M. Z. Luo, J. A. Duan, and B. Q. Kou, "Performance analysis of double-sided permanent magnet linear synchronous motor with quasi-sinusoidal ring windings," *IEEE Transactions on Energy Conversion*, vol. 35, no. 3, pp. 1465-1474, Sept. 2020.

[4] B. Rezaeealam and F. Rezaeealam, "Optimization of permanent magnet synchronous motors using conformal mappings," *Applied Computational Electromagnetic Society (ACES) Journal*, vol. 32, no. 10, pp. 915-923, Oct. 2019.

[5] G. D. Liu, Y. Z. Wang, X. P. Xu, W. Y. Ming, and X. Zhang, "The optimal design of real time control precision of planar motor," *Applied Computational Electromagnetic Society (ACES) Journal*, vol. 32, no. 10, pp. 948-954, Oct. 2019.

[6] F. F. Bian, W. X. Zhao, J. H. Ji, L. Xu, H. Chen, and M. H. Gao, "Mechanism investigation of ring type winding in linear permanent magnet vernier machine for improving force density," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 2588-2597, Mar. 2020.

[7] A. Boduroglu, Y. Demir, B. Cumhur, and M. Aydin, "A novel track structure of double-sided linear PM synchronous motor for low cost and high force density applications," *IEEE Transactions on Magnetics*, DOI 10.1109/TMAG. 3017448, 2021.

[8] Y. Zhou, C. J. Shi, R. H. Qu, D. W. Li, and Y. T. Gao, "Overview of flux-modulation linear permanent magnet machines," *Proceeding of the CSEE*, vol. 41, no. 4, pp. 1469-1484, Feb. 2021.

[9] C. J. Shi, R. H. Qu, Y. T. Gao, D. W. Li, L. B. Jing, and Y. Zhou, "Design and analysis of an interior permanent magnet linear vernier machine," *IEEE Transactions on Magnetics*, vol. 54, no. 11, pp. 1-5, June 2018.

[10] L. Y. Li, M. N. Ma, B. Q. Kou, and Q. Q. Chen, "Analysis and design of moving-magnet-type linear synchronous motor for electromagnetic launch system," *IEEE Transactions on Plasma Science*, vol. 39, no. 1, pp. 121-126, Jan. 2011.

[11] B. Q. Kou, H. X. Wu, L. Y. Li, L. L. Zhang, Z. Zhao, and H. C. Cao, "The thrust characteristics investigation of double-side plate permanent magnet linear synchronous motor for EML," *IEEE Transactions on Magnetics*, vol. 45, no. 1, pp. 501-505, Jan. 2009.

[12] W. M. Tong, S. N. Wu, and R. Y. Tang, "Research on the airflow and thermal performance in a large forced air-cooled permanent magnet synchronous machine," *IEEE Access*, DOI 10.119/ACCESS. 2951919, 2019.

[13] Y. Q. Wang, X. Y. Huang, and Y. T. Fang, "Thermal analysis and cooling structure selection of electric machine applied in EV," *Micromotors*, vol. 53, no. 1, pp. 1-5, Jan. 2020.

[14] W. B. Kong, J. Huang, R. H. Qu, M. Kang, and J. Q. Yang, "Nonsinusoidal power supply analysis for concentrated-full-pitch-winding multiphase induction motor," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 1, pp. 574-582, Jan. 2016.

[15] X. Z. Huang, T. P. Ji, L. Y. Li, B. Zhou, Z. R. Zhang, D. Gerada, and C. Gerada, "Detent force, thrust, and normal force of the short-primary double-sided permanent magnet linear synchronous motor with slot-shift structure," *IEEE Transactions on Energy Conversion*, vol. 34, no. 3, pp. 1411-1421, Sept. 2019.

[16] S. W. Seo, G. H. Jang, J. M. Kim, and J. Y. Choi, "Characteristic analysis and experimental verification for a double-sided permanent magnet linear synchronous generator according to magnetization array," *IEEE Transactions on Applied Superconductivity*, DOI 10.1109/TASC. 2790941, 2018.

[17] S. G. Lee, S. A. Kim, S. Saha, Y. W. Zhu, and Y. H. Cho, "Optimal structure design for minimizing detent force of PMLSM for a ropeless elevator," *IEEE Transactions on Magnetics*, DOI 10.1109/ TMAG. 2277544, 2013.

[18] H. Zhang, B. Q. Kou, Z. Q. Zhu, R. H. Qu, J. Luo, and Y. Shao, "Thrust ripple analysis on toroidal-winding linear permanent magnet vernier machine," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 12, pp. 9853-9862, Dec. 2018.

[19] C. X. Gao, M. Z. Gao, J. K. Si, Y. H. Hu, and C. Gan, "A novel direct-drive permanent magnet synchronous motor with toroidal windings," *Energies*, DOI 10.3390/en12030432, 2019.

[20] M. C. Sun, R. Y. Tang, X. Y. Han, and W. M. Tong, "Analysis of open circuit back electromotive force in slotless toroidal type windings axial flux permanent magnet machine," *Electric Machine and Control Application*, vol. 44, no. 9, pp. 1-8. Sep. 2017.

[21] S. L. Shi, E. L. Kang, and G. Y. Shi, "Optimization design of a permanent magnet linear motor using finite element and Taguchi method," *Micromotors*, vol. 53, no. 3, pp. 33-37, Mar. 2020.

[22] S. L. Shi and E. L. Kang, "Electromagnetic design and performance analysis of permanent magnet synchronous linear motor," *Micromotors*, vol. 48, no. 3, pp. 25-28, July 2020.

Xiaobao Chai was born in China in 1997. He received B.S. degrees in Electrical Engineering and Automation, Electrical Engineering from the Department of Electrical Engineering and Automation, Shenyang University of Technology, China, in 2019. He is currently working toward the M.S. degree in Electrical Engineering from the Department of Electrical Engineering and Automation, Zhengzhou University, China. His current research interests include design, analysis and control of double-sided permanent magnet linear synchronous motor.

Jikai Si received the B.S. degree in Electrical Engineering and Automation from the Jiaozuo Institute of Technology, Jiaozuo, China, in 1998; the M.S. degree in Electrical Engineering from Henan Polytechnic University, Jiaozuo, China, in 2005; and the Ph.D. degree in 2008 from the School of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou, China, in 2008. He is currently a Distinguished Professor at Zhengzhou University. His main research interests include the theory, application, and control of special motor. He has authored and co-authored over 160 technical papers in these areas. Si is a Member of the Green Motor System Professional Committee, China.

Yihua Hu received the B.S. degree in Electrical Motor Drives, in 2003, and the Ph.D. degree in Power Electronics and Drives, in 2011, both from China University of Mining and Technology, Jiangsu, China. Between 2011 and 2013, he was a Postdoctoral Fellow with the College of Electrical Engineering, Zhejiang University, Zhejiang, China. He is currently a Distinguished Professor at University of York. His research interests include PV generation system, power electronics converters and control, and electrical motor drives.

Yingsheng Li currently works in Zhengzhou Runhua Intelligent Eq-uipment Co., Ltd, and is the legal representative of Zhengzhou Runhua Intelligent Equipment Co., Ltd. His research interests include the application, control of motor, and power electronics converters and control, and electrical motor drives.

Dongshu Wang received the bachelor's degree in Mechanical Manufacture Technique and Equipment, the master's degree in Mechanical Manufacture and Automation, and the Ph.D. degree in Control Theory and Control Engineering from Northeastern University, Shenyang, China, in 1996, 2002, and 2006, respectively. He is currently an Associate Professor with the School of Electrical Engineering, Zhengzhou University, Zhengzhou, China. His research domains are autonomous mental development and artificial intelligence.

Electronically Reconfigurable WLAN Band-Notched MIMO Antenna for Ultra-wideband Applications

Asim Quddus [1*], Rashid Saleem [2], Salman Arain [1], Syed Rizwan Hassan [1], and M. Farhan Shafique [3]

[1] Department of Electrical Engineering, NFC Institute of Engineering and Fertilizer Research, Faisalabad

[2] Department of Telecommunication Engineering, University of Engineering and Technology, Taxila, 47050, Pakistan

[3] Center for Advanced Studies in Telecommunication
COMSATS Institute of Information Technology, Islamabad, Pakistan
*asim.quddus@iefr.edu.pk

Abstract ─ A low-profile electronically reconfigurable WLAN band-notched dual port multiple-input multiple-output (MIMO) antenna design for ultra-wideband applications has been presented. Antenna elements exhibit good impedance match ($S_{11} \leq$ -10 dB) over entire ultra-wideband (3.1 to 10.6 GHz) spectrum. The decoupling structure is used to improve isolation (S_{12}/S_{21}) above 20 dB over entire UWB band. Moreover, reconfigurable band-notching is achieved by inserting PIN diodes along the inverted L-shaped slots, in each radiator. Notch at WLAN (5.5 GHz) frequency band is achieved by switching the PIN diode to 'OFF' state. The antenna design is fabricated as well as measured, and the results suggests that the proposed design with switchable WLAN band-notch characteristics is suitable candidate for ultra-wideband applications.

Index Terms ─ Isolation, multiple-input multiple-output (MIMO), PIN diodes, reconfigurable, ultra-wideband (UWB), WLAN band-notch.

I. INTRODUCTION

Wireless communication technologies have gained much attention over last few decades. The prime focus of recent wireless technology is to use minimum resources, achieve high data rate and cause minimum interference to other existing wireless communication standards [1]. To achieve higher data rates and higher channel capacities, MIMO technology is integrated in UWB systems. However, a significant challenge in UWB-MIMO communication is the miniaturization of system. The unwanted mutual coupling is caused by miniaturization and hence the effectiveness of MIMO is compromised. Therefore, in order to provide decoupling between antenna elements, an efficient isolating/ decoupling structure is desired in MIMO systems. In existing literature, several MIMO antenna designs with decoupling structures have been reported to attain high isolation between antenna elements [2-3]. To achieve interference mitigation in UWB communication, several designs with band-notched characteristics have been reported in the existing literature [4-5]. However, these proposed antennas have permanent band-notching. For the sake of interference free communication, utilization of whole UWB spectrum may not be possible even if there is no conflicting narrow band system working in the close proximity. Therefore, for the improvement of the UWB system performance, antennas with reconfigurable/ switchable band notch performances are desirable [6-7].

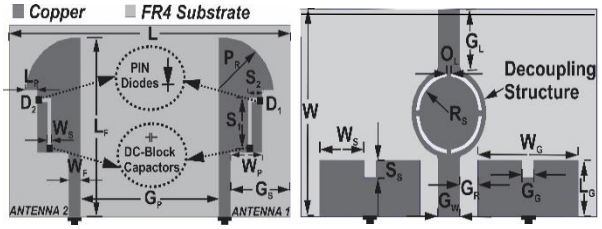


Fig. 1. Geometry of dual-port antenna system: (a) front view and (b) rear view.

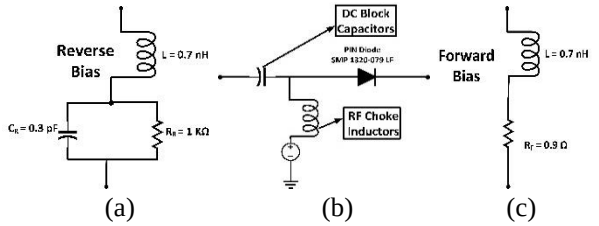
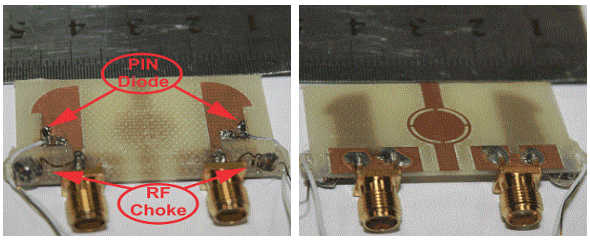


Fig. 2. Equivalent circuit model of PIN diode: (a) case I, (b) biasing network, and (c) case II.

Submitted On: June 6, 2018
Accepted On: December 28, 2019

https://doi.org/10.47037/2021.ACES.J.360821



(a) (b)

Fig. 3. Fabricated design: (a) front view and (b) back view.

A miniaturized dual port MIMO antenna exhibiting electronically reconfigurable WLAN (5.1 - 5.8 GHz) band notch capability is proposed in this research for ultra-wideband MIMO applications. More than 20 dB isolation (S_{12}/S_{21}) is achieved among radiating elements for entire ultra-wideband using a decoupling structure. The analysis and optimization of these computationally intensive structures is performed using a full-wave Finite Element Method (FEM) based electromagnetic solver (Ansys HFSS™). The whole structure is enclosed in an air box with radiation boundary condition and the antenna is fed through a wave port excitation. In order to get realistic results of simulation, whole structure is simulated at once and no symmetric boundary has been adopted for MIMO antenna designs.

II. DESIGN CONFIGURATION OF ANTENNA

The geometric structure of proposed antenna is presented in Fig. 1. Antenna design is simulated and fabricated on a 1.6 mm thick FR4 substrate. The upper layer contains two radiators. Whereas, the partial ground and circular decoupling structure is placed on the back side of substrate. The decoupling structure as shown in Fig. 1 (b) contains defected circular ring. It is placed between ground planes to isolate the side-by-side placed antenna elements. The single UWB antenna element and two port UWB MIMO antenna with decoupling structure has been designed in HFSS using finite element method. MIMO design can be extended to four port by placing symmetrical two port MIMO antenna in front of designed MIMO antenna at appropriate distance. Appropriate decoupling structure is required to achieve isolated four port operation. The overall decoupling of more than 20 dB between MIMO antenna elements is achieved.

The optimized design parameters are shown in Table 1. All parameters of the design have known effect on the antenna response like the ground related parameters affect the impedance matching thus dictates the bandwidth. The parameters associated with the decoupling structure changes the mutual coupling. The parameters associated with the slot length and width change the notching frequency. All these parameters work in groups where one group achieves one single objective, thus co-optimization is not required making things simple and easy to tune. Therefore, no automated optimization method has been used and parameters were optimized manually around their optimum value through the simulator (High Frequency Structure Simulator).

The effective length of notching structure is required to filter out the desired frequency band. Notching at desired frequency bands can be obtained using equation (1):

$$f_r = \frac{c}{4L\sqrt{\varepsilon_{eff}}}. \quad (1)$$

An inverted L-shaped slot $L_{Total} = S_1 + S_2$ is created in main radiator of each antenna element. It provides WLAN (5.5 GHz) band notch characteristics. The current around the edges of slots reverses its direction causing anti-resonance at the desired notched frequency band. Moreover, to achieve the reconfigurable band notch functionality in UWB-MIMO antenna, PIN diode is embedded in the slot of each radiator, as shown in Fig. 1 (a). The purpose of PIN diode is to switch the UWB-MIMO antenna between ultra-wideband operation and WLAN band notch functionality. The PIN diode, D_1, as shown in Fig. 1 (a), controls Antenna 1 band notching. Similarly, D_2 diode control the band notching functionality of Antenna 2. The reconfigurable characteristics of proposed UWB-MIMO antenna have two operations as below:

Case I: When diode is forward biased or is in its 'ON State'. Antenna offers no notching behavior and UWB-MIMO antenna provides matching without any band notch.

Case II: When diode is reverse biased or PIN diode is in its 'OFF State'. Antenna eliminates the WLAN band from the UWB band and operates as a band notch antenna.

The PIN diode is first modeled in HFSS™ using lumped elements according to the equivalent model for Case I and Case II, as shown in Fig. 2. SMP 1320-079 LF PIN diodes have been used for switching. The diode has low reverse bias (zero volt) capacitance of 0.3 pF at above 10 MHz frequencies and low resistance of 0.9 ohms at 10 mA during forward bias operation. For the prevention of short circuiting in PIN diode biasing circuitry, each slot is connected with a 30 pF DC block capacitor as shown in equivalent circuitry of biasing network Fig. 2 (b).

III. RESULTS AND DISCUSSIONS

Antennas are fabricated on FR4 laminate, as shown in Figs. 3 (a) and (b). As the antenna elements are identical, impedance matching is similar for each element. The results are presented in Fig. 4. Result shows that both antenna elements are well-matched over entire ultra-wideband spectrum for Case I (diodes switched 'ON'). However, for Case II (diodes-switched 'OFF') proposed UWB-MIMO antenna provides WLAN (5.15-

5.825 GHz) band notching. The DC block inline SMA module has been mounted with VNA so no onboard DC block capacitor is installed. The RF choke is added to all elements. The effect of decoupling structure on isolation of antenna elements placed side-by-side, for both Case I and II, can be observed clearly in Fig. 5. The overall isolation with decoupling structure is better than 20 dB.

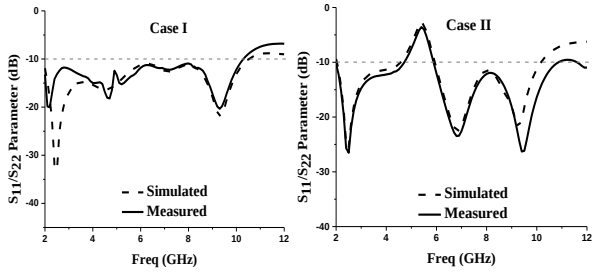


Fig. 4. Impedance matching of proposed prototype.

The effect of decoupling structure on isolation of antenna elements placed side-by-side, for both Case I and II, can be observed clearly in Fig. 5. The overall isolation with decoupling structure is better than 20 dB.

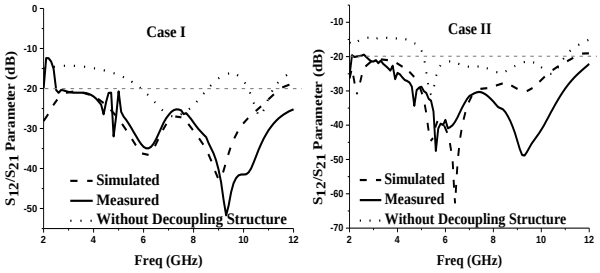


Fig. 5. Mutual coupling of proposed prototype.

Agilent N5242A PNA-X network analyzer is used for testing of proposed antenna design. Radiation patterns have been measured using Diamond Engineering setup in anechoic environment. The radiation patterns in E and H-plane of the proposed antenna (Case I and Case II) are observed at 5.5 GHz and 7 GHz and are plotted in Fig. 6. MIMO antenna system shows band notch characteristics for Case II. Therefore, in order to have good analysis of notched WLAN band, we observe gain characteristics of the proposed prototype. Figure 7 clearly depicts that the proposed antenna gives good gain results over whole UWB band except for the notched WiMAX band.

Channel Capacity Loss (CCL), Total Active Reflection Coefficient (TARC) and Envelope Correlation Coefficient (ECC) are important diversity parameters to analyze the performance of the proposed MIMO antenna. ECC can be calculated for different antenna topology (side-by-side) using equation (2):

$$\rho_e = \frac{\left|S_{ii}^{*}S_{ij} + S_{ji}^{*}S_{jj}\right|^2}{\left(1-|S_{ii}|^2+|S_{ji}|^2\right)\left(1-|S_{jj}|^2+|S_{ij}|^2\right)}. \quad (2)$$

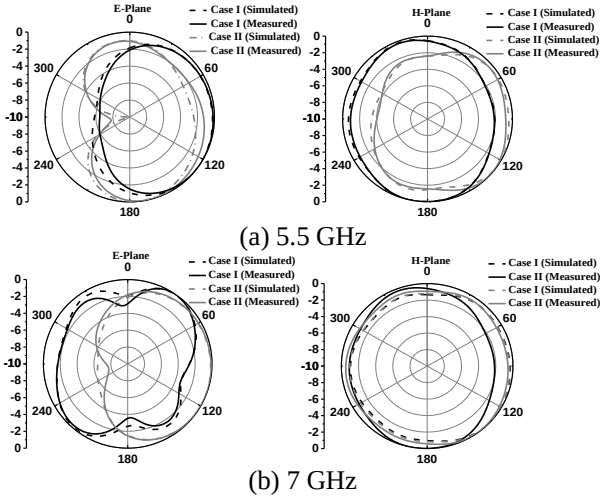


Fig. 6. Radiation plots: (a) 5.5 GHz and (b) 7 GHz.

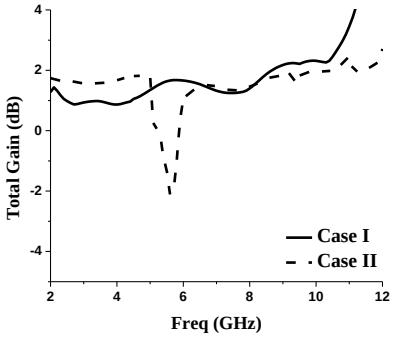


Fig. 7. Gain of proposed antenna.

Table 1: Optimized values of proposed antenna

Parameter	L	W	G_P	G_S	W_S	W_G
Value(mm)	40	23	19.8	8.5	6	13.5
Parameter	L_G	W_P	G_R	S_S	P_R	L_P
Value(mm)	6.25	4.5	2.45	2	7	1.65
Parameter	L_F	W_F	G_W	G_G	R_S	O_L
Value(mm)	21.4	1.6	3	1.6	4	0.5
Parameter	G_L	W_S	S_1	S_2	-	-
Value(mm)	6.7	0.5	4.4	1.2	-	-

Usually, an ECC value below -3 dB, TARC < 0 dB and CCL< 0.5 bits/sec/Hz in the operating band, is desired. As shown in Figs. 8 (a - c), the results are well within the allowable limits for both Cases (I and II). In Case II the CCL value at WLAN (5.15-5.825 GHz) band is above 0.5 bits/sec/Hz, because of the anti-resonant effect of band notching structure.

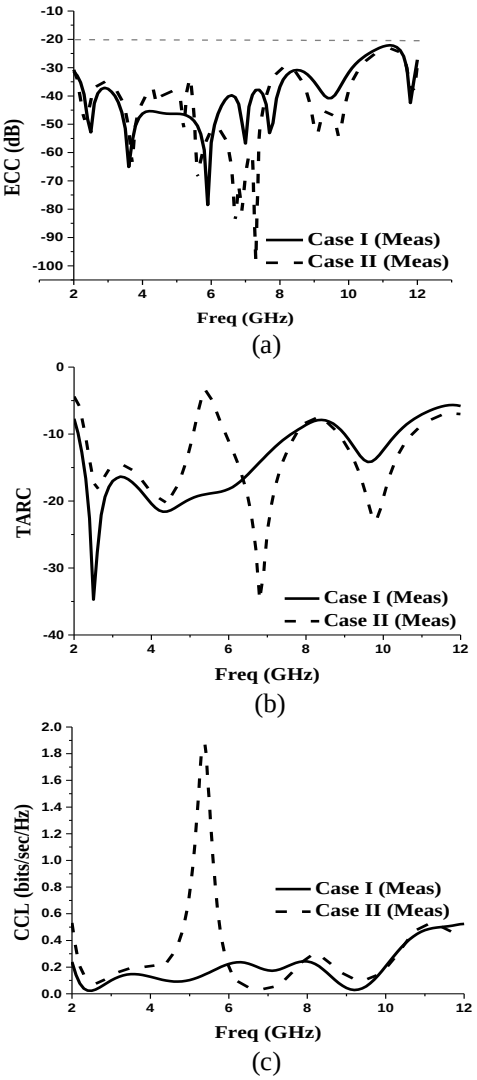


Fig. 8. MIMO/Diversity performance parameters: (a) ECC, (b) TARC, and (c) CCL.

Table 2: Comparison of proposed MIMO antenna with existing literature

Ref.	Isolation (dB)	No. of Ports	Antenna Dimension (mm^3)	Reconfigurability
[8]	> 15	2	50 x 82 x 1.6	No
[9]	> 15	2	40 x 30 x 0.8	No
[10]	> 20	2	35 x 35 x 1	No
Proposed antenna	> 20	2	40 x 23 x 1.6	Yes

IV. CONCLUSION

A miniaturized dual port MIMO antenna with electronically reconfigurable WLAN (5.5 GHz) band notch characteristics is proposed for ultra-wideband applications. The design operates over entire UWB band. Simple decoupling structure is used to achieve enhanced isolation. Diversity parameters are also within the allowed limits. More importantly simulated as well as measured results are in good agreement, suggesting that the proposed MIMO antenna design is suitable candidate for reconfigurable ultra-wideband application.

REFERENCES

[1] J. Y. Zhang, F. Zhang, W. P. Tian, and Y. L. Luo, "ACS-fed UWB-MIMO antenna with shared radiator," *Electronics Letter*, vol. 51, pp. 1301-1302, Aug. 2015.

[2] H. Huang, Y. Liu, and S. Gong, "Uniplanar differentially driven UWB polarisation diversity antenna with band-notched characteristics," *Electronics Letters*, vol. 51, pp. 206-207, Jan. 2015.

[3] S. I. Jafri, R. Saleem, M. F. Shafique, and A. K. Brown, "Compact reconfigurable multiple-input-multiple-output antenna for ultra-wideband applications," *IET Microwaves, Antennas & Propagation*, vol 10, pp. 413-419, Mar. 2016.

[4] M. S. Khan, A. D. Capobianco, S. Asif, A. Iftikhar, B. Ijaz, and B. D. Braaten, "Compact 4 × 4 UWB-MIMO antenna with WLAN band rejected operation," *Electronics Letters*, vol. 51, pp. 1048-1050, July 2015.

[5] L. Liu, S. W. Cheung, and T. I. Yuk, "Compact MIMO antenna for portable UWB applications with band-notched characteristic," *IEEE Transactions on, Antennas and Propagation*, vol. 63, pp. 1917-1924, May 2015.

[6] M. S. Khan, A. D. Capobianco, A. Naqvi, M. F. Shafique, B. Ijaz, and B. D. Braaten, "Compact planar UWB MIMO antenna with on-demand WLAN rejection," *Electronics Letters*, vol. 51, no. 13, pp. 963-964, June 2015.

[7] B. Badamchi, J. Nourinia, C. Ghobadi, and A. V. Shahmirzadi, "Design of compact reconfigurable ultra-wideband slot antenna with switchable single/ dual band notch functions," *IET Microwaves, Antennas & Propagation*, vol. 8, no. 8, pp. 541-548, June 2014.

[8] A. Toktas, "G-shaped band-notched ultra-wideband MIMO antenna system for mobile terminals," *IET Microwaves, Antennas & Propagation*, vol. 11, pp. 718-725, Dec. 2016.

[9] Y. Yin, J. Hong, C, Luo, and M. Amin, "A compact planar UWB MIMO antenna using modified ground stub structure," *IEICE Electronics Express*, vol. 14, no. 20, pp. 20170883-20170883, Oct. 2017.

[10] J. Zhu, S. Li, B. Feng, L. Deng, and S. Yin, "Compact dual polarized UWB quasi-self-complementary MIMO/diversity antenna with band-rejection capability," *IEEE Antennas and Wireless Propagation Letters*, vol. 15, pp. 905-908, Sept. 2015.

Circuit Modelling Methodology for Dual-band Planar Antennas

Kim Ho Yeap [1], Tobias Meister [2], Zi Xin Oh [1], and Humaira Nisar [1]

[1] Faculty of Engineering and Green Technology
Universiti Tunku Abdul Rahman, Jalan Universiti, Bandar Barat, 31900 Kampar, Perak, Malaysia
yeapkh@utar.edu.my, ohzixin@1utar.my, humaira@utar.edu.my

[2] Fakultät Angewandte Natur- und Kulturwissenschaften
Ostbayerische Technische Hochschule Regensburg, Postfach 12 03 27, 93025 Regensburg, Germany
tobias1.meister@st.oth-regensburg.de

Abstract ─ This paper presents a simple and systematic approach to determine the equivalent frequency-independent circuit model for a dual-band planar antenna. The Foster Canonical network synthesis technique with two RLC tanks has been employed to generate the two resonant bands of the antenna. The transfer function model is subsequently refined using a data fitting algorithm (viz the Nelder-Mead simplex algorithm). Parametric adjustments are performed at the final stage in order to further improve the accuracy of the final parameters.

Index Terms ─ Data-fitting algorithm, equivalent circuit, Foster canonical network, planar antenna, resonant bands.

I. INTRODUCTION

The continuous improvement in the design of planar printed antennas has contributed significantly to the technological advancement of wireless telecommunication. Prior to the mid-1990s, for instance, majority of the GSM handsets could only support single band operations [1]. In order to support more than one band, a few antennas are required to be installed into the device. Hence, the size of a multi-functional telecommunication device was usually bulky in those days. This is to say that; the functionalities of the device are very often traded for its size and portability. As dual-band antennas made their debut in 1996 [2 – 6], however, telecommunication devices experienced a dramatic reduction in sizes. With its ability to support multiple bands using a single antenna, much space has therefore been spared. Since then, researchers have been exploring ways to enhance the number of bands supported by an antenna; while, at the same time, making sure that its size is constrained within certain permissible dimensions [7 – 13].

When designing an antenna, it is important to derive its lumped equivalent circuit. This is because the circuit provides useful insights into the performance and operational principles of the antenna. It is also worthwhile noting that antennas are usually designed using electromagnetic field solvers, such as HFSS [14] or CST [15]; whereas, the receiver circuits are usually designed using electronic circuit simulators, such as SPICE [16]. Hence, the equivalent circuit of an antenna is useful when integrating the antenna with the front-end receiver circuits, during the validation phase. Validating both the antenna and receiver as a whole will certainly provide better accuracy to the actual performance of the system.

Antennas which generate only one resonant band can be easily modelled as a serial or parallel connection of resistance (R), inductance (L), and capacitance (C). When two or more bands are involved, however, the equivalent circuit turns out to be more complicated. In [17], Foster proposed modelling the resonant bands using *RLC* tanks, such as those shown in Fig. 1. The number of *RLC* tanks (represented by the subscript n) in the figure corresponds to the resonant bands that the circuit could generate. Thereafter, different methods, which adopt Foster's network synthesis technique as the cornerstone of their work, have been developed to derive the *RLC* equivalent circuit of antennas [18 – 22]. Since the Foster canonical network consists of various lumped passive elements, it is usually challenging to determine the exact combination of parameters which fits the return loss curve. As can be seen in the literature [18 – 22], the procedures applied to determine the parameters are usually quite laborious and mathematically involved.

In this paper, a simple and systematic approach to determine the equivalent circuit model for a dual-band planar antenna is presented. The derivation of the parameter for each element in the Foster canonical network is progressively described. We shall demonstrate that the return loss obtained from the approximate approach agrees reasonably well with that from an electromagnetic solver. In order to improve the accuracy

Submitted On: January 22, 2020
Accepted On: August 4, 2021

https://doi.org/10.47037/2021.ACES.J.360822

of the circuit, the parameters are refined using a data fitting algorithm and then parametrically fine-tuned at the final stage of the design.

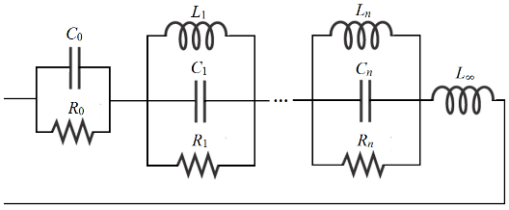


Fig. 1. General circuit topology of the Foster canonical network.

II. EQUIVALENT CIRCUIT TOPOLOGY

Besides the radiating resonant modes which are described by the *RLC* tanks in the Foster canonical network, quasi-static fields are present in an antenna as well. As can be seen in Fig. 1, these fields are characterized by the combination of resistance R_0, capacitance C_0 and inductance L_∞ in the circuit network. The resistance R_0 accounts for the conduction loss of the substrate material; whereas, the asymptotic behaviour of the feed point impedance at the higher and lower edges of the antenna frequency band is modelled by C_0 and L_∞ [23]. The quasi-static inductance L_∞ is usually infinitesimal when the order of the resonant modes is low. Hence, the circuit network for a dual-band antenna can be simplified into that in Fig. 2. The input impedance Z_{in} in the figure can be expressed as:

$$Z_{in} = \frac{1}{\frac{1}{R_0}+sC_0} + \frac{1}{\frac{1}{R_1}+\frac{1}{sL_1}+sC_1} + \frac{1}{\frac{1}{R_2}+\frac{1}{sL_2}+sC_2}, \quad (1)$$

where $s = j2\varpi f$ and f is the operating frequency.

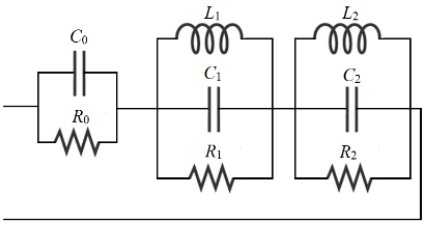


Fig. 2. The circuit topology for a dual-band antenna.

III. NUMERICAL EXAMPLE

To carefully outline the procedure in determining the lumped elements in the antenna, a dual-band monopole antenna designed using HFSS is used as an example. The planar antenna is connected to a microstrip feedline and is printed on an FR4-epoxy substrate which has a thickness of 1.6 mm, dielectric constant of 4.4 and loss tangent of 0.022. As can be seen from the layout in Fig. 3, the antenna is composed of two rectangular radiating patches. Since a monopole antenna exhibits optimal performance when its length is a quarter of its operating wavelength λ, the lengths of the top and bottom patches have been, respectively, kept to be within a quarter wavelength of the first and second resonant frequencies. As can be seen from the return loss curve in Fig. 4, the antenna resonates at f_{r1} = 2.1 GHz and f_{r2} = 5.2 GHz and it constitutes 15 dB bandwidths of 1.05 GHz (from 1.75 to 2.80 GHz) and 0.6 GHz (from 4.84 to 5.44 GHz) for the lower and upper resonant bands, respectively. Both resonant bands cover various IEEE 802.11 wireless local area network (WLAN) channels, which include those within the 2.4 GHz, 5.2 GHz, 4.9 GHz, 5.0 GHz, and 5.8 GHz bands.

Upon close inspection on the circuit topology in Fig. 2, it can be seen that the quasi-static resistance R_0 and capacitance C_0 can be readily evaluated from [19]:

$$R_0 = \lim_{f\to 0^+} Re\{Z_{in}(f)\}, \quad (2)$$

and

$$C_0 = \lim_{f\to 0^+} \frac{Im\{Y_{in}(f)\}}{2\pi f} = \frac{1}{2\pi}\lim_{f\to 0^+} \frac{\partial Im\{Y_{in}(f)\}}{\partial f}, \quad (3)$$

where Y_{in} is the input admittance of the circuit. Figures 5 and 6 depict, respectively, the simulated input resistance and susceptance of the antenna as a function of frequency. As f approaches 0, it can be seen that R_0 and C_0 can be approximated as 15 Ω and 0 F, respectively.

In comparison with the resistance, the magnitude of the susceptance turns out to be considerably smaller. This result suggests that both the non-static inductance (i.e., L_1 and L_2) and capacitances (i.e., C_1 and C_2) are much smaller than the resistance (i.e., R_1 and R_2). In order to estimate the parameters of the *RLC* tanks, values close to the 50 Ω characteristic impedance Z_0 are arbitrarily assigned to resistance R_1 (i.e., 30 Ω) and R_2 (i.e., 50 Ω); whereas, both C_1 and C_2 are arbitrarily set to 2 pF. The resonant frequency f_r of a simple parallel *RLC* network is given in (4) below:

$$f_r = \frac{1}{2\pi\times\sqrt{LC}}. \quad (4)$$

Inductance L_1 and L_2 can therefore be easily found by substituting the corresponding capacitance (i.e., C_1 and C_2) and resonant frequencies (i.e., f_{r1} and f_{r2}) into (4). The return loss *RL* of the circuit can be calculated by substituting (1) into (5) below:

$$RL = -20\log\left(\left|\frac{Z_{in}-Z_0}{Z_{in}+Z_0}\right|\right). \quad (5)$$

Figure 7 shows the comparison of the return loss curves computed from (5) and that obtained from the simulation. In order to investigate the effect of the curve at the presence of C_0, we have also plotted the return loss with C_0 arbitrarily chosen at 3 pF. Despite observing deviations among the curves, it is apparent that the theoretical results register certain resemblance with the simulation. In particular, the curve plotted based on the absence of C_0 are closer to that obtained from the simulation at the resonant frequencies f_r. To improve the accuracy of the parameters, a non-linear data fitting algorithm, such as the Nelder-Mead simplex optimization

routine, is adopted to minimize the mean-square error between both results. In this case, the approximate parameters for the lumped elements are used as the initial guesses for the search. Since eight parameters are involved in the optimization process, it is clearly not easy for the solution to converge to the required roots. Hence, it is often necessary to perform parametric adjustments on the parameters to further refine the accuracy of the results. The return loss curve, computed using the final parameters (shown in Table 1), is compared with the simulation result, as shown in Fig. 8. It is apparent from the figure that both curves are now in good agreement with each other. In fact, the *RLC* network resonates at the same frequencies and magnitudes with those generated from the planar antenna and both curves are highly correlated at the second band.

IV. CONCLUSION

A simple and systematic approach to evaluate the parameters in the Foster canonical network is presented here. The circuit network provides insight into the passive elements existing in a dual-band planar antenna. The approach starts with an approximation from the input resistance and susceptance of the antenna. The combination of parameters is subsequently refined so that the network could give a more realistic behaviour of the actual antenna.

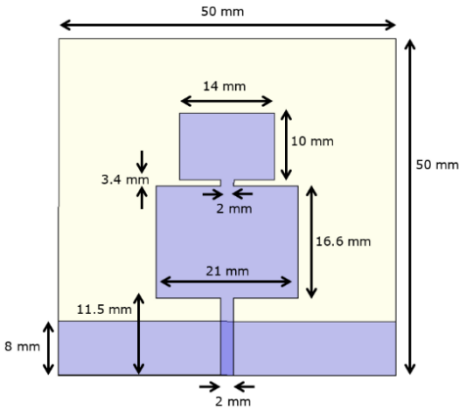


Fig. 3. The layout of a dual-band planar antenna.

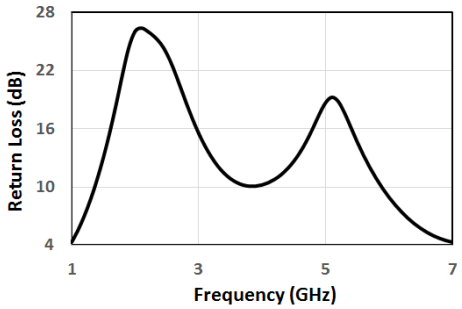


Fig. 4. The simulated return loss of a planar antenna.

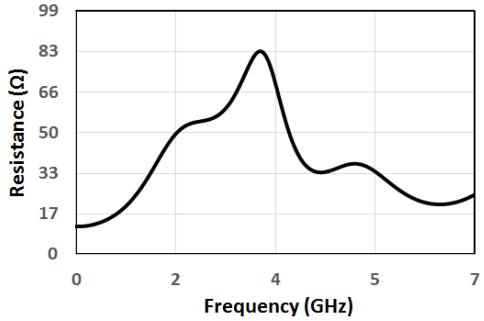


Fig. 5. The input resistance of a planar antenna.

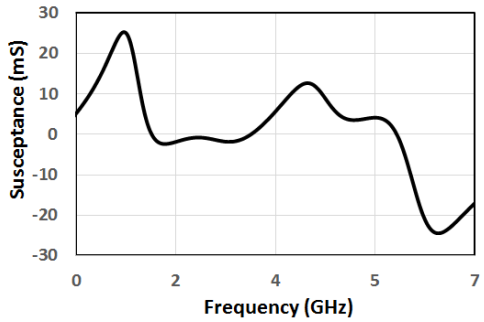


Fig. 6. The input susceptance of a planar antenna.

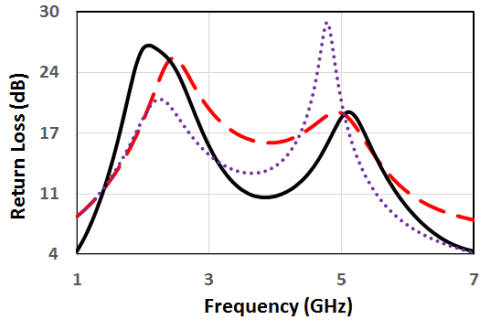


Fig. 7. The return loss obtained from the simulation (solid line) and the initial parameters with C_0 = 0 F (dashed line) and C_0 = 3 pF (dotted line).

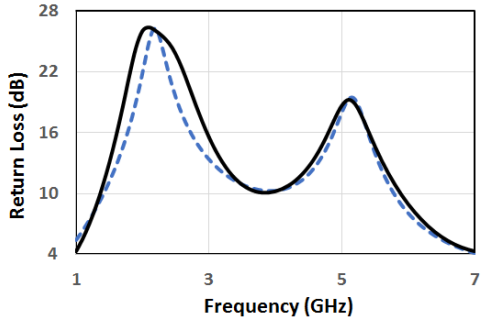


Fig. 8. The return loss obtained from the simulation (solid line) and final parameters (dotted line).

Table 1: Lumped elements in an antenna

Component	Value
R_0	1.3 Ω
C_0	0 F
R_1	54.1 Ω
C_1	1.36 pF
L_1	5.1 nH
R_2	66.5 Ω
C_2	1.66 pF
L_2	0.48 nH

REFERENCES

[1] C. Rowell and E. Y. Lam, "Mobile-phone antenna design," *IEEE Antennas and Propagat. Magazine*, vol. 54, no. 4, pp. 14-34, Aug. 2012.

[2] Z. Ying, "Progress of multiband antenna technology in mobile phone industry," *IEE Wideband and Multi-band Antennas and Arrays*, Birmingham, UK, Sep. 7, 2005.

[3] Z. Ying, "Multi-band non-uniform helical antennas," *1996-IO (granted first in USA)*, US006122102, WO-9815028.

[4] I. Egorov and Z. Ying, "A non-uniform helical antenna for dual band cellular phones," *IEEE Antennas and Propagat. Society Int. Symp.*, Salt Lake City, USA, July 16-21, 2000.

[5] P. Haapalg and P Vainikainen, "Helical antennas for multi-mode mobile phones," *26th European Microwave Conf.*, Prague, pp. 327-331, 1996.

[6] G. J. Hayes, "Dual band meander antenna," *US-0459553*, WO-WS08058, 1996.

[7] K. H. Yeap, C. S. Voon, T. Hiraguri, and H. Nisar, "A compact dual-band implantable antenna for medical telemetry," *Microwave and Opt. Technol. Lett.*, vol. 61, pp. 2105-2109, July 2019.

[8] C. S. Voon, K. H. Yeap, K. C. Lai, C. K. Seah, and H. Nisar, "A compact double-psi-shaped dual band patch antenna for WLAN/LTE applications," *Microwave and Opt. Technol. Lett.*, vol. 60, pp. 1271-1275, Apr. 2018.

[9] A. Altaf and M. Seo, "A tilted-D-shaped monopole antenna with wide dual-band dual-sense circular polarization," *IEEE Antennas and Wireless Propagat. Lett.*, vol. 17, no. 12, pp. 2464-2468, 2018.

[10] W. L. Yeo, K. C. Lai, K. H. Yeap, P. C. Teh, and H. Nisar, "A compact-dual band hook-shaped antenna for wireless applications," *Microwave and Opt. Technol. Lett.*, vol. 59, no. 8, pp. 1882-1887, Aug. 2017.

[11] L. Adam, M. N. M. Yasin, H. A. Rahim, P. J. Soh, and M. F. Abdulmalek, "A compact dual-band rectenna for ambient RF energy harvesting," *Microwave Opt. Technol. Lett.*, vol. 60, pp. 2740-2748, 2018.

[12] H. R. Cheong, K. H. Yeap, K. C. Lai, P. C. Teh, and H. Nisar, "A compact CPW-fed antenna with S-shaped patches for multiband applications," *Microwave and Opt. Technol. Lett.*, vol. 59, no. 3, pp. 541-546, Mar. 2017.

[13] K. H. Yeap, W. L. Yeo, K. C. Lai, T. Hiraguri, K. Hirasawa, and Z. X. Oh, "A compact E-shaped antenna with C-shaped slots and a back-patch for multiband applications," *J. of Electrical Engineering*, vol. 71, no. 1, pp. 49-54, Feb. 2020.

[14] High Frequency Structure Simulation (HFSS), Ansys, Inc., Canonsburg, PA.

[15] CST Studio Suite, Dassault Systèmes Simulia Corp., Providence, Rhode Island.

[16] Simulation Program with Integrated Circuit Emphasis (SPICE), University of California, Berkeley, California.

[17] R. M. Foster, "A reactance theorem," *Bell Syst. Tech. J.*, vol. 3, no. 2, pp. 259-267, Apr. 1924.

[18] K. H. Sayidmarie and L. S. Yahya, "Modeling of dual-band cresecent-shape monopole antenna for WLAN applications," *Int. J. of Electromagnetics and Applications*, vol. 4, no. 2, pp. 31-39, 2014.

[19] D. Caratelli, N. Haider, and A. Yorovoy, "Analytically based extraction of Foster-like frequency-independent antenna equivalent circuits," *Proc. of the Int. Symp. on Electromagnetic Theory*, Hiroshima, Japan, pp. 506-509, May 20-24, 2013.

[20] O. K. Heong, C. K. Chakrabarty, and G. C. Hock, "Circuit modeling for rectangular printed disc monopole antenna with slot for UWB system," *Proc. of the 3rd Int. Conf. on Intelligent Sys. Modelling and Simulation*, Kota Kinabalu, Malaysia, pp. 727-731, Feb. 8-10, 2012.

[21] S. B. T. Wang, A. M. Niknejad, and R. W. Brodersen, "Circuit modeling methodology for UWB omnidirectional small antennas," *IEEE J. on Selected Areas in Communications*, vol. 24, no. 4, pp. 871-877, Apr. 2006.

[22] T. Ali, K. D. Prasad, and R. C. Biradar, "A miniaturized slotted multiband antenna for wireless applications," *J. of Computational Electronics*, vol. 17, pp. 1056-1070, 2018.

[23] M. Ansarizadeh, A. Ghorbani, and R. A. Abd-Alhameed, "An approach to equivalent circuit modeling of rectangular microstrip antennas," *Progress in Electromagnetics Research B*, vol. 8, pp. 77-86, 2008.

Novel Methodology to Assess RF Performance of Co-located MIMO Radar Systems Transmitting Binary Phased Coded Waveforms

Nivia Colon-Diaz [1,2], Dan Janning [1], Patrick M. McCormick [1], James T. Aberle [2], and Daniel W. Bliss [2]

[1] Sensors Directorate, Air Force Research Laboratory, Wright-Patterson Air Force Base, OH
nivia.colon_diaz.3@us.af.mil

[2] Department of Electrical, Computer, and Energy Engineering, Arizona State University, Tempe, AZ

Abstract ─ In this work an experimental radar testbed and dual directional couplers (DDC) are used to collect measurements of the forward and reverse waves at each element, its mutual coupling, and beam-patterns. The collected measured data is used to validate a methodology for assessing radio frequency (RF) performance of a co-located multiple-input multiple-output (MIMO) radar system transmitting binary phased coded (BP) waveforms. The estimated and measured active reflection coefficient (ARC) and beam-patterns are presented. The effect of different excitations and mutual coupling on the radiated fields is also presented.

Index Terms ─ Active reflection coefficient, co-located MIMO radar, coupled power, mutual coupling.

I. INTRODUCTION

Radio Frequency (RF) sensors are maturing into fully digital systems characterized by a wide-band frequency range and capable of performing multiple tasks simultaneously on different elements or sub-arrays. These RF sensors are also evolving towards a capability to execute simultaneous functions such as electronic warfare, communications, and radar tasks such as search, detect, track and image a target. Early digital signal processing research on these type of sensors [1] ignored electromagnetic effects, such as interactions between elements or mutual coupling.

As commercial communication and defense systems are consuming more spectrum [2], multi-functional operation demands techniques to alleviate spectral congestion and improve the efficiency by sharing the spectrum. Multiple-input multiple-output (MIMO) emissions enable multi-functionality via simultaneous transmission of independently modulated waveforms in the same band from different radiators or sub-arrays. While MIMO techniques could be the key approach to accommodate multi-function operations, the concept demands careful understanding of how each function, subsystem, and/or sub-array, interacts with one another while ensuring enough isolation among them [3]. Previous work developed by the authors [4-5] presents a way to assess mutual coupling effects by analyzing the active reflection coefficient (ARC) and active voltage standing wave ratio (VSWR) and quantifying their impact on beam-patterns.

The benefits of improving our understanding of mutual coupling effects on MIMO radar systems are two-fold. First, the knowledge can enhance MIMO radar calibration techniques and provide key information to system designers [6]. Second, MIMO radar waveform design can take advantage of new discoveries to use the waveforms designed specifically for transmit mode and improve signal-to-noise ratio performance [7], beam-pattern and side-lobe levels [8] and other metrics.

In this paper, we present beam-pattern and ARC measurements obtained using U.S. Air Force Research Laboratory (AFRL) Baseband-digital at Every Element MIMO Experimental Radar (BEEMER) for a uniform linear array (ULA) of 6 patch antennas linearly placed with a separation of 0.5λ and transmitting binary phased coded (BP) waveforms. Even though the methodology can be applied to any transmit waveform, this initial study uses BP waveforms. The objectives of this paper are to (1) validate the methodology developed previously by directly measuring the ARC and the beam-patterns on co-located MIMO radars and (2) analyze the effects that mutual coupling and excitations have in the radiated fields on a MIMO radar transmission.

II. MUTUAL COUPLING ANALYSIS

Ideally, a signal radiated by the *m*-th element of a ULA would only radiate from the phase center of the *m*-th element without affecting adjacent elements. In reality, real antenna arrays experience mutual coupling in which the signal radiated by the *m*-th element will couple into neighboring elements. This has two consequences: the signal transmitted by the *m*-th element

Submitted On: February 19, 2020
Accepted On: January 20, 2021

https://doi.org/10.47037/2021.ACES.J.360823

1) can be observed by a receiver connected to the *n*-th element and 2) can be re-radiated from the phase center of the *n*-th element at a reduced power and coupling which will alter the array's beam-pattern.

Direct measurements of the ARC were collected using dual directional couplers (DDCs) (Hewlett Packard 777D 1.9-4GHz with -20 dB coupling factor), as described in [5]. A DDC can be equivalently modeled as two directional couplers with their isolated ports terminated as described in [9]. The majority of the signal fed to the input port will be transmitted to the output port with a portion of the signal coupled to the forward port. A signal entering through the output port will couple into the reverse port. The calibrated ratio between the reverse and forward waves can be used to compute the ARC via:

$$\Gamma_n^a = \frac{b_n}{a_n} = \frac{\sum S_{nm}\, a_m}{a_n}, \qquad (1)$$

where Γ_n^a is the ARC observed in antenna *n* due to the other *m* elements, a_n represents the amplitude and phase of the forward wave on element *n* (defined as the excitation of antenna *n*), and b_n represents the amplitude and phase of the reverse wave in antenna *n*. S_{nm} is the conventional scattering (S) parameter defined as the passive ratio of the signal coupled to port *n* from a signal incident on port *m*, when all other ports are terminated in the system impedance [10]. A block diagram showing the main components required to perform this measurement is given in Fig. 1. The DDCs were connected directly to the 6 antenna patch elements of the ULA, the other 4 antenna elements were matched terminated, as shown in Fig. 1. A description of the single multilayer patch antenna can be found in [11]. The coupled ports of each DDC were then used to measure the signals sent to and reversed from each radiator. To collect ARC measurements, BEEMER was configured to provide 6 transmit and 12 receive channels (6 forward coupled and 6 reversed coupled) so that all measurements could be collected simultaneously.

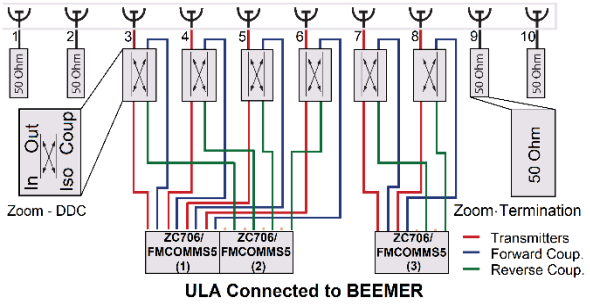


Fig. 1. BEEMER and DDC connections to directly collect ARC measurements using the FPGA ZC706 board and the 4-channel FMCOMMS5 transceiver.

The S-matrix of a 6x1 linear sub-array of horizontally polarized radiators formed from a 10x6 dual polarized patch array was measured with the unused elements terminated. An Agilent Technologies M9375A vector network analyzer with 16 channels was used to measure the S-parameters at 3.5 GHz. The measured S-parameters and Equation (1) were used to obtain the estimated ARC [12]. The estimated ARC is compared to the directly measured ARC using the DDCs following the method developed in [5] are presented in Table 1. Good agreement between estimates and measurements of the ARC can be seen in Table 1.

The BEEMER system is constructed from commercial off the shelf components [13]. The principal components are a ZC706 field-programmable gate array (FPGA) board and an FMCOMMS5 4-channel transceiver. The system has the ability to synchronize multiple ZC706's/FMCOMMS5's to provide ≥ 4 transceiver channels. Arbitrary waveforms can be sent on each channel, which allows MIMO transmission to be studied. BEEMER's maximum sampling rate is 60 Mega samples per second, its maximum transmitted signal length is 4096 samples, the center frequency selected is 3.5 GHz and the transmitted power is -8 dBm.

Table 1: Estimated and measured ARC on each channel

Channel	Estimated [dB]	Measured [dB]
1	–17.8	-17.8
2	–15.9	-15.0
3	–17.2	-19.4
4	–15.0	-16.5
5	–15.6	-16.6
6	–16.8	-18.8

To obtain beam-pattern data, BEEMER was triggered to transmit on 6 channels using BP waveforms and to record the signal observed. The waveforms were 1024 samples long. For each azimuth angle (-90 to 90), BEEMER transmitted simultaneously on each channel and the waveform on each channel was repeated 256 times at each angle. BEEMER received 2000 x 16 x 256 matrix of data at each azimuth location (where the dimensions are fast time x channel x slow time). Measurements were collected in the anechoic chamber seen in Fig. 2.

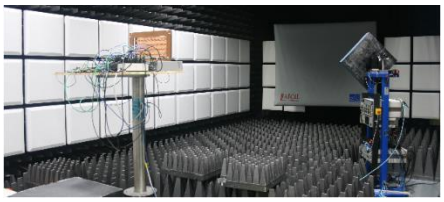


Fig. 2. BEEMER configuration in anechoic chamber.

In typical radar applications, the data collected is saved in a data cube with three dimensions: fast time (range), slow time (pulse to pulse), and channels.

Channels 1-6 collected reverse coupling, channels 7-12 recorded forward coupling, channel 13 was connected to the feed antenna (via a long SMA coaxial cable and an amplifier was placed just behind the feed antenna to compensate cables losses while minimizing the noise figure), and channels 14-16 were match terminated. The matrix was summed along the slow time dimension to coherently integrate the data while pulse compression was applied to the fast time dimension (to reduce the 2000 samples to a single sample per azimuth angle).

III. ANALYSIS ON RADIATED FIELDS

A MIMO transmission has a time-varying amplitude/phase array excitation from the superposition of independently generated waveform modulations. A MIMO radar illuminates the entire surveillance volume with each pulse, rather than sweeping a beam through the search area as a traditional phased array might. For this analysis, the fields of a MIMO transmission are characterized for a single time instant. Because each element radiates a different waveform, mutual coupling must be considered.

An approximation to the total radiated fields of a MIMO transmission, considering mutual coupling and different excitations can be represented in Equation (2) [14]. A special case of this equation is when mutual coupling is ignored ($\Gamma_n^a = 0$). This special case is known as the pattern multiplication concept and presents the approximation to the radiated fields used by early researchers. The pattern multiplication concept assumes that the element current distribution does not vary from element to element:

$$E(\theta,\emptyset) \approx f^i(\theta,\emptyset)\sum_{n=1}^{N}(1-\Gamma_n^a)a_n e^{jk(n-1)(x_n u+y_n v)}, \quad (2)$$

where $f^i(\theta,\emptyset)$ is the isolated element pattern, N is the number of elements in the array, with $u = \sin\theta\cos\emptyset$ and $v = \sin\theta\sin\emptyset$, coupling effects are captured with Γ_n^a (the ARC), the different excitations are considered with a_n, the $e^{jk(x_n u+y_n v)}$ accounts for element positions, geometrical configuration, and inter-element spacing between elements, with (x_n, y_n) the coordinates of each element on the *xy*-plane. The radiated fields of an isolated antenna element pattern were modeled using a finite element-boundary integral computational electromagnetic tool SENTRi, as shown in Fig. 3.

Figures 4-5 present predicted beam-patterns in the (u, v) space in normalized dB, using Equation (2) with the isolated patch antenna pattern $f^i(\theta,\emptyset)$ from SENTRi, as seen in Fig. 3. In Fig. 4 the fields are generated with $\Gamma_n^a = 0$ and unitary excitations ($a_n = 1$), showing that the far zone field follows the multiplication of the isolated element pattern with the array factor neglecting coupling. Figure 5 displays the fields with $a_n = [e^{j\pi}, e^{j\pi}, e^{j0}, e^{j0}, e^{j\pi}, e^{j\pi}]$ and $\Gamma_n^a \neq 0$, calculated using Equation (1). Comparing Figs. 4 and 5, one can observe the array directvity is affected by coupling and the selected excitation. These parameters must be considered in the design of the system and its waveforms.

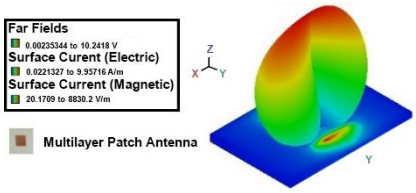


Fig. 3. Predicted fields of multilayer patch antenna.

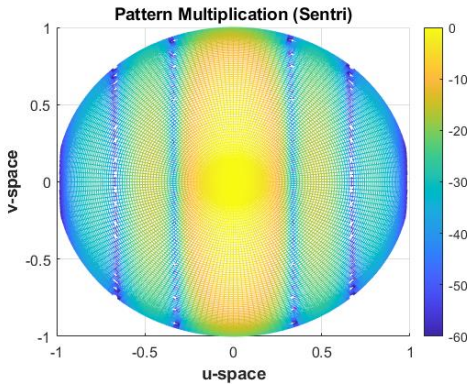


Fig. 4. Predicted radiated fields of a MIMO radar using pattern multiplication concept ($\Gamma_n^a = 0$).

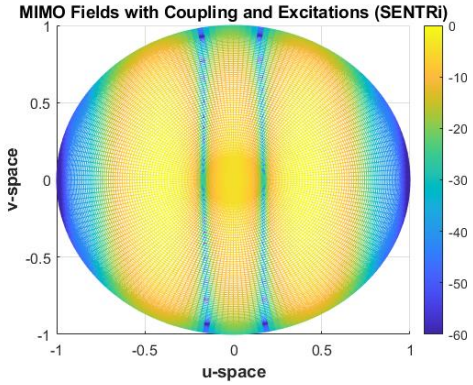


Fig. 5. Predicted fields of a MIMO radar considering coupling and binary phased excitations.

Figures 6-7 present predicted and measured fields at an instant in time. The blue-'x' and blue-solid curves depict the measured and predicted fields, respectively, when transmitting uniform excitations ($a_n = 1$). The predicted fields were obtained as previously described for Figs. 4-5 with $\Gamma_n^a \neq 0$. The measured fields were obtained using the BEEMER testbed through channel 13 as explained in Section II. The black curves on Fig. 6 present the predicted (black-dotted) and measured (black-squared) fields with excitations from Case 1 with $a_n = [e^{j0}, e^{j0}, e^{j\pi}, e^{j\pi}, e^{j\pi}, e^{j\pi}]$. The black curves

on Fig. 7 show predicted and measured fields with $a_n = [e^{j\pi}, e^{j\pi}, e^{j\pi}, e^{j0}, e^{j0}, e^{j0}]$ from Case 2. Predicted and measured fields follow the same trend and validate the methodologies used in this work. One can observe in Figs. 6-7 that the side-lobe levels and the location of the nulls vary depending on the coupling and controlled through the excitation.

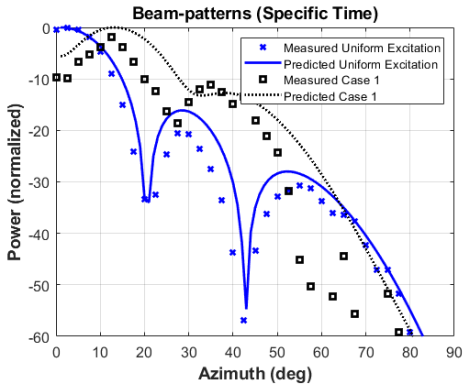


Fig. 6. Measured and predicted fields for a time instant, with uniform and Case 1 excitations.

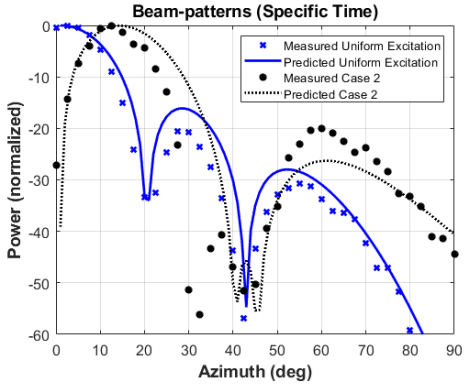


Fig. 7. Measured and predicted fields for a time instant, with uniform and Case 2 excitations.

IV. CONCLUSIONS

A methodology of co-located MIMO radars for estimating and directly measuring the ARC by using a DDC in each radiator is validated. These are the first direct measurements of reverse and forward coupled waves for co-located MIMO radars. Understanding reverse waves ensures the safety of the system and the prevention of large power reflections. This becomes important in high power applications where the reversed energy can be large enough to damage the transmitter. The impact of coupling and excitations on the beam-pattern, at an instant in time, can assist in identifying isolated effects. Figures 4-7 demonstrate that mutual coupling and the excitations of the elements play an important role in the performance of the radiation pattern and the directivity of a co-located MIMO radar system. These parameters must be considered in the design of the system and its waveforms.

REFERENCES

[1] H. V. *Trees, Optimum Array Processing (Detection, Estimation, and Modulation Theory,* Part IV), 1st ed., Wiley-Interscience, New York, 2002.

[2] A. R. Chiriyath, S. Ragi, H. D. Mittelmann, and D. W. Bliss, "Novel radar waveform optimization for a cooperative radar-communications system," *IEEE Trans. on Aerospace and Electr. Syst.*, pp. 1-1, 2019.

[3] C. I. Coman, I. E. Lager, and L. P. Ligthart, "Multifunction antennas - the interleaved sparse sub-arrays approach," *2006 European Radar Conference*, Sept. 2006.

[4] N. Colon-Diaz, J. G. Metcalf, D. Janning, and B. Himed, "Mutual coupling analysis for colocated MIMO radar applications using CEM modeling," *IEEE Radar Conf.*, May 2017.

[5] N. Colon-Diaz, D. Janning, T. Corigliano, L. Wang, and J. Aberle, "Measurement of active reflection coefficient for co-located MIMO radar using dual directional couplers," *2018 AMTA Proc.*, Nov. 2018.

[6] C. M. Schmid, C. Pfeffer, R. Feger, and A. Stelzer, "An FMCW MIMO radar calibration and mutual coupling compensation approach," *2013 European Radar Conference*, Oct. 2013.

[7] B. T. Arnold and M. A. Jensen, "The effect of antenna mutual coupling on MIMO radar system performance," *IEEE Transactions on Antennas and Prop.*, vol. 67, no. 3, pp. 1410-1416, Mar. 2019.

[8] P. M. McCormick, S. D. Blunt, and J. G. Metcalf, "Wideband MIMO frequency modulated emission design with space-frequency nulling," *IEEE J. Sel. Top Sig. Pro.*, vol. 11, no. 2, pp. 363-378, Mar. 2017.

[9] J. Jorgesen and C. Marki, "Directivity and VSWR Measurements, Understanding Return Loss Measurements," *Marki Microwave*, 2012.

[10] D. M. Pozar, *Microwave Engineering*. 4 ed., Wiley, Hoboken, 2011.

[11] B. Baumen, A. Christianson, A. J. Wegener, and W. J. Chappell, "Dynamic visualization of antenna patterns and phased-array beamsteering," *IEEE Antennas and Propagation Magazine*, vol. 54, no. 3, pp. 184-198, 2012.

[12] C. Zhang, Q. Lai, and C. Gao, "Measurement of active s-parameters on array antenna using directional couplers," *2017 IEEE Asia Pacific Microwave Conference (APMC)*, Nov. 2017.

[13] T. C. Mealey and A. J. Duly, "BEEMER a firmware-tuned, software defined MIMO radar testbed," *2016 IEEE International Symposium on Phased Array Systems and Technology*, Oct. 2016.

[14] L. Savy and M. Lesturgie, "Coupling effects in MIMO phased array," *2016 IEEE Radar Conf.*, 2016.

Model Order Reduction of Cardiac Monodomain Model using Deep Autoencoder Based Neural Networks

Riasat Khan [1] and Kwong T. Ng [2]

[1] Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh
riasat.khan@northsouth.edu

[2] Department of Electrical and Computer Engineering, New Mexico State University, Las Cruces, NM 88003, USA
ngnsr@nmsu.edu

Abstract ─ The numerical study of electrocardiology involves prohibitive computational costs because of its complex and nonlinear dynamics. In this paper, a low-dimensional model of the cardiac monodomain formulation has been developed by using the deep learning method. The restricted Boltzmann machine and deep autoencoder machine learning techniques have been used to approximate the cardiac tissue's full order dynamics. The proposed reduced order modeling begins with the development of the low-dimensional representations of the original system by implementing the neural networks from the numerical simulations of the full order monodomain system. Consequently, the reduced order representations have been utilized to construct the lower-dimensional model, and finally, it has been reconstructed back to the original system. Numerical results show that, the proposed deep learning MOR framework gained computational efficiency by a factor of 85 with acceptable accuracy. This paper compares the accuracy of the deep learning based model order reduction method with the two different techniques of model reduction: proper orthogonal decomposition (POD) and dynamic mode decomposition (DMD). The model order reduction deploying the deep learning method outperforms the POD and DMD concerning the modeling accuracy.

Index Term ─ Autoencoder, Cardiac monodomain model, deep learning technique, dynamic mode decomposition, proper orthogonal decomposition, reduced order modeling, semi-implicit scheme.

I. INTRODUCTION

Cardiac electrophysiology is a bioelectromagnetic phenomenon, where the electrical activities of the heart tissue are studied. Electrocardiological numerical simulations seek reliable and efficient mathematical models for cellular membrane dynamics. The monodomain equation used to model the cardiac electrical activity requires the solution of a nonlinear partial differential equation with appropriate boundary and initial conditions [1]. The monodomain equation leads to a complex dynamical system because of the involvement of the nonlinearity of different ionic currents and steep wavefront propagation. The solution of the monodomain equation can be obtained with various numerical techniques, e.g., finite volume method (FVM), finite difference method (FDM), and finite element method (FEM). The discretization of the monodomain equation with the finite difference method involves a large number of degrees of freedom [2]. In the literature, reduced order modeling has been studied to approximate the nonlinear dynamics of the complex systems. The most widely used order reduction method is the proper orthogonal decomposition (POD) technique, which captures the characteristic dynamics of the original discretized system [3]. POD is a projection-based model order reduction method, where a lower order basis is computed first from the original system. Finally, the reduced order model is obtained from the full order model's projection onto this small dimensional POD basis. The approximation of the obtained reduced order model requires considerably less computational complexity, with insignificant compromise on the accuracy. Different model order reduction approaches have been successfully applied in the field of cardiac electrophysiology, such as POD [4] and dynamic mode decomposition method (DMD) [5].

The neural network is a machine learning algorithm, loosely modeled from the human brain, which has been designed to detect a similar pattern in the data. Many machine learning algorithms have been successfully used in the field of dimensionality reduction to reduce overfitting, data preprocessing, and simpler data visualization [6]. Autoencoder, also known as auto-associative neural network, with three delicate layers, is an unsupervised learning algorithm. It is a neural network that learns the original system with multiple levels of representations and can predict the nonlinear dynamics of the data [7]. Autoencoder is primarily

Submitted On: July 2, 2020
Accepted On: August 4, 2021

https://doi.org/10.47037/2021.ACES.J.360824

composed of two networks, an encoder, and a decoder. The encoder is an analysis network which learns to detect all the most significant characteristics and hidden representations of the input high fidelity system. The encoder has a representation of latent-space, which contains the compressed essential properties of the input nodes [8]. The decoder is a synthesis network that recomposes the exact system from the hidden representations with minimum reconstruction error. Autoencoder with single perceptron and linear activation in the projection and reconstruction phases works almost similar to the state-of-the-art dimensionality reduction method principal component analysis (PCA).

In this work, the dimension or order of the transmembrane potential of the cardiac monodomain system has been effectively reduced with the implementation of a deep autoencoder based deep learning approach. As opposed to a traditional autoencoder, it consists of additional deep-belief layers and consequently learns more complex features. The restricted Boltzmann machine (RBM), which is the elementary element of the deep-belief networks, constructs the deep autoencoder architecture's complex layers. In the first step, the high dimensional spatiotemporal dynamics of the monodomain model has been minimized to lower order representations. Next, the lower order representations have been approximated to a reduced order dynamic model, which can replicate the dynamics of the original system. Finally, the approximated solution is reconstructed from the prediction of the lower order model. The computational efficiency and modeling accuracy of the proposed order reduction approach are examined and compared with the popular proper orthogonal decomposition and dynamic mode decomposition methods.

II. FULL ORDER CARDIAC MONODOMAIN GOVERNING EQUATION

This paper considers the nonlinear partial differential monodomain equation describing the dynamics of cardiac electric transmembrane potential $\bar{V} = V(\bar{x}, t)$, which can be formulated as:

$$\begin{cases} \frac{\delta \bar{V}}{\delta t} = \frac{1}{C_m}\left\{\frac{1}{\beta}[\nabla \cdot (\bar{\bar{\sigma}}_i \nabla \bar{V}) + I_{si}] - \sum I_{ion}(\bar{V}, w)\right\} \\ \frac{dw}{dt} + g(\bar{V}, w) = 0, \end{cases} \quad (1)$$

with Neumann boundary conditions and initial condition $V(\bar{x}, 0) = V_0(\bar{x})$ and $w(\bar{x}, 0) = w_0(\bar{x})$. The equation is considered in a spatial domain of $\Omega \in \mathbb{R}^3$, which is the considered section of the myocardium, and a temporal domain of $t \in [0, T]$. Here, C_m denotes the membrane capacitance per unit area, and β is the ratio of the membrane surface area to volume. $\bar{\bar{\sigma}}_i$ is the intracellular anisotropic conductivity tensor, which changes continuously with the fiber angle rotation, and I_{si} represents the intracellular source current, which initiates the stimulation.

The reaction element in the governing equation is the ionic current term, $\sum I_{ion}(\bar{V}, w)$, which has a nonlinear relation with the transmembrane potential and the gating parameters $w(\bar{x}, t)$. The Luo-Rudy model has been applied to obtain the ionic current, which yields the solution of eight coupled nonlinear ordinary differential equations. The finite difference method has been used to discretize the spatial derivatives of the Laplacian term of (1) [9]. The temporal discretization has been achieved by implementing the semi-implicit method [10], which leads to:

$$\underbrace{\{[I] - \frac{\theta \Delta t}{C_m \beta}[D_o]\}}_{[A]} \underbrace{\bar{V}^{n+1}}_{\bar{x}} = \underbrace{\bar{V}^n + \frac{(1-\theta)\Delta t}{C_m \beta}[D_o]\bar{V}^n + \Delta t[\frac{1}{C_m \beta}\bar{I}_{si}^{n+1} - \frac{1}{C_m}\overline{\sum I}_{ion}^{n}]}_{\bar{b}}, \quad (2)$$

where θ = 0.5 represents the Crank-Nicolson semi-implicit temporal parameter. The two major excessive computational components at each time step involve the solution of the matrix equation (2) and evaluation of the ionic current, which includes the calculation of the gating parameters and the solution of the ODEs.

III. REDUCED ORDER MODELLING WITH DEEP LEARNING APPROACH

Considering only spatial approximation, the governing nonlinear cardiac monodomain equation transforms into a system of ordinary differential equation in the time domain as:

$$\frac{d\bar{V}}{dt} = \bar{f}(\bar{V}, t) \quad (3)$$

where $\bar{f}$ is a set of nonlinear functions. The snapshots of the original spatiotemporal dynamic system have been obtained from its numerical simulations using (2). The snapshots are obtained from the finite difference solution of the full order monodomain system at the N spatial locations $x_1, \ldots, x_N$, and M temporal instances $t_1, \ldots, t_M \in [0, T]$. The ultimate goal of this work is to derive a lower order solution V_r from the high fidelity snapshot matrix of the transmembrane potential solution $\{V(x_i, t)_{i=1,t=1}^{N,M}\}$.

First, a lower order representation V_1 can be derived from the nonlinear projection $f_1: \mathbb{R}^N \rightarrow \mathbb{R}^{N_r}$ as:

$$V_1 = f_1(V), \quad (4)$$

where N_r is the reduced dimension with $N_r \ll N$. Next, the reduced order dynamic modeling of the lower dimensional representations is completed. As the dimension has been reduced, the reduced order modeling involves fewer computational resources than the original full order model. After applying the lower dimensional dynamic modeling, the reduced order form of (3) will be $\frac{d\bar{V}_r}{dt} = \bar{f}_r(\bar{V}_1, t)$, where V_r is the predicted solution of the cardiac transmembrane potential. Finally, the reduced order solution is reconstructed back to the approximated

solution by implementation the nonlinear reconstruction function $f_2: \mathbb{R}^{N_r} \rightarrow \mathbb{R}^N$ as:

$$\tilde{V} = f_2(V_r). \quad (5)$$

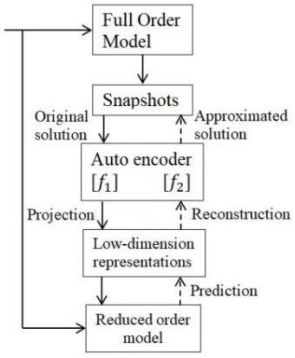


Fig. 1. Neural network based model order reduction architecture.

The modeling accuracy of the proposed order reduction method will be evaluated by minimizing the RMS error between the original solution V and approximated solution $\tilde{V}$ of the cardiac transmembrane potential. During the training stages of the deep autoencoder architecture, the RMS error of the following cost function is minimized:

$$f(V, \tilde{V}) = \sqrt{\frac{\sum_1^M (V-\tilde{V})^2}{M}}. \quad (6)$$

Here M denotes the total number of the training dataset. The projection and reconstruction should be performed simultaneously for a good approximation of the original spatiotemporal dynamics.

Figure 1 shows the architecture of the proposed neural network based model order reduction method. In Fig. 1, the hidden layers of the autoencoder are composed of the projection layer f_1, intermediate bottleneck layer for reduced order modeling, and the reconstruction layer f_2. Nonlinear sigmoid activation functions have been used for the nodes of the projection and reconstruction layers, as:

$$\begin{aligned} f_1(V) &= \sigma_n(W_n(\sigma_1(W_1 V + b_1) + \cdots + b_n)) \\ f_2(V_r) &= \sigma_1(W_1'(\sigma_n(W_n' V_r + b_n) + \cdots + b_1)), \end{aligned} \quad (7)$$

where σ_i is the sigmoid function of i^{th} layer, W_i and b_i denote the weights and bias between the layers i and $i+1$.

III. RESULTS AND DISCUSSION

In this section, numerical performances of the proposed deep learning based order reduction approach on the cardiac monodomain equation will be presented. A three-dimensional myocardial tissue of 0.5 cm × 0.1667 cm × 0.1667 cm has been considered as the computational domain. The longitudinal (x) and transverse (y and z) conductivities to the fiber have been assigned as: σ_{il} = 0.174 S/m and σ_{it} = 0.0193 S/m [9]. The tissue was stimulated by a point current source at one corner of $I_{si} = 500\beta$ with a duration of 1 ms and $\beta = 2000\ cm^{-1}$. The first order semi-implicit method (θ = 0.50) and the second order central finite difference technique have been utilized for the temporal and spatial discretizations, respectively.

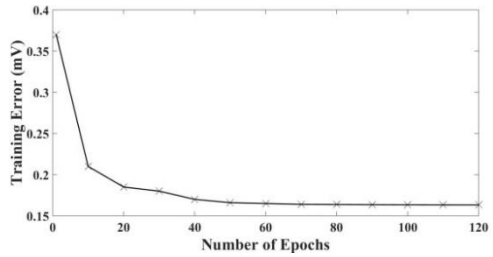


Fig. 2. RMS error (mV) vs the number of epochs during the backpropagation process.

The proposed model order reduction method initiates with the construction of the snapshots from the above-mentioned full order monodomain simulations with a time step size of Δt = 0.01 ms and for a time domain of $t \in$ [0, 20 ms]. The deep autoencoder architecture comprises two symmetrical deep-belief networks for encoding and decoding functions, with 8 RBM layers and 40 neurons in each of the hidden layers. 75% of the full order simulation data at time domain $t \in$ [0, 15 ms] was used for training, and the remaining data ($t \in$ [15, 20 ms]) was utilized for the testing of the proposed algorithm, and the training of the model has been performed on the Google Collaboratory. The backpropagation training error in Fig. 2 confirms the convergence of the deep learning process. Next, the number of reduced order modes is varied to study the efficiency of the MOR method. The RMS error decreases by almost 50% with the increase of the reduced order modes from 100 to 180. Figure 3 demonstrates that the modeling accuracy does not improve significantly if the modes number is increased after that, and hence the reduced order modes number is set to 180.

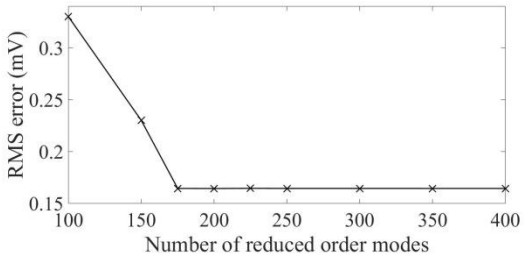


Fig. 3. RMS error (mV) for different number of reduced order modes.

Finally, the predicted solution is reconstructed back to the full order system to obtain the error between the original and reduced order solutions. The waveforms

of the original and approximated solutions of the transmembrane potential have been shown in Fig. 4 at a specific spatial point in the computational domain (x = 0.25 cm and $y = z$ = 0.083 cm). The close agreement of the full order and reduced order solutions confirms the proposed method is capable of offering an accurate reconstruction of the transmembrane potential most of the time, and subsequently, the accuracy of the neural network based order reduction method. Next, the RMS error has been calculated the solutions, and it has been compared with the error obtained from the POD and DMD methods. According to Table 1, the proposed deep learning MOR technique offers approximately 4.65 and 2.41 times better modeling accuracy in terms of RMS error than the POD and DMD methods, respectively.

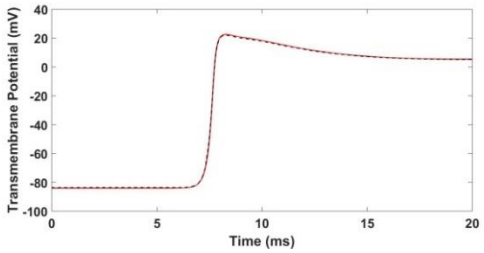


Fig. 4. Cardiac transmembrane potential waveforms of the original solution (V in red solid line) and approximated solution ($\tilde{V}$ in black dashed line).

Table 1: RMS error of the transmembrane potential between the full order and reduced order solutions for different MOR methods

Order Reduction Techniques	RMS Error (mV)
DMD	0.762
POD	0.396
Proposed deep learning	0.164

Finally, the acquired CPU time reduction factor of the proposed deep learning based MOR method has been investigated. All the simulations have been performed on a 2.4-GHz Intel Xeon E5645 processor. The required CPU time to obtain the full order and reduced order solutions are 5,399 s and 65 s, respectively. The CPU time for the proposed neural network method has a reduction factor of almost 85. The CPU time includes the required time for the online modeling and decoding steps, i.e., the prediction stage and the reconstruction phase of the original solution. It is interesting to note that, the mentioned CPU time does not include the training time, as well as, the required time to obtain the reduced order basis as this step is performed only once.

V. CONCLUSION

In this paper, a neural network method has been used for the first time to reduce the order of the cardiac complex monodomain system. An unsupervised machine learning approach, deep autoencoder, has been used for this purpose of dimensionality reduction. The autoencoder maps the full order system into a lower feature representation, as well as reconstructs the original solution from the compressed latent-space representations. Numerical results demonstrate that, the proposed MOR method achieved remarkable computational savings with a factor of almost 85. A significant contribution of this work is to compare the modeling accuracy of the proposed deep learning based technique with the conventional order reduction methods, POD, and DMD. The proposed MOR method has a better accuracy of dimension reduction than the POD and DMD. In the future, the proposed deep learning MOR strategy can be applied to the cardiac bidomain system.

REFERENCES

[1] P. C. Franzone, L. F. Pavarino, and B. Taccardi, "Simulating patterns of excitation, repolarization and action potential duration with cardiac bidomain and monodomain models," *Mathematic Biosciences*, vol. 197, pp. 35-66, 2005.

[2] S. Gandhi and B. J. Roth, "A numerical solution of the mechanical bidomain model," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 19, pp. 1099-1106, 2016.

[3] C. F. Wang, "Efficient proper orthogonal decomposition for backscatter pattern reconstruction," *Progress in Electromagnetics Research*, vol. 118, pp. 243-251, 2011.

[4] C. Corrado, J. Lassoued, M. Mahjoub, and N. Zemzemi, "Stability analysis of the POD reduced order method for solving the bidomain model in cardiac electrophysiology," *Mathematical Biosciences*, vol. 272, pp. 81-91, 2016.

[5] R. Khan and K. T. Ng, "DMD-Galerkin model order reduction for cardiac propagation modeling," *Applied Computational Electromagnetics Society Journal*, vol. 33, pp. 1096-1099, 2018.

[6] X Geng, D. C. Zhan, and Z. H. Zhou, "Supervised nonlinear dimensionality reduction for visualization and classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B* (*Cybernetics*), vol. 35, pp. 1098-1107, 2005.

[7] M. K. Lee and D. S. Han, "A numerical solution of the mechanical bidomain model," *Electronics Letters*, vol. 11, pp. 655–657, 2012.

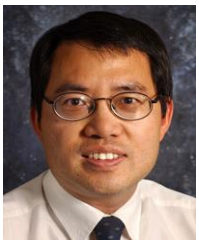
[8] Y. Wang, H. Yao, and S. Zhao, "Auto-encoder based dimensionality reduction," *Neurocomputing*, vol. 184, pp. 232-242, 2016.

[9] R. Khan and K. T. Ng, "Higher order finite difference modeling of cardiac propagation," *IEEE International Conference on Bioinformatics and Biomedicine* (*BIBM*), Kansas City, MO, pp. 1945-1951, 2017.

[10] M. Ethier and Y. Bourgault, "Semi-implicit time-

discretization schemes for the bidomain model," *SIAM Journal on Numerical Analysis*, vol. 46, pp. 2443-2468, 2008.

Riasat Khan is an Assistant Professor of the Electrical and Computer Engineering Department at North South University, Bangladesh. He obtained his M.S. and Ph.D. degrees in Electrical Engineering from New Mexico State University, Las Cruces, NM. His research interests include cardiac electrophysiology, bioelectromagnetics, computational electromagnetics, model order reduction, and power electronics.



Kwong T. Ng is a Professor of the Electrical and Computer Engineering department at New Mexico State University, Las Cruces, NM. He received the M.S. and Ph.D. degrees from The Ohio State University, Columbus, in 1981 and 1985, respectively. His current research interests include bioelectromagnetics, computational electromagnetics, and biomedical instrumentation.