APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY JOURNAL

Special Issue on ACES-China 2022 Conference

Guest Editors:

Jiaqi Liu, National Key Laboratory of Science and Technology on Test Physics & Numerical Mathematics
Jun Wang, China University of Mining and Technology
Ruofeng Xu, China University of Mining and Technology
Kuikui Fan, Hangzhou Dianzi University
Jun Hu, Hefei University of Technology
Zhuowei Miao, Southeast University
Shengjun Zhang, National Key Laboratory of Science and Technology on Test Physics & Numerical Mathematics
Lei Zhao, China University of Mining and Technology

September 2023
Vol. 38 No. 9
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 ARC research group at the Department of Electrical Engineering, Colorado School of Mines

Published, sold and distributed by: River Publishers, Alsbjergvej 10, 9260 Gistrup, Denmark

SEPTEMBER 2023 REVIEWERS

Behrokh Beiranvand
Xi Cheng
Yong-Feng Cheng
Kuikui Fan
Naixing Feng
Jun Hu
Zheng-Yu Huang
Taha Imeci
Weibin Kong
Maokun Li
Qingdong Li
Matteo Bruno Lodi
Zhuo-Wei Miao
Lin Peng
Andrew Peterson
Yi Song
Jianpeng Wang
Jun Wang
Liye Xiao
Ruofeng Xu
Peng Zhang
Hong-Xing Zheng

THE APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY JOURNAL

Vol. 38 No. 9 September 2023

TABLE OF CONTENTS

2022 International Applied Computational Electromagnetics Society (ACES) Symposium
July 28-31, 2022

ACES-China 2022

Guest Editors:
Jiaqi Liu, National Key Laboratory of Science and Technology on Test Physics & Numerical Mathematics
Jun Wang, China University of Mining and Technology
Ruofeng Xu, China University of Mining and Technology
Kuikui Fan, Hangzhou Dianzi University
Jun Hu, Hefei University of Technology
Zhuowei Miao, Southeast University
Shengjun Zhang, National Key Laboratory of Science and Technology on Test Physics & Numerical Mathematics
Lei Zhao, China University of Mining and Technology

The ACES-China 2022 symposium was successfully held in Xuzhou, China on July 28-31 2022. The conference chairs along with a dedicated team of guest editors edited this special issue to provide the whole technical community the opportunity to further explore the most significant contributions to the symposium. Seventeen papers are presented in this special issue. All have been carefully peer reviewed and we hope that you find this special issue a valuable and inspiring contribution to the development of applied computational electromagnetics.

This special issue covers the following topics:

Computational Methods and its applications
Modeling, simulation, and design of metamaterials and meta-devices
Wideband and multiband antennas; Wearable and implantable antennas; Antenna arrays
Spoof surface plasmon polaritons and its applications
Millimeter wave and terahertz devices and beamforming antennas
Metamaterials and metasurfaces for sensing applications

We'd like to also thank the Editors in Chief of ACES Journal, Professor Sami Barmadi and Professor Atef Elsherbeni for their support. And we'd like to express our thanks to the editorial and publication team at ACES Journal for their assistance.

Neural Network Modeling for the Reduction of Scattering Grating Lobes of Arrays

Zhi-Xian Liu[1], Wen-Hao Su[1], Sheng-Jun Zhang[2], and Wei Shao[1]

[1]School of Physics
University of Electronic Science and Technology of China
Chengdu, China
zxliu@std.uestc.edu.cn, suwenhao1202@163.com, weishao@uestc.edu.cn

[2]National Key Laboratory of Science and Technology on Test Physics and Numerical Mathematics
Beijing 100076, China
zhangsj98@sina.com

***Abstract* –** The monostatic radar cross-section (RCS) of an array is seriously deteriorated by the scattering grating lobe. In this paper, the scattering grating lobe of an array is suppressed by metal walls around elements. The artificial neural network with Fourier series-based transfer functions is used to accelerate the design process. A 1×8 array with the patch element operating in the range from 9.4 to 10.6 GHz is studied. The monostatic RCS of the array with designed metal walls is compared with that of the array with no metal wall. Simulated results show that the scattering grating lobe of the array with metal walls is suppressed by 5.8 dB at 12 GHz, and the change of radiation performance is acceptable. The design procedure is also available for other arrays with reduced scattering grating lobes.

***Index Terms* –** Artificial neural network, metal wall, radar cross-section, scattering grating lobe.

I. INTRODUCTION

With the development of the wave-absorbing material, the radar cross-section (RCS) of a stealth platform is suppressed effectively. The reduction of RCS of antennas still remains to be challenging work. To reduce the RCS of antennas, some effective methods have been proposed. The diffusing of scattering wave based on the metasurface is a remarkable method [1, 2]. The incident wave is diffused by the metasurface covering on the antenna based on the phase cancellation mechanism. Despite that, the diffusing method is not suitable for wide-angle RCS reduction. The absorbing material is also used to reduce RCS [3, 4]. The radiation performance of the antenna, however, is affected because the material also absorbs the radiated waves. One traditional and effective tool to suppress RCS is the frequency selective surface (FSS) covering on antennas. FSS is made as low-RCS shape to diffuse the outbound waves, and the incident inbound wave is absorbed by the loaded antenna, whereas the wave in the neighboring band is also able to penetrate FSS because of the imperfect reflection of the outbound wave. The wave with higher frequency than the inbound wave generates noticeable grating lobes, which highly deteriorate the wide-angle monostatic RCS. It remains a problem for RCS reduction of uniform arrays.

Moreover, the traditional design process is often accompanied by time-consuming full-wave simulations. The artificial neural network (ANN) with the transfer function (TF), as a substitute for the full-wave simulation, has been used for wideband electromagnetic (EM) modeling. By representing the wideband EM response as TF, ANN maps the relationship between the geometric parameters and the response. The trained ANN can obtain the EM response accurately and quickly, and it improves the whole design efficiency.

In this paper, an array whose elements are surrounded by metal walls is proposed for the design of a low RCS array. The array is modeled based on two branch ANNs that independently study the radiation and scattering performances to accelerate the design procedure. In addition, the Fourier series-based TF are used in ANN modeling [5], which has the same order for all samples and fewer coefficient orders than those of the pole-residue-based TF [6], bringing the fast convergency and robustness of the model. The monostatic RCS of the array with no metal wall is also studied for comparison. Simulated results show that the scattering grating lobes are suppressed effectively.

II. DESIGN OF ARRAY WITH LOW-GRATING LOBES

The scattering wave of a periodic structure with the illuminating of a plane wave contains Floquet-Bloch

Submitted On: February 27, 2023
Accepted On: May 16, 2023

https://doi.org/10.13052/2023.ACES.J.380901

modes, as shown in Fig. 1. The Floquet-Bloch waves have transverse wavenumbers as

$$K_s = K_0 \times \sin(\theta) + n \times \frac{2\pi}{P}, n = 0, \pm 1, \pm 2, \ldots, \quad (1)$$

where K_s is the transverse wavenumber of the scattering wave, K_0 is the wavenumber in vacuum, θ is the incident angle, and P is the period of the structure. The transverse wavenumber of the zero-order mode is the same as that of the incident wave. The first high-order mode, i.e., the -1st order mode, is usually the most obvious one, so only n = -1 is considered in this paper. The scattering grating lobe of monostatic RCS appears when the -1st order mode moves in parallel fashion in opposite directions of the incident wave. We have

$$\sin(\theta) = \frac{\lambda_0}{2P}, \quad (2)$$

where λ_0 is the wavelength in vacuum. When $\lambda_0 > 2P$, the -1st order mode is evanescent and the grating lobe is not detectable for the radar.

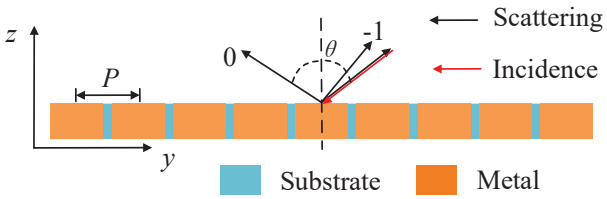


Fig. 1. Periodic structure illuminated by a plane wave.

The common phased array in the stealth platform is a periodic structure. We consider the array operating at 10 GHz as an example, as shown in Fig. 2. The linear array contains 8 elements, and its period is 15 mm. The substrate is Rogers 5880 with a relative dielectric constant of ε_r = 2.2 and a thickness of d = 3 mm.

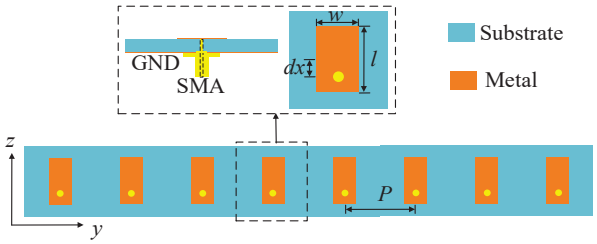


Fig. 2. Structure of the array with no metal wall, where dx = 2.4 mm, w = 6.5 mm, and l = 8.3 mm.

When a plane wave at 12 GHz illuminates the array with an angle of ±56.4°, the monostatic RCS at 12 GHz contains prominent scattering lobes, as shown in Fig. 3. The angle of the -1st scattering mode with P = 15 mm is calculated from (2), as shown in Fig. 4.

To suppress the lobes, metal walls are placed around the array elements, as shown in Fig. 5. The metal walls

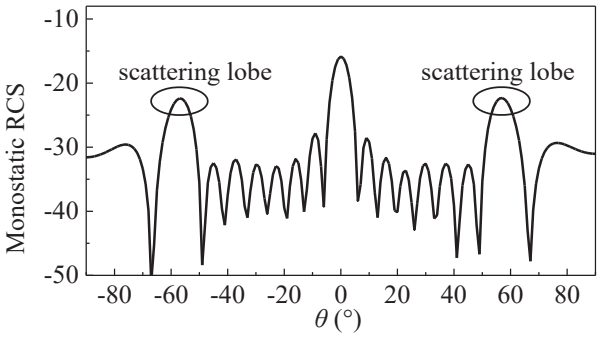


Fig. 3. Monostatic RCS of the array with no metal wall at 12 GHz.

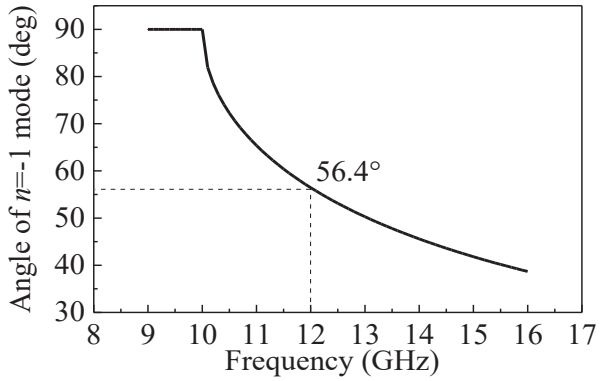


Fig. 4. Angle of the mode with the order of n = -1.

can reduce the scattering grating lobe and maintain the radiation performance of the array.

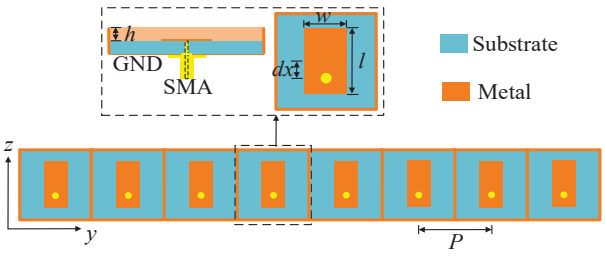


Fig. 5. Structures of the array with metal walls.

From [7], the incident light can be diffracted into the -1st order mode in a broad band by the metallic grating. When a designed metallic grating is placed around the antenna, the -1st order mode can be reduced. Here the metallic walls around the elements work as the metallic gratings. With well-designed metallic walls, the -1st order mode is reduced effectively.

III. OPTIMIZATION DESIGN

The design of the array with metal walls requires numerous full-wave simulations, leading to a time-consuming process. The master-slave boundary and Floquet-Bloch port are used to carry out the element simulation instead of the array simulation. ANN with the Fourier series-based TF is utilized to accelerate the design procedure [5]. The input and output of ANN are the element geometric parameters of $\boldsymbol{x} = [dx, h, l, w]^T$ and the EM response, respectively.

The Fourier series-based TF is written as

$$H(f) = \sum_{k=0}^{N_{\text{order}}} \left(a_k \cos(kq2\pi f) + b_k \cos(kq2\pi f)\right), \quad (3)$$

where N_{order} is the order of TF, a_k and b_k are the TF coefficients, q is the scale factor, and f is the frequency. When $k = 0$, there exists a constant term of a_0. The Fourier series-based TF is used to represent the wideband EM response. With the help of TF, the output of ANN is the coefficients of a_k and b_k instead of the EM response at multiple sampling frequency points. A trained ANN can predict a_k and b_k quickly, and then obtain the wideband EM response from (3). The learning efficiency and generalization of a wideband model can be improved with TF. The Fourier series-based TF only involves real variables and is suitable for real function fitting [5].

With the design of experiment (DOE) method [8], 49 training samples with seven-level and 16 testing samples with four-level defined in Table 1 are obtained from full-wave simulations. The EM response in this paper contains the magnitude of the -1st order Floquet-Bloch mode at 12 GHz when the incident angle is $\pm 56.4^\circ$, and the voltage standing wave ratio (VSWR) of the array is in the range of 9.5-10.5 GHz. Here, we trained two branch ANNs.

Table 1: Definition of training data and testing data for the array

Geometric Parameters	Training Data (49)			Testing Data (16)		
	Min	Max	Step	Min	Max	Step
dx (mm)	1.6	2.4	0.13	1.9	2.2	0.1
h (mm)	1.76	2.64	0.146	1.8	2.1	0.1
l (mm)	7.2	10.8	0.6	8.2	10	0.6
w (mm)	4.8	7.2	0.4	5.4	6.6	0.4

ANN_1, with the Fourier series-based TF, is used to learn the relationship between the geometric parameters and VSWR. The order of the Fourier series-based TF in ANN_1 is chosen as 4 and then the total number of Fourier series coefficients is 9 [5]. From (3), the dimension of ANN_1 output parameters is 9.

ANN_2, with no TF, is used to learn the relationship between the geometric parameters and the magnitude of the -1st order Floquet-Bloch mode. The scattering grating lobe at 12 GHz only appears when the incident angle is equal to $\pm 56.4^\circ$. Because of the axial symmetry of the array, the magnitude of the scattering grating lobe at 56.4° is equal to that at -56.4°. The output of ANN_2 represents the magnitude of the -1st order mode at the incident angle of 56.4°, so TF is not employed for ANN_2 due to the one-dimensional output. The training process of the two branch ANNs is shown in Fig. 6.

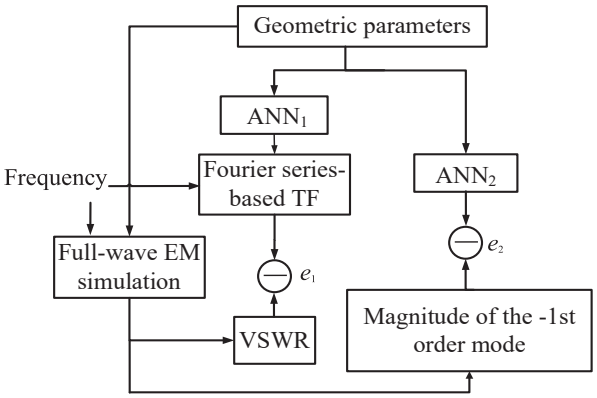


Fig. 6. Training process of the whole ANN model.

The node numbers of the hidden layers of ANN_1 and ANN_2 are determined by the Hecht–Nelson method [9]. When the node number of the input layer is n_{node}, the node number of the hidden layer is $2n_{node} + 1$. The root mean squared errors of e_1 for ANN_1 and e_2 for ANN_2 are provided below.

$$e_1 = \sqrt{\frac{\sum_{i=1}^{N_{\text{sample}}} (VSWR_i^{\text{ANN}} - VSWR_i^{\text{simulation}})^2}{N_{\text{sample}}}}, \quad (4)$$

$$e_2 = \sqrt{\frac{\sum_{i=1}^{N_{\text{sample}}} (mag_i^{\text{ANN}} - mag_i^{\text{simulation}})^2}{N_{\text{sample}}}}, \quad (5)$$

where $VSWR_i^{\text{simulation}}$ is the VSWR calculated from the full-wave simulation of the ith ($1 \leq i \leq N_{sample}$) sample, $VSWR_i^{\text{ANN}}$ is the VSWR calculated from TF whose coefficients of a_k and b_k are obtained by the trained ANN_1, and N_{sample} is the number of the samples. $mag_i^{\text{simulation}}$ is the magnitude of the -1st order Floquet-Bloch mode calculated from the full-wave simulation of the ith sample, and mag_i^{ANN} is obtained from the trained ANN_2.

After the training process, an independent dataset that is never used in training is applied to the testing process. The training and testing errors of ANN_1 are 1.78% and 1.52%, respectively, and the training and testing errors of ANN_2 are 1.04% and 0.93%, respectively. The errors of the ANNs are acceptable. Table 2 summarizes the brief information of the two ANN models.

Table 2: Brief information of two branch ANNs for array modeling

	$\mathbf{ANN}_1$	$\mathbf{ANN}_2$
Input	$\boldsymbol{x}$	$\boldsymbol{x}$
Output	Coefficients of TF (a_k and b_k)	Magnitude of the n=-1 mode
ANN structure	4-9-9	4-9-1
TF	Yes	No
Training error	1.78%	1.04%
Testing error	1.52%	0.93%

After the two ANNs are well trained, the genetic algorithm (GA) is utilized to design the antenna array. GA repeatedly calls the trained ANNs until the optimization objective is achieved, where the optimization objective is the passband of 9.8-10.2 GHz with VSWR < 2 and the magnitude of -1st order mode is small enough. The optimized geometric parameters are $\boldsymbol{x}_{opt}$ = [2.4, 2.5, 8.3, 6.5]^T. The monostatic RCS of the array, calculated by the full-wave simulation with $\boldsymbol{x}_{opt}$, at 12 GHz is shown in Fig. 7. From the figure, the grating lobe is suppressed by 5.8 dB, and the monostatic RCS at the angle of 56.4° is significantly reduced.

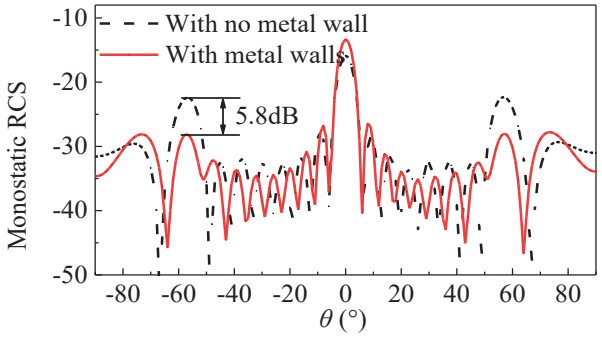


Fig. 7. Monostatic RCS of the array at 12 GHz.

The radiation performance is also considered here. The VSWRs of the optimized arrays with metal walls and without metal wall are provided in Fig. 8. The bandwidth of the array with metal walls is a little narrower than that of the array with no metal wall. Radiation patterns of the array at 10 GHz in the *xoy*-plane are also provided, as shown in Figs. 9 and 10. From Figs. 9 and 10, the gains of the arrays are reduced by 0.9 dB in the normal direction and improved by 1 dB at -50°. The metal wall works as a wide-angle scanning structure, and the change of the radiation performance is acceptable.

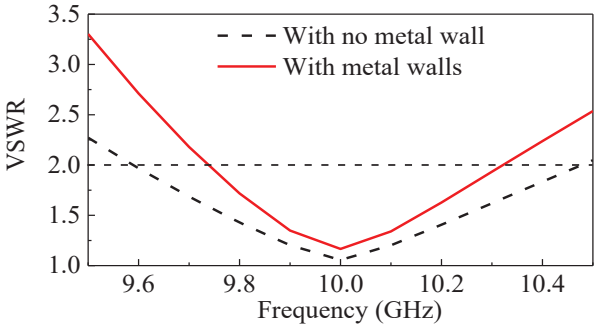


Fig. 8. Simulated VSWR of the center elements.

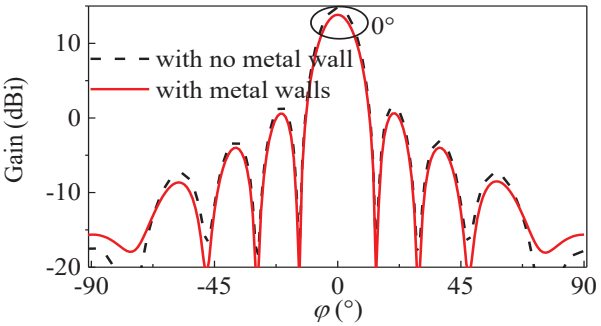


Fig. 9. Radiation patterns of arrays at 10 GHz in the *xoy*-plane with $\varphi = 0°$.

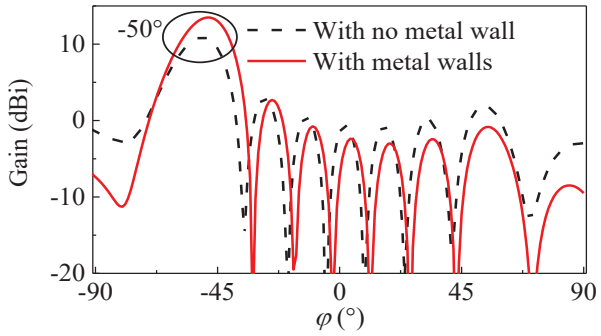


Fig. 10. Radiation patterns of arrays at 10 GHz in the *xoy*-plane with φ = -50°.

IV. CONCLUSION

In this paper, an array with low scattering grating lobes is proposed. By placing well-designed metallic walls around the array elements, the grating lobe is suppressed and the change of the radiation performance is acceptable. ANN with the Fourier series-based TF, as a substitute for the full-wave simulation, is introduced to accelerate the optimization procedure. The proposed method can also be extended to the application

of two-dimensional (2-D) arrays, and we will study 2-D planar arrays with good radiation/scattering performance in future work.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant 62171093 and by the Sichuan Science and Technology Programs under Grants 2022NSFSC0547 and 2022ZYD0109.

REFERENCES

[1] Y. Gou, Y. Chen, and S. Yang, "Radar cross section reduction of wideband Vivaldy antenna arrays with array-level scattering cancellation," *IEEE Trans. Antenna Propagat.*, vol. 70, no. 8, pp. 6740-6750, Aug. 2022.

[2] W. Zhang, Y. Liu, S. Gong, J. Wang, and Y. Jiang, "Wideband RCS reduction of a slot array antenna using phase gradient metasurface," *IEEE Antennas Wireless Propagat. Lett.*, vol. 17, no. 12, pp. 2193-2197, Dec. 2018.

[3] Y. Q. Li, H. Zhang, Y. Q. Fu, and N. C. Yuan, "RCS reduction of ridged waveguide slot antenna array using EBG radar absorbing material," *IEEE Antenna Wireless Propag. Lett.*, vol. 7, pp. 473-476, 2008.

[4] J. Ren, S. Gong, and W. Jiang, "Low-RCS monopolar patch antenna based on a dual-ring metamaterial absorber," *IEEE Antennas Wireless Propagat. Letters*, vol. 17, no. 1, pp. 102-105, Jan. 2018.

[5] Z. X. Liu, W. Shao, X. Ding, L. Peng, and B. Jiang, "Neural network with Fourier series-based transfer functions for filter modeling," *IEEE Microwave Wireless Compon. Lett.*, vol. 32, no. 7, pp. 823-826, July 2022.

[6] F. Feng, C. Zhang, J. Ma, and Q.-J. Zhang, "Parametric modeling of EM behavior of microwave components using combined neural networks and pole-residue-based transfer functions," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 1, pp. 60-77, Jan. 2016.

[7] Z. L. Deng, S. Zhang, and G. P. Wang. "A facile grating approach towards broadband, wide-angle and high-efficiency holographic metasurfaces," *Nanoscale*, vol. 8, no. 3, pp. 1588-1594, 2016.

[8] S. R. Schmidt and R. G. Launsby, *Understanding Industrial Designed Experiments*. Colorado Springs, CO, USA: Air Force Academy, 1992.

[9] R. Hecht-Nielsen, "Kolmogorov's mapping neural network existence theorem," *IEEE Int. Joint Conf. Neural Netw.*, New York, NY, USA, vol. 3, pp. 11-14, June 1987.

Zhi-Xian Liu received the B.S. degree in electronic information science and technology from the China West Normal University, Nanchong, China, in 2019. She is currently pursuing the Ph.D. degree in radio physics in University of Electronic Science and Technology of China (UESTC), Chengdu, China.

Her current research interest includes electromagnetic wave scattering and artificial intelligence.

Wen-Hao Su received the B.S. degree from UESTC, Chengdu, China, in 2021, where he is currently pursuing the master's degree in radio physics.

His current research interests include antenna design and computational electromagnetics.

Sheng-Jun Zhang received the Ph.D. in science from Beijing University of Technology in 2001. From then on he joined the team in National Key Laboratory of Science & Technology on Test Physics and Numerical Mathematics. He is now professor of the laboratory, and his research interest include scattering of EM waves, EM effects of periodic structures such as FSS, PC and gratings, as well as modulation of scattering of materials and interaction of EM waves with plasmas, and IR radiation management.

He has published some papers in journals and conferences, in addition to patents and two books.

Wei Shao received the B.E. degree in electrical engineering from UESTC in 1998, and received M.Sc. and Ph.D. degrees in radio physics from UESTC in 2004 and 2006, respectively.

He joined UESTC in 2007 and is now a professor there. From 2010 to 2011, he was a visiting scholar in the Electromagnetic Communication Laboratory, Pennsylvania State University, State College, PA. From 2012 to 2013, he was a visiting scholar in the Department of Electrical and Electronic Engineering, the University of Hong Kong. His research interests include computational electromagnetics and antenna design.

Outlier Detection-aided Supervised Learning for Modeling of Thinned Cylindrical Conformal Array

Yang Hong[1], Wei Shao[1], Yan He Lv[2], and Zhi Ning Chen[2]

[1]School of Physics
University of Electronic Science and Technology of China, Chengdu, 610054, China
yanghong@std.uestc.edu.cn, weishao@uestc.edu.cn

[2]Department of Electrical and Computer Engineering
National University of Singapore, Singapore, 117583, Singapore
aplyh@nus.edu.sg, eleczn@nus.edu.sg

Abstract – In this paper, a scheme of outlier detection-aided supervised learning (ODASL) is proposed for analyzing the radiation pattern of a thinned cylindrical conformal array (TCCA), considering the impact of mutual coupling. The ODASL model has the advantage in speed improvement and memory consumption reduction, which enables a quick generation of the synthesis results with good generalization. The utilization of the active element pattern (AEP) technique in the model also contributes to the prediction of the array performance involving mutual coupling. The effectiveness of the ODASL model is demonstrated through a numerical example of the 12-element TCCA.

Index Terms – Active element pattern (AEP), conformal array, outlier detection-aided supervised learning (ODASL), thinned array.

I. INTRODUCTION

Recently, conformal arrays have gained popularity in airborne and satellite applications for the sake of their adaptability and aerodynamic performance. However, their analysis and synthesis are particularly complicated due to the varying positions and axial directions of the elements [1]. In addition, the impact of mutual coupling between elements makes it difficult to analyze the far field of conformal arrays, using the directional product theorem applied to planar arrays [2].

The supervised learning method has been extensively recognized as a prediction tool with significant improvement and contribution to electromagnetic (EM) modeling [3–5]. It provides a fast synthesis process for array behaviors while maintaining high-level accuracy with a reduced number of full-wave simulations. However, a multitude of inner parameters of a supervised learning method needs to be determined for large-space and high dimensional problems [6, 7], which easily makes learning progress tardy. In addition, the dependence on sampling data may affect the credibility of the model, especially in the context of array modeling with a complex structure and EM environment. To solve this problem, outlier detection (OD) in data mining is explored as an effective decision-making tool [8]. The multivariate distance-based OD method is used to traverse the raw dataset and identify outlier objects, helping to construct the model effectively.

Considering the impact of mutual coupling and array environment on radiation patterns, a large-scale array can be transformed into the superposition of small sub-arrays by employing the active element pattern (AEP) technique [9]. The technique offers attractive benefits, including the avoidance of heavy calculation burden associated with the whole-array simulation.

To make better use of sampling data, this paper proposes an outlier detection-aided supervised learning (ODASL) model as an alternative to the costly measurement or full-wave simulation. Considering mutual coupling and EM environment, the AEP technique is employed to extract the patterns of sub-arrays, instead of the whole array. By filtering out the invalid sampling data, i.e., outlier data, the ODASL model can obtain satisfactory prediction results, with a strong generalization ability even for larger thinned cylindrical conformal array (TCCA) scales. A TCCA is taken as an example to demonstrate the effectiveness of the proposed model.

II. PROPOSED METHODOLOGY

A. Definition and realization of outlier identification

In the regression supervised learning method for EM modeling, the learning information is completely sourced from the sampling dataset. Hence, it is necessary to perform OD on the raw dataset for the analysis to accurately construct the ODASL framework. The training dataset $\boldsymbol{T}$, defined as $\{\boldsymbol{S}_1, \boldsymbol{S}_2, \ldots, \boldsymbol{S}_D\}$, contains a total of D labeled samples, and the EM response denoted as $\boldsymbol{y}$ in each sample is obtained from full-wave

Submitted On: February 27, 2023
Accepted On: June 13, 2023

https://doi.org/10.13052/2023.ACES.J.380902

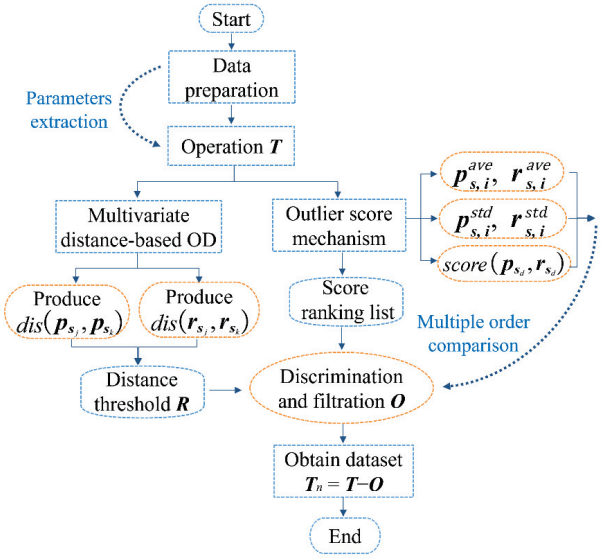


Fig. 1. Flow diagram of outlier identification mechanism.

simulations. Assuming that there are M uncertain outlier samples in the raw training dataset, they can be recorded as $\boldsymbol{O}=[\boldsymbol{O}_1,\boldsymbol{O}_2,\ldots\boldsymbol{O}_{M-1},\boldsymbol{O}_M]$. To measure the similarity in distance between a pair of data objects $\boldsymbol{S}_j$ and $\boldsymbol{S}_k$, a distance function is defined as $dis\,(\boldsymbol{S}_j,\boldsymbol{S}_k)$, satisfying the positive definiteness: $dis\,(\boldsymbol{S}_j,\boldsymbol{S}_k)\geqslant 0$, shown as

$$dis\,(\boldsymbol{S}_j,\boldsymbol{S}_k)=\sum_{1\leqslant t\leqslant L}\left(\left|y_{\boldsymbol{S}_j,t},y_{\boldsymbol{S}_k,t}\right|^q\right)^{1/q},q\geqslant 1 \quad (1)$$

where L represents the number of sampling points in $\boldsymbol{y}$ for each sample, $y_{S_{j,t}}$ and $y_{\boldsymbol{S}_k,t}$ are the the sampling point values for samples $\boldsymbol{S}_j$ and $\boldsymbol{S}_k$, respectively. Generally, the Euclidean distance, which is the second-order Minkowski distance with $q=2$ in formula (1), is adopted as the calculation method to ensure the stability of results regardless of the variation of the dataset space.

With the aid of the transfer function (TF) [10], poles $\boldsymbol{p}$ and residues $\boldsymbol{r}$, which are TF coefficients to describe behaviors of samples, are extracted as the corresponding measurement values and used to identify outliers. The implementation of the specific steps in the OD stage shown in Fig. 1, is described as follows:

1. Step 1: For each sample in $\boldsymbol{T}$, the full-wave simulation of the EM response target $\boldsymbol{y}$, representing the results obtained with AEP, is completed in the range of the specified geometric parameters. The results obtained from the frequency-domain analysis are then fitted to poles/residues-based transfer functions utilizing the vector fitting (VF) technique [11].

2. Step 2: The TF coefficients extracted from all samples have the same order N, and the poles of a sample $\boldsymbol{S}_d$ are set as $\boldsymbol{p}_{\boldsymbol{S}_d}=(p_{\boldsymbol{S}_d,1},p_{\boldsymbol{S}_d,2},\ldots,p_{\boldsymbol{S}_d,i},\ldots,p_{\boldsymbol{S}_d,N})$. Similarly, the residues of $\boldsymbol{S}_d$ are represented by $\boldsymbol{r}_{\boldsymbol{S}_d}=(r_{\boldsymbol{S}_d,1},r_{\boldsymbol{S}_d,2},\ldots,r_{\boldsymbol{S}_d,i},\ldots,r_{\boldsymbol{S}_d,N})$, where $d=1,2,\ldots,D,i=1,2,\ldots,N,p_{\boldsymbol{S}_d,i}$ and $r_{\boldsymbol{S}_d,i}$ are the pole and residue values of the ith order of $\boldsymbol{S}_d$, respectively. $p_{\boldsymbol{S}_d,i}$ and $r_{\boldsymbol{S}_d,i}$ in the dataset are clustered according to their corresponding order. For the calculation of $dis\,(\boldsymbol{S}_j,\boldsymbol{S}_k)$, it is converted to the calculation of the accumulation of two-part distances: one is related to $\boldsymbol{p}$ expressed as $dis\left(\boldsymbol{p}_{\boldsymbol{S}_j},\boldsymbol{p}_{\boldsymbol{S}_k}\right)=\sum_{1\leqslant i\leqslant N}dis\left(p_{\boldsymbol{S}_j,i},p_{\boldsymbol{S}_k,i}\right)$, and another is $dis\left(\boldsymbol{r}_{S_j},\boldsymbol{r}_{S_k}\right)=\sum_{1\leqslant i\leqslant N}dis\left(r_{S_i,i},r_{S_k,i}\right)$ for $\boldsymbol{r}$.

3. Step 3: The given dataset $\boldsymbol{T}$ is evaluated by *a*) To distinguish obvious outliers unequivocally, the distances of $dis\left(\boldsymbol{p}_{S_j},\boldsymbol{p}_{S_k}\right)$ and $dis\left(\boldsymbol{r}_{S_j},\boldsymbol{r}_{S_k}\right)$ with the same order are subject to a restrictive distance threshold $\boldsymbol{R}=\left[\boldsymbol{R}_{\boldsymbol{p}_{\text{dis}}},\boldsymbol{R}_{\boldsymbol{r}_{\text{dis}}}\right]$, with a dimension of $2N\times 1$. The threshold varies with different sampling datasets. *b*) Intrinsically, due to the interdependence among the data points, detecting micro-clusters becomes more complex as these outliers may be neglected as data points from the dense regions of data distribution. Therefore, the outlier score mechanism (OSM) is adopted, denoted as

$$score\left(\boldsymbol{p}_{\boldsymbol{S}_d},\boldsymbol{r}_{\boldsymbol{S}_d}\right)=\sum_{i}^{N}score\left(p_{\boldsymbol{S}_d,i},r_{\boldsymbol{S}_d,i}\right), \quad (2)$$

where

$$score\left(p_{\boldsymbol{S}_d,i},r_{\boldsymbol{S}_d,i}\right)=\left\|\frac{p_{\boldsymbol{S}_d,i}-p_{\boldsymbol{S},i}^{ave}}{p_{\boldsymbol{S},i}^{std}},\frac{r_{\boldsymbol{S}_d,i}-r_{\boldsymbol{S},i}^{ave}}{r_{\boldsymbol{S},i}^{std}}\right\|, \quad (3)$$

where $p_{s,i}^{ave}$ and $r_{s,i}^{ave}$ are the average values of the ith order pole and residue, respectively, $p_{s,i}^{std}$ and $r_{s,i}^{std}$ represent the standard deviation. The absolute value of the score indicates the distance between the data points and the population averages, within the scope of the standard deviation. By filtering out samples with outlier points, a collection of outlier samples $\boldsymbol{O}$ is obtained, and the remained dataset is $\boldsymbol{T}_n=\boldsymbol{T}-\boldsymbol{O}$.

To sum up, the approach for analyzing and identifying outliers can be accomplished by choosing the top M outliers with the largest outlier scores from the score ranking list and by selecting outliers from a cut-off threshold, depending on the distribution of outliers.

B. Proposed ODASL for array modeling

Taking the parametric modeling of arrays for example, the specific procedure with various stages of the ODASL model is presented in Fig. 2. The overall

input $\boldsymbol{x}$ consists of geometric parameters of the element and element spacing. In practice, we want to obtain the relationship between variables in $\boldsymbol{x}$ and EM response $\boldsymbol{y}$ that is affected by mutual coupling and array environment. For exploring the relationship of labeled data pairs of $(\boldsymbol{x}, \boldsymbol{y})$, we have

$$f : V_{\boldsymbol{x}} \to V_{\boldsymbol{y}}, \tag{4}$$

where f is the nonlinear function for the mapping of input space $V_{\boldsymbol{x}}$ to output space $V_{\boldsymbol{y}}$. In the whole set, $V_{\boldsymbol{x}}$ and $V_{\boldsymbol{y}}$ cannot be traversed completely because of the limitation of the sampling requirement, expressed by

$$f(V_{\boldsymbol{x}}) = \{f(\boldsymbol{x}) : \boldsymbol{x} \in V_{\boldsymbol{x}}\} \subseteq V_{\boldsymbol{y}}. \tag{5}$$

In Phase I of the ODASL architecture, dataset generation is the first step to provide the samples for modeling. According to the modeling requirement, the EM response is set as the target, such as S_{11} or pattern information. Then the next step is to obtain the samples with the variable geometric structure and element spacings from the full-wave simulation, and to extract the relevant TF coefficients from EM response.

In Phase II, the potential outliers in $\boldsymbol{T}$ are filtered out using the previous process, and then the input samples of $\boldsymbol{T}_n$ are obtained. For the output, the sub-array patterns based on TF coefficients are accurately captured by the AEP technique. According to the positions of elements, they are categorized into edge element, adjacent-edge element, and interior element, and their AEPs are extracted separately [12, 13].

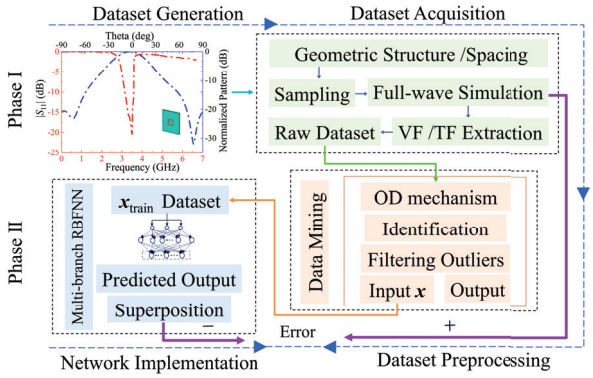


Fig. 2. Overall architecture of the ODASL model.

Subsequently, in the training process of the multi-branch radial basis function neural network (RBFNN) [14], by adjusting the network parameters of multi-branch RBFNN, including the weights of hidden/output layer and center/width of basis function, the relationships between the input and the output in three branches are established. The main purpose of training is to minimize the disparity between $\boldsymbol{y}$ and the predicted $\bar{\boldsymbol{y}}$ from the superposition of all branch results.

To test the model, its generalization ability is crucial to the stable prediction, especially for input $\boldsymbol{x}$ beyond the range of the training dataset [15]. Combined with AEP, a high-degree freedom of array design is guaranteed, and the time-consuming full-wave simulation for the whole array is substituted by the superposition of the predictions from multi-branch networks for sub-arrays. Once the proposed model is well-trained, it immediately provides an accurate response for a given input.

III. NUMERICAL RESULT

A 12-element TCCA in Fig. 3 is taken as the example to evaluate the ODASL model, where it works at 3.5 GHz [16]. The elements are placed on a cylindrical substrate with a radius $r_{\mathrm{a}} = 800$ mm and a relative dielectric constant of $\varepsilon_{\mathrm{r}} = 2.65$. To obtain a high degree of freedom, the input is denoted as $\boldsymbol{x} = [L_{\mathrm{x}}, W_{\mathrm{y}}, W_{\mathrm{l}}, L_{\mathrm{a}}, W_{\mathrm{b}}, f_{\mathrm{x}}, \boldsymbol{d}]^{\mathrm{T}}$, extracted with the design of the experiment method for sampling [17]. The circumferential distance in $\boldsymbol{d} = [d_1, d_2, \ldots, d_e]^{\mathrm{T}}$ is approximately from 0.43λ to 0.54λ during the data collection, where e is equal to 2, 3, or 4 for AEP extraction, depending on the sub-array scale. Table 1 shows the sampling data for the branch of the adjacent-edge element. Data values are standardized before they are used in the network. Similarly, 81 training samples and 36 testing samples are collected for edge elements, and 100 training samples and 64 testing samples are collected for interior elements.

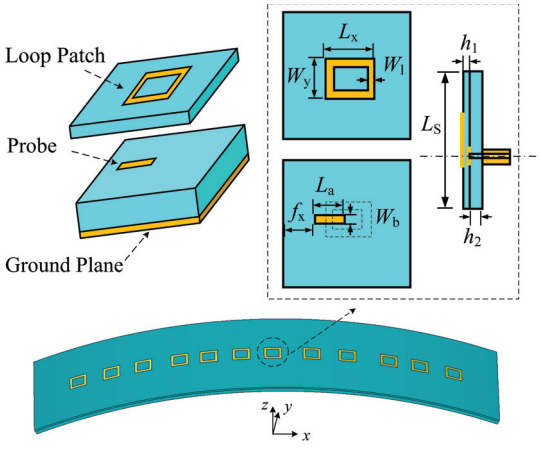


Fig. 3. Structure of the 12-element TCCA, with top, back, and cross-sectional views of the planar surface.

To predict the array pattern considering mutual coupling, firstly, the AEPs of all elements are extracted from the raw data, and the corresponding $\boldsymbol{p}_{\boldsymbol{S}_d}$ and $\boldsymbol{r}_{\boldsymbol{S}_d}$ are obtained. All collected raw datasets are operated in OD processing, and the distances of $\boldsymbol{p}$ and $\boldsymbol{r}$ for each order

Table 1: Definition of training and testing data for adjacent-edge elements (unit: mm)

Structure Parameter	Training Dataset (100 Samples)			Testing Dataset (49 Samples)		
	Min	Max	Step	Min	Max	Step
L_{x}	18	22.5	0.5	18.6	21.6	0.5
W_{y}	13	17.5	0.5	13.8	16.8	0.5
W_1	1.75	2.2	0.05	1.82	2.12	0.05
L_{a}	5.6	6.5	0.1	5.75	6.35	0.1
W_{b}	1.5	2.4	0.1	1.65	2.25	0.1
f_{x}	-8.5	-7.4	0.1	-8.75	-7.35	0.1
d_i	37	46	1	38.5	44.5	1

are acquired, with the maximum order $N = 12$. Secondly, the correlated values of the OSM in formula (3) are calculated, and the scores of each sample are ranked in terms of their degree of deviation. For outlier samples, the distance exceeds the threshold, and meanwhile, the level of outliers is relatively obvious. An example of OD processing for edge elements in the training process is displayed in Table 2, where the values of $\boldsymbol{R}$ and score ranking for each order are exhibited.

Further, Table 3 embodies the situations of $M = 13$ outlier samples, with a certain order of $\boldsymbol{p}$ or $\boldsymbol{r}$ in $\boldsymbol{S}_{dis,i}$ over $\boldsymbol{R}$ and a high degree of dispersion score, where $\boldsymbol{S}_{dis,\,i}$ is the ith distance between $dis(p_{\boldsymbol{S}_{s,i}}, p_{\boldsymbol{S}_{k},i})$ and $dis(r_{\boldsymbol{S}_{s,i}}, r_{\boldsymbol{S}_{k},i})$ for the outliers. Similarly, performing the same operations on all element categories, the resulting filtered sets are used as the branch target outputs.

Table 2: Definition of the parameters of OD processing of the training data for edge elements

Order	**1**	**2**	**3**	**4**	**5**	**6**
$R_{\boldsymbol{p}_{\mathrm{dis}},i}$	0.71	0.87	0.55	0.68	0.89	2.87
$R_{\boldsymbol{r}_{\mathrm{dis}},i}$	2.18	2.91	3.62	2.13	6.57	2.29
$p^{ave}_{\boldsymbol{S},i}$	0.35	0.46	0.69	1.61	2.42	3.54
$r^{ave}_{\boldsymbol{S},i}$	0.37	0.49	1.41	-0.98	3.77	0.21
$p^{std}_{\boldsymbol{S},i}$	0.25	0.35	0.82	0.17	0.14	0.09
$r^{std}_{\boldsymbol{S}_i}$	1.31	1.15	1.52	0.81	2.20	0.83
Order	**7**	**8**	**9**	**10**	**11**	**12**
$R_{\boldsymbol{p}_{\mathrm{dis}},i}$	0.29	0.55	0.48	1.76	0.87	0.45
$R_{\boldsymbol{r}_{\mathrm{dis}},i}$	1.87	1.48	2.12	4.61	1.12	1.83
$p^{ave}_{\boldsymbol{S},i}$	-0.31	-0.42	0.45	2.23	1.08	0.74
$r^{ave}_{\boldsymbol{S},i}$	-0.11	-0.85	1.44	1.25	0.82	0.55
$p^{std}_{\boldsymbol{S},i}$	0.13	0.06	0.05	0.08	0.14	0.06
$r^{std}_{\boldsymbol{S}i}$	1.10	0.84	0.53	2.19	0.25	0.37

Assisted by the characteristics of RBFNN, fast learning speed and high accuracy, the ODASL model predicts the pattern promptly from the three branches, avoiding full-wave EM simulations of the whole array.

Table 3: Outlier samples identified of the training dataset for edge elements

No. of Outlier	Outlier Order		$S_{dis,i}$	*Score*
	Index	p/r		
1	4	r	3.52	3.38
2	7	r	2.91	1.41
3	4	r	3.26	3.32
4	6	p	3.72	2.30
5	11	r	1.10	2.78
6	10	p	2.49	2.80
7	4	r	1.25	3.09
8	7	r	2.63	1.26
9	1	p	1.18	5.29
10	10	p	2.27	2.72
11	11	r	1.06	2.71
12	7	r	2.42	1.15
13	5	p	1.21	2.21

Furthermore, the predicted results can be written as

$$\boldsymbol{E}_{\mathrm{total}} = \boldsymbol{E}_{\mathrm{e}}(\theta,\varphi) + \boldsymbol{E}_{\mathrm{a}}(\theta,\varphi) + \boldsymbol{E}_{\mathrm{i}}(\theta,\varphi) \tag{6}$$

where $\boldsymbol{E}_{\mathrm{e}}(\theta,\varphi), \boldsymbol{E}_{\mathrm{a}}(\theta,\varphi)$, and $\boldsymbol{E}_{\mathrm{e}}(\theta,\varphi)$ represent the fields acquired from the proposed ODASL model of all the edge elements, adjacent-edge elements, and interior elements, respectively.

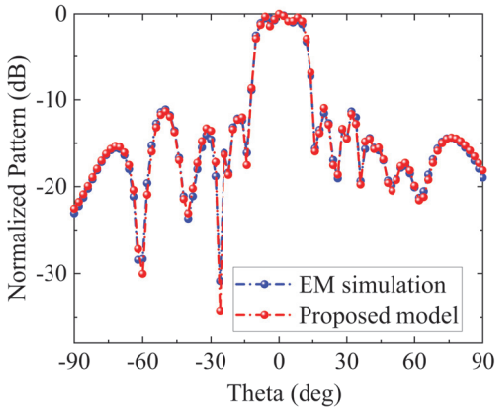
To describe the above formula, $\boldsymbol{E}_{\mathrm{e}}(\theta,\varphi)$ is shown as an example:

$$\boldsymbol{E}_{\mathrm{e}}(\theta,\varphi) = \sum_{1\leqslant s\leqslant N_{\mathrm{e}}} I_s \boldsymbol{F}^{\mathrm{e}}_s(\theta,\varphi) \mathrm{e}^{jk\hat{\boldsymbol{r}}\cdot\boldsymbol{r}s} \tag{7}$$

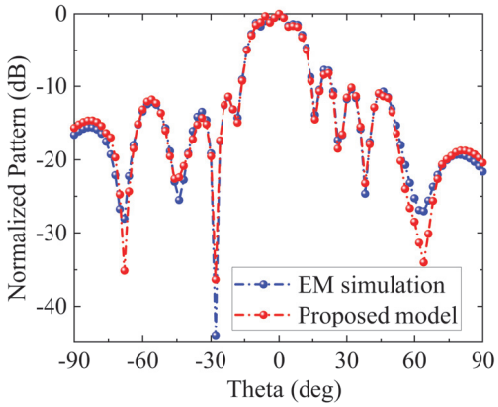
where I_{s} refers to the excitation amplitude of the sth edge element, N_{e} is the number of edge elements, $\boldsymbol{F}^{\mathrm{e}}_s(\theta,\varphi)$ is the AEP result of the sth edge element, $k = 2\pi/\lambda$ is the wavenumber in free space, where λ is the wavelength, and $\mathrm{e}^{jk\hat{\boldsymbol{r}}\cdot\boldsymbol{r}s}$ is the spatial phase factor. Accordingly, the far-field pattern of the TCCA is obtained by superimposing the extracted results.

Accordingly, after the OD process filters out 13 training and 6 testing outlier samples for edge elements, 15 training and 7 testing samples for adjacent-edge elements, and 16 training and 10 testing samples for interior elements, the construction of the ODASL model for the 12-element TCCA costs approximately 13.92 hours, and the average mean absolute percent errors (MAPEs) of the training and testing processes for the whole proposed ODASL model are 3.576% and 4.358%, respectively. All calculations are performed on an Intel i7-6700 3.40 GHz machine with 16 GB RAM.

As an example of TCCA modeling, the results of two separate arrays are shown in Fig. 4. For Array 1, the parameters of input $\boldsymbol{x}_1$ are within the training dataset range, while those of input $\boldsymbol{x}_2$ in Array 2 are outside the range: The parameters of $\boldsymbol{x}_1 = [19.2, 17.7, 1.84, 6.32, 2.03, -8.44, 37.5, 39.4, 37.7,$



(a)



(b)

Fig. 4. Pattern results of the proposed model and the full-wave simulation at 3.5 GHz: (a) Array 1 and (b) Array 2.

$40.1, 38.6, 44.2, 38.5, 37.2, 38.3, 40.7, 39.8]^{\mathrm{T}}$, with a MAPE value 3.317%, and $\boldsymbol{x}_2 = [23.2, 18.3, 2.57, 7.1, 2.26, -8.9, 36.5, 36.9, 46.4, 36.7, 46.2, 35.8, 36.2, 46.3, 36.8, 46.1, 36.6]^{\mathrm{T}}$, with a MAPE value 4.334%. From Fig. 4, the obtained agreement between the simulation and the ODASL model results proves the advantage of the proposed model in terms of accuracy for input parameters both within and outside the training dataset range.

To reflect the properties of the ODASL model with the TF coefficients as the outputs by extracting the AEPs from the sub-arrays, it is compared with the efficient extreme learning machine (ELM) [18, 19] and RBFNN, which directly output the whole array performance without involving the AEP technique. Table 4 indicates the network structure and computational accuracy of the three models. The error measurement standards include MAPE and root mean square error (RMSE), and the small MAPE and RMSE values of ODASL show its well-predicted performance in accuracy and stability. The ELM and RBFNN models collect 49 training samples and 25 testing samples within the parameter range in Table 1.

Table 4: Comparison of the three different models

	TF Order	No. of Hidden Neurons	MAPE	RMSE
RBFNN	17	23	7.92%	0.0125
	21	30	6.69%	0.0093
ELM	17	10	6.24%	0.0076
	21	15	5.71%	0.0052
ODASL	12	7;7;8	3.58%	0.0029
	14	10;11;12	3.06%	0.0018

The ELM and RBFNN models use a single hidden layer, with 17 inputs, including the structure parameters and non-uniform element spacings. Here two cases of different TF orders are employed for TCCA modeling.

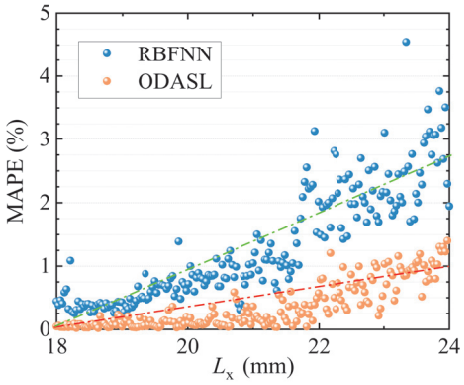
From Table 4, the proposed ODASL model shows lower errors than those of ELM and RBFNN for the two cases. In other words, more collected samples for the training of ELM and RBFNN are needed to get the same level of accuracy.

To examine the applicability of the OD process in the proposed model, Table 5 provides the error comparison for the multi-branch RBFNN model and the ODASL model, which are based on the same dataset obtained from the AEP technique. It is shown that even though the network structures for the two models are similar, the proposed model yields more accurate results.

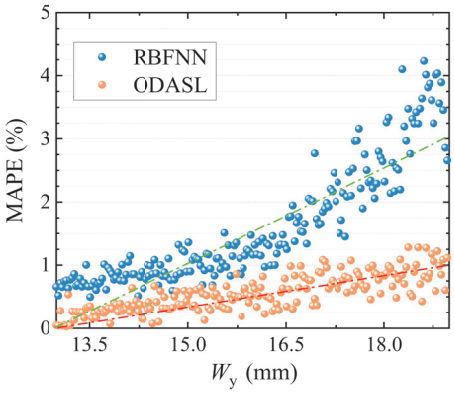
Table 5: Comparison between the proposed ODASL model and the multi-branch RBFNN model

Element Category	ODASL Model		Multi-branch RBFNN Model	
	Training Error	Testing Error	Training Error	Testing Error
Edge	0.894%	0.923%	1.272%	1.709%
Adjacent-edge	0.952%	1.106%	1.744%	2.235%
Interior	1.137%	1.194%	2.405%	2.928%

In Fig. 5 (a), the predicted MAPEs of the multi-branch RBFNN and the proposed model are compared, where the parameter of L_x is considered as a single variable of the array. Figure 5 (b) provides the results of the parameter of W_y. The results show that within the training dataset range, the proposed model gets satisfactory results. Even if the input parameter is out of the range of the training dataset, the proposed model can



(a)



(b)

Fig. 5. Comparison of MAPE results with varying parameters: (a) L_x and (b) W_y.

obtain much more accurate results than the multi-branch RBFNN model.

For further comparison, the proposed ODASL model, full-wave simulation, multi-branch RBFNN and ELM are employed to simulate a 22-element array, a 46-element array, and a 75-element array. The CPU time is listed in Table 6. Compared with the full-wave simulation, the ODASL model is constructed at a cost of 13.9 hours. For large scale arrays, however, the well-trained ODASL can be re-called to realize the fast simulation. Because ELM does not involve the AEP technique, three ELM models corresponding to the different array scales are constructed. Thus the whole modeling time with ELM is much more than that with ODASL. Compared with the multi-branch RBFNN combined with the AEP technique, the ODASL model filters out the invalid sampling data. Therefore, ODASL needs fewer training samples than the multi-branch RBFNN, and then it shows higher modeling efficiency.

Table 6: Comparison of CPU time for different arrays

Number of Elements		**22**	**46**	**75**
Full-wave Simulation	CT	–	–	–
	RT	3.3 h	8.4 h	19.0 h
	Total	30.7 h		
ELM	CT	14.3 h	16.2 h	20.2 h
	RT	1.2 m	1.5 m	1.6 m
	Total	50.8 h		
Multi-branch RBFNN	CT	26.2 h	–	–
	RT	1.8 m	2.1 m	2.7 m
	Total	26.3 h		
ODASL	CT	13.9 h	–	–
	RT	1.2 m	1.4 m	1.5 m
	Total	14.0 h		

CT/RT: Construction/Running time, h: hour, m: minute

IV. CONCLUSION

In this paper, a novel ODASL framework is proposed for efficient TCCA modeling, addressing the challenge posed by mutual coupling and data dependence, with the aim of meeting high-performance requirements for radiation pattern prediction. The proposed model provides a fast pattern realization process with an appreciable reduction of full-wave EM simulations. Combined with the AEP technique, the OD method employs multivariate distance-based clustering and OSM to enhance the discernment and quantification of outliers, constructing the operation for a highgeneralizable model. The valid samples are obtained by outlier elimination, and a numerical example demonstrates the effectiveness of the ODASL model. Additionally, the proposed model with the related data mining method can be further extended to other microwave applications.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant 62171093 and by the Sichuan Science and Technology Programs under Grants 2022NSFSC0547 and 2022ZYD0109.

REFERENCES

[1] J. Harris and H. Shanks, "A method for synthesis of optimum directional patterns from nonplanar apertures," *IRE Trans. Antennas Propag.*, vol. 10, no. 3, pp. 228-236, May 1962.

[2] L. I. Vaskelainen, "Constrained least-squares optimization in conformal array antenna synthesis," *IEEE Trans. Antennas Propag.*, vol. 55, no. 3, pp. 859-867, Mar. 2007.

[3] C. A. Oroza, Z. Zhang, T. Watteyne, and S. D. Glaser, "A machine-learning-based connectivity model for complex terrain large-scale low-power wireless deployments," *IEEE Trans. Cognit.*

Commun. Netw., vol. 3, no. 4, pp. 576-584, Dec. 2017.

[4] L.Y. Xiao, W. Shao, F. L. Jin, B. Z. Wang, and Q. H. Liu, "Inverse artificial neural network for multi-objective antenna design," *IEEE Trans. Antennas Propag.*, vol. 69, no. 10, pp. 6651-6659, Oct. 2021.

[5] P. Liu, L. Chen, and Z. N. Chen, "Prior-knowledge-guided deep-learning-enabled synthesis for broadband and large phase shift range metacells in metalens antenna," *IEEE Trans. Antennas Propag.*, vol. 70, no. 7, pp. 50245034, July 2022.

[6] A. Seretis and C. D. Sarris, "An overview of machine learning techniques for radiowave propagation modeling," *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 3970-3985, June 2022.

[7] R. Caruana, N. Karampatziakis, and A. Yessenalina, "An empirical evaluation of supervised learning in high dimensions," *Proc. 25th Int. Conf. Mach. Learn.*, pp. 96-103, 2008.

[8] H. Wang, M. J. Bah, and M. Hammad, "Progress in outlier detection techniques: A survey," *IEEE Access*, vol. 7, pp. 107964-108000, Aug. 2019.

[9] D. F. Kelley and W. L. Stutzman, "Array antenna pattern modeling methods that include mutual coupling effects," *IEEE Trans. Antennas Propag.*, vol. 41, no. 12, pp. 1625-1632, Dec. 1993.

[10] F. Feng, C. Zhang, J. Ma, and Q. J. Zhang, "Parametric modeling of EM behavior of microwave components using combined neural networks and pole-residue-based transfer functions," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 1, pp. 60-77, Jan. 2016.

[11] B. Gustavsen and A. Semlyen, "Rational approximation of frequency domain responses by vector fitting," *IEEE Trans. Power Del.*, vol. 14, no. 3, pp. 1052-1061, July 1999.

[12] Q. Q. He and B.-Z. Wang, "Design of microstrip array antenna by using active element pattern technique combining with Taylor synthesis method," *Prog. Electromagn. Res.*, vol. 80, pp. 63-76, 2008.

[13] Y. Hong, W. Shao, Y. H. Lv, B. Z. Wang, L. Peng, and B. Jiang, "Knowledge-based neural network for thinned array modeling with active element patterns," *IEEE Trans. Antennas Propag.*, vol. 70, no. 11, pp. 11229-11234, July 2022.

[14] F. Feng, W. Na, J. Jin, J. Zhang, W. Zhang, and Q. J. Zhang, "Artificial neural networks for microwave computeraided design: The state of the art," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 11, pp. 4597-4619, Nov. 2022.

[15] H. Kabir, Y. Cao, and Q. J. Zhang, "Advances of neural network modeling methods for RF/microwave applications," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 25, no. 5, pp. 423-432, May 2010.

[16] Y. Kimura, S. Saito, Y. Kimura, and T. Fukunaga, "Design of wideband multi-ring microstrip antennas fed by an L-probe for single-band and dual-band operations," *Proc. 2020 IEEE AP-S. Int. Symp*, pp. TU-A1.4A.8, July 2020.

[17] R. Schmidt and R. G. Launsby, "Understanding Industrial Designed Experiments," Colorado Springs, CO, USA: Air Force Academy, 1992.

[18] L.Y. Xiao, W. Shao, S. B. Shi, and B. Z. Wang, "Extreme learning machine with a modified flower Pollination algorithm for filter design," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 23, no. 3, pp. 279-284, 2018.

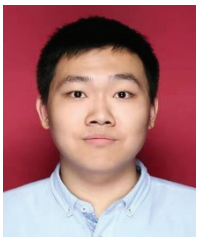
[19] B. Deng, X. Zhang, W. Gong, and D. Shang, "An overview of extreme learning machine," *Proc. 4th Int. Conf. Control Robot. Cybern. (CRC)*, pp. 189-195, Sep. 2019.

Yang Hong received the B.S. degree in electronic information science and technology from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2017. Currently, she is working toward the Ph.D. degree in physics at UESTC. In 2022, she joined the Department of Electrical and Computer Engineering, National University of Singapore, Singapore, as a visiting student.

Her research interest is neural network, antenna array, and computational electromagnetics.

Wei Shao received the B.E. degree in electrical engineering from UESTC in 1998, and received M.Sc. and Ph.D. degrees in radio physics from UESTC in 2004 and 2006, respectively.

He joined UESTC in 2007 and is now a professor there. From 2010 to 2011, he was a visiting scholar in the Electromagnetic Communication Laboratory, Pennsylvania State University, State College, PA. From 2012 to 2013, he was a visiting scholar in the Department of Electrical and Electronic Engineering, the University of Hong Kong. His research interests include computational electromagnetics and antenna design.



Yan-he Lv received the B.S. degree and Ph.D. degree in electronic information science and technology and radio physics from UESTC in 2017 and 2022, respectively.

He is currently a Research Fellow in the National University of Singapore, Singapore. His main research interests include metasurface, phased array, time-reversed electromagnetic, and computational electromagnetics.

Zhi Ning Chen received the B.Eng., M.Eng., and Ph.D. degrees in electrical engineering from the Institute of Communications Engineering (ICE), China, in 1985, 1998, and 1993, and a second Ph.D. degree from the University of Tsukuba, Tsukuba, Japan in 2003.

He joined the National University of Singapore in 2012 as a tenured full professor. From 1988 to 1995, He was a lecturer and later a professor with ICE and a post-doctoral fellow and later an associate professor with Southeast University, Nanjing, China. From 1995 to 1997, he was a research assistant and later a research fellow with the City University of Hong Kong, Hong Kong. In 2001 and 2004, he visited the University of Tsukuba twice under the JSPS Fellowship Program (senior fellow). In 2004, he joined the IBM Thomas J. Watson Research Center, Ossining, NY, USA, as an academic visitor. In 2013, he joined the "Laboratoire des SignauxetSystèmes," UMR8506 CNRS-Supelec-University Paris Sud, Gif-sur-Yvette, France, as a senior DIGITEO guest scientist. In 2015, he joined the Center for Northeast Asian Studies, Tohoku University, Sendai, Japan, as a senior visiting professor.

He was elevated a Fellow of the IEEE for the contribution to small and broadband antennas for wireless applications in 2007 and a Fellow of the Academy of Engineering, Singapore, in 2019 for the contribution to research, development, and commercialization of wireless technology.

He is pioneering in developing small and wideband/ultrawideband antennas, wearable/implanted medical antennas, package antennas, near-field antennas/coils, 3-D integrated LTCC arrays, microwave lens antennas, microwave metamaterial-metasurface (MTS)-metaline-based antennas for communications, sensing, and imaging systems. His current research interests include the translational research of electromagnetic metamaterials and the applications of prior-knowledge-guided machine learning to antenna engineering.

Parametric Modeling for Curved Slots of Vivaldi Antenna Based on Artificial Neural Network

Wen-Hao Su[1], Wei Shao[1], Haiyan Ou[1], and Sheng-Jun Zhang[2]

[1]School of Physics
University of Electronic Science and Technology of China, Chengdu, 611731, China
suwenhao1202@163.com, weishao@uestc.edu.cn, ouhaiyan@uestc.edu.cn

[2]National Key Laboratory of Science and Technology on Test Physics and Numerical Mathematics
Beijing, 100076, China
zhangsj98@sina.com

Abstract – This paper proposes an artificial neural network (ANN) model based on parametric modeling for curved slots of the Vivaldi antenna. A more effective processing method is achieved by feeding ANN with the point positions that produce curved edges via cubic spline interpolation rather than the picture of metallic patches. The predicted results of ANN, including *S*-parameter and gain, agree well with those from the full-wave simulation. With the trained model, a Vivaldi antenna with the lower cut-off frequency is optimized by the multi-objective genetic algorithm.

Index Terms – Artificial neural network, cubic spline interpolation, parametric modeling, Vivaldi antenna.

I. INTRODUCTION

Vivaldi antennas have been widely used in many ultra-wideband (UWB) applications such as ground penetrating radar [1], detection [2], communication [3], etc. The Vivaldi antenna was first introduced by Gibson in 1979 [4], while Gazit later proposed the antipodal Vivaldi antenna (AVA) in 1988 [5]. To reduce cross-polarization, Langley designed the balanced Vivaldi antenna (BAVA) [6]. With the improvement of equipment integration, BAVA needs to provide outstanding performance in a limited size. As a miniaturization approach, different types of corrugations have been extensively used in the design of BAVA. Corrugation refers to repetitive, evenly spaced and identical shaped slots made on the outer flare edge which coincides with the edge of the substrate. Corrugation helps to extend the current path to improve the bandwidth of BAVA. The current slot design, however, relies on some relatively simple curves, such as straight lines [7], elliptic curves [8], and exponential curves [9]. These curves cannot be changed arbitrarily, which constrains the design freedom of slots, leading to a limited radiation performance. More crucially, the structure of BAVA is composed of a number of curves, which will definitely increase the calculation time due to the small-grid division in the electromagnetic (EM) full-wave simulation algorithm.

In recent years, it has been demonstrated that the artificial neural network (ANN) may replace the role of full-wave simulation in the process of antenna optimization [10–11]. In parametric modeling of neural networks, the input is the parameter values of the antenna structure, and the EM response serves as the output. An ANN model with three parallel and independent branches has been presented to describe three different performance indexes of the Fabry–Perot resonator antenna [12]. Once the geometric parameters are input to the trained model, it can simultaneously predict *S*-parameter, gain, and radiation pattern. [13] presents an inverse ANN for the modeling of the multimode resonant antenna, where the input is the performance indexes, and the output is a set of related geometric parameters. The inverse model can provide antenna geometries directly without being repetitively called by an optimization process. Although ANN based on parametric modeling can effectively map the relationship from the input to the output, it is challenging to further improve the EM performance of an antenna due to its fixed topology structure using ANN.

[14] proposes a non-parametric modeling method for microwave filters using the convolutional neural network (CNN). Instead of the structural parameter values, the image of metallic patches is employed as the input of the neural network. Although the CNN model can change the component geometry flexibly and expand the solution domain, the structure of CNN is relatively complex, and the hyper-parameters in CNN are difficult to determine due to their huge number. Moreover, a large number of samples based on pixel images are required for CNN training, resulting in a time-consuming training process.

Submitted On: February 27, 2023
Accepted On: June 13, 2023

https://doi.org/10.13052/2023.ACES.J.380903

In this paper, we propose an effective ANN for the parametric modeling of curved slots of BAVA. The structural parameters of a curved slot are used as the input of the model, including the longitudinal movement value of the points, rotation angle, slot width and length, while the output is the *S*-parameter and the gain of the antenna. The cubic spline interpolation method is used to generate curved slots, and the corresponding antenna structures are input to the CST software for simulation to provide training samples for ANN. After training, the ANN prediction results are in good agreement with the CST full-wave simulation results. The resulting BAVA can obtain wider bandwidth thanks to the multi-objective optimization of the non-dominated sorting genetic algorithm-II (NSGA-II), which repeatedly calls the trained ANN.

II. PROPOSED MODEL

In this paper, the curved slots shaped by ANN are introduced to extend the current path effectively. The training process of the proposed ANN model based on parametric modeling of BAVA is shown in Fig. 1. The geometric structure of a curved slot includes the longitudinal movement value of the points, rotation angle, slot width, and length. Based on cubic spline interpolation, a curve passing through these points is formed. Through translation and rotation, curves form a slot structure. The slot generated by the proposed ANN is repeatedly copied to produce corrugation in BAVA flares. The BAVA with corrugation is sent to the CST software for calculation, and the obtained *S*-parameters and gain at each frequency point are used to train the ANN. As the *S*-parameter has large fluctuations in the frequency band, 200 points are sampled evenly across the frequency band instead of vector fitting based on the transfer function. To simplify the neural network structure and speed up training, four sub-ANNs are trained with 50 sampling points each. Finally, the outputs of the four sub-ANNs, including ANN_1, ANN_2, ANN_3, and ANN_4, are combined to produce a complete *S*-parameter. ANN_5 is used to train the gain.

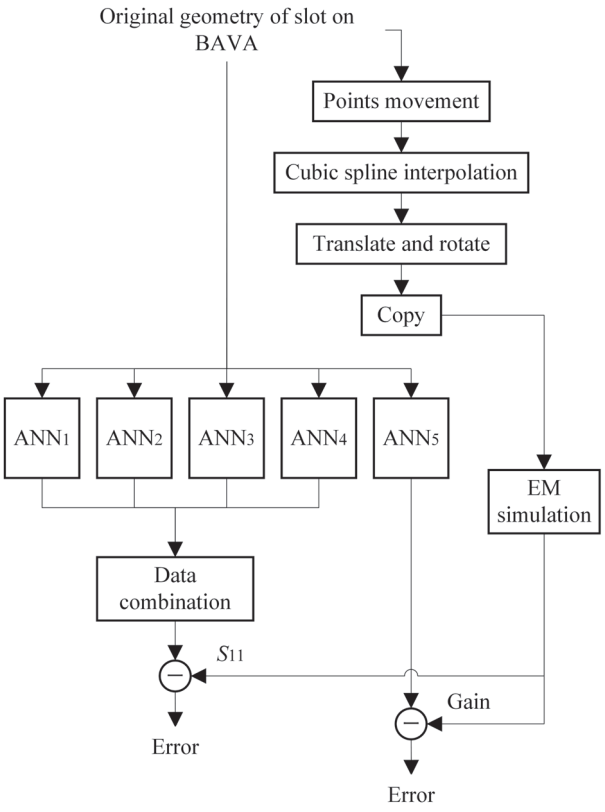


Fig. 1. Structure of the proposed model.

Cubic spline interpolation is applied to the modeling of the curved slots on the flares of BAVA, as shown in Fig. 2.

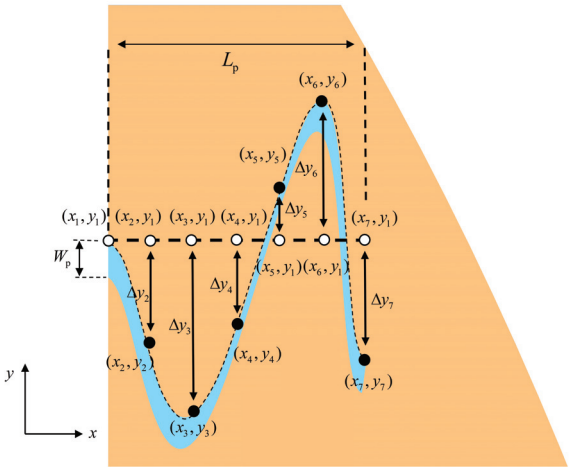


Fig. 2. Diagram of the parametric modeling for curves.

The length of the slot along the x direction is controlled by L_p. For example, seven points are set on the curve, and the position of the kth point is denoted by (x_k, y_k). The amount of space between any two adjacent points along the x-axis is the same. The first point is fixed at (x_1, y_1), and the remaining six points are moved along the y direction to control the curve shape of the slot. As a result, the curve consists of six piecewise curves. The cubic spline interpolation formula for the interval of $x_k \le x \le x_{k+1}$ is as follows

$$f(x) = y_k + \left[\frac{y_{k+1}-y_k}{x_{k+1}-x_k} - \frac{x_{k+1}-x_k}{2}m_k - \frac{x_{k+1}-x_k}{6}(m_{k+1}-m_k)\right](x-x_k) + \frac{m_k}{2}(x-x_k)^2 + \frac{m_{k+1}-m_k}{6(x_{k+1}-x_k)}(x-x_k)^3, k = 1,2,\ldots,6 \quad (1)$$

The corresponding linear equations for $m_1, m_2, \ldots,$ and m_k can be written as

$$\begin{bmatrix} 2 & 1 & & & \\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 4 & 1 \\ & & & 1 & 2 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_k \\ m_{k+1} \end{bmatrix} = \frac{6}{h} \begin{bmatrix} \frac{y_2-y_1}{x_2-x_1} \\ \frac{y_3-y_2}{x_3-x_2} - \frac{y_2-y_1}{x_2-x_1} \\ \vdots \\ \frac{y_{k+1}-y_k}{x_{k+1}-x_k} - \frac{y_k-y_{k-1}}{x_k-x_{k-1}} \\ -\frac{y_{k+1}-y_k}{x_{k+1}-x_k} \end{bmatrix}, \quad (2)$$

where $h = L_p/6$. Once m_k are solved from (2), one edge of the curved slot can be obtained based on (1). Then, the curve is shifted along the y direction by the slot width of W_p to obtain the other edge of the slot. With the above operation, the shape change of the slot is achieved through the parametric modeling.

III. APPLICATION EXAMPLE

A BAVA [15] without the slot in Fig. 3 (a) is employed as an example to evaluated the proposed model. The BAVA consists of two substrate layers and three copper layers. The first and third copper layers connected by two metallic via holes serve as ground layers, and the middle layer acts as a conductor. All copper layers are separated by substrates. The original geometric parameters are as follows: $L = 59.9$ mm, $L_1 = 22$ mm, $L_2 = 10$ mm, $L_3 = 10$ mm, $L_4 = 4$ mm, $W_a = 1$ mm, $W_g = 6.578$ mm, $W = 24$ mm, $W_{ms} = 0.4$ mm, $W_{sl} = 0.74$ mm, and $D_{vh} = 0.6$ mm.

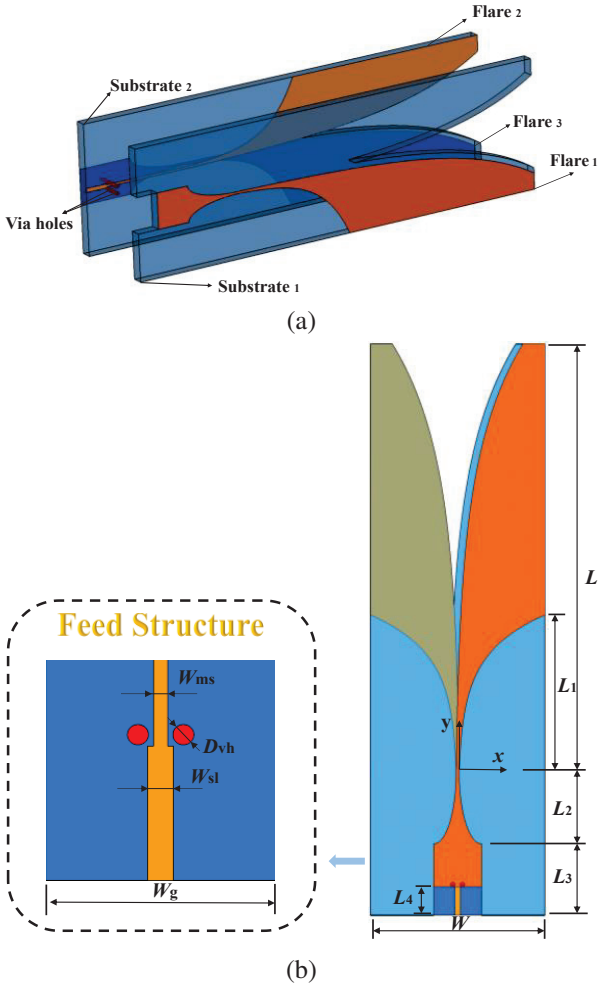


Fig. 3. Geometry of a BAVA without slots: (a) 3D view and (b) 2D view.

As shown in Fig. 3 (b), the inner and outer curved edge profiles of the flares are determined by [15]

$$\begin{cases} X_{\text{inner}} = \pm\left[-W_{ms} + (W_{ms}/2)\, e^{p_1 y}\right] \\ X_{\text{outer}} = \pm\left[(W_{ms}/2)\, e^{p_2 y}\right] \end{cases} \quad (3)$$

where $p_1 = 0.064276$ and $p_2 = 0.157475$. The asymmetric substrate cutout profile is then defined as

$$\begin{cases} X_1 = +\left[-W_{ms} + (W_{ms}/2)\, e^{p_1 y}\right] - W_a \\ X_2 = -\left[-W_{ms} + (W_{ms}/2)\, e^{p_1 y}\right] \end{cases}. \quad (4)$$

The metallic patch is printed on the substrate with a thickness of 0.254 mm and a relative dielectric constant of 2.2. Based on parametric modeling, the curved slots are introduced into the flares of BAVA. L_{up} is the distance between the slot and the top of flare, which is fixed at 3 mm. To increase the design freedom, a rotation angle is used as another variable of the slot structure as depicted in Fig. 4.

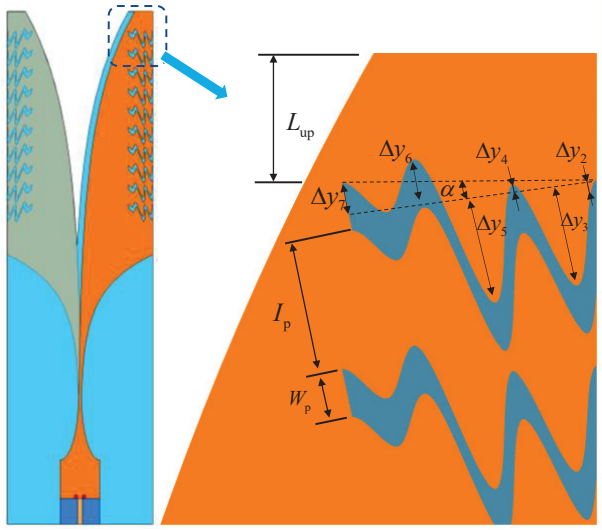


Fig. 4. Geometry of the curved slots.

The inputs of the proposed ANN model and their variation ranges are shown in Table 1, where Δy_k is the longitudinal movement of kth $(k = 2, 3, \cdots, 7)$ point, W_p and L_p represent the width and length of the slots, respectively, and α is the rotation angle of the slots. The design of experiment (DOE) method is used to collect 200 training samples and 50 testing samples.

Table 1: Definition of training and testing samples for the antenna

	Training Data (200 Samples)			Testing Data (50 Samples)		
	Min	Max	Step	Min	Max	Step
Δy_k (mm)	-1.7	1	0.3	-1.7	1	0.3
W_p(mm)	0.1	1.3	0.3	0.1	1.3	0.3
L_p(mm)	0.5	4	0.5	0.5	4	0.5
α(°)	0	20	4	0	20	4

The EM full-wave simulations with the CST software are used for the collection of training and testing samples. The neural networks designed to predict the *S*-parameter are composed of two hidden layers with 10 and 30 neurons, respectively. The neural network designed to predict the gain has a single hidden layer consisting of 15 neurons. The four sub-ANNs for the prediction of the *S*-parameter and the sub-ANN for the prediction of the gain are trained with the Levenberg-Marquardt optimization algorithm with an initial learning rate of 0.001. All the networks are trained with a set of input parameters and their corresponding *S*-parameter or gain values. The training process based on the back propagation scheme iteratively adjusts the weights of all neurons until the desired accuracy level is achieved. The mean absolute percentage error (MAPE) is used to calculate the training and testing errors. The training MAPEs of the ANN model are 4.56% for $|S_{11}|$ and 3.45% for the gain, while the testing MAPEs are 5.62% for $|S_{11}|$ and 4.12% for the gain.

Once the training of ANN is completed, it can be repeatedly called by NSGA-II for the optimization of BAVA in place of the traditional full wave simulation. The optimized variables are $\boldsymbol{x} = (\Delta y_2, \Delta y_3, \Delta y_4, \Delta y_5, \Delta y_6, \Delta y_7, L_p, W_p, \alpha)$. The initial population, which includes 500 individuals, is randomly generated within the range of the variables shown in Table 1. The objective of optimization is to obtain the best possible *S*-parameters and gain. The final optimal structural parameters are as follows: $\boldsymbol{x}_{opt1}$ = (0 mm, -1.6 mm, 0.2 mm, -1.6 mm, 0.8 mm, 0.2 mm, 4 mm, 0.8 mm, 12°). Here, the optimal parameters of $\boldsymbol{x}_{opt1}$ are input to both the ANN model and the CST software to obtain the *S*-parameter and gain of BAVA. When the simulated results from the CST software are regarded as the benchmark, the MAPEs of the *S*-parameter and gain from the ANN model are 1.5% and 0.6%, respectively, indicating its high calculation accuracy.

To highlight the advantages of the proposed approach, we set Δy_k = 0 mm (k = 1,2, …,7) and get the straight slots for comparison. The results are shown in Fig. 5. After the optimization of the BAVA with the straight slots, the optimal structural parameters are as follows: $\boldsymbol{x}_{opt2}$ = (0 mm, 0 mm, 0 mm, 0 mm, 0 mm, 0 mm, 4 mm, 1.3 mm, 0°).

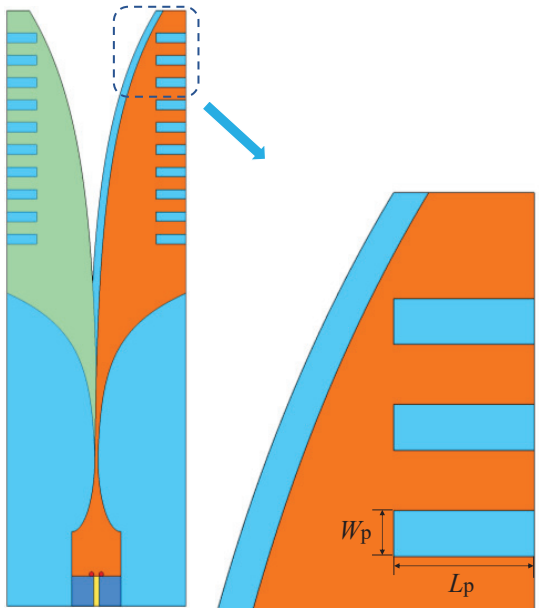
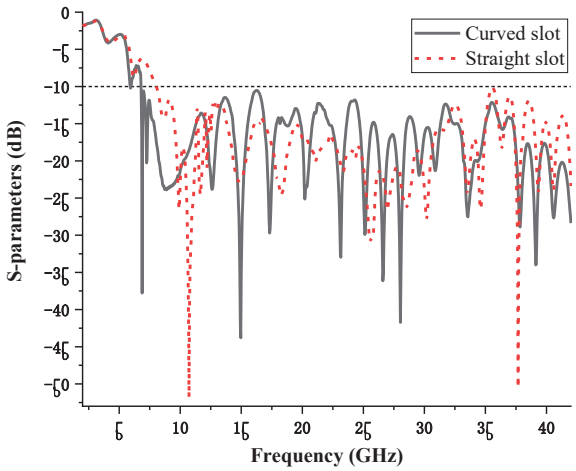


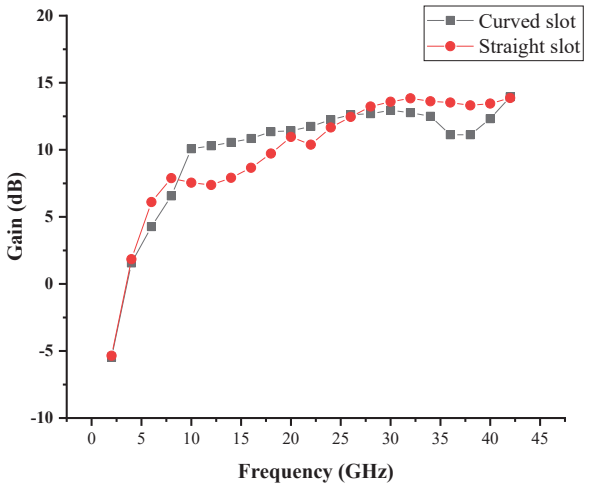
Fig. 5. BAVA with straight slots.

Figure 6 illustrates the simulated results of the BAVAs with the optimal structures of straight slots and curved slots. One can see from Fig. 6 that, the curved slots designed by the proposed model lead to a broader frequency band with nearly the same gain and cross-polarization levels as the straight slots in BAVA. According to Fig. 6 (a), the lower cut-off frequency of the BAVA with straight slots is 7.97 GHz, while that with curved

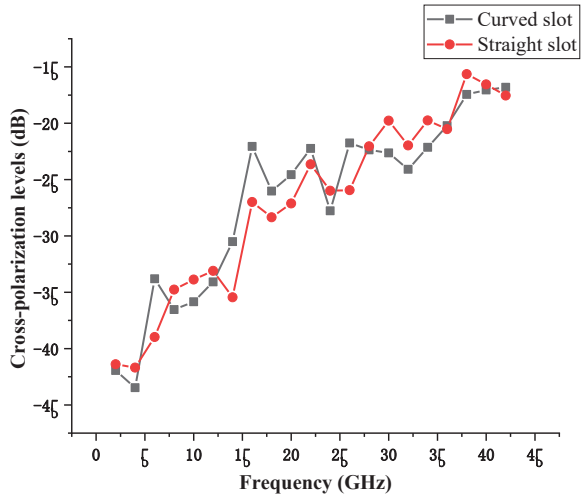


(a)

Fig. 6. *Continued*



(b)



(c)

Fig. 6. Comparison of radiation performance between BAVAs with curved slots and straight slots: (a) $|S_{11}|$, (b) gain, and (c) cross-polarization level.

slots is 6.76 GHz. This indicates that the lower cut-off frequency of the BAVA with curved slots is reduced by about 1.2 GHz.

IV. CONCLUSION

In this paper, an ANN for parametric modeling based on cubic spline interpolation is proposed to shape the slots of BAVA. The longitudinal movement values of the points, rotation angle, width, and length of the curved slot serve as the input of ANN. These input parameters of ANN help to determine the slot structure. The output of the proposed ANN includes the *S*-parameter as well as the gain of the antenna. The lower cut-off frequency is decreased without affecting BAVA's gain, thanks to the curved slots created by the proposed ANN and further optimized by NSGA-II. The proposed ANN only involves simple geometric parameters instead of image processing, leading to a convenient operation of machine learning. Moreover, the parametric modeling can also be adapted to shape design of other devices.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant 62171093 and by the Sichuan Science and Technology Programs under Grants 2022NSFSC0547 and 2022ZYD0109.

REFERENCES

[1] F. T. Wu, G. F. Zhang, X. L. Yuang, and N. C. Yuang, "Research on ultra-wide band planar Vivaldi antenna array," *Microw. Opt. Technol. Lett.*, vol. 48, no. 10, pp. 2117-2120, Oct. 2006.

[2] M. J. Horst, M. T. Ghasr, and R. Zoughi, "Design of a compact V-band transceiver and antenna for millimeter-wave imaging systems," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 11, pp. 4400-4411, Nov. 2019.

[3] M. R. Hamid, P. Gardner, P. S. Hall, and F. Ghanem, "Vivaldi with tunable narrow band rejection," *Microw. Opt. Technol. Lett.*, vol. 53, no. 5, pp. 1125-1128, May 2011.

[4] P. J. Gibson, "The Vivaldi aerial," in *Proc. 9th Eur. Microw. Conf.*, Brighton, U.K., pp. 101-105, June 1979.

[5] E. Gazit, "Improved design of the Vivaldi antenna," *IEE Proc. H-Microw. Antennas Propag.*, vol. 135, no. 2, pp. 89-92, Apr. 1988.

[6] J. D. S. Langley, P. S. Hall, and P. Newham, "Novel ultrawide-bandwidth Vivaldi antenna with low crosspolarisation," *Electron. Lett.*, vol. 29, no. 23, p. 2004, 1993.

[7] A. Dadgarpour, F. Jolani, Y. Yu, Z. Chen, B. S. Virdee, and T. A. Denidni, "A compact balanced antipodal bow-tie antenna having double notch-bands," *Microw. Opt. Technol. Lett.*, vol. 56, no. 9, pp. 2010-2014, Sep. 2014.

[8] M. C. Sai and D. Chandwani, "Balanced antipodal Vivaldi antenna design with hexagonal slots and three level geometric patches," in *2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore, India, pp. 1057-1060, June 2019.

[9] L. Juan, F. Guang, Y. Lin, and F. Demin, "A modified balanced antipodal Vivaldi antenna with improved radiation characteristics," *Micro & Optical Tech Letters*, vol. 55, no. 6, pp. 1321-1325, June 2013.

[10] Y.-F. Liu, L. Peng, and W. Shao, "An efficient knowledge-based artificial neural network for the design of circularly polarized 3-D-printed lens

antenna," *IEEE Trans. Antennas Propagat.*, vol. 70, no. 7, pp. 5007-5014, July 2022.

[11] T. N. Kapetanakis, I. O. Vardiambasis, M. P. Ioannidou, and A. Maras, "Neural network modeling for the solution of the inverse loop antenna radiation problem," *IEEE Trans. Antennas Propagat.*, vol. 66, no. 11, pp. 6283-6290, Nov. 2018.

[12] L.-Y. Xiao, W. Shao, F.-L. Jin, and B.-Z. Wang, "Multiparameter modeling with ANN for antenna design," *IEEE Trans. Antennas Propagat.*, vol. 66, no. 7, pp. 3718-3723, July 2018.

[13] L.-Y. Xiao, W. Shao, F.-L. Jin, B.-Z. Wang, and Q. H. Liu, "Inverse artificial neural network for multiobjective antenna design," *IEEE Trans. Antennas Propagat.*, vol. 69, no. 10, pp. 6651-6659, Oct. 2021.

[14] H.-Y. Luo, W. Shao, X. Ding, B.-Z. Wang, and X. Cheng, "Shape modeling of microstrip filters based on convolutional neural network," *IEEE Microw. Wireless Compon. Lett.*, vol. 32, no. 9, pp. 1019-1022, Sep. 2022.

[15] N.-N. Wang, M. Fang, H.-T. Chou, J.-R. Qi, and L.-Y. Xiao, "Balanced antipodal Vivaldi antenna with asymmetric substrate cutout and dual-scale slotted edges for ultrawideband operation at millimeter-wave frequencies," *IEEE Trans. Antennas Propagat.*, vol. 66, no. 7, pp. 3724-3729, July 2018.

Wen-Hao Su received the B.S. degree from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2021, where he is currently pursuing the master's degree in radio physics.

His current research interests include antenna design and computational electromagnetics.

Wei Shao received the B.E. degree in electrical engineering from UESTC in 1998, and received M.Sc. and Ph.D. degrees in radio physics from UESTC in 2004 and 2006, respectively.

He joined UESTC in 2007 and is now a professor there. From 2010 to 2011, he was a visiting scholar in the Electromagnetic Communication Laboratory, Pennsylvania State University, State College, PA. From 2012 to 2013, he was a visiting scholar in the Department of Electrical and Electronic Engineering, the University of Hong Kong. His research interests include computational electromagnetics and antenna design.

Haiyan Ou received the B.E. degree in electrical engineering from UESTC in 2000, and received Ph.D. degrees in optical engineering from Zhejiang University in 2009.

She joined UESTC in 2009 and is now an associate professor there. From 2010 to 2011, she was a visiting scholar in the department of Engineering, Cambridge University, UK. From 2012 to 2013, she was a post-doc in the Department of Electrical and Electronic Engineering, the University of Hong Kong. Her research interests include computational electromagnetics, microwave photonics, and digital holography.

Sheng-Jun Zhang received a Ph.D. in science from Beijing University of Technology in 2001. From then on he joined the team in National Key Laboratory of Science & Technology on Test Physics and Numerical Mathematics. He is now professor of the laboratory and his research interests include scattering of EM waves, EM effects of periodic structures such as FSS, PC, and gratings, as well as modulation of scattering of materials and interaction of EM waves with plasmas, and IR radiation management.

He has published some papers in journals and conferences, in addition to patents and two books.

A Pre-splitting Green's Function based Hybrid Fast Algorithm for Multiscale Problems

Guang-Yu Zhu[1,2], Wei-Dong Li[2], Wei E. I. Sha[1], Hou-Xing Zhou[2], and Wei Hong[2]

[1]College of Information Science and Electronic Engineering
Zhejiang University, Hangzhou, 310027, China
gyzhu_china@163.com

[2]State Key Laboratory of Millimeter Waves
Frontiers Science Center for Mobile Information Communication and Security
Southeast University, Nanjing, 210096, China

***Abstract* –** Based on the splitting form of the Green's function, a hybrid fast algorithm is proposed for efficient analysis of multiscale problems. In this algorithm, the Green's function is a priori split into two parts: a spectrally band-limited part and a spatially localized part. Then, the fast Fourier transforms (FFT) utilizing the global Cartesian grid and the matrix compression method aided by an adaptive octree grouping are implemented for these two parts, respectively. Compared with the traditional methods which only employ the FFT for acceleration, the proposed hybrid fast algorithm is capable of maintaining low memory consumption in multiscale problems without compromising time cost. Moreover, the proposed algorithm does not need cumbersome geometric treatment to implement the hybridization, and can be established in a concise and straightforward manner. Several numerical examples discretized with multiscale meshes are provided to demonstrate the computational performance of proposed hybrid fast algorithm.

***Index Terms* –** Fast algorithm, fast Fourier transform, Green's function, matrix compression, multiscale problems.

I. INTRODUCTION

Many real-world electromagnetic (EM) problems require multiscale discretization [1–4]. When the method of moments (MoM) [5], along with classical fast algorithms, is used to model the multi-scale problems, a single fast algorithm [6–12] is often insufficient to achieve satisfactory computational performance. Instead, due to capturing both the circuit physics and wave physics [1], hybrid fast algorithms [19–22] generally provide a more practical path to efficiently solve EM multiscale problems.

As we know, the multilevel fast multipole algorithm (MLFMA) is widely used for efficient simulation of electrically large problems [6–8]. However, the MLFMA encounters the sub-wavelength breakdown [8] when dense mesh occurs. As a remedy, low-frequency fast algorithms [23, 24] were studied. Correspondingly, a series of hybrid algorithms were developed [25–30], such as mixed-form FMA [19], interpolative decomposition (ID)-MLFMA [20], and adaptive cross approximation (ACA)-MLFMA [21].

Different from the MLFMA, the pre-correction-fast Fourier transforms (FFT)-based methods [9–18] (e.g., adaptive integral method (AIM) [9], pre-corrected FFT (P-FFT) [10], and integral equation FFT (IE-FFT) [11]) are free from the sub-wavelength breakdown. Nevertheless, due to the intrinsic uniformity of the global Cartesian grids, the pre-correction-FFT-based methods do not allow for sufficient and varying spatial resolutions to handle multiscale discretizations. That is, by simply tuning the grid spacing, it is hard to achieve high computational efficiency for far interactions and maintain low storage efficiency for the near matrix at the same time [31].

Recently, in order to overcome this shortcoming of the pre-correction-FFT-based methods, a hybrid fast algorithm was developed in [31]. In this algorithm, the pre-correction-FFT-based methods is augmented with the matrix compression method (ACA) [36–38]. However, to implement such a hybridization, the procedures are not as straightforward as they seem [31]. In particular, a complicated spatial grouping strategy is inherently needed, and the relations between the required auxiliary geometric data structures (such as the Cartesian grid, the expansion box, and the tree-cube) are complex and require cautious treatments [31]. Such a situation motivates us to reconsider the overall framework of the pre-correction-FFT-based methods and wonder if there exists a more elegant way to develop a hybrid algorithm that consists of

Submitted On: February 19, 2023
Accepted On: July 26, 2023

https://doi.org/10.13052/2023.ACES.J.380904

both the FFT and the matrix compression methods (e.g., ACA).

As an interesting alternative of the well-studied pre-correction-FFT-based methods, some pre-splitting-FFT-based method was developed in [32, 33]. Basically, the pre-splitting-FFT-based method bears similarities with the well-known pre-correction-FFT-based methods, since it also uses the Cartesian grid and the FFT as the essentials for the acceleration of the spatial convolutions. However, different from the widely-used pre-correction-FFT-based methods which typically utilize the Cartesian grid-based near-field corrections (based on some local expansion boxes) to compensate for the near-field contributions, the pre-splitting-FFT-based method [32, 33] instead resorts to some splitting form of the Green's function so as to realize accurate and efficient computations. Despite of its unique feature, this pre-splitting-based framework has not attracted enough attentions in the computational electromagnetics (CEM) community, and a hybrid fast algorithm based on such framework has never been explored yet.

In this work, based on the splitting form of the Green's function, a hybrid fast algorithm is proposed for efficiently analyzing multiscale problems. Specifically, the pre-splitting-FFT-based method originally devised for the problems with quasi-uniform discretizations is here further enhanced with the matrix compression method (ACA). Thus, a hybrid fast algorithm using both the FFT and the matrix compression method is established. In particular, the proposed hybrid fast algorithm is capable of maintaining low memory consumption in multiscale problems without compromising time cost, thus manifesting itself as a favorable multiscale extension of the pure FFT-based method. Furthermore, the required auxiliary geometric data structures herein (including the Cartesian grid and the spatial grouping octree) can be constructed independent of each other, and are not intertwined as in [31]. Consequently, the proposed hybrid algorithm can be implemented in a straightforward manner without any cumbersome geometric treatment.

In the following, after introducing the basic formulations and implementation key points in Section II, several numerical examples discretized with multiscale meshes are then provided in Section III to demonstrate the computational performances of the proposed hybrid fast algorithm for multiscale problems.

II. FORMULATION

A perfectly electric conductor (PEC) object illuminated by an incident plane wave is considered. The combined field integral equation (CFIE) is employed to model this typical electromagnetic scattering problem. Throughout the paper, we use λ to denote the wavelength in free space. Moreover, we use k and η to denote the wave number and the wave impedance in free space, respectively.

A. Combined field integral equation

The CFIE is the linear combination of the electric field integral equation (EFIE) and the magnetic field integral equation (MFIE), which is conventionally expressed as

$$\mathrm{CFIE} = \alpha \mathrm{EFIE} + (1-\alpha)\,\eta \mathrm{MFIE}, \tag{1}$$

where α denotes the combination factor and $0 \leqslant \alpha \leqslant 1$, the EFIE is formulated as

$$\mathbf{E}^{\mathrm{inc}}(\boldsymbol{r})\Big|_{\tan} = \left[jk\eta \int_S \left(\mathbf{J}(\boldsymbol{r}') + \frac{\nabla\nabla'\cdot}{k^2}\mathbf{J}(\boldsymbol{r}')\right) \mathrm{G}(\boldsymbol{r},\boldsymbol{r}')\, ds'\right]\Big|_{\tan}, \tag{2}$$

and the MFIE is formulated as

$$\mathbf{H}^{\mathrm{inc}}(\boldsymbol{r})\Big|_{\tan} = \left[\frac{\hat{\boldsymbol{n}} \times \mathbf{J}(\boldsymbol{r}')}{2} + P.V. \int_S \mathbf{J}(\boldsymbol{r}') \times \nabla \mathrm{G}(\boldsymbol{r},\boldsymbol{r}')\, ds'\right]\Big|_{\tan}. \tag{3}$$

In the above, $\mathbf{E}^{\mathrm{inc}}$ and $\mathbf{H}^{\mathrm{inc}}$ are the incident electric field and the incident magnetic field, respectively. Moreover, $\mathrm{G}(\boldsymbol{r},\boldsymbol{r}')$ denotes the Green's function with the field point $\boldsymbol{r}$ and the source point $\boldsymbol{r}'$. Furthermore, $\mathbf{J}(\boldsymbol{r}')$ denotes the surface current.

By using the method of moments (MoM) [5] with the RWG function [39], the CFIE can be discretized into a matrix equation

$$\boldsymbol{Z}^{[N\times N]} \cdot \boldsymbol{I}^{[N\times 1]} = \boldsymbol{V}^{[N\times 1]}, \tag{4}$$

where $\boldsymbol{Z}$ is the impedance matrix, $\boldsymbol{I}$ is the unknown surface current coefficients vector, $\boldsymbol{V}$ is the righthand side vector of the incident field, and N is the number of unknowns. More specifically, the impedance matrix $\boldsymbol{Z}$ corresponding to the CFIE is described by

$$\boldsymbol{Z}^{\mathrm{CFIE}} = \alpha \boldsymbol{Z}^{\mathrm{EFIE}} + (1-\alpha)\,\eta \boldsymbol{Z}^{\mathrm{MFIE}}, \tag{5}$$

where

$$\begin{aligned} \boldsymbol{Z}_{mn}^{\mathrm{EFIE}} &= jk\eta \int_{S_m} ds \mathbf{f}_m(\boldsymbol{r}) \cdot \int_{S_n} \mathrm{G}(\boldsymbol{r},\boldsymbol{r}')\, \mathbf{f}_n(\boldsymbol{r}')\, ds' \\ &\quad - j\frac{\eta}{k} \int_{S_m} ds \left[\nabla \cdot \mathbf{f}_m(\boldsymbol{r})\right] \cdot \int_{S_n} \mathrm{G}(\boldsymbol{r},\boldsymbol{r}') \left[\nabla \cdot \mathbf{f}_n(\boldsymbol{r}')\right] ds', \end{aligned} \tag{6}$$

and

$$\begin{aligned} \boldsymbol{Z}_{mn}^{\mathrm{MFIE}} &= \frac{1}{2} \int_{S_m} \mathbf{f}_m(\boldsymbol{r}) \cdot \mathbf{f}_n(\boldsymbol{r})\, ds \\ &\quad + \int_{S_m} ds \left[\hat{\boldsymbol{n}} \times \mathbf{f}_m(\boldsymbol{r})\right] \cdot \int_{S_n} \nabla \mathrm{G}(\boldsymbol{r},\boldsymbol{r}') \times \mathbf{f}_n(\boldsymbol{r}')\, ds'. \end{aligned} \tag{7}$$

In the above, $m = 1, \cdots, N$ and $n = 1, \cdots, N$. Moreover, $\mathbf{f}_m$ and $\mathbf{f}_n$ denote the mth and nth RWG functions, respectively. Furthermore, S_m and S_n are the supports of the corresponding RWG functions, respectively.

B. Pre-splitting the Green's function

In the following, the splitting forms of the Green's function and its gradient are introduced. First, consider the Green's function in the EFIE, namely

$$\mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)=\frac{e^{-jkR}}{4\pi R}, \tag{8}$$

where

$$\begin{aligned} R&=\left|\boldsymbol{r}-\boldsymbol{r}'\right| \\ &=\sqrt{(x-x')^2+(y-y')^2+(z-z')^2} \end{aligned} \tag{9}$$

denotes the distance between the field point $\mathbf{r}$ and the source point $\mathbf{r}'$. Clearly, the above can also be equivalently represented as

$$\mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)=\mathrm{Re}\left[\mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right]+j\,\mathrm{Im}\left[\mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right], \tag{10}$$

where

$$\mathrm{Re}\left[\mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right]=\frac{\cos kR}{4\pi R}, \tag{11}$$

$$\mathrm{Im}\left[\mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right]=-\frac{\sin kR}{4\pi R}. \tag{12}$$

A basic property of $\mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)$ is that, this function is singular for $R=0$, and its magnitude is globally nonzero for $R>0$. Also, note that the singular behavior of $\mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)$ comes from $\mathrm{Re}\left[\mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right]$.

Using the pre-splitting strategy [32, 33], the Green's function (8) can be written as two parts, namely

$$\mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)=\mathrm{G}_{\mathrm{E}}\left(\boldsymbol{r},\boldsymbol{r}'\right)+\mathrm{G}_{\mathrm{P}}\left(\boldsymbol{r},\boldsymbol{r}'\right), \tag{13}$$

where

$$\mathrm{G}_{\mathrm{E}}\left(\boldsymbol{r},\boldsymbol{r}'\right)=\begin{cases}\mathrm{Re}\left[\mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right]-\Phi\left(R\right), & R<\delta\\ 0\quad, & R\geq\delta\end{cases}, \tag{14}$$

$$\mathrm{G}_{\mathrm{P}}\left(\boldsymbol{r},\boldsymbol{r}'\right)=\begin{cases}\Phi\left(R\right)+j\mathrm{Im}\left[\mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right], & R<\delta\\ \mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)\quad, & R\geq\delta\end{cases}, \tag{15}$$

with δ being a splitting threshold, and $\Phi\left(R\right)$ being some auxiliary function. Particularly, the above two parts manifest the following properties.

- Function $\mathrm{G}_{\mathrm{E}}\left(\boldsymbol{r},\boldsymbol{r}'\right)$ is a singular and spatially localized function. It has nonzero function value only when $R<\delta$, and $\mathrm{G}_{E}\left(\boldsymbol{r},\boldsymbol{r}'\right)=0$ for $R\geq\delta$.
- Function $\mathrm{G}_{\mathrm{P}}\left(\boldsymbol{r},\boldsymbol{r}'\right)$ is a globally oscillatory and spectrally band-limited function. It is smooth and bounded for $\forall\boldsymbol{r}$ and $\forall\boldsymbol{r}'$.

To have more intuitive understanding on the properties of function $\mathrm{G}_{\mathrm{E}}\left(\boldsymbol{r},\boldsymbol{r}'\right)$ and $\mathrm{G}_{\mathrm{P}}\left(\boldsymbol{r},\boldsymbol{r}'\right)$, the corresponding function images are shown in Fig. 1. Moreover, from a physical and spectral perspective, $\mathrm{G}_{\mathrm{E}}\left(\boldsymbol{r},\boldsymbol{r}'\right)$ mainly captures the *evanescent waves*, and $\mathrm{G}_{\mathrm{P}}\left(\boldsymbol{r},\boldsymbol{r}'\right)$ mainly captures the *propagating waves*.

To achieve the above splitting and corresponding properties, the auxiliary function $\Phi\left(R\right)$ is specially designed and one simple way is to choose $\Phi\left(R\right)$ to be a polynomial of the following form [32, 33]

$$\Phi\left(R\right)=aR^3+bR^2+c, \tag{16}$$

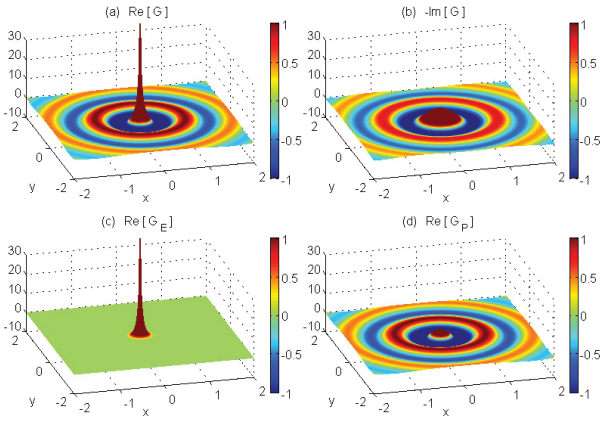


Fig. 1. The Green's function is a priori split into a singular but spatially localized part G_E and a globally oscillatory but spectrally band-limited part G_P. The relevant magnitudes of the original Green's function and of these two parts are visualized in the figure.

where a, b, and c are unknown coefficients. Clearly, the above is a three order polynomial without the linear term, and it implies

$$d\Phi\left(0\right)/dR=0, \tag{17}$$

meaning that the first order derivative function of $\Phi\left(R\right)$ is zero at $R=0$ (where $\boldsymbol{r}=\boldsymbol{r}'$).

Given the splitting threshold δ, the coefficients in (16) can be determined by some continuity conditions at $R=\delta$ as follows

$$\begin{cases}\Phi\left(\delta\right) &=& \mathrm{Re}\left[\mathrm{G}\left(\delta\right)\right]\\ d\,\Phi\left(\delta\right)/dR &=& d\,\mathrm{Re}\left[\mathrm{G}\left(\delta\right)\right]/dR\\ d^2\Phi\left(\delta\right)/dR^2 &=& d^2\mathrm{Re}\left[\mathrm{G}\left(\delta\right)\right]/dR^2\end{cases}, \tag{18}$$

or, by solving the following matrix equation

$$\begin{bmatrix}\delta^3 & \delta^2 & 1\\ 3\delta^2 & 2\delta & 0\\ 6\delta & 2 & 0\end{bmatrix}\begin{bmatrix}a\\ b\\ c\end{bmatrix}=\begin{bmatrix}\mathrm{Re}\left[\mathrm{G}\left(\delta\right)\right]\\ d\,\mathrm{Re}\left[\mathrm{G}\left(\delta\right)\right]/dR\\ d^2\mathrm{Re}\left[\mathrm{G}\left(\delta\right)\right]/dR^2\end{bmatrix}. \tag{19}$$

The derivative functions of $\mathrm{Re}\left[\mathrm{G}\left(R\right)\right]$ used in the above are given in the appendix. Thus, after $\Phi\left(R\right)$ is obtained, the splitting representation (13) is completely determined.

The polynomial choice (16), the implicit continuity condition (17), and the explicit continuity condition (18) are based on the consideration that, after the splitting, the resultant $\mathrm{G}_{\mathrm{P}}\left(\boldsymbol{r},\boldsymbol{r}'\right)$ (or, more precisely, $\mathrm{Re}\left[\mathrm{G}_{\mathrm{P}}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right]$) should be smooth everywhere for $\forall\boldsymbol{r}$ and $\forall\boldsymbol{r}'$, so that it can be interpolated accurately without much difficulty (the interpolation procedures will be discussed in the next subsection).

To make it more clear, without loss of generality, consider the following representative situation. The source point $\boldsymbol{r}'$ is fixed to the coordinate origin, and the field point $\boldsymbol{r}$ is moving along the x axis. Then, the corresponding function values of $\mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)$ and $\mathrm{G}_{\mathrm{P}}\left(\boldsymbol{r},\boldsymbol{r}'\right)$

are evaluated, as shown in Fig. 3. We can clearly see that $\mathrm{Re}\left[\mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right]$ is singular when $\boldsymbol{r}$ approaches $\boldsymbol{r}'$, while $\mathrm{Re}\left[\mathrm{G_P}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right]$ is continuous and smooth everywhere. Here, $\mathrm{Re}\left[\mathrm{G_P}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right]$ is smooth at $R=0$ (where $\boldsymbol{r}=\boldsymbol{r}'$) as a result of (17), and smooth at the splitting point $R=\delta$ as a result of (18).

Next, consider the gradient of the Green's function in the MFIE, namely

$$\nabla \mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right) = \dddot{\mathrm{G}}\left(R\right)\nabla R, \tag{20}$$

with

$$\dddot{\mathrm{G}}\left(R\right) = \mathrm{Re}\left[\dddot{\mathrm{G}}\left(R\right)\right] + j\,\mathrm{Im}\left[\dddot{\mathrm{G}}\left(R\right)\right], \tag{21}$$

where

$$\mathrm{Re}\left[\dddot{\mathrm{G}}\left(R\right)\right] = -\frac{\cos kR + kR\sin kR}{4\pi R^2}, \tag{22}$$

$$\mathrm{Im}\left[\dddot{\mathrm{G}}\left(R\right)\right] = -\frac{kR\cos kR - \sin kR}{4\pi R^2}, \tag{23}$$

and

$$\nabla R = \frac{x-x'}{R}\hat{\boldsymbol{x}} + \frac{y-y'}{R}\hat{\boldsymbol{y}} + \frac{z-z'}{R}\hat{\boldsymbol{z}}. \tag{24}$$

Similar to (8), function $\nabla \mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)$ is singular for $R=0$, and its magnitude is globally nonzero for $R>0$. Also, note that the singular behavior of $\nabla \mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)$ comes from $\mathrm{Re}\left[\dddot{\mathrm{G}}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right]$.

Again, using the pre-splitting strategy [32, 33], the gradient Green's function (20) can be written as two parts, namely

$$\nabla \mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right) = \nabla \mathrm{G_E}\left(\boldsymbol{r},\boldsymbol{r}'\right) + \nabla \mathrm{G_P}\left(\boldsymbol{r},\boldsymbol{r}'\right), \tag{25}$$

where

$$\nabla \mathrm{G_E}\left(\boldsymbol{r},\boldsymbol{r}'\right) = \begin{cases} \left\{\mathrm{Re}\left[\dddot{\mathrm{G}}\left(R\right)\right] - \Phi_{\mathrm{g}}\left(R\right)\right\}\nabla R, & R<\delta \\ 0 \quad , & R\geqslant\delta \end{cases}, \tag{26}$$

$$\nabla \mathrm{G_P}\left(\boldsymbol{r},\boldsymbol{r}'\right) = \begin{cases} \left\{\Phi_{\mathrm{g}}\left(R\right) + j\mathrm{Im}\left[\dddot{\mathrm{G}}\left(R\right)\right]\right\}\nabla R, & R<\delta \\ \nabla \mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right) \quad , & R\geqslant\delta \end{cases}, \tag{27}$$

with δ being the splitting threshold, and $\Phi_{\mathrm{g}}\left(R\right)$ being some auxiliary function.

Here, the following auxiliary function is used to achieve the desired splitting, namely

$$\Phi_{\mathrm{g}}\left(R\right) = aR^3 + bR^2 + cR. \tag{28}$$

Note that, different from (16), the above is a three order polynomial without the constant term, and it implies

$$\Phi_{\mathrm{g}}\left(0\right) = 0, \tag{29}$$

which clearly means that the function value of $\Phi_{\mathrm{g}}\left(R\right)$ is zero at $R=0$ (where $\boldsymbol{r}=\boldsymbol{r}'$).

Given the splitting threshold δ, the coefficients in (28) can be determined by some continuity conditions at $R=\delta$ as follows

$$\begin{cases} \Phi_{\mathrm{g}}\left(\delta\right) & = & \mathrm{Re}\left[\dddot{\mathrm{G}}\left(\delta\right)\right] \\ d\,\Phi_{\mathrm{g}}\left(\delta\right)/dR & = & d\,\mathrm{Re}\left[\dddot{\mathrm{G}}\left(\delta\right)\right]/dR \\ d^2\Phi_{\mathrm{g}}\left(\delta\right)/dR^2 & = & d^2\mathrm{Re}\left[\dddot{\mathrm{G}}\left(\delta\right)\right]/dR^2 \end{cases}, \tag{30}$$

or, by solving the following matrix equation

$$\begin{bmatrix} \delta^3 & \delta^2 & \delta \\ 3\delta^2 & 2\delta & 1 \\ 6\delta & 2 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \mathrm{Re}\left[\dddot{\mathrm{G}}\left(\delta\right)\right] \\ d\,\mathrm{Re}\left[\dddot{\mathrm{G}}\left(\delta\right)\right]/dR \\ d^2\mathrm{Re}\left[\dddot{\mathrm{G}}\left(\delta\right)\right]/dR^2 \end{bmatrix}. \tag{31}$$

The derivative functions of $\mathrm{Re}\left[\dddot{\mathrm{G}}\left(R\right)\right]$ used in the above are given in the appendix. Thus, after $\Phi_{\mathrm{g}}\left(R\right)$ is obtained, the splitting representation (25) is completely determined.

Similar as before, (28), (29), and (30) are based on the consideration that, after the splitting, the resultant $\nabla \mathrm{G_P}\left(\boldsymbol{r},\boldsymbol{r}'\right)$ (or, more precisely, each Cartesian component $\mathrm{Re}\left[\nabla \mathrm{G_P}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right]_\sigma$, $\sigma=\{x,y,z\}$) should be smooth everywhere for $\forall\boldsymbol{r}$ and $\forall\boldsymbol{r}'$, so that it can be interpolated accurately without much difficulty.

To make it more clear, consider the representative situation of $\boldsymbol{r}'$ and $\boldsymbol{r}$ as before. Then, the corresponding function values of $\left[\nabla \mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right]_x$ and $\left[\nabla \mathrm{G_P}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right]_x$ are evaluated, as shown in Fig. 4. Here, $\mathrm{Re}\left[\nabla \mathrm{G_P}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right]_x$ is smooth at $R=0$ (where $\boldsymbol{r}=\boldsymbol{r}'$) as a result of (29), and smooth at the splitting point $R=\delta$ as a result of (30).

Note that the auxiliary function (28) chosen for the splitting of $\mathrm{Re}\left[\nabla \mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right]_\sigma$ ($\sigma=\{x,y,z\}$) is different from the auxiliary function (16) chosen for the splitting of $\mathrm{Re}\left[\mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right]$. Such difference is due to the fact that, unlike $\mathrm{Re}\left[\mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right]$, $\mathrm{Re}\left[\nabla \mathrm{G}\left(\boldsymbol{r},\boldsymbol{r}'\right)\right]_\sigma$ ($\sigma=\{x,y,z\}$) manifest different (i.e., positive or negative) singular behaviors when $\boldsymbol{r}$ approaches $\boldsymbol{r}'$ along different directions (e.g., along $+x$ or $-x$ direction), which can be deduced from the definition of (24), or, more intuitively, observed from Fig. 3 and Fig. 4.

With the splitting form of the Green's function (13) and its gradient (25), the impedance matrix $\boldsymbol{Z}^{\mathrm{CFIE}}$ can be correspondingly split into two parts, yielding

$$\boldsymbol{Z}^{\mathrm{CFIE}} = \boldsymbol{Z}_{\mathrm{E}} + \boldsymbol{Z}_{\mathrm{P}}, \tag{32}$$

where $\boldsymbol{Z}_{\mathrm{E}}$ is related to $\mathrm{G_E}\left(\boldsymbol{r},\boldsymbol{r}'\right)$ and $\nabla \mathrm{G_E}\left(\boldsymbol{r},\boldsymbol{r}'\right)$, while $\boldsymbol{Z}_{\mathrm{P}}$ is related to $\mathrm{G_P}\left(\boldsymbol{r},\boldsymbol{r}'\right)$ and $\nabla \mathrm{G_P}\left(\boldsymbol{r},\boldsymbol{r}'\right)$. For convenience, $\boldsymbol{Z}_{\mathrm{E}}$ and $\boldsymbol{Z}_{\mathrm{P}}$ are termed here as the *evanescent matrix* and the *propagating matrix*, respectively. When the commonly-used Krylov subspace iterative method [35] is employed to solve the matrix equation (4), the matrix-vector product $\boldsymbol{Z}^{\mathrm{CFIE}}\cdot\boldsymbol{I}$ is computed. Using the splitting representation (32), this product can be equivalently implemented as

$$\boldsymbol{Z}^{\mathrm{CFIE}}\cdot\boldsymbol{I} = \boldsymbol{Z}_{\mathrm{E}}\cdot\boldsymbol{I} + \boldsymbol{Z}_{\mathrm{P}}\cdot\boldsymbol{I}. \tag{33}$$

As will be shown in the following subsections, by fully exploiting the properties of the splitting form of the Green's function and its gradient, both $\boldsymbol{Z}_{\mathrm{E}}\cdot\boldsymbol{I}$ and $\boldsymbol{Z}_{\mathrm{P}}\cdot\boldsymbol{I}$ can be evaluated in an efficient manner. For clarity, before going into further discussion, basic concepts and key elements of the presented hybrid fast algorithm are illustrated in Fig. 2.

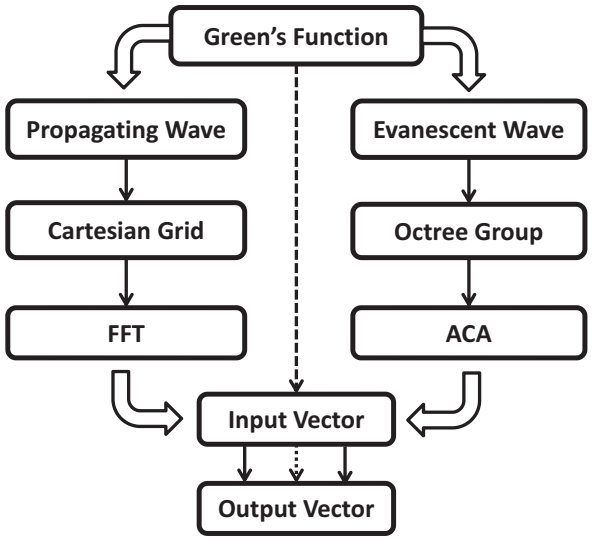


Fig. 2. Systematic diagram of the flow chart and relevant key elements of the presented hybrid fast algorithm. Particularly, the two auxiliary geometric data structures, including the Cartesian grid and the octree grouping, can be constructed independently.

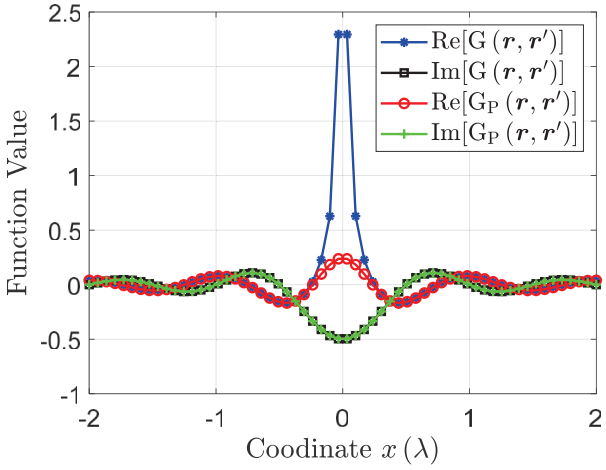


Fig. 3. The function values of $\mathrm{G}(\boldsymbol{r},\boldsymbol{r}')$ and $\mathrm{G_P}(\boldsymbol{r},\boldsymbol{r}')$. The source point $\boldsymbol{r}'$ is fixed to the coordinate origin, and the field point $\boldsymbol{r}$ is moving along the x axis.

C. Propagating matrix calculation

Due to the unique properties (smoothness and translation invariance) of $\mathrm{G_P}(\boldsymbol{r},\boldsymbol{r}')$ and $\nabla\mathrm{G_P}(\boldsymbol{r},\boldsymbol{r}')$, the matrix-vector product $\boldsymbol{Z}_\mathrm{P}\cdot\boldsymbol{I}$ can be calculated accurately and efficiently, with the aid of proper interpolation and the FFT.

To this end, some auxiliary geometric data structures need to be defined first. Specifically, the object considered is enclosed by a three dimensional (3D) rectangular bounding box. Then, this bounding box is subdivided regularly along each Cartesian dimension, yielding a cluster of global Cartesian grids [9–14]. The grid spacings along these Cartesian dimensions are here denoted by h_x, h_y, h_z, respectively. The number of grids along these Cartesian dimensions are denoted by N_x, N_x, N_x, respectively. Thus, a total of $N_\mathrm{Grid}=N_x\times N_y\times N_z$ Cartesian grids are defined.

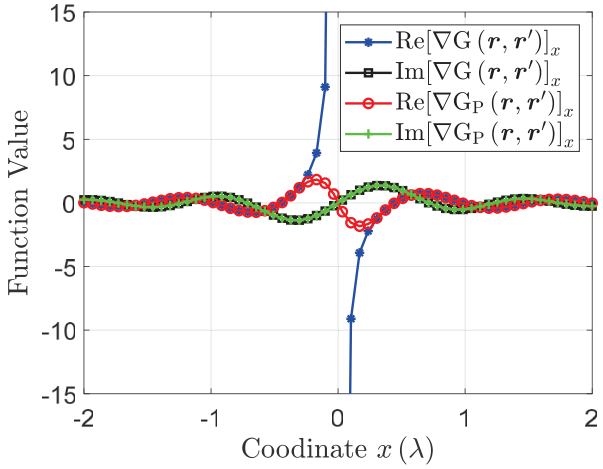


Fig. 4. The function values of $[\nabla\mathrm{G}(\boldsymbol{r},\boldsymbol{r}')]_x$ and $[\nabla\mathrm{G_P}(\boldsymbol{r},\boldsymbol{r}')]_x$. The source point $\boldsymbol{r}'$ is fixed to the coordinate origin, and the field point $\boldsymbol{r}$ is moving along the x axis.

Recall that $\mathrm{G_P}(\boldsymbol{r},\boldsymbol{r}')$ and $\nabla\mathrm{G_P}(\boldsymbol{r},\boldsymbol{r}')$ obtained from the splitting process are smooth everywhere for $\forall\boldsymbol{r}$ and $\forall\boldsymbol{r}'$. Thus, they can be well approximated using the following Cartesian grid-based sampling expansions

$$\mathrm{G_P}(\boldsymbol{r},\boldsymbol{r}')=\sum_{\boldsymbol{u}\in B_m}\sum_{\boldsymbol{v}\in B_n}\beta_{\boldsymbol{u}}(\boldsymbol{r})\,\mathrm{G_P}(\boldsymbol{u},\boldsymbol{v})\,\beta_{\boldsymbol{v}}(\boldsymbol{r}'),\tag{34}$$

$$\nabla\mathrm{G_P}(\boldsymbol{r},\boldsymbol{r}')=\sum_{\mathbf{u}\in B_m}\sum_{\boldsymbol{v}\in B_n}\beta_{\boldsymbol{u}}(\boldsymbol{r})\,\nabla\mathrm{G_P}(\boldsymbol{u},\boldsymbol{v})\,\beta_{\boldsymbol{v}}(\boldsymbol{r}'),\tag{35}$$

for $\forall\boldsymbol{r}$ and $\forall\boldsymbol{r}'$, where B_m and B_n denote the corresponding local expansion boxes that enclose $\boldsymbol{r}$ and $\boldsymbol{r}'$, respectively. Note that, depending on the relative position between $\boldsymbol{r}$ and $\boldsymbol{r}'$, B_m and B_n can be either well-separated or overlapping, as illustrated in Fig. 5. Moreover, $\boldsymbol{u}$ and $\boldsymbol{v}$ denote the coordinates of the Cartesian grids related to B_m and B_n, respectively. Furthermore, $\beta_{\boldsymbol{u}}(\boldsymbol{r})$ (or $\beta_{\boldsymbol{v}}(\boldsymbol{r}')$) denotes the 3D Lagrange interpolation basis function based on the Cartesian grids of B_m (or B_n), which is defined concretely by

$$\beta_{\boldsymbol{u}}(\boldsymbol{r})=\beta_{p_x}(x)\,\beta_{p_y}(y)\,\beta_{p_z}(z),\tag{36}$$

where

$$\beta_{p_\sigma}(\sigma)=\prod_{p_\sigma=0,p_\sigma\neq p'_\sigma}^{M_\sigma}\frac{(\sigma-\sigma_{p'_\sigma})}{(\sigma_{p_\sigma}-\sigma_{p'_\sigma})},\quad \sigma=\{x,y,z\},\tag{37}$$

with M_σ being the expansion order and σ_{p_σ} being the coordinates of the Cartesian grids involved. Clearly,

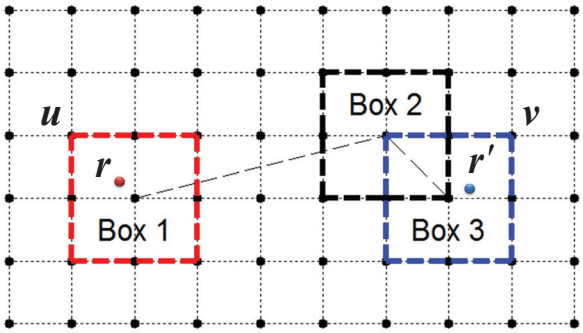


Fig. 5. Computation corresponding to the propagating matrix $\boldsymbol{Z}_{\mathrm{P}}$ can be accelerated by the FFT using the Global Cartesian grid and the local expansion boxes. Typical examples of the expansion boxes are illustrated using different colors. Here, Box 1 and Box 2 are well-separated, while Box 2 and Box 3 are overlapping.

when $M_x = M_y = M_z = M$, a total of $(M+1)^3$ Cartesian grids are involved in (36).

With the expansions (34) and (35), $\boldsymbol{Z}_{\mathrm{P}}^{\mathrm{CFIE}} \cdot \boldsymbol{I}$ can be correspondingly rewritten as

$$\begin{aligned}\boldsymbol{Z}_{\mathrm{P}}^{\mathrm{CFIE}} \cdot \boldsymbol{I} &= j\Phi k\eta \bar{\Pi} \cdot [\mathrm{G_P}(\boldsymbol{u},\boldsymbol{v})] \cdot \bar{\Pi}^T \cdot \boldsymbol{I} \\ &- j\alpha \frac{\eta}{k} \Pi_{\mathrm{d}} \cdot [\mathrm{G_P}(\boldsymbol{u},\boldsymbol{v})] \cdot \Pi_d^T \cdot \boldsymbol{I} \\ &+ (1-\alpha)\,\eta \bar{\Pi}_{\mathrm{g}} \cdot [\nabla \mathrm{G_P}(\boldsymbol{u},\boldsymbol{v})] \times \bar{\Pi}^T \cdot \boldsymbol{I},\end{aligned} \tag{38}$$

where $[\mathrm{G_P}]$ and $[\nabla \mathrm{G_P}]_\sigma$ are triple block Toeplitz matrices of dimension $N_{\mathrm{Grid}} \times N_{\mathrm{Grid}}$. Moreover, $\bar{\Pi}$, Π_{d}, and $\bar{\Pi}_{\mathrm{g}}$ are sparse matrices of dimension $N \times N_{\mathrm{Grid}}$, defined concretely by

$$\bar{\Pi} = \int_S \begin{bmatrix} \mathbf{f}_1(\boldsymbol{r}) \\ \mathbf{f}_2(\boldsymbol{r}) \\ \vdots \\ \mathbf{f}_N(\boldsymbol{r}) \end{bmatrix} \left[\, \beta_1(\boldsymbol{r}),\ \cdots,\ \beta_{N_{\mathrm{Grid}}}(\boldsymbol{r}) \,\right] ds, \tag{39}$$

$$\Pi_{\mathrm{d}} = \int_S \begin{bmatrix} \nabla\cdot\mathbf{f}_1(\boldsymbol{r}) \\ \nabla\cdot\mathbf{f}_2(\boldsymbol{r}) \\ \vdots \\ \nabla\cdot\mathbf{f}_N(\boldsymbol{r}) \end{bmatrix} \left[\, \beta_1(\boldsymbol{r}),\ \cdots,\ \beta_{N_{\mathrm{Grid}}}(\boldsymbol{r}) \,\right] ds, \tag{40}$$

$$\bar{\Pi}_{\mathrm{g}} = \int_S \begin{bmatrix} \hat{\mathbf{n}}\times\mathbf{f}_1(\boldsymbol{r}) \\ \hat{\mathbf{n}}\times\mathbf{f}_2(\boldsymbol{r}) \\ \vdots \\ \hat{\mathbf{n}}\times\mathbf{f}_N(\boldsymbol{r}) \end{bmatrix} \left[\, \beta_1(\boldsymbol{r}),\ \cdots,\ \beta_{N_{\mathrm{Grid}}}(\boldsymbol{r}) \,\right] ds. \tag{41}$$

Exploiting the block Toeplitz structure of $[\mathrm{G_P}]$ and $[\nabla \mathrm{G_P}]_\sigma$, the matrix-vector product $\boldsymbol{Z}_P \cdot \boldsymbol{I}$ can then be evaluated efficiently by the FFT as follows

$$\begin{aligned}\boldsymbol{Z}_{\mathrm{P}}^{\mathrm{CFIE}} \cdot \boldsymbol{I} &= j\alpha k\eta \bar{\Pi} \cdot \mathscr{F}^{-1}\left\{\mathscr{F}\left\{[\mathrm{G_P}]\right\} \cdot \mathscr{F}\left\{\bar{\Pi}^T \cdot \mathbf{I}\right\}\right\} \\ &- j\alpha \frac{\eta}{k} \bar{\Pi}_{\mathrm{d}} \cdot \mathscr{F}^{-1}\left\{\mathscr{F}\left\{[\mathrm{G_P}]\right\} \cdot \mathscr{F}\left\{\bar{\Pi}_d^T \cdot \mathbf{I}\right\}\right\} \\ &+ (1-\alpha)\,\eta \bar{\Pi}_{\mathrm{g}} \cdot \mathscr{F}^{-1}\left\{\mathscr{F}\left\{[\nabla \mathrm{G_P}]\right\} \times \mathscr{F}\left\{\bar{\Pi}^T \cdot \mathbf{I}\right\}\right\},\end{aligned} \tag{42}$$

where $\mathscr{F}$ and $\mathscr{F}^{-1}$ denote the forward and inverse FFT.

Notably, in the conventional pre-correction-based framework [9–13], the original Green's function $\mathrm{G}(\boldsymbol{r},\boldsymbol{r}')$ is approximated with the grid-based sampling expansion. Differently, for the pre-splitting-based framework [32, 33] adopted herein, the function $\mathrm{G_P}(\boldsymbol{r},\boldsymbol{r}')$ is instead approximated with the grid-based sampling expansion, as in (34). Since $\mathrm{G_P}(\boldsymbol{r},\boldsymbol{r}')$ is smooth everywhere without singularity, it is possible to approximate $\mathrm{G_P}(\boldsymbol{r},\boldsymbol{r}')$ with sufficiently good accuracy for arbitrary $\boldsymbol{r}$ and $\boldsymbol{r}'$ by proper interpolation. Consequently, unlike [9–13, 31], the Cartesian grid-based near-field corrections are not necessary here.

D. Evanescent matrix calculation

Referring to (14) and (26), we know that $\boldsymbol{Z}_{\mathrm{E}}$, dominated by $\mathrm{G_E}(\boldsymbol{r},\boldsymbol{r}')$ and $\nabla \mathrm{G_E}(\boldsymbol{r},\boldsymbol{r}')$, has nonzero elements only when $R < \delta$. Therefore, with quasi-uniform meshes, $\boldsymbol{Z}_{\mathrm{E}}$ is a typical sparse matrix with few nonzero elements, and thus it can be stored without much effort. However, in multiscale problems, dense mesh always occurs. Thus, near-field interactions within $R < \delta$ increase substantially. In this case, the nonzero elements of $\boldsymbol{Z}_{\mathrm{E}}$ will increase dramatically.

Fortunately, in the pre-splitting-based framework adopted, this difficulty can be trivially addressed. In a nut shell, we first construct a spatial grouping only based on the geometric position of the basis functions, and then apply the matrix compression methods based on such a grouping to achieve memory reductions of $\boldsymbol{Z}_{\mathrm{E}}$.

Here, a standard adaptive octree is utilized for accomplishing the spatial grouping, and the basic procedures for its construction are outlined as follows.

1. A root cube enclosing the whole object is constructed.
2. The nonempty cubes are subdivided into eight subcubes recursively until the number of basis functions within each cube is smaller than a prescribed constant.
3. An octree is built when the process finishes.

Accordingly, several concepts related to the octree are then defined as follows.

- Those cubes that do not need to be further subdivided are called *leaf cubes*. Here, the leaf cubes are denoted by C_s, with s being the index of some leaf cube. It should be noted that not all of the leaf cubes

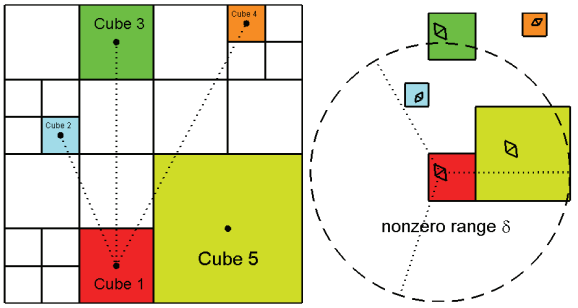


Fig. 6. The storage of $\boldsymbol{Z}_E$ is effectively compressed by using the spatial grouping octree and the matrix compression method. Illustration of the octree grouping is shown in the left subfigure, where some leaf cubes are highlighted using colors. Here, Cube 5 is neighbor cube of Cube 1, while Cube 2, Cube 3, Cube 4 are all well-separated cubes of Cube 1. Besides, the nonzero range of $\mathrm{G_E}(\boldsymbol{r},\boldsymbol{r}')$, characterized by $R<\delta$, is illustrated in the right subfigure. Note that the establishment of the grouping and the application of the compression can be realized independent of the Cartesian grids and the FFT.

are located at the same octree level. Typical examples of the leaf cubes are shown in Fig. 6 and are highlighted with colors.

- Two leaf cubes (e.g., C_s and C_t) are called *neighbor cubes* if they share at least one vertex, such as Cube 1 and Cube 5; otherwise, they are called *well-separated cubes*, such as Cube 1 and Cube 3. For convenience, we use $C_s \cap C_t \neq \emptyset$ to denote that two cubes are neighbor cubes, and use $C_s \cap C_t = \emptyset$ to denote that two cubes are well-separated cubes.

Note that, in [31], the conventional octree grouping cannot be directly applied, and a more involved grouping scheme is developed, in which the cubes for grouping the basis functions are induced from the Cartesian grids. In contrast, here the basis function grouping does not have to be dependent on the Cartesian grids, and can be implemented independently and conventionally. In other words, the commonly-used adaptive octree can be directly employed to achieve the required groupings, and the treatments of the required auxiliary geometric data structures herein are thus simple and straightforward.

On the basis of the spatial groupings above, $\boldsymbol{Z}_{\mathrm{E}}$ can then be rewritten as

$$\boldsymbol{Z}_{\mathrm{E}} = \left\{\boldsymbol{Z}_{\mathrm{E}}^{st}\right\}_{C_s\cap C_t\neq\emptyset} + \left\{\boldsymbol{Z}_{\mathrm{E}}^{st}\right\}_{C_s\cap C_t=\emptyset}, \tag{43}$$

where

- $\left\{\boldsymbol{Z}_{\mathrm{E}}^{st}\right\}_{C_s\cap C_t\neq\emptyset}$ denote those matrix subblocks corresponding to the interactions between neighbor cubes. They are directly computed and stored.
- $\left\{\boldsymbol{Z}_{\mathrm{E}}^{st}\right\}_{C_s\cap C_t=\emptyset}$ denote those matrix subblocks corresponding to the interactions between well-separated cubes. These matrix subblocks can be low-rank. They can be compressed and stored in a compact form.

Here, the ACA [36–38] is employed for accomplishing the matrix compression. Detailed procedures of the ACA can be found in, for example, [37, 38]. Correspondingly, $\left\{\boldsymbol{Z}_{\mathrm{E}}^{st}\right\}_{C_s\cap C_t=\emptyset}$ can be then factorized as

$$\left\{\boldsymbol{Z}_{\mathrm{E}}^{st}\right\}_{C_s\cap C_t=\emptyset} = \left\{\boldsymbol{U}_{\mathrm{E}}^{s}\cdot\boldsymbol{V}_{\mathrm{E}}^{t}\right\}_{C_s\cap C_t=\emptyset}. \tag{44}$$

Thus, the matrix-vector product $\boldsymbol{Z}_{\mathrm{E}}\cdot\boldsymbol{I}$ can be calculated as

$$\begin{aligned}\boldsymbol{Z}_{\mathrm{E}}\cdot\boldsymbol{I} &= \left\{\boldsymbol{Z}_{\mathrm{E}}^{st}\right\}_{C_s\cap C_t\neq\emptyset}\cdot\boldsymbol{I} + \left\{\boldsymbol{Z}_{\mathrm{E}}^{\mathbf{st}}\right\}_{C_s\cap C_t=\emptyset}\cdot\boldsymbol{I}\\ &= \left\{\boldsymbol{Z}_{\mathrm{E}}^{st}\cdot\boldsymbol{I}^{t}\right\}_{C_s\cap C_t\neq\emptyset} + \left\{\boldsymbol{U}_{\mathrm{E}}^{s}\cdot\left(\boldsymbol{V}_{\mathrm{E}}^{\mathbf{t}}\cdot\boldsymbol{I}^{t}\right)\right\}_{C_s\cap C_t=\emptyset}.\end{aligned} \tag{45}$$

Some implementation details for the compression of $\left\{\boldsymbol{Z}_{\mathrm{E}}^{st}\right\}_{C_s\cap C_t=\emptyset}$ are further clarified as follows. In particular, as shown in Fig. 6, due to the localized nature of $\boldsymbol{Z}_{\mathrm{E}}$ decided by $R<\delta$, for a basis function ϕ in Cube 1, all the basis functions in Cube 2 have full interactions with ϕ, while all the basis functions in Cube 4 have zero interactions with ϕ. Meanwhile, only parts of the basis functions in Cube 3 and Cube 5 have nonzero interactions with ϕ. Hence, for two well separated groups with index sets $\boldsymbol{s}$ and $\boldsymbol{t}$, only the submatrices representing nonzero interactions is to be compressed. Thus, the matrix $\boldsymbol{Z}_{\mathrm{E}}^{st}$ which represents the interactions between the two groups is in fact described more precisely by

$$\boldsymbol{Z}_{\mathrm{E}}^{st} = \boldsymbol{P}^{ss}\begin{bmatrix}\boldsymbol{Z}_{\mathrm{E}}^{s't'} & \mathbf{O}\\ \mathbf{O} & \mathbf{O}\end{bmatrix}\boldsymbol{Q}^{tt}, \tag{46}$$

with subsets $\boldsymbol{s}'\subset\boldsymbol{s}$ and $\boldsymbol{t}'\subset\boldsymbol{t}$. Here, $\boldsymbol{P}^{ss}$ and $\boldsymbol{Q}^{tt}$ are permutation matrices, $\boldsymbol{Z}_{\mathrm{E}}^{s't'}$ denotes the matrix subblock corresponding to nonzero interactions. Note that, depending on the relative position between the two cubes considered, there can be many zero interactions in $\boldsymbol{Z}_{\mathrm{E}}^{st}$, due to the sparsity pattern of $\boldsymbol{Z}_{\mathrm{E}}$. Therefore, the matrix compression is in fact applied to $\boldsymbol{Z}_{\mathrm{E}}^{s't'}$ rather than $\boldsymbol{Z}_{\mathrm{E}}^{st}$, and we have

$$\boldsymbol{Z}_{\mathrm{E}}^{s't'} = \boldsymbol{U}_{\mathrm{E}}\boldsymbol{V}_{\mathrm{E}}, \tag{47}$$

where $|\boldsymbol{s}'|$ and $|\boldsymbol{t}'|$ denote the sizes of index sets $\boldsymbol{s}'$ and $\boldsymbol{t}'$, respectively. Meanwhile, the dimension of the matrix $\boldsymbol{U}_{\mathrm{E}}$ is $|\boldsymbol{s}'|\times\Gamma$, and the dimension of the matrix $\boldsymbol{V}_{\mathrm{E}}$ is $\Gamma\times|\boldsymbol{t}'|$. Here, Γ denotes the compression rank of $\boldsymbol{Z}_{\mathrm{E}}^{s't'}$ for a given tolerance. Ultimately, the compressed matrices $\boldsymbol{U}_{\mathrm{E}}$ and $\boldsymbol{V}_{\mathrm{E}}$ in (47) are stored, and are applied to the surface current vector $\boldsymbol{I}$ in an efficient manner. For many practical problems that involve multiscale discretization, the matrix block $\boldsymbol{Z}_{\mathrm{E}}^{s't'}$ can be effectively compressed, and the memory consumption for $\boldsymbol{U}_{\mathrm{E}}$ and $\boldsymbol{V}_{\mathrm{E}}$ are greatly reduced compared to that for $\boldsymbol{Z}_{\mathrm{E}}^{st}$.

III. NUMERICAL RESULTS

In this section, the computational performance of the proposed hybrid algorithm is demonstrated through sev-

eral numerical examples. The conjugate gradient method (CG) [35] is used as the iterative solver, and the tolerance is set to 1.0E-3. The diagonal preconditioner is used to improve the convergence. Unless otherwise stated, the grid spacings (i.e., h_x, h_y, h_z) for the FFT acceleration are by default set to the common value $h = 0.1\ \lambda$, and the expansion orders (i.e., M_x, M_y, M_z) are by default set to the common value $M = 3$. Furthermore, the well-known FFTW [34] is used to perform the FFT. For clarity, in the following, the traditional pre-splitting-FFT-based algorithm which employs the uncompressed $\boldsymbol{Z}_{\mathrm{E}}$ will be termed as the PSG-FFT algorithm, and the proposed hybrid fast algorithm (enhanced with the ACA) will be referred to as the PSG-FFT-ACA algorithm.

A. Sphere

A PEC sphere of radius 1.0 λ is first considered. The surface of the sphere is discretized with multiscale triangular meshes, resulting in 34,764 RWG functions. In particular, as shown in Fig. 7, about a quarter of the surface is discretized with dense meshes of edge lengths about 0.02 λ, and the rest is discretized with regular meshes of edge lengths about 0.1 λ.

For this example, the standard MoM (without using any acceleration technique), the traditional PSG-FFT algorithm, and the proposed PSG-FFT-ACA algorithm are adopted as the the solution methods. After solving the corresponding matrix equations as formulated in (4), the surface current coefficients (i.e, the solutions of equation (4)) $\boldsymbol{I}^{\mathrm{MoM}}$, $\boldsymbol{I}^{\mathrm{PSG\text{-}FFT}}$, and $\boldsymbol{I}^{\mathrm{PSG\text{-}FFT\text{-}ACA}}$ are then obtained, respectively. Here, it should be noticed that both $\mathbf{I}^{\mathrm{PSG\text{-}FFT}}$ and $\boldsymbol{I}^{\mathrm{PSG\text{-}FFT\text{-}ACA}}$ depend on the choice of the pre-splitting threshold δ. However, when the standard MoM is used, δ is not involved, and $\boldsymbol{I}^{\mathrm{MoM}}$ is therefore independent of δ. Thus, the solution $\boldsymbol{I}^{\mathrm{MoM}}$ can be used as a proper reference, to evaluate the accuracy of $\boldsymbol{I}^{\mathrm{PSG\text{-}FFT}}$ and $\boldsymbol{I}^{\mathrm{PSG\text{-}FFT\text{-}ACA}}$ for varying δ. The relative errors of the surface current coefficients are defined as follows

$$\mathscr{E} = \frac{\left\|\boldsymbol{I}^{\mathrm{Test}} - \boldsymbol{I}^{\mathrm{Reference}}\right\|_2}{\left\|\boldsymbol{I}^{\mathrm{Reference}}\right\|_2}, \tag{48}$$

where $\boldsymbol{I}^{\mathrm{Test}}$ denotes the surface current coefficients to be tested, $\boldsymbol{I}^{\mathrm{Reference}}$ denotes some specified surface current coefficients to be used as the reference, and $\|\cdot\|_2$ denotes the 2-norm of vector. Based on (48), the relative errors of both the traditional PSG-FFT algorithm and the proposed PSG-FFT-ACA algorithm for several different δ are calculated and shown in Fig. 8. We can see that both algorithms can achieve accurate results and show similar accuracy level. In particular, within the range of the δ between 0.15 λ and 0.65 λ, the errors of both algorithms decrease with increasing δ. Furthermore, the accuracies of both algorithms are not strongly sensitive with respect to the δ considered.

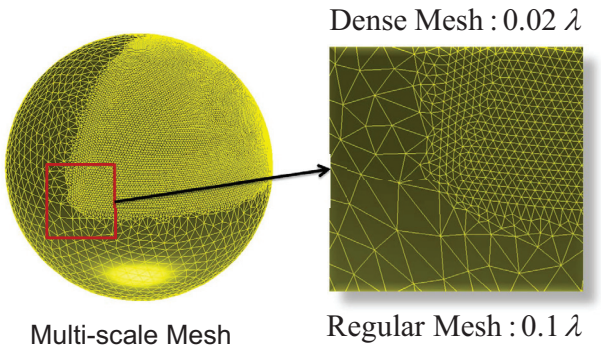


Fig. 7. Mesh configurations of the PEC sphere. The radius of the sphere is 1.0 λ. Here, about a quarter of the spherical surface is discretized with a relatively dense mesh of 0.02 λ, and the rest of the surface is discretized with a regular mesh of 0.1 λ. For clarity, the enlarged view of the multiscale mesh is also visualized in the above.

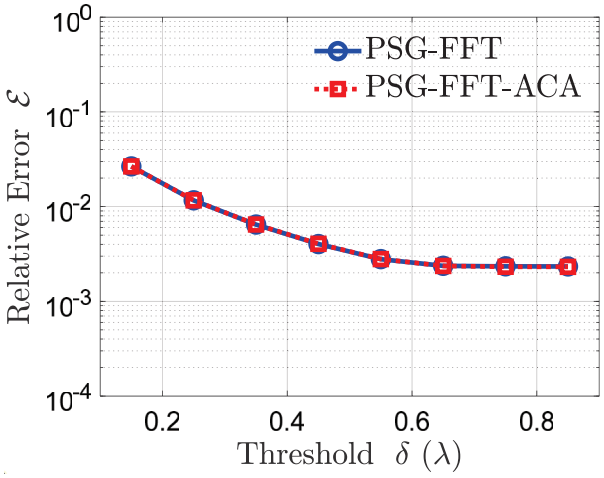


Fig. 8. Relative errors $\mathscr{E}$ of the surface currents coefficients $\boldsymbol{I}$ with respect to different splitting threshold δ. The relative errors of both the traditional PSG-FFT algorithm and the proposed PSG-FFT-ACA algorithm are illustrated. The surface currents coefficients obtained using the standard MoM are used as the reference.

In the following, the choice of δ is discussed. In principle, δ can be of any value. However, in practice, a moderate value of δ is suggested, due to the following considerations.

- If δ is chosen be relatively large, the number of nonzero elements of the evanescent matrix $\boldsymbol{Z}_{\mathrm{E}}$ will then become relatively large. Clearly, in this case, the corresponding computational cost will increase.

In Table 1, the computational costs with respect to δ are recorded. It can be observed that the time and memory consumptions of both the traditional PSG-FFT algorithm and the proposed PSG-FFT-ACA algorithm increase as δ increase. Hence, due to efficiency considerations, δ is not suggested to be too large.

- If δ is chosen to be relatively small, the number of nonzero elements of the evanescent matrix $\boldsymbol{Z}_{\mathrm{E}}$ will become relatively small, and the computational cost is correspondingly reduced. However, from Fig. 8, it is observed that the accuracy of the solution is slightly reduced when δ becomes smaller. Hence, due to accuracy considerations, δ is not suggested to be too small as well.

Based on the above considerations, the δ between 0.15 λ and 0.65 λ can be proper choice. In practice, by tuning the splitting threshold δ within this range, we can trade off between accuracy and computational cost.

Besides, the well-studied IE-FFT algorithm [17], being a typical instance of the pre-correction-based methods, is also used to calculate this example. Similar to the configurations of the pre-splitting-based methods employed, the grid spacing h for the IE-FFT is set to 0.1 λ and the expansion order M is set to 3. Then, using $\boldsymbol{I}^{\mathrm{MoM}}$ as the reference, the relative error of the obtained surface current coefficients $\boldsymbol{I}^{\mathrm{IE\text{-}FFT}}$ is calculated, which is 2.96E-3. Referring to Fig. 8, we can see that such error is close to the error of the PSG-FFT(-ACA) algorithm with δ about 0.45 λ. Moreover, for the IE-FFT algorithm, the memory requirement is 3.09718 G and the CPU time required is 1.08727 m. Referring to Table 1, we can see that such cost is slightly lower than that of the PSG-FFT algorithm with δ about 0.45 λ. From the above numerical results, it can be known that the computational performance of the IE-FFT algorithm during the solution stage is slightly better than that of the PSG-FFT algorithm. However, it should be emphasized that such slight advantage of the IE-FFT algorithm is at the price of the additional grid-based numerical correction operations during the setup (precomputation) stage.

Table 1: Time and memory for the sphere model with different splitting threshold δ

$\delta(\lambda)$	PSG-FFT		PSG-FFT-ACA	
	Memory (GB)	CPU Time (**m**)	Memory (GB)	CPU Time (**m**)
0.15(λ)	0.9924	0.6601	0.9917	0.7087
0.25(λ)	1.9707	0.8078	1.6581	0.8167
0.35(λ)	3.1454	1.1445	2.2623	0.8679
0.45(λ)	4.6662	1.2814	2.9167	0.9177
0.55(λ)	6.3642	2.0607	3.5851	0.9812
0.65(λ)	8.3831	1.8004	4.2083	1.0390
0.75(λ)	10.514	2.1802	4.3177	1.0939
0.85(λ)	11.528	2.4548	5.1668	1.2197

B. Monopole antenna on top of large platform

In the following, a more realistic example is considered, i.e., a monopole antenna on top of a large platform, as shown in Fig. 9. Clearly, this is a typical multiscale structure and often encountered in many practical scenarios. The width and height of the platform are 4.0 λ and 0.3 λ, respectively. Moreover, the width and height of the monopole antenna are 0.06 λ and 0.6 λ, respectively. The object surface is discretized with the multiscale triangular meshes. Specifically, the monopole antenna is discretized with meshes of edge length about 0.01 λ, while the large platform is discretized with meshes of edge length about 0.1 λ. With such mesh, a total of 24,807 RWG functions are generated.

The standard MoM and the proposed PSG-FFT-ACA algorithm are employed to calculate this example. Here, the pre-splitting threshold δ is chosen to be 0.35 λ. Using $\boldsymbol{I}^{\mathrm{MoM}}$ as the reference, the relative errors of $\boldsymbol{I}^{\mathrm{PSG\text{-}FFT\text{-}ACA}}$ are calculated for different grid spacings h and different expansion orders M. For clarity, the relative errors are illustrated in Fig. 10. The corresponding time and memory consumptions are recorded in Table 2. It can be clearly seen that, for decreasing grid spacing h and increasing expansion order M, the relative error decreases and the corresponding computational cost increases. In other words, by tuning h and M, we can trade off between accuracy and computational cost.

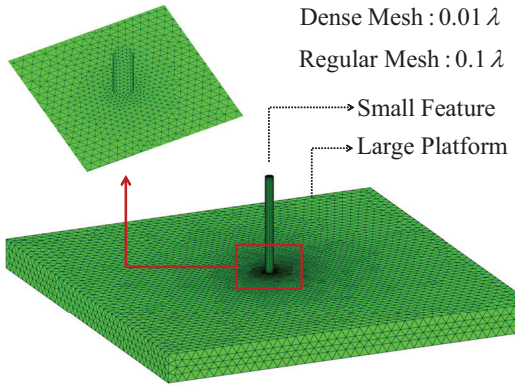


Fig. 9. Mesh configurations of a small monopole antenna on top of a large platform. A multiscale mesh is used to discretize this object. The monopole antenna is discretized using dense mesh of about 0.01 λ, while the large platform is discretized with regular mesh of about 0.1 λ.

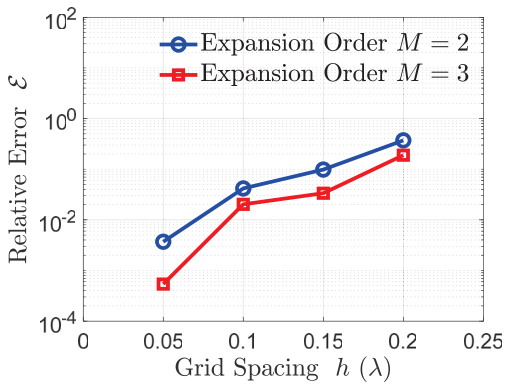


Fig. 10. Relative errors $\mathscr{E}$ of the surface currents coefficients with respect to different grid spacings h and expansion orders M. The relative errors of the proposed PSG-FFT-ACA algorithm are illustrated. The surface currents coefficients obtained using the standard MoM are used as the reference.

Table 2: Time and memory for the monopole-platform model with different grid spacing h and expansion order M

$h(\lambda)$	$M = 2$		$M = 3$	
	Memory (GB)	CPU Time (**m**)	Memory (GB)	CPU Time (**m**)
$0.05(\lambda)$	1.0074	0.7863	1.1094	0.8256
$0.10(\lambda)$	0.8643	0.2280	0.9662	0.2792
$0.15(\lambda)$	0.8311	0.1554	0.9304	0.2106
$0.20(\lambda)$	0.8243	0.1155	0.9261	0.1801

C. Cone-shape object

To further study the effectiveness of the matrix compression enhancement, a PEC cone model is then considered, as shown in Fig. 11. The height of the cone is 3.0 λ. The radii of the top and the bottom faces are 0.03 λ and 0.6 λ, respectively. The surface is discretized with a gradually-varied meshes. In particular, three mesh configurations are considered. For the three cases, the edge length for the regular part of the mesh is fixed to 0.1 λ, while the edge lengths for the dense part of the mesh are set to 0.01 λ, 0.005 λ, and 0.0025 λ, respectively. For clarity, the mesh settings considered are illustrated in Fig. 11.

For each of these mesh configurations, we solve the corresponding equations first using the PSG-FFT algorithm which employs the uncompressed $\boldsymbol{Z}_{\mathrm{E}}$. Then, we recalculate these problems using the PSG-FFT-ACA algorithm which compresses $\boldsymbol{Z}_{\mathrm{E}}$ with the ACA. Here, the

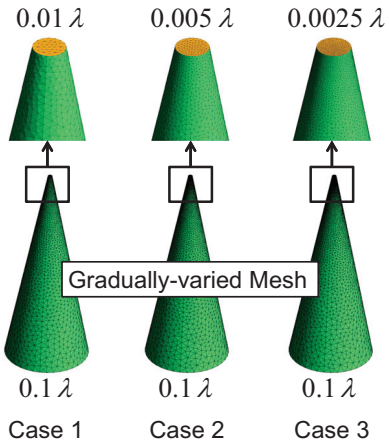


Fig. 11. Mesh configurations of the PEC cone. A gradually-varied mesh is used to discretize the cone-shape object. The edge length for the regular part of the mesh is 0.1 λ. The edge lengths for the dense part of the mesh are 0.01 λ, 0.005 λ, and 0.0025 λ, respectively. For clarity, the enlarged views are also shown in the above.

splitting threshold δ is chosen to be 0.35 λ. Notice that, with such configuration of δ, all the interactions corresponding to $\boldsymbol{Z}_{\mathrm{E}}$ occur within the subwavelength regime. Thus, the ACA compression can be applied effectively [21, 37]. In Table 3, 4, 5, we tabulate the time and memory consumptions of both the PSG-FFT algorithm and the PSG-FFT-ACA algorithm for the three mesh cases considered. It can be seen that, for all the cases, the PSG-FFT-ACA algorithm shows improved computational performances relative to the PSG-FFT algorithm.

Table 3: Time and memory for the cone model of case 1

Algorithm	Memory (GB)		CPU Time	
	Z_{E}	Total	$Z_{\mathrm{E}} \cdot I$ (s)	Total (**m**)
PSG-FFT	0.3313	0.6286	12.128	0.6955
PSG-FFT-ACA	0.2418	0.4738	8.3050	0.6194

Table 4: Time and memory for the cone model of case 2

Algorithm	Memory (GB)		CPU Time	
	Z_{E}	Total	$Z_{\mathrm{E}} \cdot I$ (s)	Total (m)
PSG-FFT	0.9611	1.6176	69.751	2.3978
PSG-FFT-ACA	0.4935	0.8778	35.127	1.8475

Table 5: Time and memory for the cone model of case 3

Algorithm	Memory (GB)		CPU Time	
	Z_E	Total	$Z_E \cdot I$ (m)	Total (m)
PSG-FFT	2.6465	4.2615	181.47	4.4874
PSG-FFT-ACA	1.1579	2.0570	78.678	2.8009

Further inspections show that the degree of improvements becomes more prominent when the dense mesh is finer. To make this more clear, we calculate the compression ratios of the evanescent matrix $\boldsymbol{Z}_E$ and use this as an indicator to measure the degree of improvements of the matrix compression enhancement. In Fig. 12, the compression ratios with respect to different mesh configurations are illustrated. We can clearly see that the compression ratio becomes higher when the dense mesh becomes finer. That is, the effectiveness of the proposed hybrid algorithm (with matrix compression enhancement) becomes more prominent when the ratio between the regular mesh size and the dense mesh size is larger.

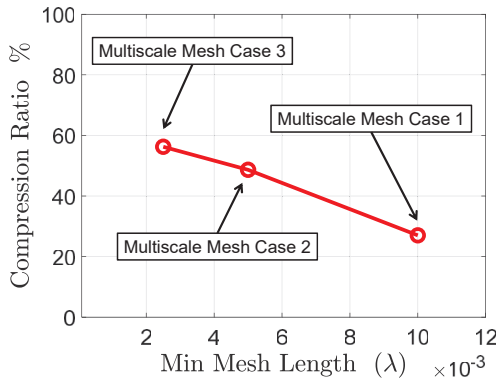


Fig. 12. Compression ratios of the cone with different mesh configurations.

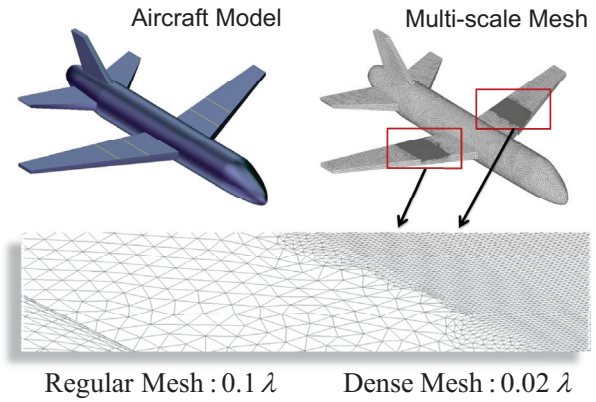


Fig. 13. Geometry model and corresponding multiscale mesh of the PEC aircraft. Here, a part of the aircraft surface on the wings are discretized with a relatively dense mesh of 0.02 λ, while the rest of the aircraft surface is discretized with a regular mesh of 0.1 λ. The enlarged view of the multiscale mesh is also highlighted in the above.

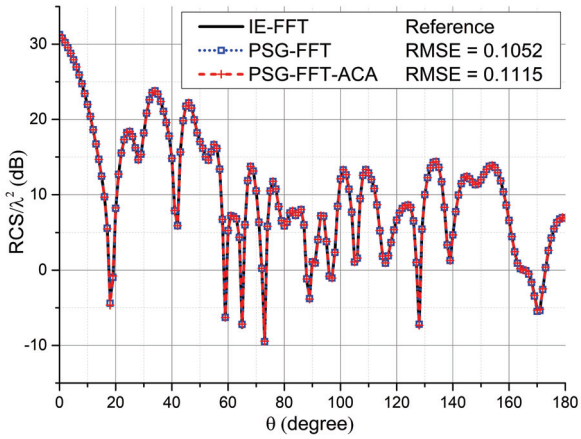


Fig. 14. Bistatic RCS curves of the PEC aircraft model. The result obtained by the IE-FFT is used as the reference. The RMSE of the PSG-FFT and the PSG-FFT-ACA are calculated.

D. Aircraft model

A relatively large PEC aircraft model with multiscale meshes is finally considered, as shown in Fig. 13. In this model, both the length and the wingspan of the aircraft are about 16 λ. The surface is discretized with triangular meshes, resulting in 120,072 RWG functions. Here, a part of the aircraft surface on the wings is discretized with dense meshes of edge lengths about 0.02 λ, and the rest of the aircraft surface is discretized with regular meshes of edge lengths about 0.1 λ.

The bistatic RCS curves obtained by the PSG-FFT algorithm and the PSG-FFT-ACA algorithm are illustrated in Fig. 14. The pre-splitting threshold δ is here chosen to be 0.35 λ. The bistatic RCS curve obtained by the well-tested IE-FFT algorithm [17] is also shown in the figure and used as the reference to calculate the root mean square error (RMSE). We can see that these curves agree well with each other.

The time and memory consumptions of both the PSG-FFT algorithm and the PSG-FFT-ACA algorithm are tabulated in Table 6. From the table, we can see that the memory requirement of the evanescent matrix $\boldsymbol{Z}_E$ in the PSG-FFT-ACA algorithm is reduced by about 34.1218%, compared with that of the PSG-FFT

Table 6: Time and memory for the aircraft model

Algorithm	Memory (GB)		CPU Time	
	Z_{E}	Total	$Z_{\mathrm{E}} \cdot I$ (s)	Total (**m**)
PSG-FFT	3.5558	7.7408	76.910	23.580
PSG-FFT-ACA	2.3425	5.4033	48.053	21.431

algorithm. Meanwhile, the time cost for performing $\boldsymbol{Z}_{\mathrm{E}} \cdot \boldsymbol{I}$ in the whole solution process is reduced by about 37.52%. Thus, the time and memory cost related to the evanescent matrix $\boldsymbol{Z}_{\mathrm{E}}$ can be greatly reduced, as a result of the matrix compression enhancement.

Nevertheless, the total cost of the overall algorithm are determined not only by the cost related to the evanescent matrix $\boldsymbol{Z}_{\mathrm{E}}$, but also determined by the cost related to the propagating matrix $\boldsymbol{Z}_{\mathrm{P}}$. For this relatively large object, the dense mesh region only occupies a relatively small portion of the overall object surface, and the ratio between the edge length (0.1 λ) of the regular mesh and the edge length (0.02 λ) of the dense mesh is here not extremely large. Thus, the time cost for performing $\boldsymbol{Z}_{\mathrm{P}} \cdot \boldsymbol{I}$ still dominates the time cost for performing $\boldsymbol{Z} \cdot \boldsymbol{I}$, and the reduction of the time cost for performing $\boldsymbol{Z}_{\mathrm{E}} \cdot \boldsymbol{I}$ therefore only has a minor influence on the total time cost for performing $\boldsymbol{Z} \cdot \boldsymbol{I}$. As a result, although the total time can be reduced due to the reduced cost related to $\boldsymbol{Z}_{\mathrm{E}}$, the reduction of the total time does not look very striking for this example.

IV. CONCLUSIONS

In this paper, a hybrid fast algorithm has been established for efficiently analyzing multiscale problems. In this algorithm, the Green's function is a priori split into a singular but spatially localized part, and a smooth, oscillatory but bandlimited part. Then, the fundamental blocks for acceleration, i.e., the FFT and the matrix compression method (ACA), are applied to these two parts, respectively. Compared with the traditional methods which only employ the FFT for acceleration, the proposed hybrid algorithm can maintain low memory consumption in multiscale cases without compromising time cost, thus manifesting itself as a favorable multiscale extension of the traditional pure FFT-based methods. Furthermore, the required auxiliary geometric data structures herein can be constructed independent of each other, thus permitting the two kinds of acceleration methods to be connected seamlessly in a concise and elegant manner.

On the basis of the present work, several future directions may be explored. First, except for the Lagrange interpolation used in this work, other more accurate interpolation schemes (such as the Gaussian interpolation [15] or the fitting techniques [12]) can be employed to further improve the computational performance of the overall algorithm. Moreover, although only the PEC problems are considered here, extensions to the dielectric problems (especially inhomogeneous cases [16]) are clearly feasible. Furthermore, effective parallelization (especially on heterogeneous architectures [13]) of the proposed hybrid algorithm can be interesting subject of future research.

APPENDIX

The derivative functions of $\mathrm{Re}\left[\mathrm{G}\left(R\right)\right]$, which are used in (18) and (19), are given below

$$\frac{d}{dR}\mathrm{Re}\left[\mathrm{G}\left(R\right)\right] = -\frac{\cos\left(kR\right)+kR\sin\left(kR\right)}{4\pi R^2}, \tag{49}$$

$$\frac{d^2}{dR^2}\mathrm{Re}\left[\mathrm{G}\left(R\right)\right] = -\frac{\left(k^2R^2-2\right)\cos\left(kR\right)-2kR\sin\left(kR\right)}{4\pi R^3}. \tag{50}$$

The derivative functions of $\mathrm{Re}\left[\ddot{\mathrm{G}}\left(R\right)\right]$, which are used in (30) and (31), are given below

$$\frac{d}{dR}\mathrm{Re}\left[\ddot{\mathrm{G}}\left(R\right)\right] = -\frac{\left(k^2R^2-2\right)\cos\left(kR\right)-2kR\sin\left(kR\right)}{4\pi R^3}, \tag{51}$$

$$\frac{d^2}{dR^2}\mathrm{Re}\left[\ddot{\mathrm{G}}\left(R\right)\right] = -\frac{3\left(k^2R^2-2\right)\cos\left(kR\right)-kR\left(k^2R^2-6\right)\sin\left(kR\right)}{4\pi R^4}. \tag{52}$$

ACKNOWLEDGMENT

This work was supported in part by the Fundamental Research Funds for the Central Universities (No. 2242022k30003), in part by the National Key Research and Development Project (No. 2019YFB1803202), and in part by the National Science Foundation of China (Nos. 62293492, 62131008, 61975177, U20A20164).

REFERENCES

[1] W. C. Chew and L. J. Jiang, "Overview of large-scale computing: the past, the present, and the future," *Proceedings of the IEEE*, vol. 101, no. 2, pp. 227-241, Feb. 2013.

[2] M.-K. Li and W. C. Chew, "Multiscale simulation of complex structures using equivalence principle algorithm with high-order field point sampling scheme," *IEEE Trans. Antennas Propag.*, vol. 56, no. 8, pp. 2389-2397, 2007.

[3] Z.-G. Qian and W. C. Chew, "Fast full-wave surface integral equation solver for multiscale structure modeling," *IEEE Trans. Antennas Propag.*, vol. 57, no. 11, pp. 3594-3601, Nov. 2009.

[4] J. Zhu, S. Omar, and D. Jiao, "Solution of the electric field integral equation when it breaks down," *IEEE Trans. Antennas Propag.*, vol. 62, no. 8, pp. 4122-4134, Aug. 2014.

[5] R. F. Harrington, *Field Computation by Moment Methods*. New York: The MacMillian Co., 1968.

[6] J. M. Song and W. C. Chew, "Multilevel fast-multipole algorithm for solving combined field integral equations of electromagnetic scattering," *Microw. Opt. Technol. Lett.*, vol. 10, no. 1, pp. 14-19, Sep. 1995.

[7] J.-M. Song, C.-C. Lu, and W. C. Chew, "Multilevel fast multipole algorithm for electromagnetic scattering by large complex objects," *IEEE Trans. Antennas Propag.*, vol. 45, no. 10, pp. 1488-1493, Oct. 1997.

[8] W. C. Chew, J.-M. Jin, E. Michielssen, and J. Song, *Fast and Efficient Algorithms in Computational Electromagnetics*. Boston, MA: Artech House, 2001.

[9] M. Bleszynski, E. Bleszynski, and T. Jaroszewicz, "AIM: Adaptive integral method for solving large-scale electromagnetic scattering and radiation problems," *Radio Sci.*, vol. 31, no. 5, pp. 1225-1251, 1996.

[10] J. R. Phillips and J. K. White, "A precorrected-FFT method for electrostatic analysis of complicated 3-D structures," *IEEE Trans. Computer-Aided Design*, vol. 16, no. 10, pp. 1059-1072, Oct. 1997.

[11] S. M. Seo and J.-F. Lee, "A fast IE-FFT algorithm for solving PEC scattering problems," *IEEE Trans. Magn.*, vol. 41, no. 5, pp. 1476-1479, May 2005.

[12] J.-Y. Xie, H.-X. Zhou, W. Hong, W.-D. Li, and G. Hua, "A highly accurate FGG-FG-FFT for the combined field integral equation," *IEEE Trans. Antennas Propag.*, vol. 61, no. 9, pp. 4641-4652, Sep. 2013.

[13] S. Peng and C.-F. Wang, "Precorrected-FFT method on graphics processing units," *IEEE Trans. Antennas Propag.*, vol. 61, no. 4, pp. 2099-2107, Apr. 2013.

[14] L.-W. Li, Y.-J. Wang, and E.-P. Li, "MPI-based parallelized precorrected FFT algorithm for analyzing scattering by arbitrarily shaped three-dimensional objects," *Electromagn. Waves Appl.*, vol. 17, pp. 1489-1491, Jan. 2003.

[15] B. Lai, X. An, H.-B. Yuan, Z.-S. Chen, C.-M. Huang, and C.-H. Liang, "A novel Gaussian interpolation formula-based IE-FFT algorithm for solving EMscattering problems," *Microw. Opt. Tech. Lett.*, vol. 51, no. 9, pp. 2233-2236, Sep. 2009.

[16] X.-C. Nie, L.-W. Li, N. Yuan, T. S. Yeo, and Y.-B. Gan, "Precorrected-FFT solution of the volume integral equation for 3-D inhomogeneous dielectric objects," *IEEE Trans. Antennas Propag.*, vol. 53, no. 1, pp. 313-320, Jan. 2005.

[17] J.-Y. Xie, H.-X. Zhou, W.-D. Li, and W. Hong, "IE-FFT for the combined field integral equation applied to electrically large objects," *Microw. Opt. Tech. Lett.*, vol. 54, no. 4, pp. 891-896, Apr. 2012.

[18] S. Peng, C.-F. Wang, and J.-F. Lee, "Analyzing PEC scattering structure using an IE-FFT algorithm," *ACES Journal*, vol. 24, no. 2, Apr. 2009.

[19] L. J. Jiang and W. C. Chew, "A mixed-form fast multipole algorithm," *IEEE Trans. Antennas Propag.*, vol. 53, no. 12, pp. 4145-4156, Dec. 2005.

[20] X.-M. Pan, J.-G. Wei, Z. Peng, and X.-Q. Sheng, "A fast algorithm for multiscale electromagnetic problems using interpolative decomposition and multilevel fast multipole algorithm," *Radio Sci.*, vol. 47, no. 1, Feb. 2012.

[21] L.-F. Ma, Z. Nie, J. Hu, and S.-Q. He, "Combined MLFMA-ACA algorithm application to scattering problems with complex and fine structure," *Asia Pacific Microwave Conference*, pp. 802-805, 2009.

[22] W.-B. Kong, H.-X. Zhou, K.-L. Zheng, and W. Hong, "Analysis of multiscale problems using the MLFMA with the assistance of the FFT-based method," *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 9, pp. 4184-4188, 2015.

[23] J. S. Zhao and W. C. Chew, "Three dimensional multilevel fast multipole algorithm from static to electrodynamic," *Microw. Opt. Technol. Lett.*, vol. 26, no. 1, pp. 43-48, 2000.

[24] L. J. Jiang and W. C. Chew, "Low-frequency fast inhomogeneous plane-wave algorithm," *Microw. Opt. Tech. Lett.*, vol. 40, no. 2, pp. 117-122, Jan. 2004.

[25] T. Xia, L. L. Meng, Q. S. Liu, H. H. Gan, and W. C. Chew, "A low-frequency stable broadband multilevel fast multipole algorithm using plane wave multipole hybridization," *IEEE Trans. Antennas Propag.*, vol. 66, no. 11, pp. 6137-6145, Nov. 2018.

[26] H. Wallen and J. Sarvas, "Translation procedures for broadband MLFMA," *Prog. Electromagn. Res.*, vol. 55, pp. 47-78, 2005.

[27] D. Wulf and R. Bunger, "An efficient implementation of the combined wideband MLFMA/LF-FIPWA," *IEEE Trans. Antennas Propag.*, vol. 57, no. 2, pp. 467-474, Feb. 2009.

[28] D. T. Schobert and T. F. Eibert, "Fast integral equation solution by multilevel Green's function interpolation combined with multilevel fast multipole method," *IEEE Trans. Antennas Propag.*, vol. 60, no. 9, pp. 4458-4463, Sep. 2012.

[29] I. Bogaert, J. Peeters, and F. Olyslager, "A nondirective plane wave MLFMA stable at low frequencies," *IEEE Trans. Antennas Propag.*, vol. 56, no. 12, pp. 3752-3767, Dec. 2008.

[30] O. Ergul and B. Karaosmanoglu, "Broadband multilevel fast multipole algorithm based on an approximate diagonalization of the Green's function," *IEEE Trans. Antennas Propag.*, vol. 63, no. 7, pp. 3035-3041, July 2015.

[31] W.-B. Kong, H.-X. Zhou, K.-L. Zheng, X. Mu, and W. Hong, "FFT-based method with near-matrix compression," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 11, pp. 5975-5983, 2017.

[32] A. Bespalov, "Cost-effective solution of the boundary integral equations for 3D Maxwell problems," *Russ. J. Numer. Anal. Math. Modelling*, vol. 14, no. 5, pp. 403-428, 1999.

[33] A. Bespalov, "On the use of a regular grid for implementation of boundary integral methods for wave problems," *Russ. J. Numer. Anal. Math. Modelling*, vol. 15, no. 6, pp. 469-488. 2000.

[34] M. Frigo and S. Johnson, FFTW Manual [Online]. Available: http://www.fftw.org/

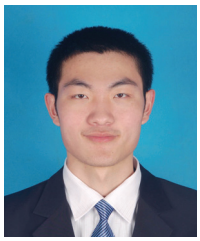
[35] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. M. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. Van der Vorst, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods.* Philadelphia, PA, USA: SIAM, 1994.

[36] M. Bebendorf and S. Rjasanow, "Adaptive low-rank approximation of collocation matrices," *Computing*, vol. 70, no. 1, pp. 1-24, 2003.

[37] K. Zhao, M. N. Vouvakis, and J.-F. Lee, "The adaptive cross approximation algorithm for accelerated method of moments computations of EMC problems," *IEEE Trans. Electromag. Compat.*, vol. 47, no. 4, pp. 763-773, Nov. 2005.

[38] H.-X. Zhou, G.-Y. Zhu, W.-B. Kong, and W. Hong, "An upgraded ACA algorithm in complex field and its statistical analysis," *IEEE Trans. Antennas Propag.*, vol. 65, no. 5, pp. 2734-2739, May 2017.

[39] S. M. Rao, D. R. Wilton, and A. W. Glisson, "Elettromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas Propag.*, vol. 30, no. 3, pp. 409-418, May 1982.



Guang-Yu Zhu Guang-Yu Zhu received the B.D. in electrical engineering from Nanjing Agricultural University, Nanjing, China, in 2012, and the Ph.D. degree in radio engineering from Southeast University, Nanjing, China, in 2020, respectively. He is currently a postdoctoral research fellow with the Department of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China. His research interests are in computational electromagnetics with focus on fast algorithms and domain decomposition methods.

Wei-Dong Li Wei-Dong Li received the M.S. degree in mathematics and the Ph.D. degree in radio engineering from Southeast University, Nanjing, China, in 2003 and 2007, respectively. From January 2008 to January 2009, he was a visiting scholar with the Technische University Darmstadt, Germany. He is currently an associate professor with the State Key Laboratory of Millimeter Waves and the School of Information Science and Engineering, Southeast University. He has authored or coauthored over 40 technical articles. He serves as a reviewer for the *IEEE Transactions on Antennas and Propagation* and *IET Microwave, Antennas and Propagation*. His research interests are in computational EM with focus on integral-equation-based overlapped domain decomposition method, multilevel fast multipole algorithm, fast and accurate inter/extrapolation techniques, and basic functions.

Wei E. I. Sha Wei E.I. Sha received the B.S. and Ph.D. degrees in electronic engineering from Anhui University, Hefei, China, in 2003 and 2008, respectively. From July 2008 to July 2017, he was a postdoctoral research fellow and then a research assistant professor with the Department of Electrical and Electronic Engineering, University of Hong Kong, Hong Kong. From March 2018 to March 2019, he worked as a Marie Skodowska-Curie Individual Fellow with University College London, London, U.K. From October 2017, he joined the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China, where

he is currently a tenure-tracked assistant professor. He has authored or coauthored 180 refereed journal articles, 150 conference publications (including five keynote talks and one short course), nine book chapters, and two books. His Google Scholar citation is 8193 with H-index of 45. His research interests include theoretical and computational research in electromagnetics and optics, focusing on the multiphysics and interdisciplinary research. His research involves fundamental and applied aspects in computational and applied electromagnetics, nonlinear and quantum electromagnetics, micro- and nano-optics, optoelectronic device simulation, and multiphysics modeling. Dr. Sha is a member of Optical Society of America (OSA). He served as a reviewer for 60 technical journals and technical program committee member for ten IEEE conferences. He was a recipient of the Applied Computational Electromagnetics Society (ACES) Technical Achievement Award in 2022 and PIERS Young Scientist Award in 2021. In 2015, he was awarded Second Prize of Science and Technology from Anhui Province Government, China. In 2007, he was awarded the Thousand Talents Program for Distinguished Young Scholars of China. He also received six Best Student Paper Prizes and one Young Scientist Award with his students. He also served as an associate editor for *IEEE Journal on Multiscale and Multiphysics Computational Techniques*, *IEEE Open Journal of Antennas and Propagation*, and *IEEE Access*.

Hou-Xing Zhou Hou-Xing Zhou received the M.S. degree in mathematics from Southwest Normal University, Chongqing, China, in 1995, and the Ph.D. degree in radio engineering from Southeast University, Nanjing, China, in 2002. Since 2002, he has been with the State Key Laboratory of Millimeter Waves, Southeast University. He is currently a professor with the School of Information Science and Engineering, Southeast University. He has authored or coauthored 40 journal articles. His main research interests include numerical algorithms in computational EM, including the fast algorithm for spatial domain dyadic Green's functions of stratified media, the multilevel fast multiple algorithm, the FFT-based fast algorithm, the IE-based domain decomposition method, the FEM-BI-based domain decomposition method, the higher order method of moments, and the parallel computation based on GPU/multicore-CPU platform. In addition, the full-wave simulation technology for the electromagnetic field distribution around the base-station is one of his current interests.

Wei Hong Wei Hong received the B.S. degree from the University of Information Engineering, Zhengzhou, China, in 1982, and the M.S. and Ph.D. degrees from Southeast University, Nanjing, China, in 1985 and 1988, respectively, all in radio engineering. Since 1988, he has been with the State Key Laboratory of Millimeter Waves and serves for the director of the lab, since 2003. In 1993, 1995, 1996, 1997, and 1998, he was a short-term visiting scholar with the University of California at Berkeley and at Santa Cruz, respectively. He is currently a professor with the School of Information Science and Engineering, Southeast University. He has been involved in numerical methods for electromagnetic problems, millimeter wave theory and technology, antennas, RF technology for wireless communications, and so on. He has authored and coauthored over 300 technical publications and two books. Dr. Hong was an elected IEEE MTT-S AdCom Member, from 2014 to 2016. He is a fellow of CIE, the vice president of the CIE Microwave Society and Antenna Society, and the chair of the IEEE MTT-S/AP-S/EMC-S Joint Nanjing Chapter. He was twice awarded the National Natural Prizes, thrice awarded the first-class Science and Technology Progress Prizes issued by the Ministry of Education of China and Jiangsu Province Government. He also received the Foundations for China Distinguished Young Investigators and for Innovation Group issued by NSF of China. He served as the associate editor for the *IEEE Transactions on Microwave Theory and Techniques* from 2007 to 2010, and one of the guest editors for the 5G special issue of the *IEEE Transactions on Antennas and Propagation*, in 2017.

Physics-informed Deep Learning to Solve 2D Electromagnetic Scattering Problems

Ji-Yuan Wang[1] and Xiao-Min Pan[2*]

[1]School of Integrated Circuit and Electronics
Beijing Institute of Technology, Beijing, 100081, China
3120210673@bit.edu.cn

[2]School of Cyberspace Science and Technology
Beijing Institute of Technology, Beijing, 100081, China
xmpan@bit.edu.cn
*Corresponding Author

Abstract **–** The utilization of physics-informed deep learning (PI-DL) methodologies provides an approach to augment the predictive capabilities of deep learning (DL) models by constraining them with known physical principles. We utilize a PI-DL model called the deep operator network (DeepONet) to solve two-dimensional (2D) electromagnetic (EM) scattering problems. Numerical results demonstrate that the discrepancy between the DeepONet and conventional method of moments (MoM) is small, while maintaining computational efficiency.

Index Terms **–** Electromagnetic scattering, physics-informed deep learning, the method of moments.

I. INTRODUCTION

Machine learning (ML) has emerged as a promising alternative for solving electromagnetic (EM) scattering problems [1–4] through an offline training - online prediction pattern. In the current landscape, two primary categories of ML approachs have been identified for EM scattering problems. The first approach treats ML as a "black box" and employs it as an alternative to traditional solvers [5–7]. Alternatively, the second approach involves replacing a module of a traditional solver with an ML-based solution [8–10]. Among these methodologies, physics-informed ML (PI-ML) [11] has attracted significant attention since it can gracefully incorporate both empirical data and prior physical knowledge. This integration of physical laws and data-driven approaches has been shown to improve the accuracy, reliability, and interpretability of predictions in various applications related to physical mechanisms. In [12], the physics-informed neural network (PINN) [13] was developed to solve time domain EM problems, where initial conditions, boundary conditions, as well as Maxwell's equations, were encoded as the constraints during the training process of the network.

Among many deep learning (DL) networks, the deep operator network (DeepONet) has the attractive ability to break the "curse of dimensionality" [14]. For the first time, to the best of our knowledge, this work utilizes the DeepONet to solve dynamic EM scattering problems. The performance of the DeepONet is investigated in terms of different configurations.

The rest of the paper is as follows. In Section II, we review the process of EM scattering, using a two-dimensional (2D) dielectric scatterer as an illustrative example. In Section III, we show how to use the DeepONet to solve EM scattering problems. In Section IV, several numerical results are given to verify the good performance of the DeepONet. Conclusions are given in Section V.

II. PROBLEM STATEMENT

In this section, we consider a 2D nonmagnetic transverse magnetic (TM) case, as shown in Fig. 1. The unknown scatterers are positioned in a domain $D \in \mathbb{R}^2$ with homogeneous background, i.e., the free space. Harmonic incident waves are excited by the transmitter with the time-factor being e^{jwt}. The scattered fields can be measured by the receivers in the domain $S \in \mathbb{R}^2$.

The EM scattering problem aims to determine the scattered fields generated by the illumination of incident fields on scatterers. The method of moments (MoM) [1, 4] is widely adopted to solve this problem. After discretizing the computational domain D into $M = M_1 \times M_2$ sub-domains, we can obtain the matrix system

$$\bar{\mathbf{J}} = \bar{\bar{\chi}} \cdot \left(\bar{\mathbf{E}}^i + \bar{\bar{\mathbf{G}}}_D \cdot \bar{\mathbf{J}} \right), \tag{1}$$

where $\bar{\mathbf{J}}$ and $\bar{\mathbf{E}}^i$ are vectors of the equivalent electric currents and incident fields, $\bar{\bar{\mathbf{G}}}_D$ is the matrix of the free space Green's function in the domain D, and $\bar{\bar{\chi}}$ is the diagonal matrix storing the contrast of each sub-domain.

Submitted On: February 28, 2023
Accepted On: September 20, 2023

https://doi.org/10.13052/2023.ACES.J.380905

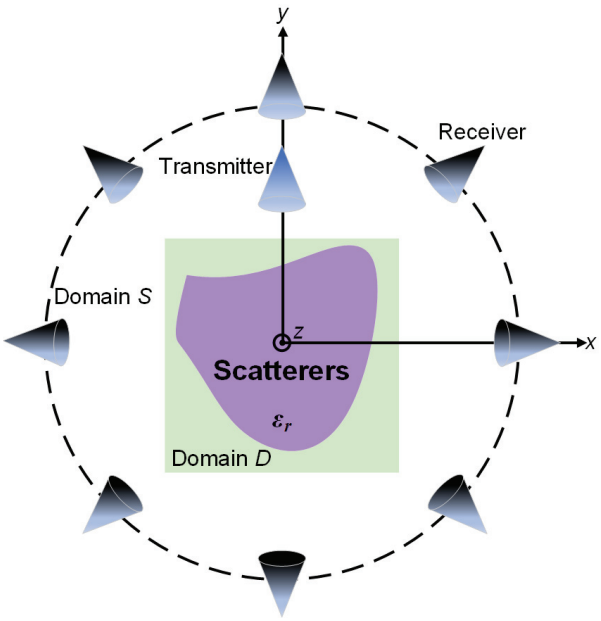


Fig. 1. Setup of EM scattering problems in the 2D case.

After $\bar{\mathbf{J}}$ is obtained by solving Eq. (1), the discrete total fields $\bar{\mathbf{E}}^t$ in the domain D can be obtained:

$$\bar{\mathbf{E}}^t = \bar{\mathbf{E}}^i + \bar{\bar{\mathbf{G}}}_D \cdot \bar{\mathbf{J}}. \tag{2}$$

Similarly, the scattered fields of any observation region can be calculated:

$$\bar{\mathbf{E}}^s = \bar{\bar{\mathbf{G}}}_S \cdot \bar{\mathbf{J}}, \tag{3}$$

where $\bar{\bar{\mathbf{G}}}_S$ is the matrix of the free space Green's function in the domain S.

III. PHYSICS-INFORMED DEEP LEARNING

The DeepONet [15] is a physics-informed DL (PI-DL) model, learning a variety of explicit operators that map from one function space to another, which is based on the universal approximation theorem for operators [16]. The theorem portrays a structure that consists of two sub-networks: one is the branch network for encoding the discrete input function space while the other is the trunk network for encoding the output function space, and shows that this particular network structure is capable of approximating a class of physical operators with arbitrary accuracy, which suggests its significant implications for the accurate modeling of complex physical systems. Essentially, the underlying mathematical foundation distinguishes the DeepONet from other neural networks.

The overall flow of employing the DeepONet to solve EM scattering problems is shown in Fig. 2. Firstly, we approximate the unknown current generation operator, i.e., the map from $\chi(\mathbf{r})$ to $\mathbf{J}(\mathbf{r})$, utilizing the DeepONet. Then the total and scattered fields are computed through a simple matrix-vector multiplication process using the predicted results acquired by the DeepONet. Specifically, the branch network of the DeepONet is a complex-valued (CV) ResNet [17] consisting of a convolutional input layer, followed by four residual blocks, and a fully connected output layer to generate modules of the real and imaginary parts. The setup of the CV network is similar to that in [18]. Here, we are dealing with images, or data matrices, of size 64×64. In the forward calculation, the number of channels is first changed to 64 through the input layer while the size of each channel is halved. The four residual blocks are then connected in turn. Assume that the input of one residual block has N_{in} channels and the size of each channel is $M_3 \times M_3$. After the forward calculation of this block, the number of channels becomes $2N_{in}$, and the size of each channel is halved to $\frac{M_3}{2} \times \frac{M_3}{2}$. After that, we flatten the 1×1 matrix over 1024 channels and connect a fully connected layer to output the result. The trunk networks are two real-valued fully connected networks to encode respectively the real and imaginary parts of the equivalent electric current function. They both have six hidden layers.

The input of the branch network is $\bar{\bar{\chi}} \oplus \bar{\bar{\chi}} \circ \bar{\bar{\mathbf{E}}}^i$, where $\oplus$ means to put the matrices on both sides of the symbol into the two channels of the tensor, and $\circ$ denotes the Hadamard product. The inputs to both trunk networks are the coordinates of one sub-domain, i.e. (x, y), for the 2D problem in this paper.

The output of the DeepONet is the real and imaginary part of the equivalent electric current within the sub-domain based on the input coordinates of the trunk networks, derived by computing the dot product between the real and imaginary components of the branch network output with the corresponding elements of the two trunk network outputs.

Trainable biases θ_{bias} are added to the final output of the DeepONet to improve generalization performance. The number and size of channels or the number of neurons is indicated below each layer.

According to the universal approximation theorem, the output layer of the branch network does not need to employ an activation function, while the output layer of the trunk network requires the use of it. Apparently, the background region with a contrast of zero does not contribute to equivalent electric currents. We incorporate this prior knowledge into the training process of the DeepONet, i.e., we only need to predict equivalent electric currents on sub-domains with non-zero contrasts.

The loss function of the DeepONet can be expressed in terms of the mean absolute error as

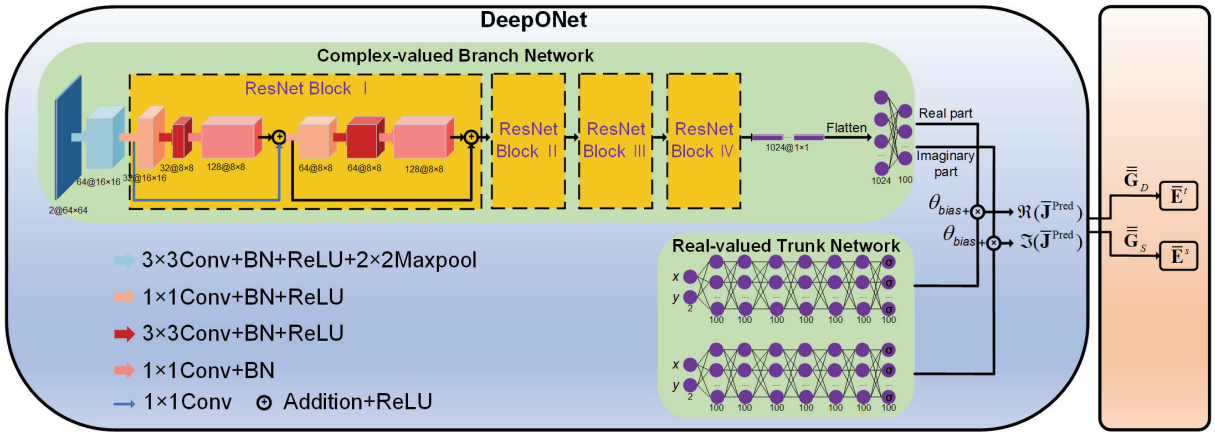


Fig. 2. The overall flow of employing the DeepONet to solve EM scattering problems.

$$\mathscr{L} = \frac{1}{B}\sum_{b=1}^{B}\left(\frac{\left|\Re\left(\bar{\mathbf{J}}_b^{\text{Pred}}\right)-\Re\left(\bar{\mathbf{J}}_b^{\text{MoM}}\right)\right|}{N_b} + \frac{\left|\Im\left(\bar{\mathbf{J}}_b^{\text{Pred}}\right)-\Im\left(\bar{\mathbf{J}}_b^{\text{MoM}}\right)\right|}{N_b}\right), \tag{4}$$

where B is the batch size, N_b is the number of non-zero contrast sub-domains for the bth scatterer, $\bar{\mathbf{J}}_b^{\text{Pred}}$ is equivalent electric currents for the bth scatterer obtained from the DeepONet while $\bar{\mathbf{J}}_b^{\text{MoM}}$ is equivalent electric currents calculated by the MoM, and $\Re(\cdot)$ and $\Im(\cdot)$ denote taking the real and imaginary part respectively.

IV. NUMERICAL RESULTS

In this section, we use several numerical examples to verify our approach. The DeepONet in this paper is built by Pytorch [19] and written in Python. We use the MoM to generate labels of training, validation and testing sets, i.e., equivalent electric currents. Specifically, the generalized minimal residual (GMRES) algorithm [20] is employed, with the aid of the fast Fourier transform (FFT) acceleration. The performance of the DeepONet is investigated on a server with two Intel Xeon CPUs and NVIDIA RTX 3090 GPU.

The data employed in our work is obtained by the MoM. In particular, 2D dielectric scatterers are generated according to handwritten numbers from the MNIST [21], which range from 0 to 9. Each scatterer has the computational domain D with a dimension of 2×2 m^2, centering at (0, 0). The real and imaginary parts of the relative permittivity of the training samples are randomly selected from 1.10-2.00 and 0.00-1.00, respectively. The number of scatterers in the training, validation and testing set are respectively, 20,000, 500, and 1000. We limit our analysis to the scenario of a single incident angle of 90°. The incident field operates at a frequency of 300 MHz. The scattered fields are sampled at 360 points that are uniformly located on a circle with a radius of 5 m. Figure 3 depicts some inputs of the branch network as well as their corresponding labels.

In the training stage, the adaptive moment estimation method (Adam) [22] is employed to minimize the loss function shown in Eq. (4). The learning rates of all sub-networks are set to 0.0002 at the beginning and are halved sequentially at the 100th, 140th, and 170th iterations. Each training batch contains 40 samples. The total number of iterations is 201. In the testing stage, to quantify the difference between the DeepONet predictions and true values generated by the MoM, we define the relative error (RE_q) for the qth obstacle as follows:

$$\text{RE}_q = \frac{\left\|\bar{\mathbf{J}}_q^{\text{Pred}} - \bar{\mathbf{J}}_q^{\text{MoM}}\right\|_2}{\left\|\bar{\mathbf{J}}_q^{\text{MoM}}\right\|_2}. \tag{5}$$

For the entire testing set, we define the mean relative error (MRE)

$$\text{MRE} = \frac{1}{Q}\sum_{q=1}^{Q}\text{RE}_q, \tag{6}$$

where Q is the number of samples in the testing set.

A. Feasibility validation

The testing set is set up in the same manner as the training set. Trajectories of losses are shown in Fig. 4 (a) and the histogram of REs for the testing set is shown in Fig. 4 (b). The MRE is found to be 0.043 while the variance of REs is estimated to be 4.9e-04. The predicted results are illustrated in Figs. 5 (a) and 5 (b), comprising a test instance from the MNIST test set and another from a distinct graphic type, respectively, to exhibit the feasibility of the DeepONet. It can be found that the predicted values match well with the true ones.

B. Generalization testing

In the context of the DeepONet framework, the size of the input function space and the size of the output

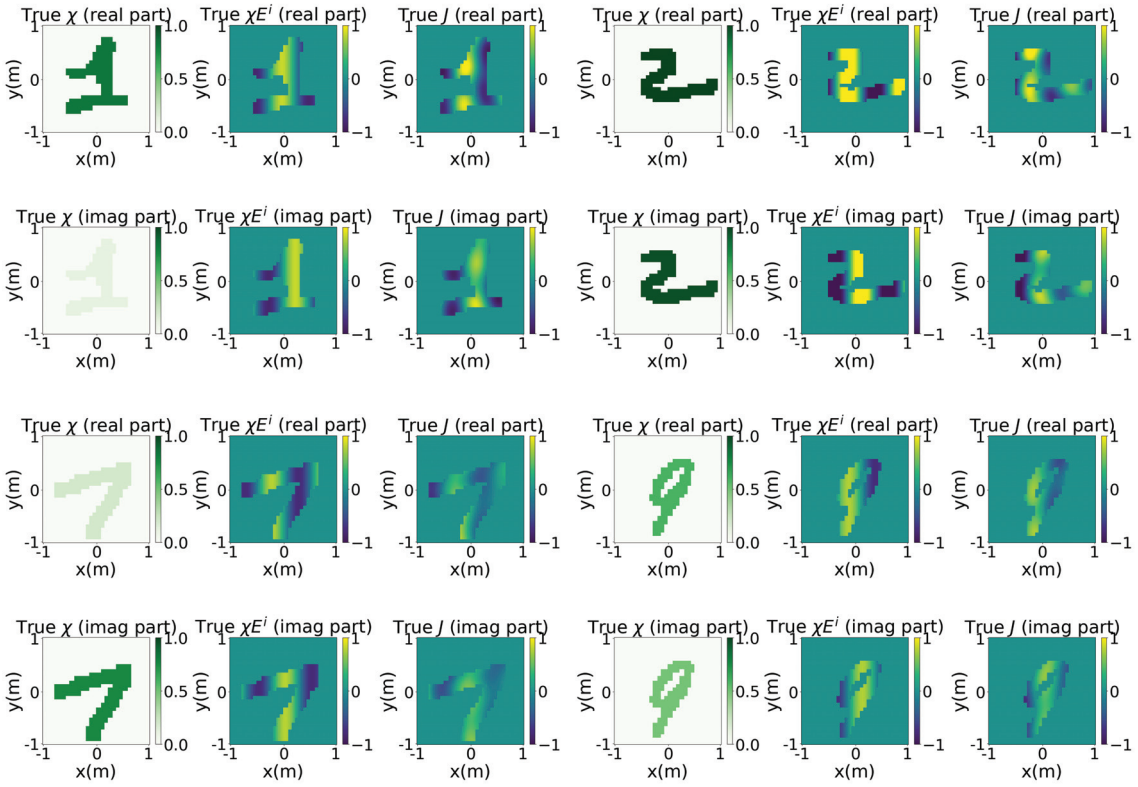
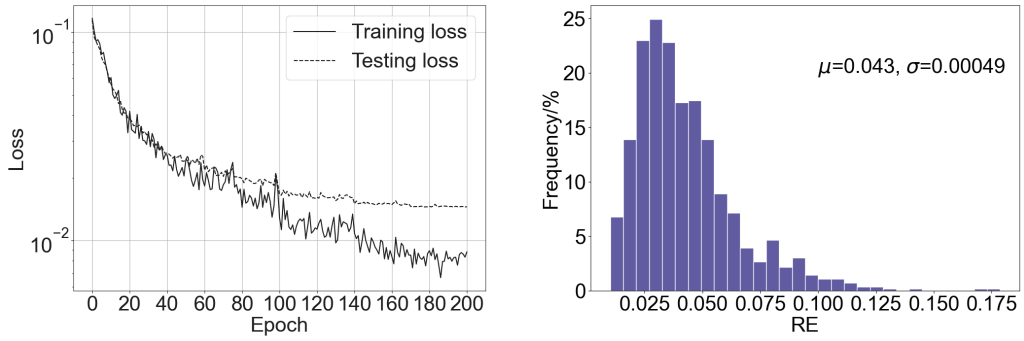


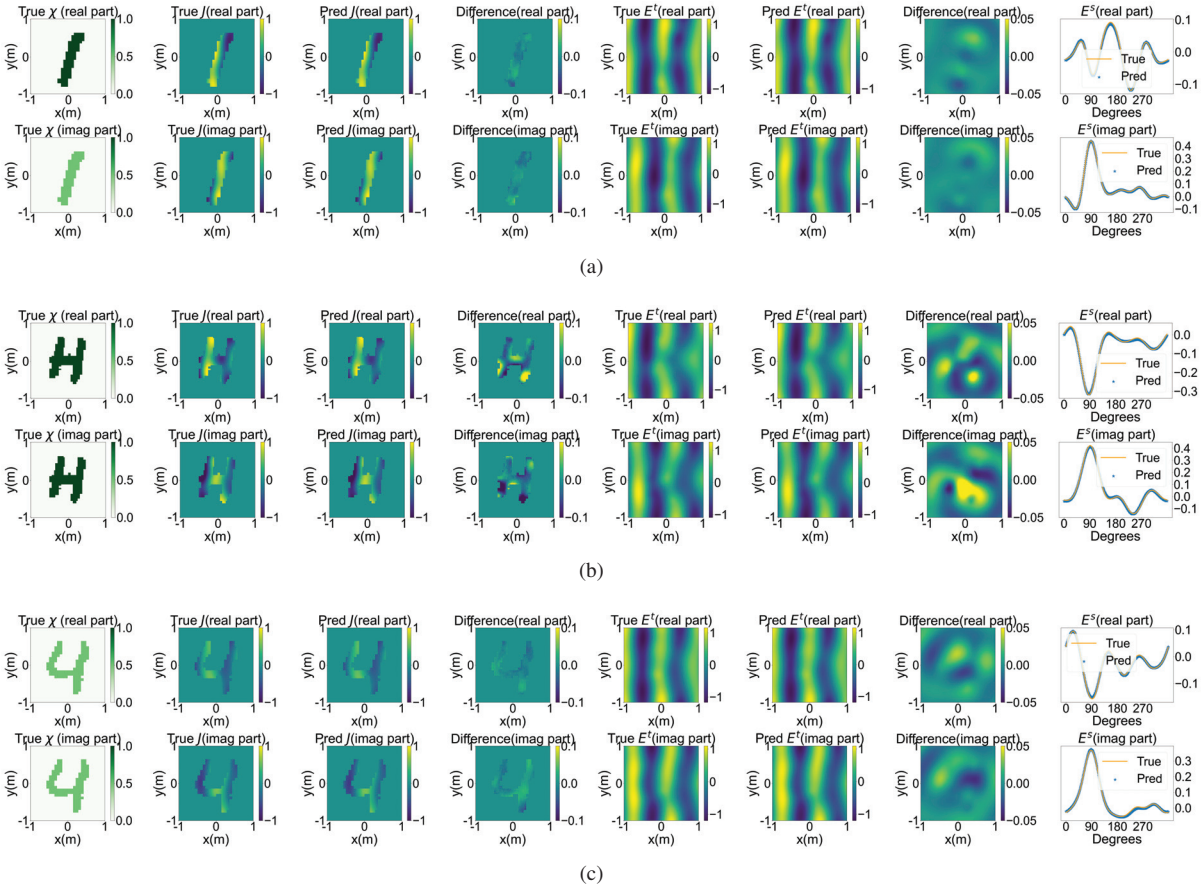
Fig. 3. Examples of the branch network input $\overline{\overline{\chi}} \oplus \overline{\overline{\chi}} \circ \overline{\overline{\mathbf{E}}}^{i}$ and corresponding equivalent electric current labels.



(a) (b)

Fig. 4. (a) Trajectories of losses as a function of the number of epochs. (b) The RE histogram of the testing set with 1000 samples.

function space after discretization are found to be independent. While fixing the input size of the branch network, we can flexibly adjust the output of the DeepONet by varying the coordinate input of the trunk network during the testing stage. In this part, we explore the potential of predicting equivalent electric currents with grids of 32×32 and 96×96 using the DeepONet trained on the training set of grid size 64×64, with the position and dimension of the domain D being held constant. For each of the two aforementioned cases, we generate 1,000 scatterers for the testing set. The configuration of the scatterers are set similarly to the previous example. Figure 6 (a) and 6 (b) give histograms of REs for the two cases. It is indicated that the DeepONet successfully

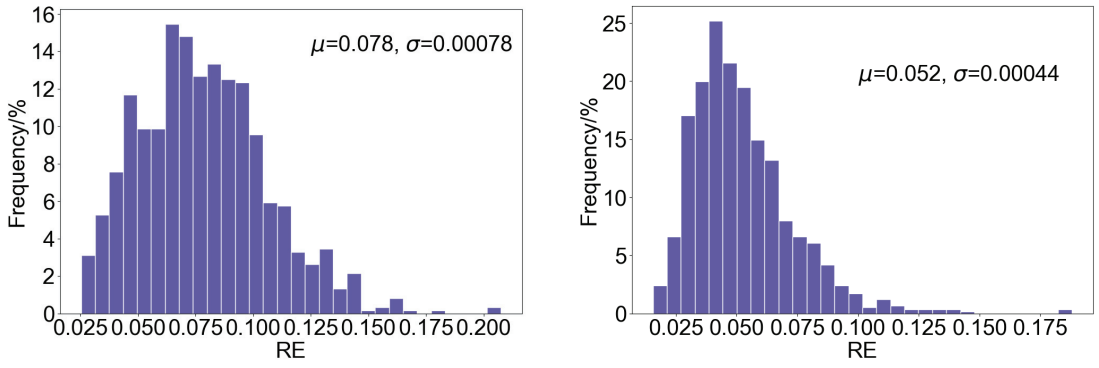


(a)

(b)

(c)

Fig. 5. The predicted results. The 1st column is the real and imaginary parts of contrasts. The 2nd column gives the real and imaginary parts of true equivalent electric currents computed by the MoM. The 3rd column is the real and imaginary parts of predicted equivalent electric currents. The 4th and 7th columns show the differences between true and predicted results. The 5th and 6th columns are the total fields calculated with the true and predicted equivalent electric currents, respectively. The last column are the true and predicted scattered fields in $\phi \in [0^\circ, 359^\circ]$. (a) A sample from the MNIST testing set with a grid size of 64×64, (b) a sample in a different graphic type with a grid size of 64×64, (c) a sample from the MNIST testing set with a grid size of 1024×1024.



(a) (b)

Fig. 6. The RE histograms of the testing sets with different grids of the equivalent electric currents. Each testing set contains 1000 samples. (a) 32×32 case. (b) 96×96 case.

predicted multi-grid results, confirming its nice generalization capability as a grid-free EM numerical solver.

C. Computational time

Although the DeepONet is a picture-to-pixel model that yields the equivalent electric current on a single sub-domain per forward calculation, we can supply all the coordinate data concurrently and utilize the graphics processing unit (GPU) to expedite computations, thus significantly reducing the computation time. Additionally, the prediction of equivalent electric currents on subdomains with non-zero contrasts leads to further reduction of total prediction time. Notably, as the domain D is divided into finer partitions, the traditional numerical solver suffers from the "curse of dimensionality." resulting in an exponential increase in computation time. Conversely, the DeepONet's computation time remains relatively unchanged. The reason is that the time increase in the trunk network and dot product during the forward calculation due to the increase in coordinate inputs is negligible, while the cost of the branch network remains constant.

Table 1: Comparison of computation time between the DeepONet and MoM

Grid Size	DeepONet (s)	MoM (s)
64×64	0.018	0.014
256×256	0.018	0.38
1024×1024	0.024	6.39

Table 1 compares the DeepONet and MoM on different grids in terms of the computational time to obtain equivalent electric currents. The test obstacle is configured in an identical manner as in previous experiments. Using an example of grid size 1024×1024, the DeepONet predicts equivalent electric currents in just 0.024 s, compared to 6.39 s for the MoM to complete the corresponding calculation. The predicted results of grid size 1024×1024 are shown in Fig. 5 (c). In this case, both the total and scattered fields calculated with the true equivalent electric currents and the predicted ones match well.

V. CONCLUSION

Based on PI-DL, this paper utilizes the DeepONet to solve 2D dynamic EM scattering problems from the perspective of approximating the unknown current generation operator. The underlying mathematical foundation distinguishes the DeepONet from other neural networks. The accuracy, efficiency, and generalization of the DeepONet are verified through numerical examples. Furthermore, the DeepONet exhibits the ability to perform multi-grid prediction and overcome the "curse of dimensionality." Since the DeepONet can be extended to three-dimensional cases, the proposed approach has the potential to solve 3D dynamic EM scattering problems.

ACKNOWLEDGMENT

This work was partly supported by NSFC under grant 62171033.

REFERENCES

[1] W. C. Chew, M. S. Tong, and B. Hu, "Integral equation methods for electromagnetic and elastic waves," *Synthesis Lectures on Computational Electromagnetics*, vol. 3, no. 1, pp. 1–241, 2008.

[2] D.-M. Yu, X. Ye, X.-M. Pan, and X.-Q. Sheng, "Fourier bases-expansion for three-dimensional electromagnetic inverse scattering problems," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[3] X. Chen, *Computational Methods for Electromagnetic Inverse Scattering*, John Wiley & Sons, 2018.

[4] Y.-N. Liu and X.-M. Pan, "Solution of volume-surface integral equation accelerated by MLFMA and skeletonization," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 7, pp. 6078–6083, 2022.

[5] L. Li, L. G. Wang, F. L. Teixeira, C. Liu, A. Nehorai, and T. J. Cui, "DeepNIS: Deep neural network for nonlinear electromagnetic inverse scattering," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 3, pp. 1819–1825, 2018.

[6] S. Qi, Y. Wang, Y. Li, X. Wu, Q. Ren, and Y. Ren, "Two-dimensional electromagnetic solver based on deep learning technique," *IEEE Journal on Multiscale and Multiphysics Computational Techniques*, vol. 5, pp. 83–88, 2020.

[7] Z. Ma, K. Xu, R. Song, C.-F. Wang, and X. Chen, "Learning-based fast electromagnetic scattering solver through generative adversarial network," *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 4, pp. 2194–2208, 2020.

[8] R. Guo, T. Shan, X. Song, M. Li, F. Yang, S. Xu, and A. Abubakar, "Physics embedded deep neural network for solving volume integral equation: 2-D case," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 8, pp. 6135–6147, 2021.

[9] Y. Hu, Y. Jin, X. Wu, and J. Chen, "A theory-guided deep neural network for time domain electromagnetic simulation and inversion using a differentiable programming platform," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 1, pp. 767–772, 2021.

[10] B.-W. Xue, D. Wu, B.-Y. Song, R. Guo, X.-M. Pan, M.-K. Li, and X.-Q. Sheng, "U-net conjugate gradient solution of electromagnetic

scattering from dielectric objects," in *2021 International Applied Computational Electromagnetics Society (ACES-China) Symposium*, pp. 1–2, IEEE, 2021.

[11] Z. Hao, S. Liu, Y. Zhang, C. Ying, Y. Feng, H. Su, and J. Zhu, "Physics-informed machine learning: a survey on problems, methods and applications," *arXiv preprint arXiv:2211.08064*, 2023.

[12] P. Zhang, Y. Hu, Y. Jin, S. Deng, X. Wu, and J. Chen, "A Maxwell's equations based deep learning method for time domain electromagnetic simulations," *IEEE Journal on Multiscale and Multiphysics Computational Techniques*, vol. 6, pp. 35–40, 2021.

[13] L. Lu, X. Meng, Z. Mao, and G. E. Karniadakis, "DeepXDE: A deep learning library for solving differential equations," *SIAM Review*, vol. 63, no. 1, pp. 208–228, 2021.

[14] S. Lanthaler, S. Mishra, and G. E. Karniadakis, "Error estimates for deeponets: A deep learning framework in infinite dimensions," *Transactions of Mathematics and Its Applications*, vol. 6, no. 1, p. tnac001, 2022.

[15] L. Lu, P. Jin, G. Pang, Z. Zhang, and G. E. Karniadakis, "Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators," *Nature Machine Intelligence*, vol. 3, no. 3, pp. 218–229, 2021.

[16] T. Chen and H. Chen, "Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems," *IEEE Transactions on Neural Networks*, vol. 6, no. 4, pp. 911–917, 1995.

[17] C. Trabelsi, O. Bilaniuk, Y. Zhang, D. Serdyuk, S. Subramanian, J. F. Santos, S. Mehri, N. Rostamzadeh, Y. Bengio, and C. J. Pal, "Deep complex networks," *arXiv preprint arXiv:1705.09792*, 2017.

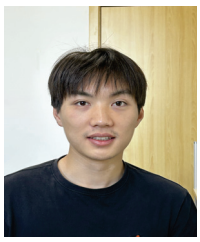
[18] X.-M. Pan, B.-Y. Song, D. Wu, G. Wei, and X.-Q. Sheng, "On phase information for deep neural networks to solve full-wave nonlinear inverse scattering problems," *IEEE Antennas and Wireless Propagation Letters*, vol. 20, no. 10, pp. 1903–1907, 2021.

[19] N. Ketkar, J. Moolayil, N. Ketkar, and J. Moolayil, "Introduction to pytorch," *Deep Learning with Python: Learn Best Practices of Deep Learning Models with PyTorch*, pp. 27–91, 2021.

[20] Y. Saad and M. H. Schultz, "GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems," *SIAM Journal on Scientific and Statistical Computing*, vol. 7, no. 3, pp. 856–869, 1986.

[21] The MNIST database can be downloaded at http://yann.lecun.com/exdb/mnist/.

[22] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.



Ji-Yuan Wang received the B.S. degree from Communication University of China, Beijing, China, in 2021. He is currently pursuing the M.S. degree with the School of Integrated Circuit and Electronics, Beijing Institute of Technology, Beijing. His current research interests include deep learning and computational electromagnetics.

Xiao-Min Pan received the B.S. and M.S. degrees from Wuhan University, Wuhan, China, in 2000 and 2003, respectively, and the Ph.D. degree from the Institute of Electronics, Chinese Academy of Sciences, Beijing, China, in 2006. He is currently a professor with the School of Cyberspace Science and Technology, Beijing Institute of Technology, Beijing. He has authored or coauthored over 60 papers in refereed journals. His current research interests include high-performance methods in computational electromagnetics, artificial intelligence in electromagnetics, and electromagnetic compatibility analysis of complex systems.

Prof. Pan was the third recipient of the First Prize of Beijing Science and Technology Awards in 2011, the recipient of the New Century Excellent Talents in University in 2012 and of Young Elite Talents in Beijing in 2012. He received the Ulrich L. Rohde Innovative Conference Paper Award in 2016 and several other international academic awards. He served as a technical program committee co-chair, special sessions co-chairs, and technical program committee member for several international conferences and symposiums.

A Hybrid QOGWO-GPR Algorithm for Antenna Optimization

Hao-Yun Zhu, Jia-Wei Qian, Xiao-Hui Tang, and Wei-Dong Li*

School of Information Science and Engineering, State Key Laboratory of Millimeter Waves
Southeast University, Nanjing, 210096, China
zhuhaoyun@seu.edu.cn, 220211003@seu.edu.cn, 220221032@seu.edu.cn, wdli@seu.edu.cn
*Corresponding Author

Abstract **–** Optimization of antenna performance is a non-linear, multi-dimensional and complex issue, which entails a significant investment of time and labor. In this paper, a hybrid algorithm of quasi-opposition grey wolf optimization (QOGWO) and Gaussian process regression (GPR) model is proposed for antenna optimization. The QOGWO is prone to global optimality, high precision for complex problems, and fast convergence rate at the later stage. The GPR model can reduce time cost of antenna samples generation. After being optimized by the proposed approach, a stepped ultrawideband monopole antenna and a dual-band MIMO antenna for WLAN can achieve wider bandwidth and higher gain or isolation at low time cost, compared to other intelligent algorithms and published literatures.

Index Terms **–** Antenna optimization, Gaussian process regression, grey wolf optimization, quasi-opposition.

I. INTRODUCTION

Since their invention, antennas have become indispensable devices in electric equipment, and they are also important parts of wireless communication systems. In microwave engineering, antenna optimization is a non-linear, multi-dimensional, and highly complex problem. Antennas are usually analyzed by electromagnetic simulation software, and their optimization and design require extensive professional experience. Therefore, antenna design entails a significant investment of time and labor. However, emerging intelligent algorithms provide practical approaches to fast optimization and efficient design.

In recent years, many scholars have explored abundant intelligent optimization algorithms. Nysaeter employed the NSGA-III algorithm to optimize two-dimensional antenna locations for MIMO two-way antenna patterns [1]. Singh applied the cat-swarm-based genetic optimization (CSGO) to the design of a miniaturized multiband antenna [2]. Mirjalili proposed a grey wolf optimization (GWO) in 2014 [3]. It has been applied in versatile fields because of few parameters to be set, robustness, and high performance in some cases [4, 5]. It has potential applications in antenna optimization, which motivates our work in this paper.

Based on the GWO algorithm [3], a quasi-opposition grey wolf optimization (QOGWO) algorithm is built. It is prone to global optimality, high-precision solution for complex problems, and fast convergence rate. For reducing time cost of antenna samples generation, the Gaussian process regression (GPR) model is introduced and combined with the QOGWO into the QOGWO-GPR algorithm. It is validated by a stepped ultrawideband monopole antenna and a dual-band MIMO antenna for WLAN, and the results show that it has made the two antennas achieve wider bandwidth and higher gain or isolation at low time cost compared to the other intelligent algorithms and the published literature [14].

II. IMPROVED GWO

The GWO algorithm imitates the hunting behaviors of a wolf pack in nature. According to the distances between individual grey wolves and the prey, the three closest wolves are named α, β, and δ. The remaining wolves are denoted as ω. The predation process is led by α wolf and consists of three steps: tracking the prey, surrounding the prey, and attacking the prey. The GWO algorithm [3] is outlined as follows:

Step 1: The dimensions of hunting space and their bounds, the number of individuals in the grey wolf population N are set.

Step 2: The grey wolf population are initialized randomly.

Step 3: In the population, the three wolf individuals closest to the prey are selected as α, β, and δ respectively. In the specific problem, the distance corresponds to a pre-defined fitness function value.

Step 4: Convergence factor a is expressed, with the tth iteration, as

$$a = 2 - \frac{2t}{Max\_iter}, \tag{1}$$

where Max_iter is the maximum number of iterations. Obviously, a decreases linearly from 2 to 0 as t increases.

Submitted On: February 28, 2023
Accepted On: August 23, 2023

https://doi.org/10.13052/2023.ACES.J.380906

Step 5: Each individual wolf ω updates its position in the $(t+1)$th iteration by its relative position to α, β, and δ as following:

$$\overline{X}_{\omega}(t+1)=\frac{\overline{P}_{\alpha}(t)+\overline{P}_{\beta}(t)+\overline{P}_{\delta}(t)}{3}, \tag{2}$$

where $\overline{P}_{\alpha}(t)$, $\overline{P}_{\beta}(t)$, and $\overline{P}_{\delta}(t)$ denote the forward vectors of ω toward α, β, and δ in the tth iteration.

Step 6: If the iteration number reaches *Max_iter*, the iteration process will stop; otherwise, go to Step 3.

The other details of the GWO algorithm are referred to [3]. The GWO algorithm has some advantages such as few parameters to be set, high robustness, and good performance in some cases [4, 5]. However, it easily falls into local optimal solutions, and has low precision in solving complex problems, and is slowly convergent in the later stage [6]. In what follows, the GWO algorithm will be improved in three aspects: population initialization in Step 2, convergence factor in Step 4, and individual position updated in Step 5.

A. Quasi-opposition population initialization

In the GWO algorithm, the initial population is generated randomly. This does not ensure a stable and rich population diversity and lowers the solution precision for complex problems [7]. The quasi-opposition principle [8] serves as a potential approach to improving the population diversity.

Each individual is a multidimensional vector. For the sake of clarity, its component is considered as a one-dimensional example to illustrate the concept of quasi-opposition. The optimal solution to be searched is supposed in the interval $[a_0,b_0]$. For any $x\in[a_0,b_0]$, its symmetry point x' is a_0+b_0-x, the probability that x is close to the optimal solution is the same as x'. The midpoint of $[a_0,b_0]$ is denoted by x_m, and the quasi-opposition point x_{qop} can be generated by

$$x_{qop}=\begin{cases} rand\,(x_m,x'),x'>x_m \\ rand\,(x',x_m),x'\le x_m. \end{cases} \tag{3}$$

As shown in [9], x_{qop} has a higher probability to be closer to the optimal solution than x'. Obviously, applying (3) to each component of vector generates the quasi-opposition point of an individual.

First, the population A is generated randomly, and the quasi-opposition population A_{qop} is obtained by generating the quasi-opposition points of all individuals in A. Second, the fitness values of individuals in these two populations are calculated and ranked. Finally, the top N best individuals are chosen as the initial population for subsequent iterations.

B. Non-linear convergence factor

In each iteration of the GWO algorithm, the individual grey wolf ω dynamically changes its direction according to the vector $\overline{u}$, which is expressed as

$$\overline{u}=2a\overline{r}_1-a\left[1\ \cdots\ 1\right]^T, \tag{4}$$

where $\overline{r}_1$ is a random vector whose components are randomly located in $[0,1]$. Since its module (i.e., the max absolute value of components) is a random number in $[0,1]$, $|\overline{u}|$ is a random value in $[0,a]$. When $a>1$, the algorithm will be more prone to global optimization search; when $a<1$, it will instead be more likely to perform local search. To avoid the GWO algorithm getting trapped early in the local optimum solution, a is chosen as

$$a=2cos\left(\frac{\pi t}{2Max\_iter}\right). \tag{5}$$

In such way, a decreases non-linearly from 2 to 0, which increases the proportion of global search in the iterations.

C. Updating positions by weighting coefficients

Eq. (2) is used to update each individual's position in the next iteration, and implies the equal contributions from α, β, and δ. This mean way leads to a slow convergence rate in the later stage toward the optimal solution. Since α is the approximate solution closest to the optimal one, its contribution is dominant. Hence, according to the contribution of α, β, and δ, the strategy of weighting coefficients is adjusted as

$$\overline{X}_{\omega}(t+1)=\frac{f(\alpha)\overline{P}_{\alpha}(t)+f(\beta)\overline{P}_{\beta}(t)+f(\delta)\overline{P}_{\delta}(t)}{f(\alpha)+f(\beta)+f(\delta)}, \tag{6}$$

where the weighting function $f(\cdot)$ is the fitness value of the individual.

After the improvement in Subsections A-C, the GWO algorithm is modified into the QOGWO version, whose flowchart is demonstrated in Fig. 1.

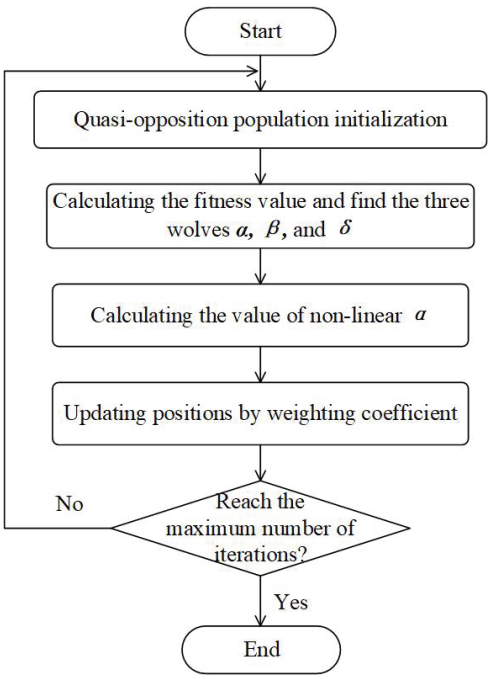


Fig. 1. The flowchart of the QOGWO algorithm.

III. HYBRID ALGORITHM OF QOGWO AND GPR

In the QOGWO algorithm, each individual denotes a vector whose components are dimension parameters of the antenna and randomly valued within the ranges. Calculating the fitness value of each individual requires one full wave simulation. For complex antennas, the time cost may be unacceptable. For reducing the times of full wave simulations, the GPR model is introduced to predict the fitness values of individuals in this section.

In the GPR model, the mean function and the covariance function determine its prediction accuracy. The mean function often takes a value of 0. The suited kernel function serves as the covariance function [12]. For improving its performance, the hyperparameters of the kernel function are usually optimized by the conjugate gradient (CG) algorithm [10]. However, the CG algorithm has the shortcomings of over-dependence on the initial value, easily falling into local optimal solution, and poor convergence. As a remedy, the QOGWO algorithm is preferably chosen as the optimizer of hyperparameter due to its merits, shown in Section II A-C.

The QOGWO and the above improved GPR model are combined into the QOGWO-GPR algorithm. It consists of two steps. The first step is training the GPR prediction model, where the QOGWO is used to optimize the hyperparameters of the different kernel functions for obtaining the best-fitting model. The second step is optimizing individuals, i.e., sets of dimension parameters of antenna, where the QOGWO invokes the trained GPR model to calculate the fitness values of individuals. The process will be detailed in the simulation experiments below.

IV. EXAMPLES AND ANALYSIS

A stepped ultrawide-band monopole antenna [11] and a dual-band MIMO antenna for WLAN [14] are used to validate the hybrid QOGWO-GPR algorithm. The simulations are performed on a computer with Core i7-8700K CPU and 32 GB memory.

A. A stepped ultrawideband monopole antenna

A.1. Configuration of antenna

Figure 2 shows the configuration of the antenna. The middle layer is the dielectric substrate FR4 (ε_r = 4.5, h = 1.6 mm). The top surface of the dielectric substrate is partially covered with a stepped monopole radiator patch, which is fed by 50 Ω microstrip line. On the bottom surface is a corner-truncated ground plate. The stepped design of both layers is capable of decreasing the discontinuity of the structure at the antenna feed source and improving the impedance matching of the antenna [11]. The frequency fre varies from 2 to 12 GHz.

For broadening the bandwidth of the antenna and improving its directivity by the QOGWO-GPR algorithm, the performance targets here are S_{11} and *Gain*. Dimension parameters to be optimized and their ranges are listed in Table 1.

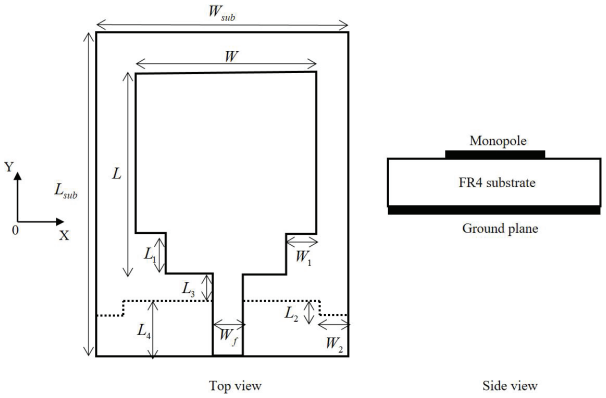


Fig. 2. Configuration of the stepped ultrawideband monopole antenna (W_{sub} = 20 mm, L_{sub} = 30 mm, W = 16 mm, L = 18 mm, L_4 = 5 mm).

Table 1: Dimension parameters and their ranges of the antenna

Parameters	L_1	W_1	L_2	W_2	L_3
Ranges (mm)	4-8	2-5	0.5-3	5-8	1.5-2.5

A.2. Training GPR prediction model

Firstly, we use HFSS simulation to calculate S_{11} and *Gain* of the samples. Then, samples are put into the GPR model. In the GPR model, the vector $[L_1, W_1, L_2, W_2, L_3, fre]$ is the input variable, and S_{11} and *Gain* are the output values. For each dimension parameter, three values are sampled uniformly within the range, and 243 sets are yielded. The frequency step is 500 MHz, so there are 5103 input samples.

The input sample data is divided into two parts: 70% samples randomly chosen as the training set and 30% samples as the test set. Based on the two sets, the kernel function candidates such as RQ, SE and Matern functions are built for the GPR model [12]. Their hyperparameters are optimized by the QOGWO algorithm, where N is set to 24 and Max_iter set to 30.

To reduce the bias in experimental results caused by randomness, the above test is conducted 10 rounds for each kernel function. The performance of the GPR model resulted from the kernel function is assessed by the root mean square error (RMSE) [13] as

$$RMSE = \sum_{d=1}^{D} \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(y_i^d - \hat{y}_i^d\right)^2}, \tag{7}$$

where n is the number of test samples data. D is the dimensionality of the output sample. y_i^d is the output sample value, and $\hat{y}_i^d$ is the predicted output sample value

by the GPR model. Obviously, the smaller the RMSE, the better the prediction.

Our simulation results show that the Matern kernel function (with hyperparameters $l = 3.713$ and $\sigma_f^2 = 1.654$) can achieve the best training effect. Thus, it serves as the covariance function in the GPR model. The RMSE changes with iterations, as shown in Fig. 3.

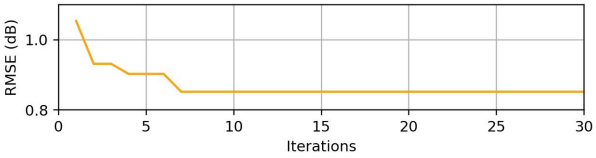


Fig. 3. RMSE changes with iterations.

A.3. Simulation of QOGWO-GPR algorithm

In the QOGWO-GPR algorithm, one individual denotes the vector $[L_1, W_1, L_2, W_2, L_3]$. N is set to 100 and Max_iter set to 500. The dimension parameters and the ranges are listed in Table 1. Once the initial population and the quasi-opposition counterpart are generated, a trained GPR model is invoked to predict S_{11} and *Gain* of each individual at arbitrary frequencies.

For achieving the multi-objective optimization, weighting S_{11} and *Gain* yields the fitness value as

$$fitness = \frac{w_1}{Num}\sum_{i=1}^{Num}\frac{\mathrm{S}_{11}(i)}{S_{\mathrm{tol}}} + \frac{w_2}{Num}\sum_{i=1}^{Num}\frac{Gain_{\mathrm{tol}}}{\mathrm{Gain}(i)}, \quad (8)$$

where Num is the number of frequency points. The weight coefficients w_1 and w_2 are set to 0.5. S_{tol} and $Gain_{\mathrm{tol}}$ are the tolerance values of two optimization targets, taken as -10 dB and 3 dBi respectively. $\mathrm{S}_{11}(i)$ is the return loss at the ith frequency point after preprocessing by

$$\mathrm{S}_{11}(i) = \begin{cases} -12, S_{11}(i) < -12 \\ S_{11}(i), S_{11}(i) \geq -12. \end{cases} \quad (9)$$

$\mathrm{Gain}(i)$ is the gain at the ith frequency point after normalization by

$$\mathrm{Gain}(i) = \begin{cases} \mid Gain(i) \mid, Gain(i) < -3 \\ 3, Gain(i) \geq -3. \end{cases} \quad (10)$$

Finally, the QOGWO algorithm iterates until the individual with the best fitness value (i.e., the optimal solution) is generated. The optimized parameters of the antenna are: L_1 = 4.4 mm, W_1 = 5 mm, L_2 = 1.16 mm, W_2 = 8 mm and L_3 = 2 mm.

The hybrid QOGWO-GPR algorithm is compared with the genetic algorithm (GA), the particle swarm optimisation algorithm (PSO) and the QOGWO algorithm. In the later three algorithms, N = 100, and Max_iter = 100. The variance factor p_m in the GA is set to 0.1 and the crossover probability factor p_c set to 0.8. The learning factors c_1 and c_2 are set to 0.5, and the inertia factor ω set to 0.8 for the PSO. In these three algorithms, S_{11} and *Gain* for each individual are obtained from the HFSS software simulation, and the corresponding fitness value is generated via (8). When the maximum number of iterations is reached, the dimension parameters corresponding to the best fitness individual are put into the HFSS software for simulation. Figures 4 and 5 compare the results of S_{11} and the normalized *Gain* obtained from the different algorithms. Tables 2 and 3 list the statistical values, where Hybrid denotes the QOGWO-GPR, and $\bar{S}_{11}$ the average S_{11}.

Seen from the above figures and tables, the antenna optimized by the hybrid QOGWO-GPR algorithm has the best overall performance, and fully satisfies the design requirements. The one by the QOGWO algorithm has the second best overall performance, and the average

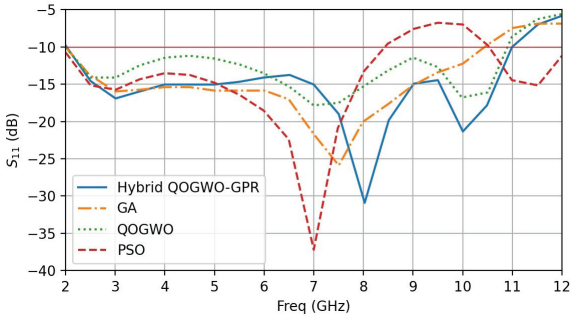


Fig. 4. Comparison of S_{11} curves from different algorithms.

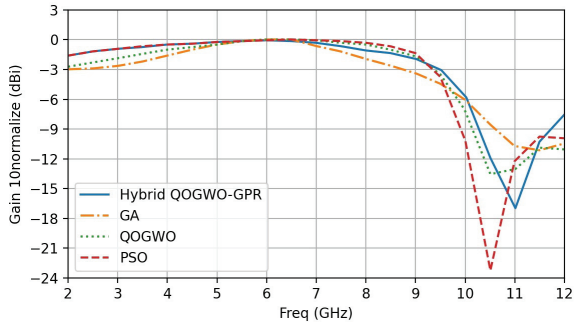


Fig. 5. Comparison of normalized *Gain* curves from different algorithms.

Table 2: Statistical values of impedance bandwidths from different algorithms

Algorithm	Bandwidth	$\bar{S}_{11}$
GA	107.56% (2 GHz-10.39 GHz)	-14.67 dB
PSO	82.31% (2 GHz-8.42 GHz)	-14.56 dB
QOGWO	113.84% (2 GHz-10.88 GHz)	-12.77 dB
Hybrid	113.97% (2.06 GHz-10.95 GHz)	-15.30 dB

Table 3: Statistical values of gains from different algorithms

Algorithm	-3dB Gain Bandwidth
GA	6.69 GHz (2.03 GHz-8.72 GHz)
PSO	7.31 GHz (2 GHz-9.31 GHz)
QOGWO	7.45 GHz (2 GHz-9.45 GHz)
Hybrid	7.53GHz (2 GHz-9.53 GHz)

S_{11} is 2 dB higher than these by the GA and PSO algorithms.

In this example, the time cost of the hybrid QOGWO-GPR algorithm consists of three parts: generating the sample data, yielding the best-fitting GPR model, and optimizing the antenna size parameters. It takes a total of 7.47 hours. In the other three algorithms, the time cost of the individual updating position is negligible, and the time is mainly consumed by HFSS simulation for yielding the fitness values. It takes a total of 25.93 hours. The hybrid QOGWO-GPR algorithm takes about 28.8% of the time of the other three algorithms, respectively.

B. Dual-band MIMO antenna for WLAN

B.1. Configuration of antenna

Figure 6 shows the configuration of the antenna [14]. It generates two separate resonant modes to cover 2.45 and 5 GHz WLAN bands. The middle layer is the dielectric substrate chosen as FR4 (ε_r = 4.4, h = 0.8 mm). The antenna consists of two radiating units, which are the same and located symmetrically. The F-type antenna with two branches resonates at low frequency and high frequency, respectively. Two I-type floor branches and a floor gap act as an isolator.

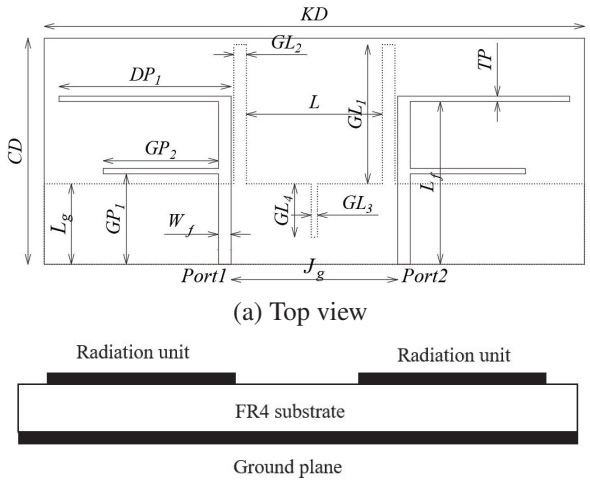
In the QOGWO-GPR algorithm, the performance targets here are S_{11} and S_{21}. The dimension parameters determining S_{11} and S_{21} are listed in Table 4. They are optimized by the QOGWO-GPR algorithm for broadening the bandwidth of the antenna and improving its isolation.

Table 4: Dimension parameters and their ranges of the dual-band MIMO antenna

Parameters	DP_1	GL_1	GL_3
Ranges (mm)	17.37-21.23	13.86-16.94	0.5-0.9
Parameters	LF	TP	GP_2
Ranges (mm)	16.2-19.8	0.5-0.7	11.52-14.08

B.2. Training GPR prediction model

In the GPR model, S_{11} and S_{21} are the output values, and the vector $[DP_1, GL_1, GL_3, GP_2, LF, TP, fre]$ is the input variable. For each dimension parameter, three values are sampled uniformly within the range, and



(a) Top view

(b) Side view

Fig. 6. Configuration of the dual-band MIMO antenna for WLAN (KD = 63 mm, CD = 28 mm, W_f = 1.4 mm, L_f = 18 mm, GP_1 = 9 mm, J_g = 18.5 mm, L_g = 8.9 mm, GL_2 = 1.4 mm, GL_4 = 5.9 mm, L = 15.1 mm).

729 sets are yielded. The frequencies are 2, 2.3, 2.45, 2.8, 4.4, 5, 5.3, and 5.9 GHz, so there are 5832 input samples. Then, putting samples into the GPR model, we optimize their hyperparameters by the QOGWO algorithm, where N is set to 24 and Max_iter set to 30. The GPR model simulation results show that the Rational Quadratic kernel function(with hyperparameters l = 4.79, α = -0.76 and σ_f^2 = 0.06) can achieve the best training effect. Thus, it serves as the covariance function in the GPR model. The RMSE changes with iterations, as shown in Fig. 7.

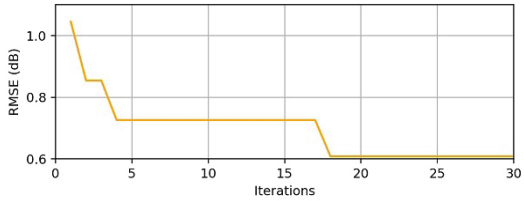


Fig. 7. RMSE changes with iterations.

B.3. Simulation of QOGWO-GPR algorithm

In the QOGWO-GPR algorithm, one individual denotes the vector $[DP_1, GL_1, GL_3, GP_2, LF, TP]$. N is set to 30, and Max_iter set to 50. The dimension parameters and the ranges are the same as those in Table 4. For achieving the multi-objectives optimization, weighting S_{11} and S_{21} yields the fitness value as

$$fitness = \frac{w_1}{Num}\sum_{i=1}^{Num}\frac{S_{11}(i)}{S_{11\text{tol}}} + \frac{w_2}{Num}\sum_{i=1}^{Num}\frac{S_{21}(i)}{S_{21\text{tol}}}, \qquad (11)$$

where *Num* is the number of frequency points. The weight coefficients w_1 and w_2 are set to 0.5 each. $S_{11\text{tol}}$ and $S_{21\text{tol}}$ are the tolerance values of two optimization targets, taken as -10 dB and -15 dB, respectively. $S_{11}(i)$ is the modified return loss at the ith frequency point after pre-processing by (9). $S_{21}(i)$ is the modified isolation between two ports at the ith frequency point after normalization by

$$S_{21}(i) = \begin{cases} -15, S_{21}(i) < -15 \\ S_{21}(i), S_{21}(i) \geq -15. \end{cases} \quad (12)$$

Finally, the QOGWO algorithm iterates until the individual corresponding to the best fitness value (i.e., the optimal solution) is generated. The optimized parameters of the antenna are: DP_1 = 21.23 mm, GL_1 = 16.48 mm, GL_3 = 0.5 mm, GP_2 = 14.08 mm, LF = 16.2 mm, and TP = 0.5 mm. The hybrid QOGWO-GPR algorithm takes a total of 20.18 hours.

The values are input into HFSS simulation, and the results are shown in Figs. 8 and 9. After optimization, two operation frequency bands are 2.40 GHz – 2.50 GHz and 4.36 GHz – 6.39 GHz, and the isolation of all operation frequency bands is less than -15 dB. In [14], the two bandwiths at 2.45 and 5 GHz are 36.7% and 23.8%, respectively. And these are 40.8% and 40.6%, respectively, in the hybrid QOGWO-GPR. Table 5 shows the comparison between the Ref. [14] and the hybrid algorithm.

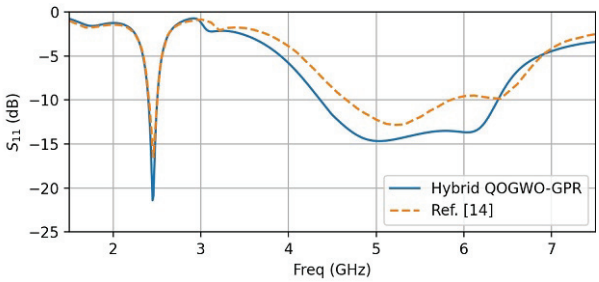


Fig. 8. Comparison of S_{11} curves between Ref. [14] and the hybrid algorithm.

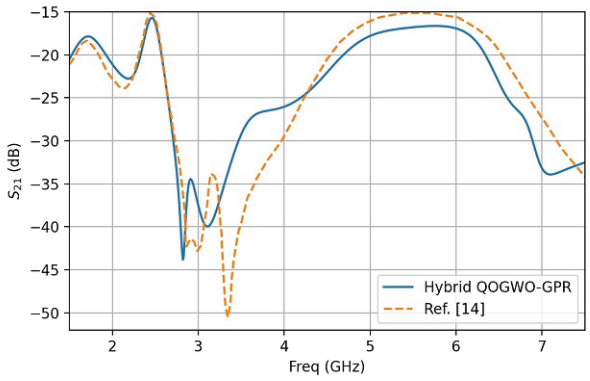


Fig. 9. Comparison of S_{21} curves between Ref. [14] and the hybrid algorithm.

Table 5: Comparison between Ref. [14] and the hybrid algorithm

	2.45GHz	5GHz
Ref. [14]	2.39 GHz-2.48 GHz	4.64 GHz-5.83 GHz
Hybrid	2.40 GHz-2.50 GHz	4.36 GHz-6.39 GHz

V. CONCLUSION

In this paper, the hybrid QOGWO-GPR algorithm is presented to improve the GWO algorithm in antenna optimization. It is prone to global optimality, high precision for complex problems and fast convergence rate at the later stage. It can also reduce the optimization time of the antenna. We have given two examples to validate the proposed algorithm in this work. At low time cost, for the stepped ultrawideband monopole antenna, the impedance bandwidth can reach 113.97% and the -3 dB gain bandwidth can reach 7.53 GHz; for the dual-band MIMO antenna for WLAN, the impedance bandwidths can reach 40.8% and 40.6% at 2.45 and 5 GHz, respectively, and the isolation of all operation frequency bands is less than -15 dB.

ACKNOWLEDGMENT

This work was supported in part by the NKRDP (2019YFB1803202) and the NSFC (62131008 and 62293492).

REFERENCES

[1] A. Nysaeter, "Two-way MIMO sparse array antenna optimization with NSGA-III," *2020 IEEE Radar Conference (Radar-Conf20)*. IEEE, pp. 1-6, 2020.

[2] A. Singh, R. Mehra, and V. Pandey, "A hybrid approach for antenna optimization using cat swarm based genetic optimization," *Advanced Electromagnetics*, vol. 7, no. 3, pp. 23-34, 2018.

[3] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46-61, 2014.

[4] Y. Zhang, J. Li, and M. Zhang, "Motion key frame extraction based on grey wolf optimization algorithm," *MATEC Web of Conferences*, vol. 232, no. 03032, pp. 1-5, 2018.

[5] C. Han, M. Chen, L. Pan, and X. Chen, "A community detection algorithm by utilizing grey wolf optimization," *2017 9th International Conference on Modelling, Identifcation and Control (ICMIC)*, IEEE, pp. 567-572, 2017.

[6] H. Faris, I. Aljarah, M. A. Al-Betar, and S. Mirjalili, "Grey wolf optimizer: A review of

recent variants and application," *Neural Computing and Applications*, vol. 30, no. 2, pp. 413-435, 2018.

[7] C.-F. Wang, K. Liu, and P.-P. Shen, "A novel genetic algorithm for global optimization," *Acta Mathematicae Applicatae Sinica, English Series*, vol. 36, no. 2, pp. 482-491, 2020.

[8] L. Peng and Y. Wang, "Differential evolution using uniform-quasiopposition for initializing the population," *Information Technology Journal*, vol. 9, no. 8, pp. 1629-1634, 2010.

[9] S. Rahnamayan, H. R. Tizhoosh, and M. M. Salama, "Quasi-oppositional differential evolution," *2007 IEEE Congress on Evolutionary Computation*, pp. 2229-2236, 2007.

[10] J. R. Shewchuk, "An introduction to the conjugate gradient, method without the agonizing pain," Ph.D. dissertation, Carnegie-Mellon University, 1994.

[11] S. Sun, "Compact wideband microstrip antenna design based on improved genetic algorithm," Ph.D. dissertation, Beijing University of Posts and Telecommunications, 2011.

[12] E. Schulz, M. Speekenbrink, and A. Krause, "A tutorial on gaussian process regression: Modelling, exploring, and exploiting functions," *Journal of Mathematical Psychology*, vol. 85, pp. 1-16, 2018.

[13] T. O. Hodson, "Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not," *Geoscientifc Model Development Discussions*, pp. 1-10, 2022.

[14] T. Ma, C. Du, and Z. Jiao, "Miniaturized dual-band MIMO antenna for WLAN," *Instrumentation Technology*, pp. 7-10, 2019.

Hao-Yun Zhu received the M.S. degree in computer technology in 2022 from Southeast University, Nanjing, China. During the graduate period, he majored in electromagnetic computation, and his research interest focuses on optimization algorithms for electromagnetic simulations.

Jia-Wei Qian received the B.S. degree in electronic and information engineering in 2020 from the Nanjing University of Posts and Telecommunications, Nanjing, China. He is currently working toward for the master's degree in electronic and information engineering at Southeast University. His current research interest is full-wave algorithm in time domain and optimization algorithms.

Xiao-Hui Tang received the B.S. degree in electromagnetic fields and wireless technology in 2021 from the Nanjing University of Posts and Telecommunications, Nanjing, China, where she is currently working toward the master's degree in electronic and information engineering at Southeast University. Her current research interest is optimization algorithms for antenna array.

Wei-Dong Li received the M.S. degree in mathematics and the Ph.D. degree in radio engineering from Southeast University, Nanjing, China, in 2003 and 2007, respectively. He is currently an associate professor of School of Information Science and Engineering. From January 2008 to January 2009, he was a visiting scholar with the Technische Universität Darmstadt, Germany. His research interests include optimization of sparse antenna array, integral equation numerical modeling and fast algorithm, fast and accurate inter/extrapolation techniques, and DGTD in computational EM. He has authored or co-authored over 40 technical papers. He serves as reviewers for *IEEE Transactions on Antennas and Propagation* and *IET Microwave, Antennas and Propagation*.

Realization of an Optimum Load for Wireless Power Transfer System

Chaoling Wang[1] and Qi Wu[1,2]

[1]School of Electronics and Information Engineering
Beihang University, Beijing, 100191, China
wangchaoling@buaa.edu.cn

[2]Department of Fundamental Research
Zhongguancun Laboratory, Beijing, China
qwu@buaa.edu.cn

***Abstract* –** Wireless power transfer (WPT) system has been an integral part of personal living since its regained interest, especially the magnetic resonance (MR) scheme. MR-WPT scheme suffers, however, change of the coil separation distance and various alignment errors. This paper reports a realization of optimum load for MR-WPT system, which can change the loading impedance accordingly for different coupling coefficients between the Tx and Rx coils. A simple varactor circuit is adopted to realize the optimum load curve. Usefulness of this is demonstrated through both the steady and transient analysis. The proposed realization relies on an open-circuit scheme, and hence it is suitable for scenarios with low-cost and small-size requirements.

***Index Terms* –** Coupling coefficient, load impedance, magnetic resonance, wireless power transfer.

I. INTRODUCTION

Wireless power transfer (WPT) technology has regained interest since the demonstration of magnetic resonance (MR) WPT scheme [1–4]. The basic elements of the MR-WPT system consist of a power source, a Tx coil, a Rx coil, and a termination load [5–7] The Tx and Rx coils of the MR-WPT system must be resonant at the same frequency to achieve the optimum power transfer efficiency [8]. In addition, the load impedance must comply with the coupling coefficient of the Tx and Rx coils in order to obtain the maximum transfer efficiency or the highest output voltage after rectification [9–11]. The requirement of an optimum load impedance poses a strict limitation on the scenarios for practical use. For instance, if a fixed load impedance is adopted in the MR-WPT system, the users must make a very precise alignment of the Tx and Rx coils with appropriate separation distance. It is difficult to achieve such an accurate alignment in a WPT system, charging a mobile phone or an electric vehicle, for example.

To overcome the limitations of accurate alignment of Tx and Rx coils in the MR-WPT system, a variable load impedance approach can be adopted. A strict derivation of an optimum load impedance has been discussed in a few works [6], [12]. In addition, a massive number of measurements were performed and a detailed diagram of the transfer efficiency and separation distance was obtained [13]. Those works build a solid background of the optimum load in MR-WPT systems.

Realization of optimum load impedance has been an active research topic in the past decade. In general, two different approaches have been extensively discussed: Closed-loop and open-loop scenarios. The closed-loop scenarios usually enjoy better flexibility and also slightly higher efficiency. For instance, two switch-controlled capacitors were added to the Tx and Rx coils for optimizing the transmission efficiency [14]. The operating frequency of MR-WPT can be also alternated in a closed-loop fashion through variable compensation capacitors with a constant load [15]. Several other closed-loop scenarios were also reported using reconfigurable structures to achieve high energy efficiency over different load impedances [16–19]. One drawback of the closed-loop WPT system is high complexity in the whole system. Therefore, the open-loop WPT system is favored for low-cost applications. In addition, for the case with multiple Rx coils, the open-loop architecture is more flexible to use [20].

In this paper, a new realization of an optimum load for WPT systems is reported, which can change the loading impedance accordingly for different coupling coefficients between the Tx and Rx coils. A simple varactor circuit is adopted to realize the optimum load curve. Usefulness of this realization is demonstrated through both the steady and transient analysis. The proposed realization relies on an open-circuit scheme, and hence it is suitable for the scenarios with low-cost and small-size requirements.

Submitted On: February 22, 2023
Accepted On: April 27, 2023

https://doi.org/10.13052/2023.ACES.J.380907

II. STATEMENT OF THE PROBLEM

A two-coil WPT system shown in Fig. 1 is discussed in this paper. The Tx and Rx coils are identical, and their structural parameters are the same as the case studied in [12]. The coils have a diameter of 8 cm and 10 turns. For practical use, diameter of the Tx and Rx coils should be determined by the allowed space for accommodating these coils. It is also noted that the optimum separation distance between the Tx and Rx coils is usually in the same order as the diameter of the coils [21]. Therefore, there is a fundamental trade-off between the size and coverage area for the coils. Moreover, turns of the coils would determine their resonant frequency. If the resonant frequency is lower than the desired one, an additional capacitor may be used to tune the resonance, as illustrated in Fig. 1 (a). In the presented configuration, the first anti-resonance frequency is approximately 29.5 MHz. Therefore, a capacitor is added in parallel to the Tx and Rx coils, respectively, and the coils are resonant at ~13.6 MHz.

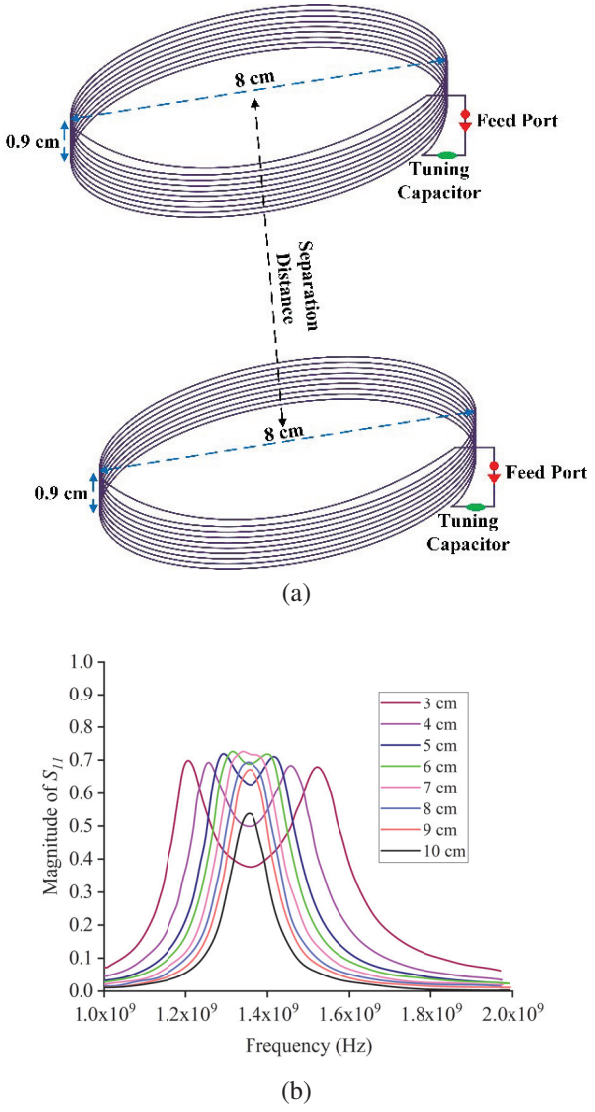


Fig. 1. (a) Structural diagram of the WPT system, and (b) power transfer efficiency versus the separation distance between the Tx and Rx coils.

If a constant load impedance of 50 ohms is adopted, the power transfer efficiency depends largely on the separation distance between the two Tx and Rx coils, as demonstrated in Fig. 1 (b). Some frequency splitting phenomena can be also seen when the separation distance is too small. Generally, if the separation distance is approximately equal to the diameter of the coils, the WPT system yields a high transfer efficiency [21]. In addition, the resonant frequency is hardly alternated in this separation distance. The predefined settings and obtained results serve as a reference for the presented investigation on the realization of an optimum load impedance.

III. ANALYSIS OF THE OPTIMUM LOAD IMPEDANCE

A. Expression of the optimum load impedance

In order to achieve a maximum transfer efficiency, the load impedance of the Rx coil should be alternated accordingly to the separation distance between the Tx and Rx coils. In practice, the load impedance is related to the magnetic coupling coefficient of the Tx and Rx coils, which can be read as [6], [16]

$$Z_L = \left(\omega L_2 \sqrt{1+\Delta}\right) \Big/ Q_2 - j\omega L_2, \qquad (1)$$

where L_2 is the self-inductance of the Rx coil, Q_2 is the quality factor of the Rx coil. Since $\Delta = K_{12}^2 Q_1 Q_2$, K_{12} is the magnetic coupling coefficient of the Tx and Rx coils, and Q_1 is the quality factor of the Tx coil.

For the WPT system shown in Fig. 1 (a), the optimum load impedance can be evaluated through (1) and the curves are shown in Fig. 2. It is seen that the optimum load impedance is a complex number. In addition, both the real and imaginary parts of the optimum

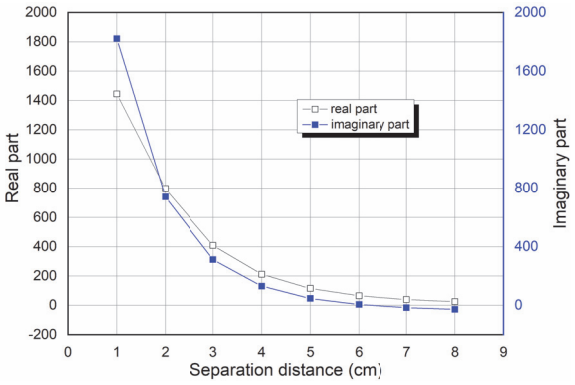


Fig. 2. The optimum load impedance versus the separation gap between the Tx and Rx coils.

load impedance tend to decrease with the separation gap. This phenomenon is related to the factor that the coupling coefficient K_{12} is reduced for a large separation distance. Before we proceed to simulate this optimum load impedance, it is interesting to investigate the behaviors of the WPT system with such an optimum load impedance.

B. *V-I* characteristics of the optimum load impedance

Based on the analysis in Part A, the optimum load impedance for the WPT system can be evaluated as shown in Fig. 2. However, the inherent characteristics of the optimum load impedance is not clear. The voltage-current (*V-I*) characteristic curve is a useful tool for analyzing the load impedance. It this case, we assume that an excitation voltage is applied to the Tx coil, and the open-circuit (OC) voltage of the Rx coil can be obtained for different separation distances as illustrated in Fig. 3. With the OC voltage applied to the optimum load impedance, we can further obtain the *V-I* curve of the optimum load impedance, as shown in Fig. 4.

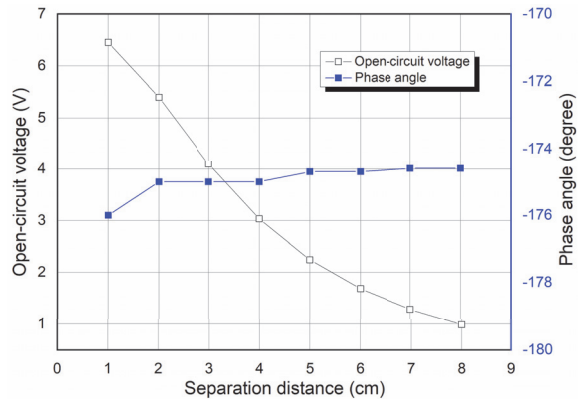


Fig. 3. The open-circuit voltage of the Rx coil versus the separation gap between the Tx and Rx coils.

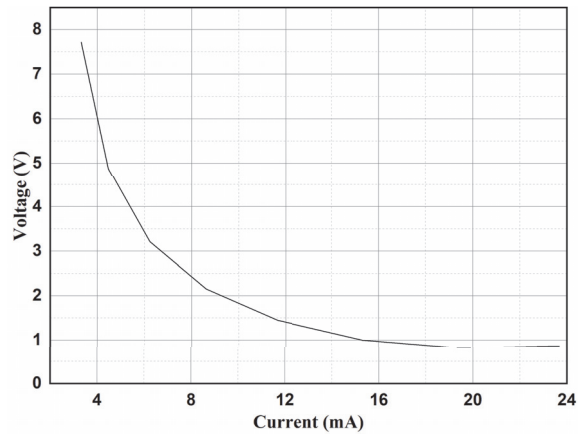


Fig. 4. *V-I* curve of the optimum load.

The *V-I* curve shows clearly that the optimum load behaves like a voltage-controlled resistor, in which its resistance reduces rapidly if the OC voltage increases.

One may denote such an ideal *V-I* curve using the curve fitting technology. Based on our investigation, a five-order polynomial expression is sufficient for depicting such a *V-I* curve, which reads

$$I = \sum_n a_n \cdot V^n = a_5 \cdot V^5 + a_4 \cdot V^4 + a_4 \cdot V^4 + a_3 \cdot V^3 + a_2 \cdot V^2 + a_1 \cdot V + a_0. \quad (2)$$

Equation (2) yields a good accuracy for this problem.

Table 1: Curve-fitting coefficients of the *V-I* curve

Coefficient	Value	Coefficient	Value
a_5	$-9.150\text{E}-5$	a_4	**1.767E − 3**
a_3	$-1.256\text{E}-2$	a_2	$4.164\text{E}-2$
a_1	$-6.740\text{E}-2$	a_0	$5.250\text{E}-2$

IV. REALIZATION OF THE OPTIMUM LOAD IMPEDANCE

The optimum load impedance yields certain nonlinear responses with regard to the OC voltage across the load of the Rx coil. A possible scheme for realizing such optimum load impedance is illustrated in Fig. 5. It consists of a fixed resistor and a variable capacitor. It is noted that capacitance of the variable capacitor should be a function of the applied voltage U as

$$C = f(U). \quad (3)$$

The current induced by the applied voltage U across the optimum load is then obtained by

$$I = \frac{d[f(U) \cdot U]}{dt} + \frac{U}{R}. \quad (4)$$

We can further expand (4) into three items as

$$I = U \cdot \frac{df(U)}{dt} + f(U) \cdot \frac{dU}{dt} + \frac{U}{R}. \quad (5)$$

Consider a time-harmonic excitation as

$$U = U_0 \cdot \cos(\omega t + \theta_0). \quad (6)$$

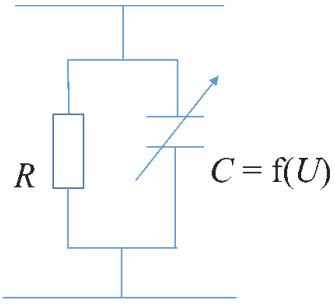


Fig. 5. The proposed realization of the optimum load.

By comparing (2) and (5), it can be concluded that the unknown function $C = f(U)$ must be a fourth-order polynomial of U to achieve the required fifth-order approximation in (2).

For verification of the proposed realization of the optimum load impedance, the MR-WPT system is modeled with an equivalent circuit model (ECM) [10], [11], as shown in Fig. 6. The ECM consists of three parts: the Tx coil, the Rx coil, and the realized load. Transient simulations of the ECM are performed, and the results are shown in Fig. 7. A unit step is added to the feeding point of the Tx coil, and an ordinary 50 Ω load and the optimum load are used to connect the Rx coil, respectively. It is seen that in the beginning the realized load is over-voltaged and then its capacitance starts to change. After a short time, the whole system is stable with a constant output current. This shows that the proposed realization of the optimum load impedance can compensate the WPT system with different separation distances. On the other hand, if a 50 Ω load is adopted to connect the Rx coil, the induced current across the load is much smaller in magnitude. The current is faded into zero in a short time, as the WPT system yield low quality factor with a 50 Ω load.

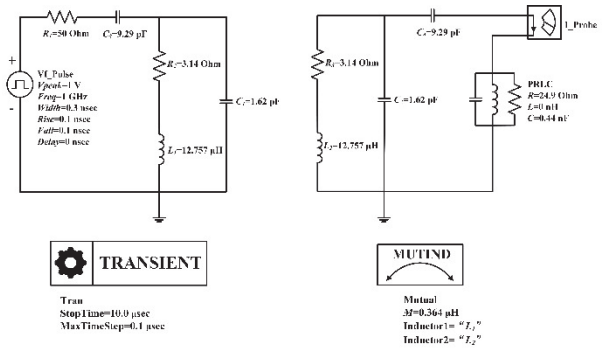


Fig. 6. ECM of the WPT system with optimum load.

Time-harmonic solution may be also applied to the ECM model in Fig. 6. For the same separation distance (i.e., 4 cm), the induced currents at the Rx at 13.56 MHz are shown in Fig. 8. It is clearly seen that the optimum load reaches stable power transmission in a short time. But the WPT system with 50 Ω load yields some fluctuations in a much longer time, which affects the transmission efficiency at last. Figure 9 compares the transmission efficiency of the same WPT system with the optimum loads and 50 Ω load at 13.56 MHz. The transmission efficiency of the WPT system with 50 Ω load decreases significantly with a reduced separation distance, as the resonant frequency splits as shown in Fig. 1 (b). The usefulness of the optimum loads can be comprehended from those results.

The cost of the whole system would not increase by loading extra impedance. We may use a varactor to realize the optimum loads, which is cost efficient if

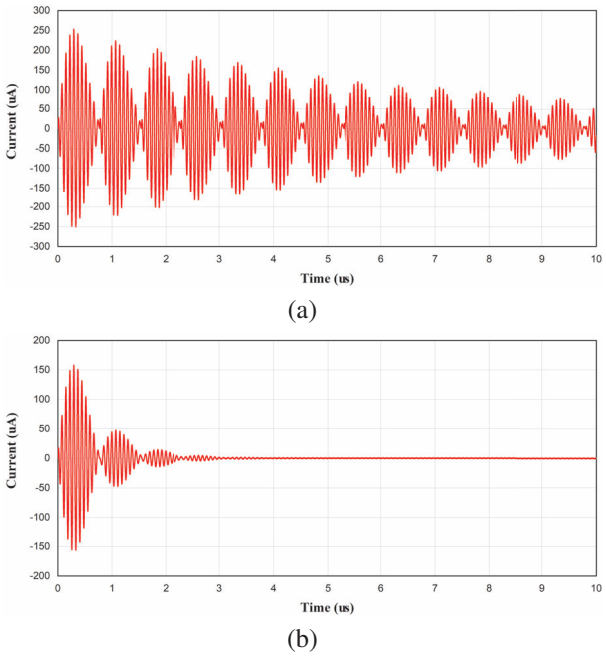


Fig. 7. Induced current across the load of the Rx coil (separation distance equals 4 cm, unit pulse is applied): (a) optimum load and (b) 50 Ω load.

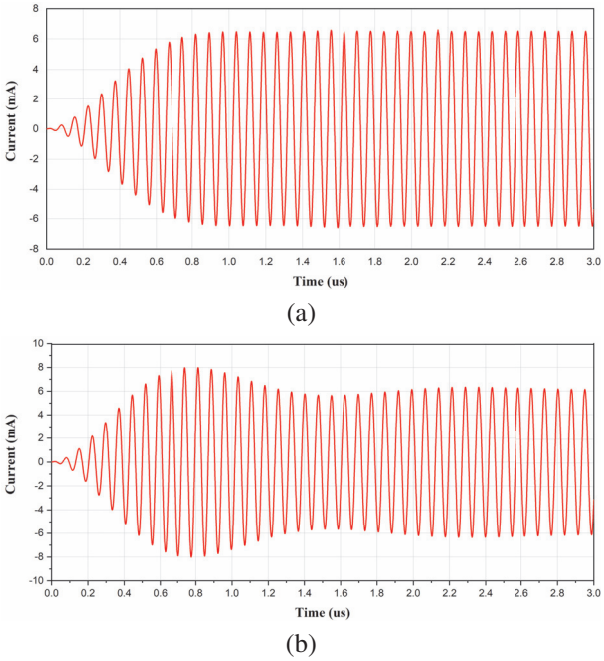


Fig. 8. Induced current across the load of the Rx coil (separation distance equals 4 cm, time-harmonic excitation at 13.56 MHz): (a) optimum load and (b) 50 Ω load.

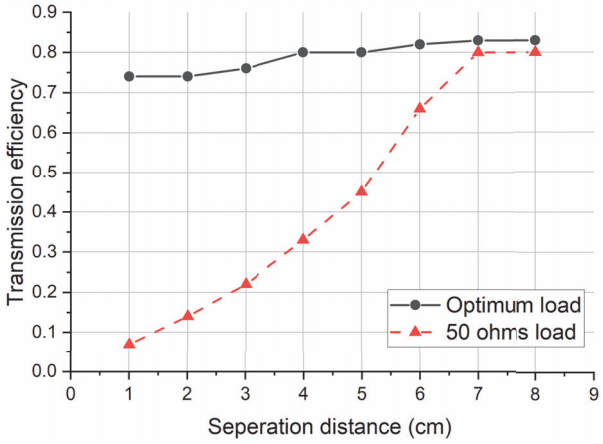


Fig. 9. Transmission efficiency of the WPT system with the optimum load and 50 Ω load.

compared with other methods. This is an advantage of the presented method. Since the selection of a proper varactor is another important topic, the experimental results are not available in this work.

The Rx coil is a passive device and hence one may not consider its sensitivity. The Rx coil would drive a rectifier circuit for extracting power from the WPT system, and the output voltage of the Rx coil is an important factor. The loading varactor would not affect the output voltage of the Rx coil; instead it can keep the whole system stable and more efficient.

V. CONCLUSION

In this paper, a new realization of an optimum load for a WPT system is reported, which can change the loading impedance accordingly for different coupling coefficients between the Tx and Rx coils. A simple varactor circuit is adopted to realize the optimum load curve. Usefulness of this realization is demonstrated through both the steady and transient analyses. The proposed realization relies on an open-circuit scheme, and hence it is suitable for the scenarios with low-cost and small-size requirements.

ACKNOWLEDGMENT

This work is supported in part by the National Natural Science Foundation of China under grant U2141230.

The authors thank Mr. Zhongkui Wen and Yijie Yang for assisting some numerical simulations.

REFERENCES

[1] A. Kurs, A. Karalis, R. Moffatt, J. D. Joannapoulos, P. Fisher, and M. Soljacic, "Wireless power transfer via strongly coupled magnetic resonances," *Science*, vol. 317, pp. 83-86, 2007. https://www.science.org/doi/10.1126/science.1143254

[2] Y. Guo, L. Wang, and C. Liao, "A general equivalent model for multi-coil wireless power transfer system analysis and its applications on compensation network design," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 33, no. 6, pp. 648-646, June 2018. https://journals.riverpublishers.com/index.php/ACES/article/view/9121

[3] G. Perez-Greco, J. Barreto, A.-S. Kaddour, and S. V. Georgakopoulos, "Effects of the human body on wearable wireless power transfer system," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 35, no. 11, pp. 1454-1456, Nov. 2020. https://journals.riverpublishers.com/index.php/ACES/article/view/7655

[4] D. Kim, A. T. Sutinjo, and A. Abu-Siada, "Near-field analysis and design of inductively-coupled wireless power transfer system in FEKO," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 35, no. 1, pp. 82-93, Jan. 2020. https://journals.riverpublishers.com/index.php/ACES/article/view/8043

[5] D. Schneider, "Wireless power at a distance is still far away [Electrons Unplugged]," *IEEE Spectrum*, vol. 47, no. 5, pp. 34-39, May 2010. https://ieeexplore.ieee.org/document/5453139

[6] H. Ron, W. Zhong, and C. K. Lee, "A critical review of recent progress in mid-range wireless power transfer," *IEEE Transactions on Power Electronics*, vol. 29, no .9, pp. 4500-4511, Sep. 2014. https://ieeexplore.ieee.org/document/6472081/

[7] N. Kyungmin, J. Heedon, M. Hyunggun, and B. Franklin, "Tracking optimal efficiency of magnetic resonance wireless power transfer system for biomedical capsule endoscopy," *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 1 pp. 295-304, Jan. 2015. https://ieeexplore.ieee.org/document/6957601/

[8] H. Nguyen and I. A. Johnson, "Splitting frequency diversity in wireless power transmission," *IEEE Transactions on Power Electronics*, vol. 30, no. 11, pp. 6088-6096, Nov. 2015. https://ieeexplore.ieee.org/document/7089278/

[9] S. Cheon, Y.-H. Kim, S.-Y. Kang, M. L. Lee, J.-M. Lee, and T. Zyung, "Circuit-model-based analysis of a wireless energy-transfer system via coupled magnetic resonances," *IEEE Trans. Industrial Electron.*, vol. 58, no. 7, pp. 2906-2914, July 2011. https://ieeexplore.ieee.org/document/5560805/

[10] J. Bito, S. Jeong, and M. M. Tentzeris, "A real-time electrically controlled active matching circuit utilizing genetic algorithms for wireless power transfer to biomedical implants," *IEEE Trans. Microw. Theo. Techniq.*, vol. 64, no. 2, pp. 365-374,

Feb. 2016. https://ieeexplore.ieee.org/document/7384758

[11] Z. Zhang, H. Pang, A. Georgiadis, and C. Cecati, "Wireless power transfer – An overview," *IEEE Trans. Industrial Electron.*, vol. 66, no. 2, pp. 1044-1058, Feb. 2019.

[12] P. Liang, Q. Wu, H. Bruens, and C. Schuster, "Efficient modeling of multi-coil wireless power transfer systems using combination of full-wave simulation and equivalent circuit modeling," *2018 IEEE Int. Symp. Electromag. Compat. and 2018 IEEE Asia-Pacific Symp. Electromag. Compat. (EMC/APEMC)*, Singapore, pp. 466-471, 2018. https://ieeexplore.ieee.org/document/8393822

[13] Z. Wen, Q. Wu, O. F. Yildiz, and C. Schuster, "Design of experiments for analyzing the efficiency of a multi-coil wireless power transfer system using polynomial chaos expansion," *2019 Joint Int. Symp. Electromag. Compat., Sapporo and Asia-Pacific Int. Symp. Electromag. Compat. (EMC Sapporo/APEMC)*. https://ieeexplore.ieee.org/document/8893695

[14] J. Zhang, J. Zhao, Y. Zhang, and F. Deng, "A wireless power transfer system with dual switch-controlled capacitors for efficiency optimization," *IEEE Trans. Power Electron.*, vol. 35, no. 6, pp. 6091-6101, June 2020. https://ieeexplore.ieee.org/document/8892619

[15] F. Grazian, T. B. Soeiro, and P. Bauer, "Inductive power transfer based on variable compensation capacitance to achieve an EV charging profile with constant optimum load," *IEEE Journal Emerging Selected Topics in Power Electron.*, vol. 11, no. 1, pp. 1230-1244, Feb. 2023. https://ieeexplore.ieee.org/document/9813693

[16] D. Ahn, S. Kim, J. Moon, and I.-K. Cho, "Wireless power transfer with automatic feedback control of load resistance transformation," *IEEE Trans. Power Electron.*, vol. 31, no. 11, pp. 7876-7886, Nov. 2016. https://ieeexplore.ieee.org/document/7368178

[17] W. Zhong, and S.Y. Hui, "Reconfigurable wireless power transfer systems with high energy efficiency over wide load range," *IEEE Trans. Power Electron.*, vol. 33, no. 7, pp. 6379-6390, July 2018. https://ieeexplore.ieee.org/document/8024081

[18] D.-G. Seo, S.-H. Ahn, J.-H. Kim, S.-T. Khang, S.-C. Chae, J.-W. Yu, and W.-S. Lee, "Power transfer efficiency for distance-adaptive wireless power system," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 33, no. 10, pp. 1171-1174, Oct. 2018. https://journals.riverpublishers.com/index.php/ACES/article/view/8973

[19] N. Fortana, S. Barmada, M. Raugi, D. Brizi, and A. Monorchio, "Spiral resonator arrays for misalignment compensation in wireless power transfer system," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 37, no. 7, pp. 765-773, July 2022. https://journals.riverpublishers.com/index.php/ACES/article/view/18111

[20] M. Fu, H. Yin, M. Liu, Y. Wang, and C. Ma, "A 6.78 MHz multiple-receiver wireless power transfer system with constant output voltage and optimum efficiency," *IEEE Trans. Power Electron.*, vol. 33, no. 6, pp. 5330-5340, June 2018. https://ieeexplore.ieee.org/document/7976358/

[21] M. Fu, H. Yin, X. Zhu, and C. Ma, "Analysis and tracking of optimal load in wireless power transfer systems," *IEEE Trans. Power Electron.*, vol. 30, no. 7, pp. 3952-3963, July 2015. https://ieeexplore.ieee.org/document/6876181

Chaoling Wang received the B.E. Degree from Airforce Engineering University in communication engineering. His research interest includes fiber communications and photoelectronic imaging technology. He is currently pursuing the Ph.D. degree at Beihang University, Beijing, China.

Qi Wu received the B.S. degree from East China Normal University, Shanghai, China, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, both in electrical engineering, in 2004 and 2009, respectively.

He joined the faculty of School of Electronics and Information Engineering, Beihang University, Beijing, China, in 2009 and now is a full professor. During 2011 and 2012, he was a visiting scholar in the Department of Electrical Engineering, University of California, Los Angeles. During 2014 and 2016, he was an Alexander von Humboldt Fellow in the Institute of Electromagnetic Theory, Technical University of Hamburg, Germany. He has authored over 40 journal papers and two books and holds 20 patents as the first inventor. His research interests include broadband antennas, computational electromagnetics, and related EMC topics.

Dr. Wu received the Young Scientist Award from the International Union of Radio Science (URSI) in 2011, the Nominee Award for Excellent Doctoral Dissertation from the National Minister of Education in 2012, Young Scientist Award of APEMC in 2016, and Excellent Researcher from Chinese institute of Electronics in 2020.

Antenna Shape Modeling based on Histogram of Oriented Gradients Feature

Hai-Ying Luo[1], Wen-Hao Su[1], Haiyan Ou[1], Sheng-Jun Zhang[2], and Wei Shao[1]

[1]School of Physics
University of Electronic Science and Technology of China, Chengdu, 611731, China
hyluo@std.uestc.edu.cn, suwenhao1202@163.com, ouhaiyan@uestc.edu.cn, weishao@uestc.edu.cn

[2]National Key Laboratory of Science and Technology on Test Physics and Numerical Mathematics
Beijing 100076, China
zhangsj98@sina.com

***Abstract* –** A least square support vector machine (SVM) model is proposed for shape modeling of slot antennas. The slot image is mapped into the electromagnetic response by the SVM model. A modified shape-changing technique is also proposed to describe the antenna geometry by the quadratic uniform B-spline curve and generate the slot images. In the model, the histogram of oriented gradients feature is extracted from the slot images to show the appearance and shape of the slot. The relationship between the histogram of oriented gradient features and the electromagnetic responses is preliminarily built on SVM and the transfer function. Then a radial basis function network is used for error correction. The effectiveness of the proposed model is confirmed with an example of a tri-band microstrip-fed slot antenna. Compared with the convolutional neural network (CNN), the feature extracted by CNN is substituted by the histogram of oriented gradients feature, and the proposed model shows the same accuracy and the improvement of training efficiency.

***Index Terms* –** B-spline curve, microstrip antenna, shape modeling, support vector machine.

I. INTRODUCTION

In recent years, the artificial neural network has been widely applied to the modeling of antennas, speeding up the design process [1–3]. The mapping relation between the antenna geometry and the electromagnetic response is learned by the neural network model. Although the generation of training and testing samples needs a certain number of full-wave simulations, the trained model can predict the response quickly and accurately.

A parametric model based on the artificial neural network is proposed for predicting the input resistance of a broadband antenna [4]. The frequency-dependent resistance envelope of the antenna is parametrized by a Gaussian model, and the neural network maps the geometrical parameter of the antenna to the Gaussian parameters. To build a neural network with high dimension of geometrical parameter space and large geometrical variations, the transfer function is employed to extract the feature of *S*-parameters, representing electromagnetic responses versus frequency [5, 6]. In this model, the neural network predicts the transfer function coefficients as a function of geometrical parameters. In [2], an artificial neural network (ANN) model with three parallel and independent branches is proposed. This model describes the antenna performance with various parameters and simultaneously output *S*-parameters, gain, and radiation pattern of a Fabry-Perot resonator antenna. [3] proposes a support vector machine (SVM) model to learn the mapping relation from the slot-position and slot-size to electromagnetic responses. Compared with the ANN model, the SVM model costs less time in training and predicting.

For antennas with special shapes, their geometries cannot be readily parametrized. Therefore, it is difficult to model these antennas. In [7], a convolutional neural network (CNN) model is proposed for predicting the resonant frequency of pixelated patch antennas. The input to CNN is not the geometric parameters but the image of the pixelated patch antenna. [8] proposed a CNN model for the shape modeling of a metallic strip for the microstrip filter. The input to CNN is the image of the metallic strip with different shapes. The metallic strip with different shape is generated by the shape-changing technique, which is based on the cubic spline interpolation. Although the CNN model can change the component/antenna geometry flexibly and expand the solution domain, it is difficult to determine CNN hyper-parameters due to their huge number, resulting in a time-consuming training process of CNN. The performance of a machine learning model is controlled by the hyper-parameters, such as the learning rate and the number of layers. The traditional model (such as the

Submitted On: February 27, 2023
Accepted On: May 1, 2023

https://doi.org/10.13052/2023.ACES.J.380908

least square SVM) performs well with several tuned hyper-parameters, while there are forty or more hyper-parameters in the neural network.

This paper proposes an improved least square SVM model for the shape modeling of microstrip-fed slot antennas based on the histogram of oriented gradients feature [9]. Based on the quadratic uniform B-spline curve, the training and testing samples with different structures are generated. The electromagnetic responses are preliminarily predicted by the combination of the histogram of oriented gradients feature, SVM, and transfer function. The input for the least square SVM is the histogram of oriented gradients feature extracted from slot images. The output for SVM is the coefficients of the pole-residue-based transfer function. Then a radial basis function network is employed as error correction. An example of the tri-band microstrip-fed slot antenna is selected to confirm the validity of the proposed model. In this example, the *S*-parameter and radiation pattern are predicted.

II. PROPOSED MODEL

As shown in Fig. 1, the proposed SVM model integrates the histogram of oriented gradients, transfer function, and radial basis function network. The model learns the mapping relation between the slot image and electromagnetic responses (*S*-parameter or radiation patterns). The slot images are generated from the defined control points with a modified shape-changing technique. Once the slot images are given, the trained model, which substitutes the full-wave simulation, can predict S_{11} and radiation patterns accurately and quickly.

A. Modified shape-changing technique

In [7], a shape-changing technique is applied one by one to the sides which need to be changed, and the contour of the metallic strip is modeled by the spline curve. To simplify the process, a modified shape-changing technique based on the quadratic uniform B-spline curve is proposed here to change the contour of the metallic patch or slot more flexibly. The function of contour and its digital image, obtained by the modified shape-changing technique, are used for the full-wave simulation and the model input, respectively.

With the modified shape-changing technique, the contour of slot is determined by an iteration process as shown in Fig. 2. An ellipse slot is taken as an example to show the iteration.

First, the slot domain is discretized with a coarse square gird with the side length *d,* as shown in Fig. 2 (a), where d is defined based on the complexity of the slot contour and the training time. The positions of the grid lines are $x = -d/2 + i_x d$ and $y = -d/2 + i_y d$, where $i_x \in \mathbb{Z}$ and $i_y \in \mathbb{Z}$. Second, a set of the special grid centers are selected as the control

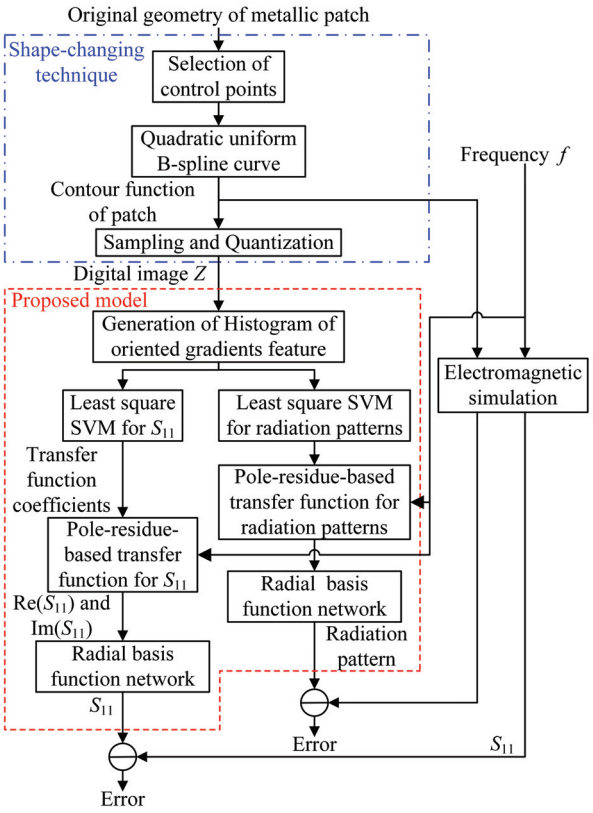


Fig. 1. Structure of the proposed model and the shape-changing technique.

points. At least one edge of the special grids intersects with the contour. Here, the control points are denoted counterclockwise by (x_1, y_1), (x_2, y_2), …, (x_n, y_n) in Fig. 2 (a). Then, the control points are shifted within the corresponding grid, i.e., the kth shifted control point $(x'_k, y'_k) = (x_k + \Delta x_k, y_k + \Delta y_k)$, where $-d/2 \le \Delta x_k \le d/2$ and $-d/2 \le \Delta y_k \le d/2$. To generate a closed curve, $(x'_n, y'_n) = (x'_0, y'_0)$ and $(x'_{n+1}, y'_{n+1}) = (x'_1, y'_1)$ should be satisfied. For the ellipse slot, shifted control points are shown in Fig. 2 (b). Fourth, the new contour based on the quadratic uniform B-spline curve can be defined in segments by n parametric curves with control points (x'_k, y'_k) for $k = 0, 1, 2, \ldots, n+1$. The kth parametric curve between $(0.5x'_{k-1}+0.5x'_k, 0.5y'_{k-1}+0.5y'_k)$ and $(0.5x'_k+0.5x'_{k+1}, 0.5y'_k+0.5y'_{k+1})$ over the local parameter interval $\{t \mid 0 \le t \le 1\}$ is

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix}^{\mathrm{T}} = \frac{1}{2}\begin{bmatrix} 1 & t & t^2 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ -2 & 2 & 0 \\ 1 & -2 & 1 \end{bmatrix}\begin{bmatrix} x'_{k-1} & y'_{k-1} \\ x'_k & y'_k \\ x'_{k+1} & y'_{k+1} \end{bmatrix}, \quad k = 1, 2, \ldots, n. \tag{1}$$

For the ellipse slot, the corresponding curve for the circular control points is also shown in Fig. 2 (b).

In the next iteration, the coarse grid is shifted in both x- and y-directions by $d/2$. The control points are selected

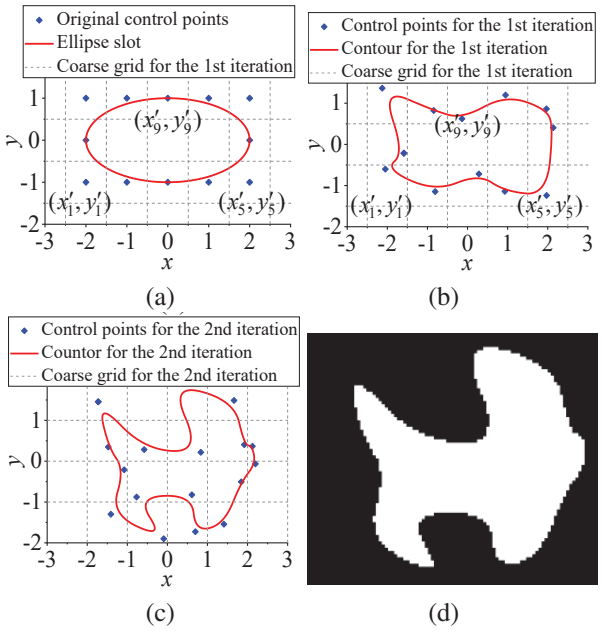


Fig. 2. Diagram of the shape-changing technique for an ellipse slot: (a) Ellipse slot and the original control points, (b) control points and corresponding contour for the first iteration, (c) control points and corresponding contour for the second iteration, and (d) binary image *Z*.

and shifted in a similar way again. In Fig. 2 (c), the positions of the grid lines are $x = i_x d$ and $y = i_y d$, and a new contour is generated by (1) with square control points. The iteration process ends after the maximum number of iterations.

In sampling and quantization [10], the coordinate values are digitized to determine the pixel positions, and each material is denoted by specified numbers. In other words, the slot domain is discretized with a fine square gird, i.e., pixel, with the side length of d_{pixel}. The value of the pixels whose center lies inside the slot is 0, and the value of the others is 1.

B. Histogram of oriented gradients feature

Because the local object appearance and shape can be characterized well by the distribution of local intensity gradients or edge position, the histogram of oriented gradients feature is often used for human detection. In our model, the feature is used as the input of SVM to simply the relationship learned by the neural network.

The feature of the example is visualized over the slot image in Fig. 3. The extraction of the feature can be implemented by the MATLAB function *extractHOGFeatures*.

First, the gradient of each pixel is calculated. Second, the image is divided into small connected regions, i.e., cells, as shown in Fig. 3 (a). The size

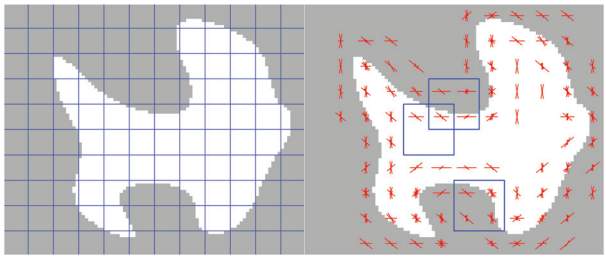


Fig. 3. Extraction of histogram of oriented gradients features: (a) Binary image *Z* with cells, (b) visualization of the histogram of oriented gradients feature and blocks.

of cell is also defined based on the complexity of the slot contour and the training time. Then a local one-dimensional histogram of gradient directions over the pixels of the cell is accumulated for each cell. Nine bins are evenly spaced over the range of 0°-180°. To reduce aliasing, votes are interpolated bilinearly between the neighboring bin centers in both orientation and position. The vote of each pixel is the gradient magnitude. Fourth, the local histogram is contrast-normalized by a measure of the intensity across several cells, called a block, as shown in Fig. 3 (b). The number of overlapping cells between horizontally or vertically adjacent blocks is $h_{\text{overlap}} \times v_{\text{block}}$ or $h_{\text{block}} \times v_{\text{overlap}}$, i.e., the block stride for the horizontal direction is h_{overlap} cells, and the one for the vertical direction is v_{overlap} cells. The block normalization scheme is Lowe-style clipped L2 norm. Then the histogram of oriented gradients feature vector is gotten. For an image with the pixel size of $h_{\text{image}} \times v_{\text{image}}$, the length N of the histogram of oriented gradients feature vector is

$$N = 9h_{\text{block}}v_{\text{block}}\left\lfloor \frac{\frac{h_{\text{image}}}{h_{\text{cell}}} - h_{\text{block}}}{h_{\text{block}} - h_{\text{overlap}}} + 1 \right\rfloor \left\lfloor \frac{\frac{v_{\text{image}}}{v_{\text{cell}}} - v_{\text{block}}}{v_{\text{block}} - v_{\text{overlap}}} + 1 \right\rfloor, \tag{2}$$

where the cell contains $h_{\text{cell}} \times v_{\text{cell}}$ pixels and the block contains $h_{\text{block}} \times v_{\text{block}}$ cells.

C. Multi-output least square SVM

To improve the reliability and accuracy of the whole model, the multi-output least square SVM [11] learns the mapping relation from the histogram of oriented gradients feature $\boldsymbol{h}_i$ to the transfer function coefficients $\boldsymbol{c}_i$ for electromagnetic responses (S_{11} or radiation patterns). The objective function of SVM and corresponding constrains are

$$\begin{aligned} &\min_{\boldsymbol{w}_0,\boldsymbol{v}_j,\boldsymbol{b},\xi_i} \frac{1}{2}\boldsymbol{w}_0^{\mathrm{T}}\boldsymbol{w}_0 + \frac{\gamma_v}{2n_c}\sum_{j=1}^{n_c}\boldsymbol{v}_j^{\mathrm{T}}\boldsymbol{v}_j + \frac{\gamma_\xi}{2}\sum_{i=1}^{m_{\text{sample}}}\xi_i^{\mathrm{T}}\xi_i, \\ &s.t.\ \boldsymbol{c}_i = \boldsymbol{w}^{\mathrm{T}}\varphi(\boldsymbol{h}_i) + \boldsymbol{b} + \boldsymbol{\xi}_i, i = 1,2,\ldots,m_{\text{sample}}, \end{aligned} \tag{3}$$

where a radial basis function is used as the kernel function, i.e., $\varphi(\boldsymbol{h}_i)^T\varphi(\boldsymbol{h}_j) = \exp(-p||\boldsymbol{h}_i - \boldsymbol{h}_j||^2)$, p is a positive kernel scale parameter, weight matrix $\boldsymbol{W} = [\boldsymbol{w}_0 + \boldsymbol{v}_1, \boldsymbol{w}_0 +$

$\boldsymbol{v}_2, \ldots, w_0+v_{n_c}]$ (the smaller the $\boldsymbol{w}_0$ is, the more similar the transfer function coefficients are to each other), $\boldsymbol{b}$ is the bias vector, ξ_i is the slack variable, γ_v and γ_ξ are the positive real regularized parameters, and m_{sample} represents the number of training samples. Compared with CNN, there are much fewer hyper-parameters for the multi-output least square SVM, i.e., p, γ_v, and γ_ξ.

D. Pole-residue-based transfer function

The transfer function is employed to extract the feature of electromagnetic responses [5, 6]. The pole-residue-based transfer function coefficients have low sensitivities with respect to geometrical parameters, and they are a reasonable representation of electromagnetic responses. Thus, the neural network model based on the transfer function is accurate with high dimension of geometrical parameter space and large geometrical variations. The transfer function coefficients can be obtained with vector fitting [12]. The pole-residue-based transfer function is presented as

$$H(s) = \sum_{i=1}^{Q} \left(\frac{r_i}{s - p_i} + \frac{r_i^*}{s - p_i^*} \right), \tag{4}$$

where p_i and r_i represent the pole and residue coefficients of transfer function, respectively, s is the frequency in Laplace domain, Q represents the order of transfer function, and the superscript * represents complex conjugate. The transfer function coefficient $\boldsymbol{c}_i$ is a vector of the real and imaginary part of p_i and r_i.

E. Radial basis function network

The radial basis function network is used for error correction. It learns the relationship between the electromagnetic response (S_{11} or radiation patterns) obtained from the transfer function and the simulated one. As shown in Fig. 4, there are two layers in the radial basis function network, i.e., a hidden radial basis layer with R_1 neurons and an output linear layer with R_0 neurons. The mathematical form is

$$\boldsymbol{y}_{\text{RBF}} = \boldsymbol{w}_2' \exp\left(-\left(\left\| \boldsymbol{w}_1' - \boldsymbol{x}_{\text{RBF}} \right\| \circ \boldsymbol{b}_1' \right)^2 \right) + \boldsymbol{b}_2', \tag{5}$$

where $\boldsymbol{y}_{\text{RBF}}$ is a vector of real and imaginary parts of electromagnetic responses, $\boldsymbol{x}_{\text{RBF}}$ is a vector of real and imaginary parts of $H(s)$, $\circ$ denotes the Hadamard product, $\boldsymbol{w}_1'$ and $\boldsymbol{w}_2'$ are the weight vectors of the two layers, and $\boldsymbol{b}_1'$ and $\boldsymbol{b}_2'$ are the bias vectors of the two layers. This can be implemented by the MATLAB function *newrb*. The function *newrb* iteratively creates one more neuron for the radial basis function network. Neurons are added to the network until the mean squared error falls beneath a preset error goal or a maximum number of neurons is reached.

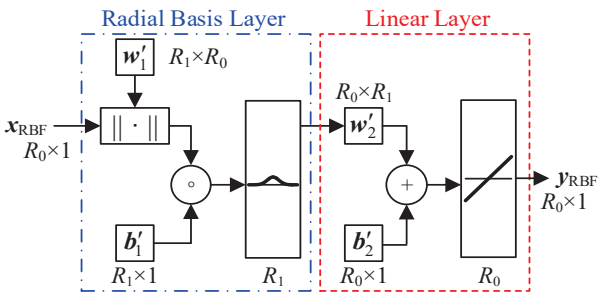


Fig. 4. Structure of the radial basis function network.

F. Training and optimization

In the training process, if the accuracy of the model is not acceptable, the hyper-parameters are adjusted or more training samples are added to retrain the model. The mean absolute percentage errors of the frequencies at $|S_{11}|$ = -10 dB and the radiation pattern are adopted to measure calculation precision.

Once the model is trained, it can substitute the full-wave simulation and speed up the process of optimization. It is difficult to directly optimize the contour of slot in the image domain. The optimized variable is a vector of shifting distance in x- and y-directions for the control points.

III. APPLICATION EXAMPLE

A tri-passband microstrip-fed slot antenna [13] (see Fig. 5) is employed as an example to evaluate the proposed model. The original geometric parameters are as follows: W = 31 mm, L = 41 mm, W_{f} = 3.14 mm, L_{f} = 13.5 mm, l_1 = 6.7 mm, l_2 = 2.6 mm, t = 0.5 mm, s_1 = 0.3 mm, D_1 = 18 mm, D_2 = 20 mm, d_1 = 2.3 mm, d_2 = 3.5 mm, w_1 = 1.7 mm, and w_2 = 4.1 mm. The metallic strips are printed on a substrate with a thickness of 1.59 mm and a relative permittivity of 4.4.

The slot shown in Fig. 5 (a) is considered for modeling, i.e., the first two resonant frequencies are modeled. The slot keeps mirror symmetry in both horizontal and vertical directions. Therefore, one-eighth of the slot is presented by a binary image, which contains 168 × 144 square pixels, i.e., $h_{\text{image}} \times v_{\text{image}}$ = 168 × 144, with d_{pixel} = 0.05 mm. The metal and the substrate are denoted by 0 and 1, respectively. For the histogram of oriented gradients feature, h_{cell} = 8, v_{cell} = 8, h_{block} = 2, v_{block} = 2, h_{overlap} = 1, and v_{overlap} = 1. The length of the histogram of oriented gradients feature vector is 12,240.

HFSS 17.0 performs the full-wave simulation and generates training and testing samples; 400 training samples and 100 testing samples are defined randomly in Table 1. The control points are shifted in the blue polygon in Fig. 5 (c), and the geometric parameters of the blue polygon are l_3 = 0.7 mm, l_4 = 2 mm, l_5 = 0.95 mm, l_6 = 0.84 mm, l_7 = 7.96 mm, and θ = 35.16°. In the shape-changing process, the maximum iteration is 4 and the side length d of the coarse square gird for the samples is 0.8 mm.

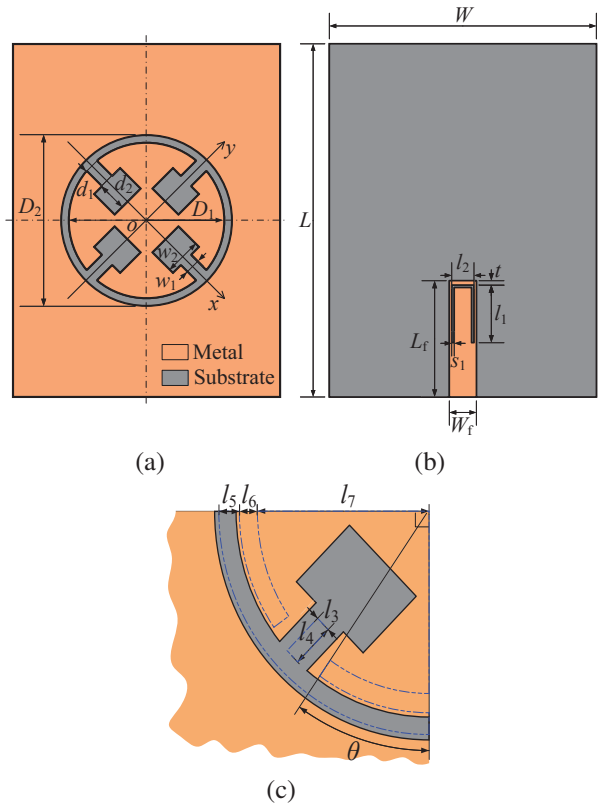


Fig. 5. Geometry of the tri-passband microstrip-fed slot antenna: (a) Top view, (b) back view, (c) quarter slot.

Table 1: Definition of training and testing samples

	Training Data (400 Samples)			**Testing Data (100 Samples)**		
	Min.	Max.	Step	Min.	Max.	Step
Δx_k (mm)	-0.35	0.35	0.1	-0.3	0.4	0.1
Δy_k (mm)	-0.35	0.35	0.1	-0.3	0.4	0.1

Table 2: Hyper-parameters of the trained SVM

	$\|S_{11}\|$	**Patterns at the First Resonant Frequency**		
		E Plane (θ = [0°, 180°])	**E Plane (θ = [180°, 360°])**	**H Plane**
p	6.17×10^{-7}	5.90×10^{-3}	1.62×10^{1}	8.84×10^{-7}
γ_v	1.91×10^{-2}	2.42×10^{-10}	1.17×10^{-4}	1.05×10^{-4}
γ_ξ	8.93×10^{13}	1.62×10^{15}	1.98×10^{15}	8.05×10^{-2}

The hyper-parameters of the trained least square SVM are listed in Table 2, where the hyper-parameters are obtained with Bayesian optimization [14]. The training and testing errors for the frequencies at $|S_{11}|$ = -10 dB and the radiation pattern (taking the pattern at the first resonant frequency for example) are given in Table 3. For comparison, a CNN model is also trained for the tri-passband antenna. The training and testing

Table 3: Training and testing errors for the frequencies at $|S_{11}|$ = -10 dB and the radiation patterns at the first resonant frequency

	Frequencies at $\|S_{11}\|$ = -10 dB		**Patterns at the First Resonant Frequency**	
	Proposed Model	**CNN Model**	**Proposed Model**	**CNN Model**
Training error	0.13%	0.48%	0.96%	0.85%
Testing error	0.57%	0.62%	1.18%	0.89%

errors of the CNN model are at the same level as the proposed model. The training costs 1.1 mins for SVM, and 56.6 mins for CNN. Much less time is spent to train the proposed model, and fewer hyper-parameters need to be determined. For this paper, the calculations are performed on an Intel i5-1135G7 (2.4 GHz) machine with 16 GB RAM.

Two examples out of the training range are chosen to test the model, as shown in Figs. 6 and 7. The predicted electromagnetic responses of the proposed model and CNN model both agree well with the full-wave simulated ones from HFSS. The corresponding control points for the sample are shown in Figs. 6 (c) and 7 (c).

Once the model is trained, it can be applied to the optimization design as a substitute for the full-wave simulation. Then a tri-passband antenna is optimized

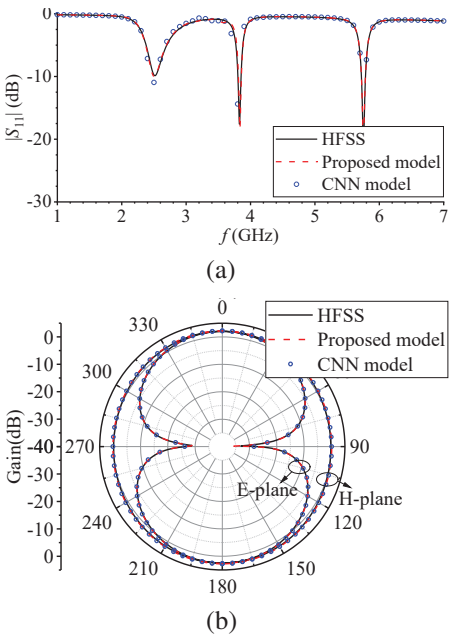
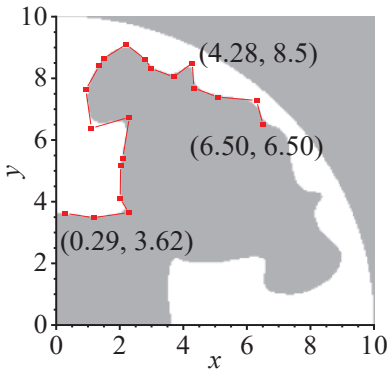
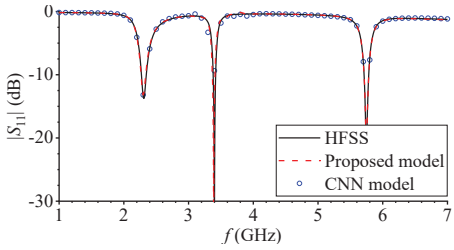


Fig. 6. *Continued*

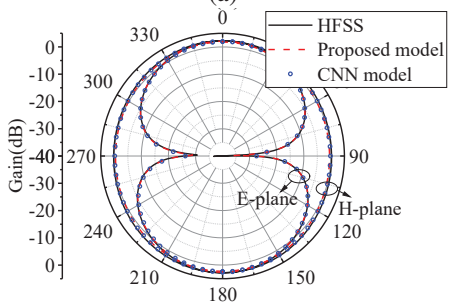


(c)

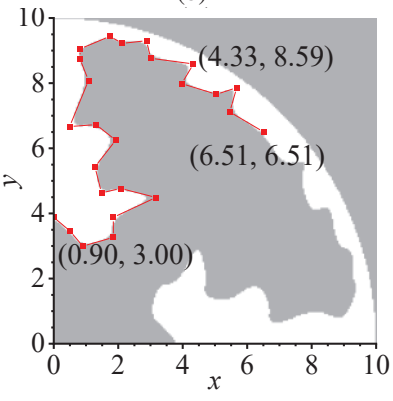
Fig. 6. Comparison of the first testing sample: (a) $|S_{11}|$, (b) radiation patterns at the first resonant frequency, (c) image of a quarter of slot.



(a)

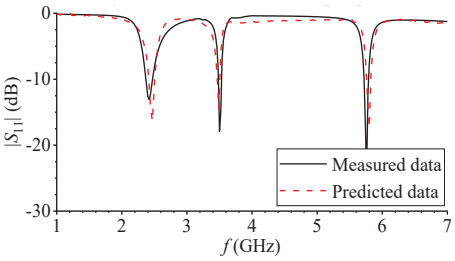


(b)

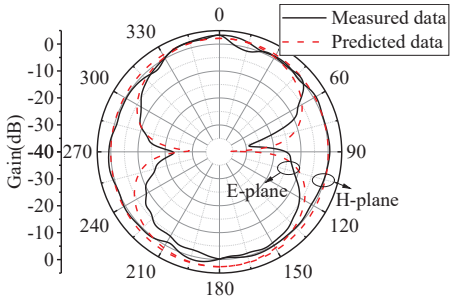


(c)

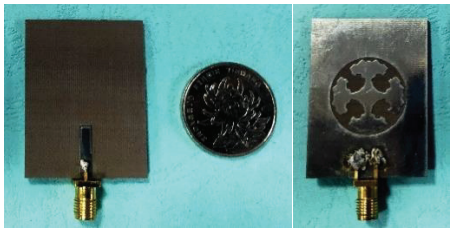
Fig. 7. Comparison of the second testing sample: (a) $|S_{11}|$, (b) radiation patterns at the first resonant frequency, (c) image of a quarter of slot.



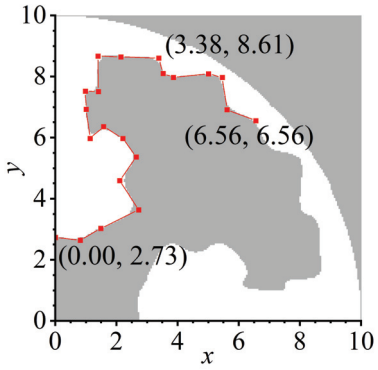
(a)



(b)



(c) (d)



(e)

Fig. 8. Optimization results for the example and photograph of the fabricated antenna: (a) $|S_{11}|$, (b) radiation patterns at the first resonant frequency, (c) top view, (d) back view, (e) image of a quarter of slot.

with the genetic algorithm [15] to reach the design specification of $|S_{11}| \leq$ -10 dB (2.4-2.48 GHz, 3.49-3.54 GHz, and 5.72-5.79 GHz). From Fig. 8, the curves of predicted *S*-parameter and radiation pattern for the first resonant frequency agree well with the measured ones.

IV. CONCLUSION

In this paper, a new least square SVM model for shape modeling of slot antennas is proposed. In the modified shape-changing technique, the B-spline interpolation curve is used to describe the slot shape, and the corresponding slot image is used as the model input. The model consists of histogram of oriented gradients feature, transfer function, and the radial basis function network. The histogram of oriented gradients feature is extracted to obtain the distribution of local intensity gradients or edge position from the slot images. Then, the least square SVM maps the histogram of oriented gradients features into transfer function coefficients, and there are only three hyper-parameters that need to be determined in the training of SVM. The transfer function provides a preliminary prediction about electromagnetic response. In the end, the radial basis function network is applied to the error correction. A tri-passband microstrip-fed slot antenna is employed as an example to confirm the effectiveness of the proposed model. The proposed model gets the same precision as the CNN model while it costs much less training time.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant 62171093 and by the Sichuan Science and Technology Programs under Grants 2022NSFSC0547 and 2022ZYD0109.

REFERENCES

[1] H. M. E. Misilmani, T. Naous, and S. K. A. Khatib, "A review on the design and optimization of antennas using machine learning algorithms and techniques," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 30, no. 10, p. e22356, Oct. 2020.

[2] L.-Y. Xiao, W. Shao, F.-L. Jin, and B.-Z. Wang, "Multiparameter modeling with ANN for antenna design," *IEEE Trans. Antennas Propag.*, vol. 66, no. 7, pp. 3718-3723, July 2018.

[3] T. Khan and C. Roy, "Prediction of slot-position and slot-size of a microstrip antenna using support vector regression," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 29, no. 3, p. e21623, Mar. 2019.

[4] Y. Kim, S. Keely, J. Ghosh, and H. Ling, "Application of artificial neural networks to broadband antenna design based on a parametric frequency model," *IEEE Trans. Antennas Propag.*, vol. 55, no. 3, pp. 669-674, Mar. 2007.

[5] F. Feng, C. Zhang, J. Ma, and Q.-J. Zhang, "Parametric modeling of EM behavior of microwave components using combined neural networks and pole-residue-based transfer functions," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 1, pp. 60-77, Jan. 2016.

[6] F. Feng, V.-M.-R. Gongal-Reddy, C. Zhang, J. Ma, and Q.-J. Zhang, "Parametric modeling of microwave components using adjoint neural networks and pole-residue transfer functions with EM sensitivity analysis," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 6, pp. 1955-1975, June 2017.

[7] J. P. Jacobs, "Accurate modeling by convolutional neural-network regression of resonant frequencies of dual-band pixelated microstrip antenna," *IEEE Antennas Wireless Propag. Lett.*, vol. 20, no. 12, pp. 2417-2421, Dec. 2021.

[8] H.-Y. Luo, W. Shao, X. Ding, B.-Z. Wang, and X. Cheng, "Shape modeling of microstrip filters based on convolutional neural network," *IEEE Microw. Wireless Compon. Lett.*, vol. 32, no. 9, pp. 1019-1022, Sep. 2022.

[9] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, San Diego, CA, USA, vol. 1, pp. 886-893, 2005.

[10] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, USA: Prentice Hall, Upper Saddle River, NJ, 2002.

[11] S. Xu, X. An, X. Qiao, L. Zhu, and L. Li, "Multi-output least-squares support vector regression machines," *Pattern Recognit. Lett.*, vol. 34, no. 9, pp. 1078-1084, July 2013.

[12] B. Gustavsen and A. Semlyen, "Rational approximation of frequency domain responses by vector fitting," *IEEE Trans. Power Del.*, vol. 14, no. 3, pp. 1052-1061, July 1999.

[13] K. D. Xu, Y. H. Zhang, R. J. Spiegel, Y. Fan, W. T. Joines, and Q. H. Liu, "Design of a stub-loaded ring-resonator slot for antenna applications," *IEEE Trans. Antennas Propag.*, vol. 63, no. 2, pp. 517-524, Feb. 2015.

[14] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms" in *Adv. in Neural Inf. Process. Syst. 25*, USA: Curran Associates Inc., Red Hook, NY, pp. 2951-2959, 2012.

[15] R. L. Haupt and D. H. Werner, *Genetic Algorithms in Electromagnetics*, USA: Wiley, Hoboken, NJ, 2007.

Hai-Ying Luo received the B.S. degree from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2017, where he is currently pursuing the Ph.D. degree in radio physics.

His current research interest is computational electromagnetics.

Wen-Hao Su received the B.S. degree from UESTC, Chengdu, China, in 2021, where he is currently pursuing the master's degree in radio physics.

His current research interests include antenna design and computational electromagnetics.

Haiyan Ou received the B.E. degree in electrical engineering from UESTC in 2000, and received Ph.D. degrees in optical engineering from Zhejiang University in 2009.

She joined UESTC in 2009 and is now an associate professor there. From 2010 to 2011, she was a visiting scholar in the department of Engineering, Cambridge University, UK. From 2012 to 2013, she was a post-doc in the Department of Electrical and Electronic Engineering, the University of Hong Kong. Her research interests include computational electromagnetics, microwave photonics, and digital holography.

Sheng-Jun Zhang received Ph.D. in science from Beijing University of Technology in 2001. From then on he joined the team in National Key Laboratory of Science & Technology on Test Physics and Numerical Mathematics. He is now professor of the laboratory and his research interests include scattering of EM waves, EM effects of periodic structures such as FSS, PC and gratings, as well as modulation of scattering of materials and interaction of EM waves with plasmas, IR radiation management.

He has published some papers in journals and conferences, in addition to patents and two books.

Wei Shao received the B.E. degree in electrical engineering from UESTC in 1998, and received M.Sc. and Ph.D. degrees in radio physics from UESTC in 2004 and 2006, respectively.

He joined UESTC in 2007 and is now a professor there. From 2010 to 2011, he was a visiting scholar in the Electromagnetic Communication Laboratory, Pennsylvania State University, State College, PA. From 2012 to 2013, he was a visiting scholar in the Department of Electrical and Electronic Engineering, the University of Hong Kong. His research interests include computational electromagnetics and antenna design.

Wideband Multi-polarization Reconfigurable Antenna based on Non-uniform Polarization Convert AMC Reflector

Long Li[1], Jia-Jun Liang[1,2], Xiaoxiao Liu[3], Tiejun Chen[2], Jier Lv[1], and Zhao Wu[1,2*]

[1]Research Center of Intelligent Information and Communication Technology

[2]Guangxi Colleges and Universities Key Lab of Complex System Optimization and Big Data Processing
Yulin Normal University, Yulin 537000, China
longli@ylu.edu.cn, shuigpjd@163.com, chentj@ylu.edu.cn, jielv@ylu.edu.cn, kianty@163.com*

[3]Wuhan Maritime Communication Research Institute
Wuhan 430000, China.
xxliu1215@163.com

Abstract **–** A novel design of wideband multi-polarization reconfigurable antenna is proposed, based on a non-uniform polarization convert artificial magnetic conductor (AMC) reflector. The proposed antenna consists of a radiator element and an AMC reflector. Firstly, a modified polarization convert AMC reflector is designed. The non-uniform AMC reflector causes an enhancement of 3 dB axial ratio (AR) performance. Secondly, a wideband linearly polarized monopole antenna is presented as the main radiator, utilizing the broadband characteristic of a C-shaped monopole. The polarization reconfigurability of the proposed antenna can be achieved by properly rotating the AMC reflector, which can be switched between linear polarization (LP), left-hand circular polarization (LHCP), and right-hand circular polarization (RHCP). A prototype of the proposed antenna is fabricated and experimented with to validate the theoretical performance. The measured results show a -10 dB impedance bandwidth of 42.7% and 44.4% for LP and CP modes, respectively, and a 3 dB AR bandwidth of 20% for CP modes. In addition, the measured peak gain reaches 8 dBi/dBic. A good agreement is shown between the simulation and measurement, pointing to the good performance of the proposed antenna.

Index Terms **–** Non-uniform metasurface (MS), polarization convert AMC reflector, polarization reconfigurability, wideband.

I. INTRODUCTION

With the rapid development of mobile and satellite communications, lots of multifunctional antennas have been investigated in recent years, such as frequency reconfigurable antennas, pattern reconfigurable antennas, polarization reconfigurable antennas, or hybrid reconfigurable antennas [1]–[25]. Since circular polarization has the characteristic of reducing the effect of multipath loss, modern wireless communications require antenna polarization diversity to strengthen the communication quality, especially for satellite communications [3], [4]. Moreover, to meet more wireless communication applications needs, a wideband characteristic of the antenna is needed. Thus, the antenna, which combines performances of wideband and multi-polarization simultaneously, has attracted more and more attention.

Conventionally, the approach of antenna reconfigurability includes embedding RF or optical switches in slots on radiator patches or the ground plane to change the antenna current distribution and make different polarization modes [5]–[14]. In [5]–[7], three tri-polarization reconfigurable patch antennas are presented; they have simple geometries, but all obtain a narrow operating bandwidth of 2%, 8%, and 3%, respectively. On the contrary, a tri-reconfigurable antenna makes a large 3 dB axial ratio (AR) bandwidth of 50% caused by a C-shaped radiator [8], but the optical switch is not easy to apply to modern communication systems. Some other works unfortunately have polarization reconfigurability, switching only within linearly polarized (LP) modes or circularly polarized (CP) modes [9]–[12]. Other works realize polarization diversity by controlling switches inserted in the slot on the ground plane [13], [14].

Another method for a polarization reconfigurable antenna is to design a multipath phase-shift network, which can change the phase difference between the feeding points and lead to polarization diversities [15]–[10]. In this method, it is easy to offer a phase difference and then achieve polarization diversity. For example, in [15] and [18], a tri-polarization reconfigurable antenna and a quad-polarization reconfigurable antenna are proposed,

Submitted On: February 1, 2023
Accepted On: June 25, 2023

https://doi.org/10.13052/2023.ACES.J.380909

but they obtain a narrow 3 dB AR bandwidth of 17% and 12%, respectively. Obviously, the shortcoming of this method is that a wideband phase-shift network is challenging to implement, and many lumped components are used, which causes cost increases.

It's worth noting that, as a novel design, some multi-polarization antenna investigations about metasurface (MS) for polarization diversity have been presented [20]–[25]. Compared with conventional works, the polarization reconfigurable antenna based on MS has the advantage of a low profile. In [20], a CP reconfigurable antenna based on non-uniform MS with a low profile and broadband is investigated, controlled by RF switches. Unfortunately, this antenna can't work in LP mode. In [23], a multi-polarization reconfigurable antenna based on polarization convert MS is proposed with a bandwidth of 30%, switching polarization states by rotating the MS. Compared with RF switches, this method can reduce the loss of bias circuits and lumped components.

In this paper, a wideband multi-polarization reconfigurable antenna based on a polarization convert artificial magnetic conductor (AMC) reflector is proposed. Different from the previously published design, this antenna is beneficial for a simple polarization switching strategy and a reduction of loss, selecting polarization modes by properly rotating the AMC reflector without any switches. Meanwhile, compared to the conventional metal reflector, the AMC reflector leads to an improvement of the antenna gain with a low profile. Eventually, the proposed antenna acquires multi-polarization working modes and a profile decline of 33.3% compared to conventional mental reflectors with a profile of $1/4\lambda 0$ (at $f0 = 5$ GHz).

II. ANTENNA DESIGN PRINCIPLE

As depicted in Fig. 1, the proposed antenna includes an AMC reflector and a C-shaped radiator. The wideband performance of the C-shaped radiator has already been validated by previous work [8]. There is a modified C-shaped monopole to offer a wideband incident wave source. The modified C-shaped patch and ground plane are printed on top and bottom of a 1.524 mm Rogers 4003C substrate with a relative permittivity of 3.38. In others, an improved AMC reflector is placed $H = 10$ mm below the radiator to achieve polarization reconfigurability and radiation directionality. The AMC reflector consists of 6×6 MS units fabricated on a 2 mm R4 substrate with a relative permittivity of 4.4. To explain the design principle of the proposed antenna, a detailed analysis of the AMC reflector and C-shaped monopole are discussed as follows.

A. Polarization convert AMC reflector analysis

The geometry of the proposed AMC unit is illustrated in Fig. 2, in which its cell is a shuttle-shaped patch etched with a crossed slot. An equivalent circuit is provided to explain how the AMC reflector realizes polarization conversion. Along the diagonal corners, the orthogonal surface impedance component, marked Z_u and Z_v, of the AMC unit can be approximately calculated as

$$\begin{aligned} Z_u &= 2R_1 + 2j\omega(2L_1) + 2/(j\omega C_1) \\ &\quad + 1/(j\omega C_2) \\ &= R_u + jX_u \text{ and} \end{aligned} \tag{1a}$$

$$\begin{aligned} Z_v &= 2R_2 + 2j\omega(2L_2) + 2/(j\omega C_3) \\ &\quad + 1/(j\omega C_4) \\ &= R_v + jX_v, \end{aligned} \tag{1b}$$

where $R1$, $L1$, $C1$, and $C2$, are the resistance, inductance, and distributed capacitance of the AMC unit in the u-direction, respectively, and $R2$, $L2$, $C3$, and C4 are in the v-direction. Further, according to the Euler formula, (1) can be simplified as follows

$$Z_u = |Z_u|\angle\varphi 1 = |Z_u|e^{j\varphi 1} \text{ and} \tag{2a}$$

$$Z_v = |Z_v|\angle\varphi 2 = |Z_v|e^{j\varphi 2}. \tag{2b}$$

What can be known is that the resistance, inductance, and distributed capacitance of the AMC unit are related to the triangle truncation and the crossed slot. Thus, by adjusting the size of the triangle truncated and the crossed slot, a phase difference can be achieved between Z_u and Z_v.

Assuming that the incident source is a linearly polarized plane wave along the x-axis, so the incident $\boldsymbol{E}$-field is also along the x-axis named $\boldsymbol{E}_{ix}$. Here, $\boldsymbol{E}_{ix}$ is broken into two orthogonal components, $\boldsymbol{E}_{iu}$ and $\boldsymbol{E}_{iv}$. By supposing that the magnitude of the two orthogonal $\boldsymbol{E}$-field

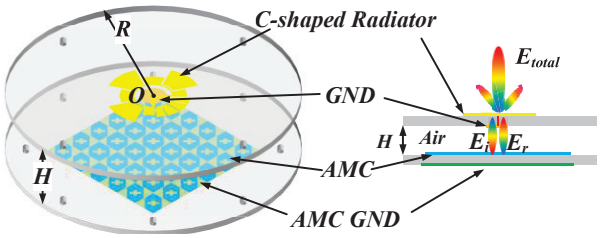


Fig. 1. The overall prototype of the proposed antenna.

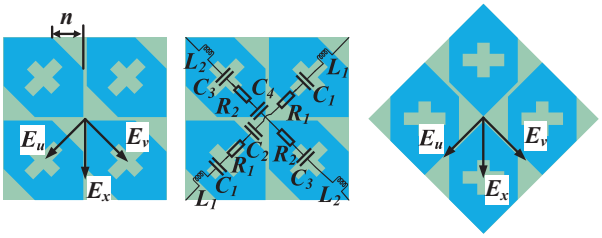


Fig. 2. The geometry of the proposed AMC.

components is $|E_m|$ and the phase is $\ddot{o}3$ and $\ddot{o}4$, $\boldsymbol{E}_{iu}$ and $\boldsymbol{E}_{iv}$ can be denoted by the following:

$$E_{\mathrm{iu}} = |Zm|e^{j\varphi 3} \text{ and} \tag{3a}$$

$$E_{\mathrm{iv}} = |Zm|e^{j\varphi 4}. \tag{3b}$$

When $\boldsymbol{E}$-field components $\boldsymbol{E}_{iu}$ and $\boldsymbol{E}_{iv}$ are incident on the AMC reflector, respectively, numerous electrons on the surface of the AMC unit will be excited and move along the u-direction and v-direction, respectively, generating induced currents, thus forming the reflected $\boldsymbol{E}$-fields $\boldsymbol{E}_{ru}$ and $\boldsymbol{E}_{rv}$ individually. If there is no energy loss in reflection, the reflected and incident $\boldsymbol{E}$-field are equal in magnitude. Furthermore, the phase of the reflected wave is equivalent to the combination of the phase with the incident wave, the phase of the AMC surface impedance along the incident wave direction, and the phase difference generated by the air gap. Therefore, the reflected $\boldsymbol{E}$-field components can be roughly described as follows

$$E_{ru} = |Zm|e^{j(\varphi 1+\varphi 3+\Delta\varphi)} \text{ and} \tag{4a}$$

$$E_{rv} = |Zm|e^{j(\varphi 2+\varphi 4+\Delta\varphi).} \tag{4b}$$

Eventually, the composite $\boldsymbol{E}$-field components can be roughly described as follows:

$$E_u = |Zm|\left[e^{j\varphi 3} + e^{j(\varphi 1+\varphi 3+\Delta\varphi)}\right] \text{ and} \tag{5a}$$

$$E_v = |Zm|\left[e^{j\varphi 4} + e^{j(\varphi 2+\varphi 4+\Delta\varphi)}\right]. \tag{5b}$$

Since $\boldsymbol{E}_{iu}$ and $\boldsymbol{E}_{iv}$ are symmetric to $\boldsymbol{E}_{ix}$, these two orthogonal $\boldsymbol{E}$-field components have the same phase, meaning that $\ddot{o}3$ is equal to $\ddot{o}4$. And appropriately optimizing the AMC unit, a 90° phase difference can be obtained between Z_u and Z_v, assuming $\ddot{o}1$ leads $\ddot{o}2$ by 90°. As a result, (5a) and (5b) can be simplified as (6a) and (6b).

$$E_u = |Zm|\left[e^{j\varphi 4} + je^{j(\varphi 2+\varphi 4+\Delta\varphi)}\right] \text{ and} \tag{6a}$$

$$E_v = |Zm|\left[e^{j\varphi 4} + e^{j(\varphi 2+\varphi 4+\Delta\varphi)}\right]. \tag{6b}$$

Thus, the composite $\boldsymbol{E}$-field $\boldsymbol{E}_{total}$ is a CP wave, which is combined by two $\boldsymbol{E}$-field components with identical magnitudes and a 90° phase difference. In addition, because the phase of the $\boldsymbol{E}$-field component in the u-direction leads the v-direction, the antenna will work in left-handed circularly polarized (LHCP) mode. As shown in Fig. 2, when the AMC unit is rotated by -45° or 45°, the AMC is symmetric about the x-axis. Thus, there is an in-phase or anti-phase of the surface impedance between the u-direction and v-direction. Accordingly, the composite $\boldsymbol{E}$-field $\boldsymbol{E}_{total}$ is an LP wave. Moreover, when the AMC unit is rotated by 90°, the polarization of the proposed antenna will be in right-handed circularly polarized (RHCP) mode.

Ansoft HFSS simulates the AMC unit with the Floquet port to verify the analysis. The width of the triangle truncation on AMC units is defined as parameter n. Its effect on the reflected characteristic of the proposed AMC unit and the polarization convert performance is given in Fig. 3. With the increase of n, the high-frequency resonance point of the AMC unit will move toward the high frequency, whereas the low frequency remains almost constant. Meanwhile, the impedance matching would deteriorate. In addition, for better CP performance, the magnitude ratio between the incident $\boldsymbol{E}$-field and the reflected must be within ±3 dB. As shown in Fig. 3 (b), while n increases, the magnitude ratio varies, and the broadest 3 dB magnitude ratio bandwidth is obtained when n is equal to 6 mm. Considering the performance of the impedance matching and the CP, there are two AMC units with different truncations, n = 4.5 mm and 6 mm, respectively, to build a non-uniform AMC reflector.

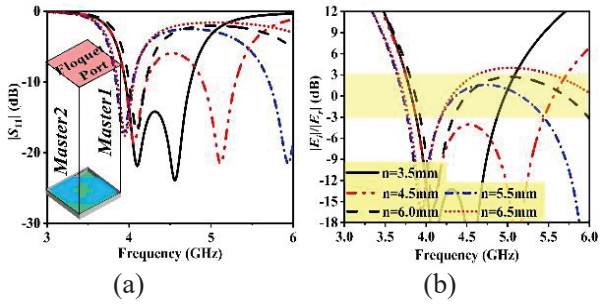


Fig. 3. The effect of parameter n on the AMC unit: (a) Reflection coefficient $|S_{11}|$, (b) the magnitude ratio between the incident E-field and reflected E-field.

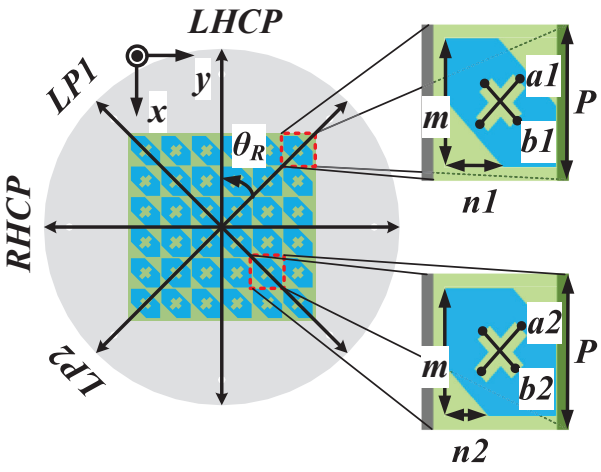
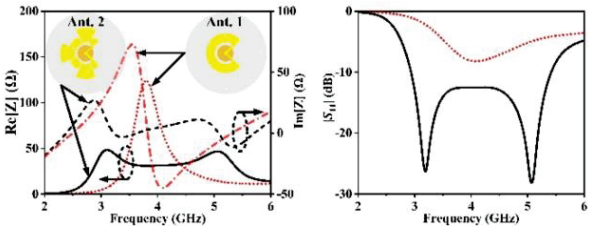


Fig. 4. The phototype of the proposed AMC reflector (design parameters: m = 11.5 mm, P = 12 mm, $n1$ = 6 mm, $n2$ = 4.5 mm, $a1$ = 5.5 mm, $a2$ = 6 mm, $b1$ = 4.5 mm, $b2$ = 5 mm).

B. C-shaped monopole antenna with the AMC reflector

Given the polarization conversion characteristic of the proposed AMC reflector, a modified C-shaped monopole is designed and introduced to offer an incident *x*-LP wave source, positioned at the top of the reflector.

As shown in Fig. 5, for Ant. 1, the impedance matching is poor, in which the real part of the impedance is not close to 50 Ù, and the imaginary part is not near 0 Ù within a certain bandwidth. There is a resonance only near 4 GHz, noted f_1. To improve impedance matching and bring multi-resonance, the C-shaped patch edge is slotted and placed into three parasitic stubs in a fan shape, labeled Ant. 2. Due to the slotting, the current length of the resonant frequency f_1 will be expanded so that the resonance frequency is scaled down and recorded as f_2. The parasitic stubs, with the induced current, will create a new resonance, marked f_3, achieving multi-frequency resonance. As a result, two resonant frequency points near 3 and 5 GHz are generated and bring a wide impedance bandwidth ($|S_{11}| <$ -10 dB) of about 60%.



(a) (b)

Fig. 5. The performance of two C-shaped monopoles: (a) Impedance, (b) $|S_{11}|$.

Then the proposed AMC reflector is applied below the C-shaped monopole antenna, about 0.15ë, to realize directionality and polarization reconfigurability. To reduce the crossed polarization, especially in CP modes, a set of fan-shaped parasitic stubs are added at the notch of Ant.2. The added parasitic stubs will be excited and generated with a polarization mode consistent with the antenna, so the main polarization is enhanced, which implies a reverse weakening of the cross-polarization, as shown in Fig. 6. The final design of the C-shaped monopole antenna, shown in Fig. 7, contains a C-shaped patch with seven parasitic fan stubs and a circular ground plane with two additional rectangle stubs.

Eventually, when holding the C-shaped monopole antenna still and rotating the AMC reflector counterclockwise by -45°, 45°, 0°, and 90°, the antenna can work on *x*-LP1, *x*-LP2, LHCP, and RHCP modes, respectively.

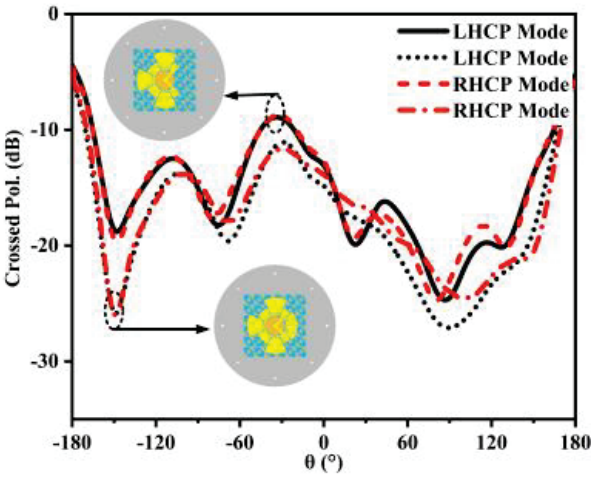


Fig. 6. The effect of the small fan parasitic stubs.

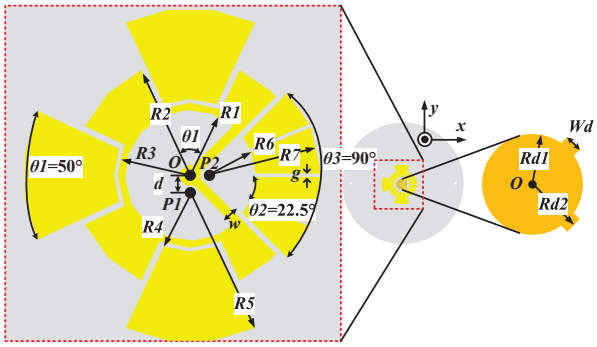


Fig. 7. The phototype of the C-shaped monopole (design parameters [millimetres]: *R1* = 9, *R2* = 16, *R3* = 10, *R4* = 8.5, *R5* = 21, *R6* = 7, *R7* = 16, *Rd1* = 6.5, *Rd2* = 7.5, *w* = 1.3, *d* = 2, *g* = 0.5, *Wd* = 2.3).

III. SIMULATED AND EXPERIMENTAL RESULTS

The proposed antenna is designed and optimized by Ansys HFSS and then fabricated and measured to verify the performance by the ROHDE&SCHWARZ ZVB-8 vector network analyzer and multi-probe antenna testing system. The manufactured antenna and far-field experiment setup are shown in Fig. 8.

Figure 9 illustrates simulated and experimental reflection coefficients $|S_{11}|$ in different polarization modes. The experimental -10 dB overlapped impedance bandwidth of 42.7%, covers 3.5 GHz to 5.4 GHz in both LP modes. It obtains 44.4% of the measured -10 dB impedance bandwidth for the CP modes and covers between 3.5 GHz and 5.5 GHz. The measured impedance matching of the antenna decreases slightly at high frequency due to manufacturing error. The experimental results are in good accordance with the simulation results.

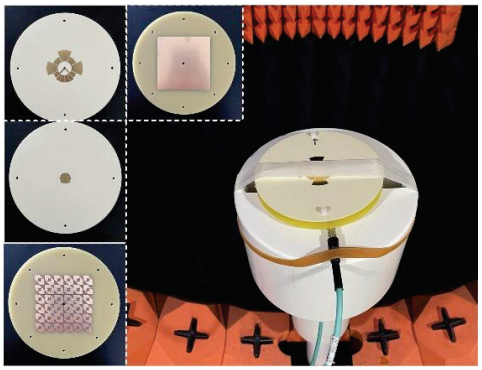


Fig. 8. Photographs of the manufactured antenna and experiment setup.

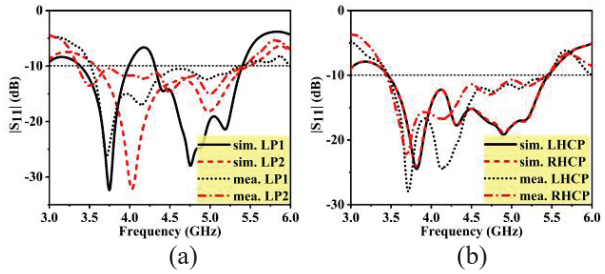


Fig. 9. The simulated and measured reflection coefficient $|S_{11}|$: (a) LP modes and (b) CP modes.

Figure 10 shows the simulated and measured axial ratio and gain in different polarization modes. When working in LP modes, the axial ratio of the antenna is larger than 25 dB within operating bandwidth. In addition, the measured peak gain reaches about 8 dBi. When working in CP modes, an overlapped 3 dB axial ratio bandwidth of 20%, covering 4.5 GHz to 5.5 GHz, shows good agreement with the simulation. Similarly, the measured peak gain is about 8 dBic for CP modes. Compared with the simulations, there is a 2 dB loss of the measured gain, which is caused by the error of fabrication and the loss of dielectric material. The measured and simulated normalized radiation patterns of the proposed antenna at

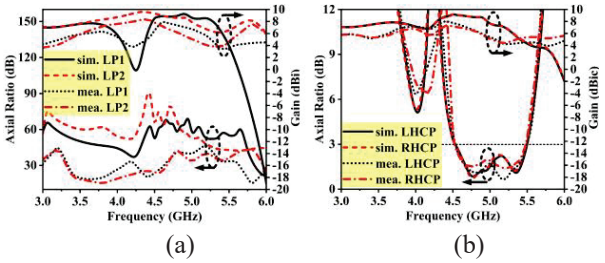


Fig. 10. The simulated and measured results of axial ratio and gain: (a) LP modes and (b) CP modes.

4.8 and 5 GHz are illustrated in Figs. 11 and 12 for CP and LP modes, respectively. The simulated results match well for all states with good directivity.

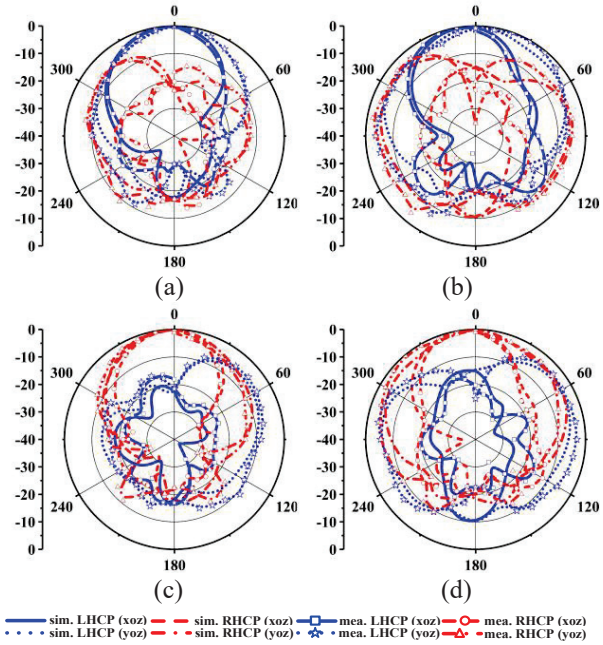


Fig. 11. The simulated and measured normalized radiation pattern: (a) LHCP at 4.8 GHz, (b) 5.2 GHz, (c) RHCP at 4.8 GHz, and (d) 5.2 GHz.

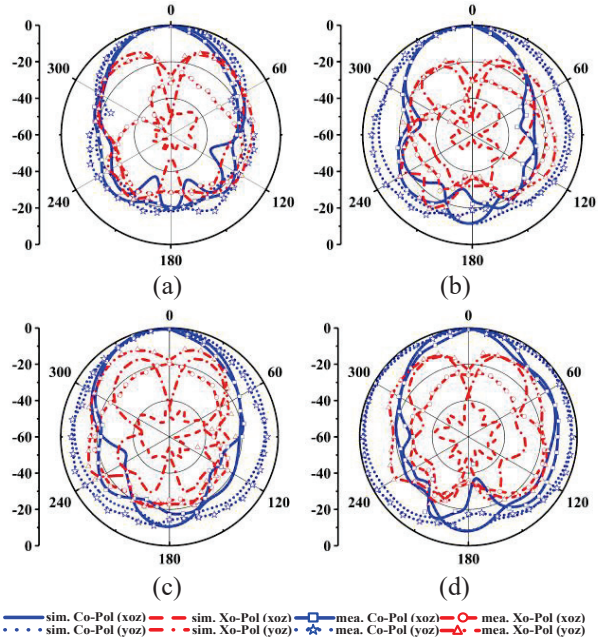


Fig. 12. The simulated and measured normalized radiation pattern: (a) LP1 at 4.8 GHz, (b) 5.2 GHz, (c) LP2 at 4.8 GHz, and (d) 5.2 GHz.

Table 1 shows a performance comparison with other polarization reconfigurable works. The proposed antenna has a wider -10 dB impedance bandwidth of over 40% and more operating modes. Meanwhile, this work has a simple polarization switch strategy by rotating the MS, with no RF switches applied. In addition, the 3 dB axial ratio bandwidth and peak gain of the proposed work have shown good performance.

Table 1: Performance comparison with other works

Ref.	Modes	Impedance Bandwidth (%)	AR Bandwidth (%)	Peak Gain (dBi/c)
[11]	2 CP	41.1	34	12
[12]	2 CP	5	4	5.2
[21]	2 CP	36	21.5	15.5
[22]	1 LP2 CP	LP: 17.1 CP: 8.7	8.7	LP: 9 CP: 8.3
This work	2 LP2 CP	LP: 42.7 CP: 44.4	20	LP: 8 CP: 8

IV. CONCLUSION

This paper proposes a design for a polarization reconfigurable antenna based on a rotating AMC reflector. The proposed antenna is composed of two parts, an AMC reflector and a C-shaped monopole. By turning the AMC reflector at $\pm 45°$, $0°$, and $90°$ clockwise, respectively, the antenna could switch the polarization state among x-LP, LHCP, and RHCP. Compared to other previous polarization reconfigurable antennas, it reduces the application of the DC bias circuits with a simple polarization switching strategy. The proposed antenna has a wide operating bandwidth of 42.7% and 20% for LP and CP modes, approximately covering 3.5 to 5.4 GHz and 4.5 to 5.5 GHz, respectively. Furthermore, the maximum measured gain reaches 8 dBi(c). It could be applied to 5G mobile communication systems, satellite communications, and other polarization-diverse applications.

ACKNOWLEDGMENT

This work was supported in part by the Natural Science Foundation Fund Project in Guangxi, China, under Grant 2021GXNSFBA220003, in part by General Project of Guangxi Natural Science Foundation Project (Guangdong-Guangxi Joint Fund Project) 2021GXNSFAA075031.

REFERENCES

[1] R. L. Haupt and M. Lanagan, "Reconfigurable antennas," *IEEE Antennas and Propagation Magazine*, vol. 55, no. 1, pp. 49-61, Feb. 2013.

[2] K. Ramahatla, M. Mosalaosi, A. Yahya, and B. Basutli, "Multiband reconfigurable antennas for 5G wireless and CubeSat applications: A review," *IEEE Access*, vol. 10, pp. 40910-40931, 2022.

[3] W. Yang, W. Che, H. Jin, W. Feng, and Q. Xue, "A polarization-reconfigurable dipole antenna using polarization rotation AMC structure," *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 12, pp. 5305-5315, Dec. 2015.

[4] B. S. Qiu, Y. F. Xia, and Y. S. Li, "Gain-enhanced wideband circularly polarized antenna with a non-uniform metamaterial reflector," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 37, no. 3, pp. 281-286, Mar. 2022.

[5] R. K. Singh, A. Basu, and S. K. Koul, "A novel reconfigurable microstrip patch antenna with polarization agility in two switchable frequency bands," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 10, pp. 5608-5613, Oct. 2018.

[6] S. G. Zhou, G. L. Huang, H.-Y. Liu, A. S. Lin, and C. Y. D. Sim, "A CPW-fed square-ring slot antenna with reconfigurable polarization," *IEEE Access*, vol. 6, pp. 16474-16483, 2018.

[7] Q. Chen, J. Y. Li, G. Yang, B. Cao, and Z. Zhang, "A polarization-reconfigurable high-gain microstrip antenna," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 5, pp. 3461-3466, May 2019.

[8] G. Jin, L. Li, W. Wang, and S. W. Liao, "A wideband polarization reconfigurable antenna based on optical switches and C-shaped radiator," *Microwave Optical Technology Lett.*, vol. 22, pp. 2415-2422, 2020.

[9] A. Bhattacharjee and S. Dwari, "A monopole antenna with reconfigurable circular polarization and pattern tilting ability in two switchable wide frequency bands," *IEEE Antennas and Wireless Propagation Letters*, vol. 20, no. 9, pp. 1661-1665, Sep. 2021.

[10] K. D. Hong, X. Chen, X. Zhang, L. Zhu, and T. Yuan, "A slot-loaded high-gain circular patch antenna with reconfigurable orthogonal polarizations and low cross polarization," *IEEE Antennas and Wireless Propagation Letters,* vol. 21, no. 3, pp. 511-515, Mar. 2022.

[11] Q. C. Ye, J. L. Li, and Y. M. Zhang, "A circular polarization-reconfigurable antenna with enhanced axial ratio bandwidth," *IEEE Antennas and Wireless Propagation Letters*, vol. 21, no. 6, pp. 1248-1252, June 2022.

[12] A. Priya, S. K. Mohideen, and M. Saravanan, "Design of polarization reconfigurable patch antenna for wireless communications," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 35, no. 8, pp. 893-898, Aug. 2020.

[13] W. W. Yang, X. Y. Dong, W. J. Sun, and J. X. Chen, "Polarization reconfigurable broadband dielectric resonator antenna with a lattice structure," *IEEE Access,* vol. 6, pp. 21212-21219, 2018.

[14] M. Li, Z. Zhang, M. C. Tang, L. Zhu, and N. W. Liu, "Bandwidth enhancement and size reduction of a low-profile polarization-reconfigurable antenna by utilizing multiple resonances," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 2, pp. 1517-1522, Feb. 2022.

[15] W. Sun, S. Liu, X. Zhu, X. Zhang, P. L. Chi, and T. Yang, "A novel 1.05 GHz to 1.25 GHz filtering antenna feeding network with reconfigurable frequency and polarization," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 1, pp. 156-166, Jan. 2022.

[16] C. Liu, Y. Li, T. Liu, Y. Han, J. Wang, and S. Qu, "Polarization reconfigurable and beam-switchable array antenna using switchable feed network," *IEEE Access*, vol. 10, pp. 29032-29039, 2022.

[17] W. Li, Y. M. Wang, Y. Hei, B. Li, and X. Shi, "A compact low-profile reconfigurable metasurface antenna with polarization and pattern diversities," *IEEE Antennas and Wireless Propagation Letters*, vol. 20, no. 7, pp. 1170-1174, July 2021.

[18] L. Kang, H. Li, B. Tang, X. Wang, and J. Zhou, "Quad-polarization-reconfigurable antenna with a compact and switchable feed," *IEEE Antennas and Wireless Propagation Letters*, vol. 20, no. 4, pp. 548-552, Apr. 2021.

[19] K. Li, Y. Shi, H. Shen, and L. Li, "A characteristic-mode-based polarization-reconfigurable antenna and its array," *IEEE Access*, vol. 6, pp. 64587-64595, 2018.

[20] H. H. Tran, C. D. Bui, N. Nguyen-Trong, and T. K. Nguyen, "A wideband non-uniform metasurface-based circularly polarized reconfigurable antenna," *IEEE Access*, vol. 9, pp. 42325-42332, 2021.

[21] J. Liu, J. Y. Li, J. J. Yang, Y. X. Qi, and R. Xu, "AMC-loaded low-profile circularly polarized reconfigurable antenna array," *IEEE Antennas and Wireless Propagation Letters*, vol. 19, no. 7, pp. 1276-1280, July 2020.

[22] D. Chen, W. Yang, W. Che, Q. Xue, and L. Gu, "Polarization-reconfigurable and frequency-tunable dipole antenna using active AMC structures," *IEEE Access*, vol. 7, pp. 77792-77803, 2019.

[23] C. Ni, M. S. Chen, Z. X. Zhang, and X. L. Wu, "Design of frequency- and polarization-reconfigurable antenna based on the polarization conversion metasurface," *IEEE Antennas and Wireless Propagation Letters,* vol. 17, no. 1, pp. 78-81, Jan. 2018.

[24] H. L. Zhu, S. W. Cheung, X. H. Liu, and T. I. Yuk, "Design of polarization reconfigurable antenna using metasurface," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 6, pp. 2891-2898, June 2014.

[25] J. Hu, Z.-C. Hao, and W. Hong, "Design of a wideband quad-polarization reconfigurable patch antenna array using a stacked structure," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 6, pp. 3014-3023, June 2017.

Long Li was born in Guangxi, China, in 1993. He received his B.E. degree in electronic and information engineering from Hohai University, Nanjing, China, in 2017. He received his M.E degree in electronic and communication engineering from South China University of Technology, Guangzhou, China, in 2020. He is now working with the School of Physics and Telecommunication Engineering, Yulin Normal University, Yulin, China. His research interests include reconfigurable antennas, metasurface antennas, and wideband antennas.

Jia-Jun Liang received his B.E. degree in electronic science and technique from Guilin University of Electronic Technology (GUET), Guilin, China, in 2012, and the M.E. degree in radio physics from the University of Electronic Science and Technology of China, Chengdu, China, in 2015. He received the Ph.D. degree in information and communication engineering at Shenzhen University, Shenzhen, China, in 2018. He is now with the School of Physics and Telecommunication Engineering, Yulin Normal University. His current research interests include MIMO antennas, 3D printing antennas, and millimetre wave antennas.

Jier Lv was born in 1964. He received his B.E. degree in physics form Guangxi Normal University, Guilin, China, in 1988. His research interests include nonlinear complex systems and electromagnetic computation.

Xiaoxiao Liu is currently an engineer with Wuhan Maritime Communication Research Insititute. She is mainly engaged in the design of ship antennas, wideband VHF/UHF antennas, reconfigurable antennas, and related technology research.

Tiejun Chen received the B.S. degree from University of Electronic Science and Technology of China, China, in 1988, and received the M.Sc. degree from Guilin University of Electronic Technology, China, in 2007. He is currently a professor in the School of Physics and Telcommunication Engineering, Yulin Normal University, Yulin, China, and he is the senior member of China Electronics Society. His research interests include wireless communications, signal processing, and embedded systems.

Zhao Wu was born in Guangxi, China, in 1987. He received the B.E. degree in electronic and information engineering and the Ph.D. degree in electro-magnetic fields and microwave technology from Xidian University, Xi'an, China, in 2011 and 2016, respectively. From October 2016 to March 2017, he was with Huawei Technologies Co. Ltd. Since April 2017, he has been working with the School of Physics and Telecommunication Engineering, Yulin Normal University. His research interests include metamaterials, novel antennas, reconfigurable antenna design and applications.

A Compact Low-profile 5G Millimeter-wave Circularly Polarized Antenna Based on LTCC

Ting Wang[1], Jun Wang[1], Chenyu Ding[2], Zhuowei Miao[2], Jie Wang[3], and Lei Zhao[1]

[1]School of Information and Control Engineering
China University of Mining and Technology, Xuzhou 221116, China
twangmax@163.com, jun-wang@cumt.edu.cn, leizhao@cumt.edu.cn

[2]State Key Laboratory of Millimeter Waves, School of Information Science and Engineering
Southeast University, Nanjing 211189, China
cyding_aly@seu.eu.cn, zwmiao@seu.edu.cn

[3]Suzhou Bohai Microsystem Co. Ltd.
Suzhou, China
jwang@bmsltcc.com

Abstract **–** In this paper, a circularly polarized millimeter-wave L-shaped dipole antenna based on low temperature cofired ceramic (LTCC) technology is proposed, which realizes compact size and low-profile performance. The designed antenna consists of radiation patches and the grounded coplanar waveguide-substrate integrated waveguide (GCPW-SIW) feeding structure, which connects each other by two via holes. The radiation patches include a pair of L-shaped patches and four parasitic patches. The simulated results show that the proposed antenna operates from 26.5 to 29.5 GHz for $|S_{11}| < -10$ dB and AR < 3 dB with a peak gain of 6.7 dBic. The antenna element size is only $0.58\lambda_0 \times 0.58\lambda_0 \times 0.056\lambda_0$, where λ_0 is free-space wavelength at the center frequency of 28 GHz. A sample of the antenna is fabricated and measured to verify the proposed design, which has a good agreement with the simulated ones, indicating that the antenna has potential applications for the fifth generation (5G) mm-Wave n257 (26.5 - 29.5 GHz) frequency band communications and satellite communication systems.

Index Terms **–** 5G millimeter wave (mm-Wave), circularly polarized (CP), low-profile, low temperature cofired ceramic (LTCC), substrate integrated waveguide (SIW).

I. INTRODUCTION

To meet the demands of users for high-capacity and high data transmission rate of the fifth generation (5G) mobile communication, the 5G millimeter-wave (mm-Wave) band is being extensively studied and applied [1–2]. As one of the 5G commercial millimeter-wave bands, n257 (26.5-29.5 GHz) frequency band is of great practical significance in the research of the antenna.

The traditional processes of manufacturing circular polarization antennas mainly include printed circuit board (PCB) technology and low temperature cofired ceramic (LTCC) technology. With the rapid development of fabrication technology, miniaturization and integration have become a research hotspot. Especially, LTCC technology has become a good candidate for designing miniaturization and integration of electronic components [3].

Recently, different kinds of LTCC and PCB mm-Wave antennas have been reported for various circularly polarized applications [4–21]. For example, an s-dipole based on PCB is employed to constitute an 8×8 broadband circularly polarized (CP) antenna array, which has an impedance bandwidth of 27.6% and axial ratio bandwidth up to 32.7% [4]. Nevertheless, the antenna size is $0.71\lambda_0 \times 0.71\lambda_0 \times 0.46\lambda_0$, which needs to be further reduced. Similarly, a 4×4 magnetoelectric dipole array is devised in [5], which uses sequential rotary feed network to obtain the wide bandwidth. However, it needs to be further miniaturized. Because the LTCC has unique multilayer technology and high dielectric constant performance compared to the PCB process, it can be used to design miniaturization and integration antennas. Accordingly, a 4×4 60 GHz LTCC helical antenna array is proposed, which achieves a bandwidth of 20% with a small plane size [6]. However, its profile and the antenna structure need to be further reduced and simplified. Meanwhile, an antenna-in-package array with relatively simple structure based on LTCC technology with low-profile has been proposed in [7]. Unfortunately, the antenna sacrifices bandwidth to obtain low-profile characteristics. Accordingly, a LTCC low-profile and wide bandwidth helical antenna is shown in [8]. Moreover, a

Submitted On: February 16, 2023
Accepted On: December 7, 2023

https://doi.org/10.13052/2023.ACES.J.380910

SIW cavity and L-shaped planar probe were combined to form circularly polarized radiation, which realizes high gain performance [9]. However, the axial ratio bandwidth still needs to be enhanced. Therefore, designing a compact, low-profile, and easy to manufacture circularly polarized antenna is a challenging task.

In this paper, a circularly polarized L-shaped dipole antenna with four parasitic patches based on LTCC technology is proposed, which realizes compact size and low-profile performance. The antenna operates at 26.5-29.5 GHz with $|S_{11}|$ < -10 dB and AR < 3 dB. The peak gain value within the operating frequency band is 6.7 dBic. A prototype is fabricated and measured to verify the simulated results, which are basically consistent with the simulated ones.

II. DESIGN OF CP ANTENNA

A. Antenna Geometry

Figure 1 (a) presents the geometry of the proposed antenna, which mainly includes two parts: one is radiation patches with six layers LTCC and the other is the GCPW-SIW feeding structure with two layers LTCC. The L-shaped dipole patches at the top layer are connected with the SIW-based rectangular slot through two via holes. Moreover, four patches with square chamfer are placed around the L-shaped patch as parasitic elements to improve the AR bandwidth. It should be noted that an extra via hole is inserted into SIW to improve the impedance matching. Furthermore, an extra structure is added at the end of the antenna to install the microwave connector.

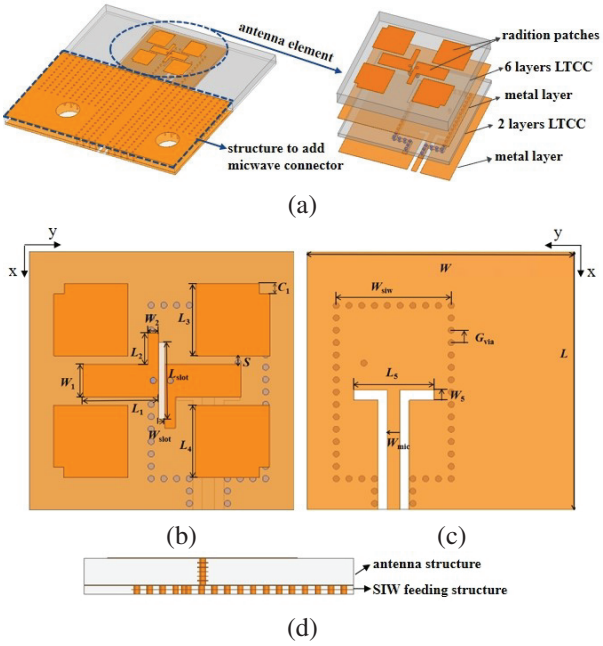


Fig. 1. (a) 3D view, (b) top view, (c) bottom view of the bottom layer, and (d) side view of the antenna.

Table 1: Dimensions of the antenna (unit: mm)

L	W	$Wsiw$	$L1$	$W1$
6.25	6.25	2.7	1.61	0.7
W_2	L_3	L_4	C_1	L_{slot}
0.25	1.65	1.61	0.2	1.9
S	L_5	W_5	W_{mic}	G_{via}
0.2	2.02	0.45	0.29	0.3

B. Design Theory of the Proposed Antenna

Figure 2 gives the improvement process of the proposed antenna structure. Type 1 is the initial structure with a thickness of h = 1.2 mm, and two centrosymmetric L-shaped patches are employed to form orthogonal line current for the sake of producing CP radiation. However, the AR bandwidth cannot meet the requirement of 26.5-29.5 GHz, and the profile needs to be further reduced. Hence, our proposed LTCC employs eight layers to attain a lower profile, as shown in the type 2 structure. Concurrently, the performance of the antenna deteriorates, especially the axial ratio performance. Therefore, four parasitic patches with square chamfer are introduced to compensate the degradation results, as shown in type 3 structure.

Figure 3 displays the simulated results of $|S_{11}|$ and axial ratio of types 1 - 3 antennas. The type 1 has excellent impedance bandwidth according to Fig. 3 (a). However, the axial ratio bandwidth of type 1 does not satisfy the desired bandwidth, and the antenna profile is high as displayed in Fig. 3 (b). Therefore, the type 2 antenna

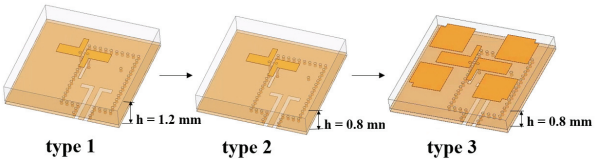


Fig. 2. The process of the antenna design.

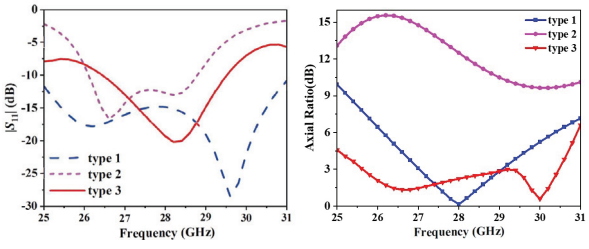


Fig. 3. (a) Simulated $|S_{11}|$ of types 1-3 structure and (b) simulated axial ratio of types 1-3 structure.

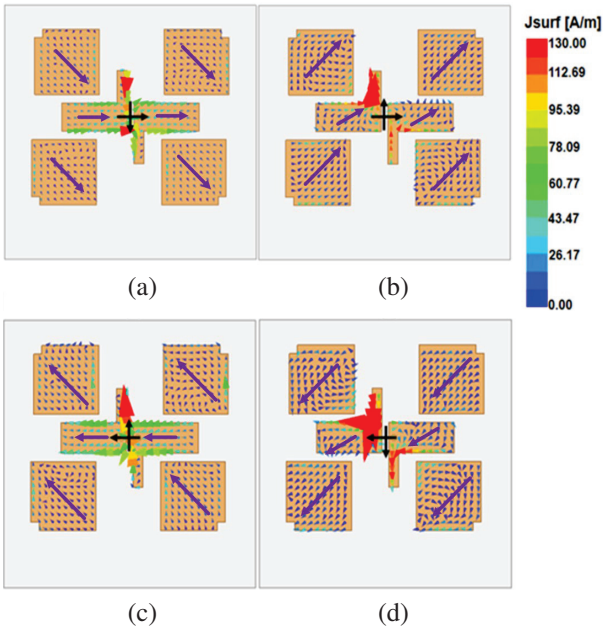


Fig. 4. Current distributions of the proposed CP antenna at 28 GHz: (a) t = 0(T), (b) t = T/4, (c) t = T/2, and (d) t = 3T/4.

is designed to decrease the profile compared to type 1. Nevertheless, the bandwidth of type 2 still cannot meet the requirements of the n257 frequency band, as the reduction of profile affects the performance of the antenna. Hence, to obtain the desired bandwidth, type 3 is proposed, based on type 2, which commendably covers the 26.5-29.5 GHz frequency band whether impedance or axial ratio bandwidth.

In order to explain the operating principle of the proposed antenna, Fig. 4 shows the simulated surface current distribution of the radiation elements at 28 GHz. According to the change of the surface current on the L-shaped dipole, the orientation of the surface current rotates 360° within one period, which reveals that right-handed circular polarization is formed. Moreover, the orientation of the surface current of the additional chamfered parasitic patches changes counter-clockwise in one period, which will form another circular polarization resonance. As a result, the axial bandwidth is expanded to cover the 26.5-29.5 GHz frequency band.

Figures 5 (a) and (b) show the field distributions of the SIW cavity without and with via hole, respectively, which indicates that the via hole disrupts the field distribution in the SIW cavity. Moreover, the via hole is placed at the site of the weak electric field, which is equivalent to that the cavity wall moves inward. Therefore, the operating AR bandwidth frequency shifts to the high frequency and realizes the required n257 frequency band, as shown in Fig. 5 (c).

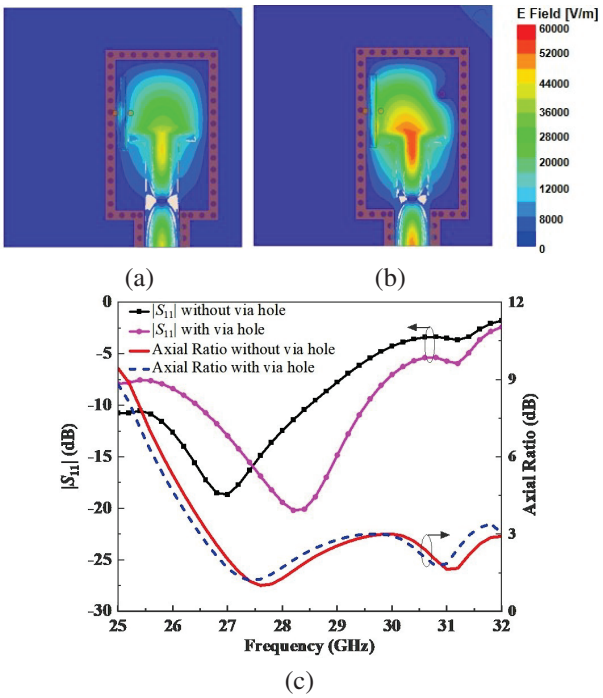


Fig. 5. (a) Electric field distribution without via hole, (b) electric field distribution with via hole, and (c) simulated $|S_{11}|$ and axial ratio in both (a) and (b) cases.

III. SIMULATION RESULTS AND EXPERIMENTAL VERIFICATION

In order to verify the validity of simulated results, a physical model is manufactured and measured. Figures 6 (a) and (b) show the photographs of the

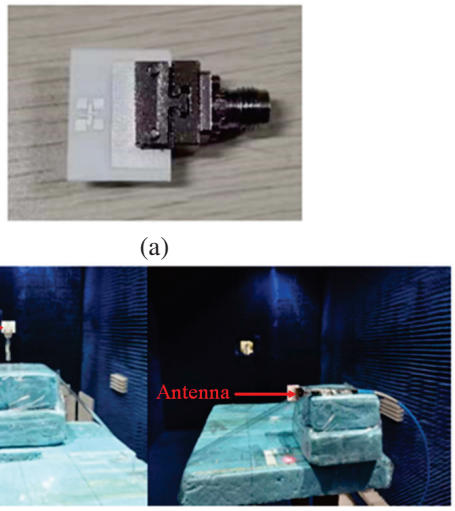


Fig. 6. Fabricated prototype of the CP antenna: (a) Prototype with connectors installed and (b) antenna under test in the anechoic chamber.

fabricated sample based on the LTCC process and the arrangements of the far-field measurement, respectively. It should be noted that an additional structure is added at the end of antenna feeding structure, as shown in Fig. 6 (a), which is convenient for the installation of the microwave connector.

Figure 7 presents the simulated and measured results of $|S_{11}|$ and RHCP gain. It indicates that the simulation and measured impedances bandwidth of the antenna are 25.9-29.5 GHz and 26.3-29.5 GHz for $|S_{11}| < -10$ dB, which illustrates that the simulation and measured $|S_{11}|$ results have good consistency. In addition, compared to the simulated RHCP gain of the proposed antenna, the measured result has a slight deviation owing to the slight variation in dielectric constant of LTCC. The measured peak gain of the proposed CP antenna is 6.7 dBic within the operating frequency band.

The comparison of axial ratio between measured and simulated results are shown in Fig. 8, which reveals that the simulated axial ratio is less than 3 dB within 26.5-29.5 GHz. However, the measured axial ratio is deteriorated to about 6 dB within 26.5-29.5 GHz, which is caused by the test environment. During the test process of axial ratio, the linear polarization test scheme is adopted due to the lack of a circularly polarized horn antenna. Using the measured horizontal and vertical polarization results, the axial ratio of a circularly polarized antenna can be calculated. Moreover, it should be noted that the measured results of pitch angle deviation may be about 5 degrees residual. Considering the measurement tolerance error, the deviation between the simulation and the test results is within a reasonable range (AR $<$ 4 dB). Furthermore, Fig. 9 shows the simulated and measured co-polarized and cross-polarized radiation patterns of the antenna at 27 GHz, 28 GHz, and 29 GHz. It can be observed that measured results of the primary

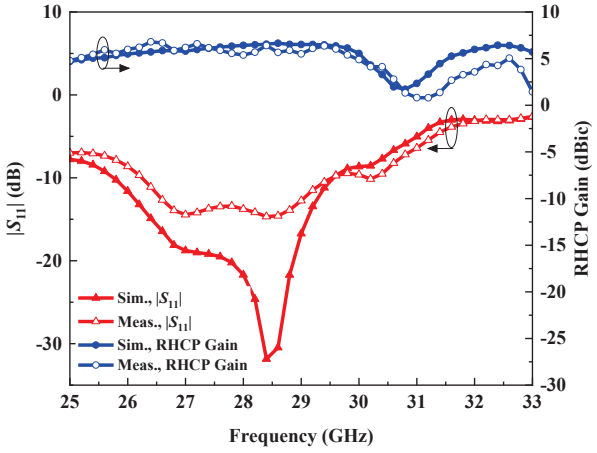


Fig. 7. Simulated and measured $|S_{11}|$ and RHCP gain of the fabrication model.

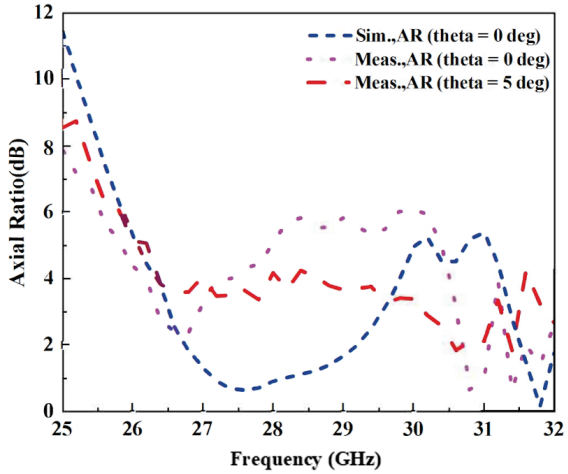


Fig. 8. Simulated and measured axial ratio of the fabrication model.

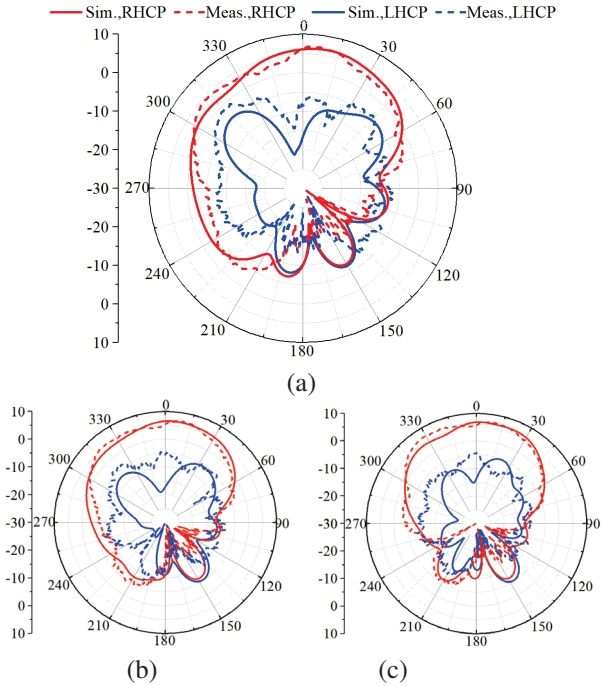


Fig. 9. Simulated and measured radiation patterns: (a) 27 GHz, (b) 28 GHz, and (c) 29 GHz.

and cross polarization patterns are basically consistent with the simulated results.

To further illustrate the features of the proposed antenna, the performance comparison of the proposed CP antenna with the existing CP antennas are given in Table 2. References [19], [22], [23], and [24] present polarized antennas based on PCB technology, which can achieve wide bandwidth. However, their superiority in miniaturization is not outstanding comparing to the other circularly polarized antennas based on LTCC

Table 2: Comparison of existing CP mm-Wave antennas

Ref.	Process Technology	f_0 (GHz)	Antenna Element Size (mm)	Thickness* λs $(\lambda s = \lambda 0/\sqrt{\varepsilon r})$	Impedance Bandwidth	Axial Ratio Bandwidth
[19]	PCB	28.63	$8+8$ $(0.84\cdot 0.84\lambda_0)$	0.25	21.83%	5.9%
[22]	PCB	60	$5.6+5.6$ $(1.12*1.12\lambda_0)$	0.22	23.8%	23.4%
[23]	PCB	30.5	$9.5+9.5$ $(0.99\cdot 0.99\lambda_0)$	0.3	27.7%	28.5%
[24]	PCB	28.35	$12+12$ $(1.13\cdot 1.13\lambda_0)$	0.31	$>14\%$	14%
[17]	LTCC	35	$3.83+3.83$ $(0.45*0.45\lambda_0)$	0.24	29.6%	$>26\%$
[18]	LTCC	60	$3.5*4$ $(0.7*0.8\lambda_0)$	0.19	16.5%	11.5%
Thiswork	LTCC	28	$6.25+6.25$ $(0.58+0.58\lambda_0)$	0.18	$>10\%$	$>10\%$

technology proposed in [17] and [18]. Moreover, the profile of the proposed antenna in this paper is lower than the other antenna structures (see Table 2) with the bandwidth exactly covering the required n257 operating frequency band.

IV. CONCLUSION

In this paper, a compact low-profile circularly polarized mm-Wave L-shaped dipole antenna with four parasitic patches is proposed. Four parasitic patches with square chamfer are placed around the centrosymmetric L-shaped patches to improve the axial ratio bandwidth. The designed antenna bandwidth is more than 10%, which can meet the required n257 operation frequency band. Additionally, it has the feature of smaller size and lower profile. The measured results have good agreement with the simulated ones. Hence, the proposed antenna can be an appropriate candidate for the applications of 5G millimeter-wave n257 (26.5 - 29.5 GHz) frequency band communications and satellite communications systems.

ACKNOWLEDGMENT

This work was supported in part by the National Science Foundation of China under Grant 62201575. (Corresponding authors: Jun Wang; Lei Zhao).

REFERENCES

[1] L. Wei, R. Q. Hu, Y. Qian, and G. Wu, "Key elements to enable millimeter wave communications for 5G wireless systems," *IEEE Wireless Commun.*, vol. 21, no. 6, pp. 136-143, Dec. 2014.

[2] W. Hong, K. Baek, Y. Lee, Y. Kim, and S. Ko, "Study and prototyping of practically large-scale mm-Wave antenna systems for 5G cellular devices," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 63-69, Sep. 2014.

[3] U. Ullah, N. Mahyuddin, Z. Arifin, M. Z. Abdullah, and A. Marzuki, "Antenna in LTCC technologies: A review and the current state of the art," *IEEE Antennas and Propagation Magazine*, vol. 57, no. 2, pp. 241-260, Apr. 2015.

[4] L. Zhang, K. Wu, S. W. Wong, Y. J. He, P. Chu, W. Y. Li, K. X. Wang, and S. Gao, "Wideband high-efficiency circularly polarized SIW-fed S-dipole array for millimeter-wave applications," *IEEE Trans. Antennas Propag.*, vol. 68, no. 3, pp. 2422-2427, Mar. 2020.

[5] Z. Gan, Z. Tu, Z. Xie, Q. Chu, and Y. Yao, "Compact wideband circularly polarized microstrip antenna array for 45 GHz application," *IEEE Trans. Antennas Propag.*, vol. 66, no. 11, pp. 6388-6392, Nov. 2018.

[6] C. Liu, Y. Guo, X. Bao, and S. Xiao, "60-GHz LTCC integrated circularly polarized helical antenna array," *IEEE Trans. Antennas Propag.*, vol. 60, no. 3, pp. 1329-1335, Mar. 2012.

[7] W. Zhang, Y. P. Zhang, M. Sun, C. Luxey, D. Titz, and F. Ferrero, "A 60-GHz circularly polarized array antenna-in-package in LTCC technology," *IEEE Trans. Antennas Propag.*, vol. 61, no. 12, pp. 6228-6232, Dec. 2013.

[8] M. Du, J. Xu, X. Ding, J. P. Cao, J. H. Deng, and Y. L. Dong, "A low-profile wideband LTCC integrated circularly polarized helical antenna array for millimeter-wave applications," *Radioengineering*, vol. 27, no. 2, pp. 455-462, 2018.

[9] B. Cao, Y. Shi, and W. Feng, "W-band LTCC circularly polarized antenna array with mixed U-type substrate integrated waveguide and ridge gap waveguide feeding networks," *IEEE Antennas and Wireless Propag. Lett.*, vol. 18, no. 11, pp. 2399-2403, Nov. 2019.

[10] Y. Li, Z. N. Chen, X. Qing, Z. Zhang, J. Xu, and Z. Feng, "Axial ratio bandwidth enhancement of 60-GHz substrate integrated waveguide-fed circularly polarized LTCC antenna array," *IEEE Trans. Antennas Propag.*, vol. 60, no. 10, pp. 4619-4626, Oct. 2012.

[11] H. T. Chou, S. J. Chou, J. D. S. Deng, C. H. Chang, and Z.-D. Yan, "LTCC-based Antenna-in-Package array for 5G user equipment with dual-polarized endfire radiations at millimeter-wave frequencies," *IEEE Trans. Antennas Propag.*, vol. 70, no. 4, pp. 3076-3081, Apr. 2022.

[12] M. A. Ahmad and L. J. A. Olule, "Meshed stacked LTCC antenna for space application," *IEEE Access*, vol. 10, pp. 29473-29481, 2022.

[13] H. Xu, J. Zhou, K. Zhou, Q. Wu, Z. Yu, and W. Hong, "Planar wideband circularly polarized cavity-backed stacked patch antenna array for millimeter-wave applications," *IEEE Trans. Antennas Propag.*, vol. 66, no. 10, pp. 5170-5179, Oct. 2018.

[14] A. B. Guntupalli and K. Wu, "60-GHz circularly polarized antenna array made in low-cost fabrication process," *IEEE Antennas Wireless Propag. Lett.*, vol. 13, pp. 864-867, 2014.

[15] D. Guan, C. Ding, Z. Qian, Y. Zhang, Y. J. Guo, and K. Gong, "Broadband high-gain SIW cavity-backed circular-polarized array antenna," *IEEE Trans. Antennas Propag.*, vol. 64, no. 4, pp. 1493-1497, Apr. 2016.

[16] Q. Wu, H. Wang, C. Yu, and W. Hong, "Low-profile circularly polarized cavity-backed antennas using SIW techniques," *IEEE Trans. Antennas Propag.*, vol. 64, no. 7, pp. 2832-2839, July 2016.

[17] M. Du, Y. Dong, J. Xu, and X. Ding, "35-GHz wideband circularly polarized patch array on LTCC," *IEEE Trans. Antennas Propag.*, vol. 65, no. 6, pp. 3235-3240, June 2017.

[18] H. Sun, Y. Guo, and Z. Wang, "60-GHz circularly polarized U-slot patch antenna array on LTCC," *IEEE Trans. Antennas Propag.*, vol. 61, no. 1, pp. 430-435, Jan. 2013.

[19] B. Feng, J. Lai, K. L. Chung, T.-Y. Chen, Y. Liu, and C.-Y.-D. Sim, "A compact wideband circularly polarized magneto-electric dipole antenna array for 5G millimeter-wave application," *IEEE Trans. Antennas Propag.*, vol. 68, no. 9, pp. 6838-6843, Sep. 2020.

[20] J. Xu, W. Hong, Z. H. Jiang, and H. Zhang, "Low-cost millimeter-wave circularly polarized planar integrated magneto-electric dipole and its arrays with low-profile feeding structures," *IEEE Antennas Wireless Propag. Lett.*, vol. 19, no. 8, pp. 1400-1404, Aug. 2020.

[21] X. Bai, S. Qu, S. Yang, J. Hu, and Z. Nie, "Millimeter-wave circularly polarized tapered-elliptical cavity antenna with wide axial-ratio beam width," *IEEE Trans. Antennas Propag.*, vol. 64, no. 2, pp. 811-814, Feb. 2016.

[22] X. Ruan, S. W. Qu, Q. Zhu, K. B. Ng, and C. H. Chan, "A complementary circularly polarized antenna for 60-GHz applications," *IEEE Antennas and Wireless Propag. Lett.*, vol. 16, pp. 1373-1376, 2017.

[23] C. M. Zhu, G. H. Xu, D. W. Ding, J. Wu, W. Wang, Z. X. Huang, and X. L. Wu, "Low-profile wideband millimeter-wave circularly polarized antenna with hexagonal parasitic patches," *IEEE Antennas and Wireless Propag. Lett.*, vol. 20, no. 9, pp. 1651-1655, Sep. 2021.

[24] J. Wu, Y. J. Cheng, and Y. Fan, "Millimeter-wave wideband high-efficiency circularly polarized planar array antenna," *IEEE Trans. Antennas Propag.*, vol. 64, no. 2, pp. 535–542, Feb. 2016.

Ting Wang was born in Xuzhou, China, in 1998. She is currently working toward the master's degree with China University of Mining and Technology, Xuzhou, China. Her research interest is circularly polarized antenna theory and application.

Jun Wang was born in Jiangsu, China. He received the B.Eng. and M.S. degrees from Jiangsu Normal University, Xuzhou, China, in 2013 and 2017, respectively, and the Ph.D. degree in electromagnetic field and microwave technology from Southeast University, Nanjing, in 2021. From 2015 to 2016, he was with the Department of Electronic and Electrical Engineering, Nanyang Technological University of Singapore, as a research associate.

He joined the China University of Mining and Technology, Xuzhou, China, in 2021. He has authored or co-authored over 30 referred journal and conference papers. His current research interests include the design of RF/microwave antennas and components.

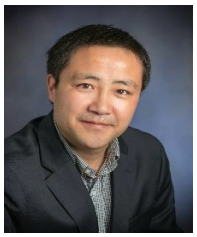
Chen-Yu Ding received the B.S. degree in electromagnetic fields and wireless technology from Northwestern Polytechnical University, Xi'an, China, in 2020. He is currently pursuing the Ph.D. degree in electromagnetic field and microwave technology at the State Key Laboratory of Millimeter Waves, Southeast University, Nanjing, China., His current research interests include high-gain antenna array, multibeam antennas, and terahertz technology.

Zhuo-Wei Miao received the B.S. degree in electronics and information engineering from Nanjing Normal University, Nanjing, China, in 2014, and the Ph.D. degree in electromagnetic field and microwave technology from Southeast University, Nanjing, in 2020. From 2019 to 2020, he was a visiting Ph.D. student with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore.

He is currently a research fellow with the State Key Laboratory of Millimeter Waves, School of Information Science and Engineering, Southeast University. His current research interests include millimeter-wave and terahertz antennas and circuits, terahertz systems, and metamaterials.

Jie Wang Deputy General Manager of Suzhou Bohai Entrepreneurial Microsystems Co., LTD. He received degree in electromagnetic field and microwave technology from Southeast University, Nanjing and as a postdoctoral fellow with Department of Electronic Engineering, Chinese University of Hong Kong, and adjunct professor with School of Electronic Information, Soochow University.

His research interests include electromagnetic field numerical calculation, LTCC multilayer circuit numerical calculation and equivalent circuit derivation, LTCC microwave device design, microwave millimeter wave component design based on LTCC, etc. He has published more than 10 papers such as IEEE-MTT, Journal of Electronics, Journal of Infrared and Millimeter Wave, Journal of Radio Wave Science, etc.



Lei Zhao received the B.S. degree in mathematics from Jiangsu Normal University, China, in 1997, and the M.S. degree in computational mathematics and the Ph.D. degree in electromagnetic fields and microwave technology from Southeast University, Nanjing, China, in 2004 and 2007, respectively.

He joined the China University of Mining and Technology, Xuzhou, China, in 2019, where he is currently a full professor. From September 2009 to December 2018, he worked with Jiangsu Normal University, Xuzhou, China. From August 2007 to August 2009, he worked with the Department of Electronic Engineering, The Chinese University of Hong Kong, as a research associate. From February 2011 to April 2011, he worked with the Department of Electrical and Computer Engineering, National University of Singapore, as a research fellow. From September 2016 to September 2017, he worked with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Champaign, IL, USA, as a visiting scholar. He has authored or coauthored more than 100 refereed journal and conference papers. His current research interests include spoof surface plasmon polaritons theory and its applications, RF/microwave antenna and filter design, computational electromagnetics, and effects of electromagnetic radiation on the human body.

Dr. Zhao serves as an associate editor for *IEEE Access*, an associate editor-in-chief for *ACES Journal* and a reviewer for multiple journals and conferences including the *IEEE Trans. on Microwave Theory and Techniques*, *IEEE Trans. Antennas and Propagation*, *IEEE Antennas and Wireless Propagation Letters*, *ACES Journal*, and other primary electromagnetics and microwave related journals.

Research on Quasi-isotropic Radiation of Small Circular Arc Antenna

Hailong Liu[1], Jinbo Liu[1], Xiaoxia Nie[1], Jiming Song[2], and Zengrui Li[1]

[1]School of Information and Communication Engineering
Communication University of China, Beijing, 100024, China
liuhailong@cuc.edu.cn, liuj@cuc.edu.cn, nxiaoxia@cuc.edu.cn, zrli@cuc.edu.cn

[2]Department of Electrical and Computer Engineering
Iowa State University, Ames, IA 50011, USA
jisong@iastate.edu

***Abstract* –** The radiation characteristics of circular arc antennas are studied by applying the method of moments (MoM) to solve the electric field integral equation (EFIE). Based on the triangular current distribution of small circular arc antennas, the analytical expression of radiation fields of circular arc antennas is derived. It is found that the circular arc antenna is equivalent to a superposition of an electric dipole and a magnetic dipole, resulting in near isotropic radiation pattern. Finally, both MoM and CST simulations show that the small arc antenna can realize the near isotropic radiation pattern.

***Index Terms* –** Circular arc antenna, electric field integral equation (EFIE), method of moments (MoM), quasi-isotropic radiation.

I. INTRODUCTION

Isotropic antennas can radiate electromagnetic energy equally in all directions [1]. However, isotropic antennas are impossible in theory, because the transverse electric field in the far field region cannot be uniform over a sphere if the field is linearly polarized everywhere [2], [3]. So quasi-isotropic antennas were proposed and commonly used in applications such as radio frequency identification, radio frequency energy harvesting, and wireless access points [4].

Design approaches for quasi-isotropic antennas include folded dipoles [5], magnetic dipoles [6], orthogonal dipoles [2], combination of multiple dipoles [7], split ring resonators [8], and so on. This paper will introduce a new approach based on circular arc antennas.

An arc antenna is a usual deformation of a dipole antenna, which is the simplest and most basic of various antenna structures. Although the circular arc antenna is simple, it still has characteristics that are worth study for applications. For example, using a circular antenna in a logging tool as a receiving antenna has great advantages in ultra-deep boundary detection [9]. It is more convenient to analyze the circular arc antenna by applying the method of moments (MoM) to solve the electric field integral equation (EFIE) than the traditional analytical method [10]. In this paper, based on the conclusion that the small circular arc antenna can be equivalent to a superposition of an electric dipole and a magnetic dipole in [11], it is proposed that the small circular arc antenna can realize quasi-3D omnidirectional radiation. As a result, the arc antenna can be used to overcome the zero-reading problem of dipole antennas [12].

II. THEORY AND FORMULA DERIVATION

A. Derivation of circular arc antenna EFIE

The EFIE for the antennas or scatterers which are perfect electric conductor (PEC) can be expressed as [10]

$$\mathrm{j}\omega\mu_0\hat{t}\bullet\iint_S\left[\mathbf{J}_S(\mathbf{r}')+\tfrac{1}{k^2}\nabla'\bullet\mathbf{J}_S(\mathbf{r}')\nabla\right]g(\mathbf{r},\mathbf{r}')\,dS' = \hat{t}\bullet\mathbf{E}^i(\mathbf{r}), \tag{1}$$

where $\mathbf{J}_S$ is the induced surface current, g is the Green's function, k and μ_0 are the wave number and permeability of the free space, and $\hat{t}$ denotes any tangential unit vector to the PEC surface.

The geometry for the circular arc antenna of radius b and line radius a is depicted in Fig. 1. If $b << \lambda$, where λ is the wavelength, the antenna can be regarded as a

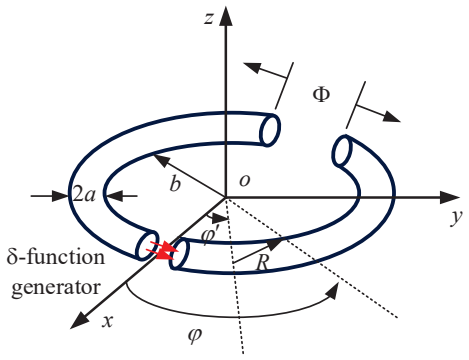


Fig. 1. Circular arc antenna geometry.

Submitted On: February 26, 2023
Accepted On: June 19, 2023

https://doi.org/10.13052/2023.ACES.J.380911

small circular arc antenna. And if $a << 2\pi b$, the antenna can be regarded as a thin wire antenna. Considering the thin wire antenna approximations, the source point $\mathbf{r}'$ and field point $\mathbf{r}$ can be represented by source point $\boldsymbol{\rho}'$ and field point $\boldsymbol{\rho}$. In addition,

$$\mathbf{J}_S(\mathbf{r}') = I(\varphi')\,\hat{\varphi}'/(2\pi a), \tag{2}$$

$$\hat{t} = \hat{\varphi}, \tag{3}$$

$$\hat{\varphi} \bullet \hat{\varphi}' = \cos(\varphi - \varphi'), \tag{4}$$

$$g(\varphi, \varphi') = \frac{e^{-\mathrm{j}kR}}{4\pi R}, \tag{5}$$

where R is the distance from the source point $\boldsymbol{\rho}'$ to the field point $\boldsymbol{\rho}$, expressed as

$$R = \sqrt{(a+b)^2 + b^2 - 2(a+b)b\cos(\varphi - \varphi')}$$
$$= \sqrt{a^2 + 4b(a+b)\sin^2[(\varphi - \varphi')/2]}. \tag{6}$$

By substituting (2)-(6) into (1), the EFIE suitable for the circular arc antenna is rewritten as

$$\int_{-\Phi/2}^{\Phi/2} I(\varphi')\,K(\varphi, \varphi')\,d\varphi' = -\mathrm{j}\frac{kb}{\eta}E_{\varphi}^{i}, \tag{7}$$

with

$$K(\varphi, \varphi') = \left[(kb)^2\cos(\varphi - \varphi') + \frac{\partial^2}{\partial \varphi^2}\right] g(\varphi, \varphi'). \tag{8}$$

B. Radiation field and equivalent model of circular arc antenna

The relationship between magnetic potential A and current vector $\mathbf{I}(\mathbf{r}')$ of the circular arc antenna is

$$\mathbf{A} = \frac{\mu}{4\pi}\int_C \mathbf{I}(\mathbf{r}')\frac{e^{-\mathrm{j}kR}}{R}dl', \tag{9}$$

where $\mathbf{I}(\mathbf{r}') = I(\varphi')\,\hat{\varphi}'$ and $dl' = bd\varphi'$. Therefore, (9) can be changed to

$$\mathbf{A} = \frac{b\mu}{4\pi}\int_{-\Phi/2}^{\Phi/2} I(\varphi')\,\hat{\varphi}'\frac{e^{-\mathrm{j}kR}}{R}d\varphi'. \tag{10}$$

In the spherical coordinate system, $\mathbf{A}$ can be expressed by the sum of the components in r, θ, and φ directions, that is

$$\mathbf{A}(r, \theta, \varphi) = A_r(r, \theta, \varphi)\,\hat{r} + A_\theta(r, \theta, \varphi)\,\hat{\theta} + A_\varphi(r, \theta, \varphi)\,\hat{\varphi}. \tag{11}$$

The vector potential is expressed as [13]

$$\begin{cases} A_r = \frac{b\mu_0}{4\pi}\sin\theta\int_{-\Phi/2}^{\Phi/2} I(\varphi')\sin(\varphi - \varphi')\frac{e^{-\mathrm{j}kR}}{R}d\varphi' \\ A_\theta = \frac{b\mu_0}{4\pi}\cos\theta\int_{-\Phi/2}^{\Phi/2} I(\varphi')\sin(\varphi - \varphi')\frac{e^{-\mathrm{j}kR}}{R}d\varphi' \\ A_\varphi = \frac{b\mu_0}{4\pi}\int_{-\Phi/2}^{\Phi/2} I(\varphi')\cos(\varphi - \varphi')\frac{e^{-\mathrm{j}kR}}{R}d\varphi' \end{cases}. \tag{12}$$

According to the far field approximation [14], $R \approx r - b\sin\theta\cos(\varphi - \varphi')$, the magnetic potential $\mathbf{A}$ is then transformed into

$$\begin{cases} A_r = \frac{b\mu_0 e^{-\mathrm{j}kr}}{4\pi r}\sin\theta\int_{-\Phi/2}^{\Phi/2} I(\varphi')\sin(\varphi - \varphi')\,e^{\mathrm{j}kb\sin\theta\cos(\varphi - \varphi')}d\varphi' \\ A_\theta = \frac{b\mu_0 e^{-\mathrm{j}kr}}{4\pi r}\cos\theta\int_{-\Phi/2}^{\Phi/2} I(\varphi')\sin(\varphi - \varphi')\,e^{\mathrm{j}kb\sin\theta\cos(\varphi - \varphi')}d\varphi' \\ A_\varphi = \frac{b\mu_0 e^{-\mathrm{j}kr}}{4\pi r}\int_{-\Phi/2}^{\Phi/2} I(\varphi')\cos(\varphi - \varphi')\,e^{\mathrm{j}kb\sin\theta\cos(\varphi - \varphi')}d\varphi' \end{cases}. \tag{13}$$

For a small circular arc antenna, the current distribution obtained by the MoM is shown in Fig. 2. Obviously,

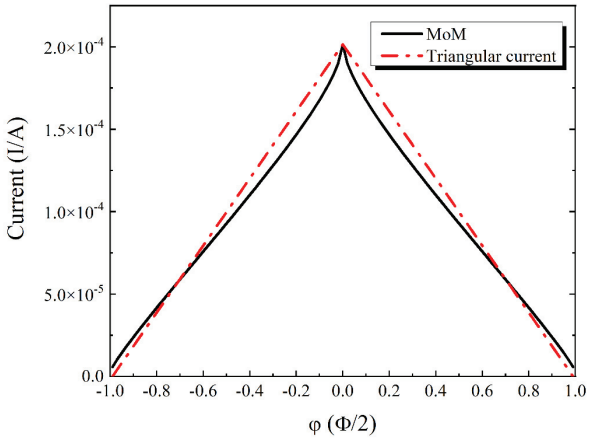


Fig. 2. Circular arc antenna current when $kb = 2\pi\times10^{-4}$.

the current distribution is close to the triangular current and can be written as

$$I(\varphi') = I_A(1 - 2|\varphi'|/\Phi) \quad -\Phi/2 \le \varphi' \le \Phi/2, \tag{14}$$

where I_A is the peak value of current amplitude calculated with MoM.

With small arc approximation, $e^{\mathrm{j}kb\sin\theta\cos(\varphi - \varphi')}$ can be expanded to $1 + \mathrm{j}kb\sin\theta\cos(\varphi - \varphi')$. Substituting (14) into (13) as well as integrating them respectively yields

$$\begin{cases} A_r = \frac{b\mu_0 e^{-\mathrm{j}kr}\sin\theta I_A}{4\pi r\Phi}\left[\begin{array}{l} 8\sin\varphi\sin^2(\Phi/4) \\ +\mathrm{j}kb\sin\theta\sin(2\varphi)\sin^2(\Phi/2)\end{array}\right] \\ A_\theta = \frac{b\mu_0 e^{-\mathrm{j}kr}I_A}{8\pi r\Phi}\left[\begin{array}{l} 16\cos\theta\sin\varphi\sin^2(\Phi/4) \\ +\mathrm{j}kb\sin(2\theta)\sin(2\varphi)\sin^2(\Phi/2)\end{array}\right] \\ A_\varphi = \frac{b\mu_0 e^{-\mathrm{j}kr}I_A}{16\pi r\Phi}\left\{\begin{array}{l} 32\cos\varphi\sin^2(\Phi/4) \\ +\mathrm{j}kb\sin\theta\left[\Phi^2 + 4\cos(2\varphi)\sin^2(\Phi/2)\right]\end{array}\right\} \end{cases}. \tag{15}$$

In the spherical coordinate system, the expression of electric field $\mathbf{E}^S$ in far field region is

$$\mathbf{E}^S = -\mathrm{j}\omega\left(\hat{\theta}A_\theta + \hat{\varphi}A_\varphi\right) \tag{16}$$

with

$$\begin{cases} E_\theta = \frac{-\mathrm{j}\omega b\mu_0 e^{-\mathrm{j}kr}I_A}{8\pi r\Phi}\left[\begin{array}{l} 16\cos\theta\sin\varphi\sin^2(\Phi/4) \\ +\mathrm{j}kb\sin(2\theta)\sin(2\varphi)\sin^2(\Phi/2)\end{array}\right] \\ E_\varphi = \frac{-\mathrm{j}\omega b\mu_0 e^{-\mathrm{j}kr}I_A}{16\pi r\Phi}\left\{\begin{array}{l} 32\cos\varphi\sin^2(\Phi/4) + \mathrm{j}kb\sin\theta \\ \left[\Phi^2 + 4\cos(2\varphi)\sin^2(\Phi/2)\right]\end{array}\right\} \end{cases}. \tag{17}$$

The small circular arc antenna can be characterized by an electric dipole moment $I_e l$ and a magnetic dipole moment $I_m l$ [15]. Next, we will discuss the radiation field expressions of the electric dipole and magnetic dipole.

The position of the electric dipole along the y-direction located at $x = x_e$ along x-axis is shown in Fig. 3 (a). Then, the magnetic vector potential $\mathbf{A}$ is

$$\mathbf{A}(\mathbf{r}) = \hat{y}\frac{\mu I_e l}{4\pi}\frac{e^{-\mathrm{j}kR}}{R}, \tag{18}$$

where

$$R = \sqrt{r^2 - 2x_e r\sin\theta\cos\varphi + x_e^2}. \tag{19}$$

Considering $r >> x_e$, R can be written as

$$R = r\sqrt{1 - 2\frac{x_e}{r}\sin\theta\cos\varphi + \left(\frac{x_e}{r}\right)^2} \approx r - x_e\sin\theta\cos\varphi, \quad (20)$$

thus,

$$\frac{e^{-\mathrm{j}kR}}{R} \approx \frac{e^{-\mathrm{j}k(r - x_e\sin\theta\cos\varphi)}}{r}. \quad (21)$$

Therefore, (18) can be simplified as

$$\mathbf{A}(\mathbf{r}) = \hat{y}\frac{\mu I_e l}{4\pi r}e^{-\mathrm{j}k(r - x_e\sin\theta\cos\varphi)}, \quad (22)$$

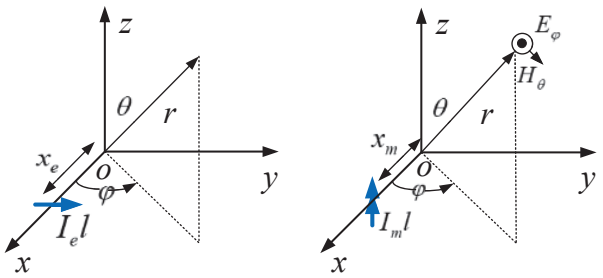
where $\hat{y} = \hat{r}\sin\theta\sin\varphi + \hat{\theta}\cos\theta\sin\varphi + \hat{\varphi}\cos\varphi$.

The radiation field expressions of the electric dipole along the y-direction located at $x = x_e$ are then

$$\begin{cases} E_\theta^e = -\mathrm{j}\eta\frac{kI_e l\cos\theta\sin\varphi}{4\pi r}e^{-\mathrm{j}k(r - x_e\sin\theta\cos\varphi)} \\ E_\varphi^e = -\mathrm{j}\eta\frac{kI_e l\cos\varphi}{4\pi r}e^{-\mathrm{j}k(r - x_e\sin\theta\cos\varphi)} \end{cases}. \quad (23)$$

The position of the magnetic dipole along the z-direction located at $x = x_m$ along the x-axis is shown in Fig. 3 (b). Its radiation field expression that can be derived in the same way is

$$\begin{cases} E_\theta^m = 0 \\ E_\varphi^m = -\frac{\mathrm{j}kI_m l\sin\theta}{4\pi r}e^{-\mathrm{j}k(r - x_m\sin\theta\cos\varphi)} \end{cases}. \quad (24)$$



(a) (b)

Fig. 3. Schematic diagrams of (a) an electric dipole and (b) a magnetic dipole.

By comparing the radiation field expression (17) with (23) and (24), it is concluded that the small circular arc antenna can be equivalent to a superposition of an electric dipole along the y direction located at $x_e = b\cos^2(\Phi/4)$ and a magnetic dipole along z direction located at $x_m = x_e$ as shown in Fig. 4.

The electric dipole moment is expressed as

$$I_e l = \left[8bI_A\sin^2(\Phi/4)\right]/\Phi, \quad (25)$$

and the magnetic moment is given as

$$I_m l = \frac{\mathrm{j}\omega\mu I_A b^2\Phi}{4}\left[1 - \frac{4\sin^2(\Phi/2)}{\Phi^2}\right]. \quad (26)$$

For a very small arc antenna with $\Phi << 1$, we have

$$x_e = x_m = b,\ I_e l = \frac{I_A b}{2}\Phi,\ I_m l = \frac{\mathrm{j}\omega\mu I_A b^2}{12}\Phi^3 \to 0. \quad (27)$$

For one half circular arc antenna with $\Phi = \pi$,

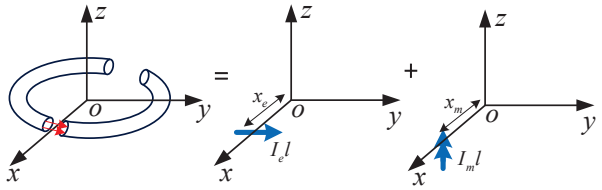
$$x_e = x_m = \frac{b}{2},\ I_e l = \frac{4I_A b}{\pi},\ I_m l = \frac{\mathrm{j}\omega\mu I_A\pi b^2}{4}\left(1 - \frac{4}{\pi^2}\right). \quad (28)$$

Similarly, for a full loop with a gap,

$$x_e = x_m = 0,\ I_e l = \frac{4I_A b}{\pi},\ I_m l = \frac{\mathrm{j}\omega\mu I_A\pi b^2}{2}. \quad (29)$$

For a full loop without a gap, we have the well-known results [16]

$$x_e = x_m = 0,\ I_e l = 0,\ I_m l = \mathrm{j}\omega\mu I_A\pi b^2. \quad (30)$$



(a) (b) (c)

Fig. 4. Equivalent model of a circular arc antenna: (a) Circular arc antenna with angle Φ, (b) electric dipole $I_e l$, and (c) magnetic dipole $I_m l$.

C. Electric dipole and magnetic dipole

The normalized pattern of an antenna is defined as

$$F(\theta,\varphi) = \frac{|E(\theta,\varphi)|}{E_{\max}}. \quad (31)$$

According to the radiation field expressions (23) and (24), the normalized pattern of an electric dipole and a magnetic dipole are

$$\begin{cases} F_\theta^e = \cos\theta\sin\varphi \\ F_\varphi^e = \cos\varphi \end{cases}, \quad (32)$$

and

$$\begin{cases} F_\theta^m = \sin\theta \\ F_\varphi^m = 0 \end{cases}, \quad (33)$$

respectively, while their radiation patterns are shown in Fig. 5.

From the patterns of the electric dipole and magnetic dipole, their radiation intensity near the central axis is very low or even zero, but these axes are not the same. If the electric dipole and magnetic dipole are superimposed, the combined patterns can be complementary and will have no direction where the normalized pattern is zero.

Where the ratio of the electric dipole to magnetic dipole $I_e l/I_m l = n_e/\eta\ (0 < n_e < \infty)$ and where n_e is the normalized ratio, the electric field obtained by superposition of the electric dipole and magnetic dipole is

$$\begin{cases} E_\theta = n_e E_\theta^e + E_\theta^m \\ E_\varphi = n_e E_\varphi^e + E_\varphi^m \end{cases}. \quad (34)$$

According to equation (31), the normalized pattern of the superposition of an electric dipole and a magnetic dipole can be obtained, and the minimum can be found.

Fig. 5. Radiation patterns of (a) an electric dipole along the y direction and (b) a magnetic dipole along the z direction; (c) the color map.

III. NUMERICAL RESULTS AND ANALYSIS

The minimum value of the normalized pattern as a function of n_e is shown in Fig. 6. When n_e = 1, that is, $I_e l/I_m l = 1/\eta$, the minimum is $\sqrt{2}/2$, which is the largest, and the pattern is the closest to 3D omnidirectional, as shown in Fig. 7.

In Section II A, it is concluded that the small arc antenna can be equivalent to the superposition of an electric dipole and a magnetic dipole. In Section II B, it is shown that the superposition of an electric dipole and magnetic dipole can realize quasi-isotropic radiation. Therefore, it is predictable that the small arc antenna can realize quasi-isotropic radiation.

In order to obtain the gain deviation, namely, the difference between the maximum gain and the minimum gain, and the radius of the arc when the pattern of the arc antenna with different angles Φ is the closest to 3D omni-directional, the MoM and the electromagnetic simulation software CST [17] are used to build the simulation model. The line radius of the arc is λ/1000 in both methods.

The gain deviation, radius, and arc length calculated by the two methods are compared and verified. Figure 8 shows the relationship between gain deviation and arc antenna angle. The difference in results between MoM and CST may be due to inconsistent meshing. In

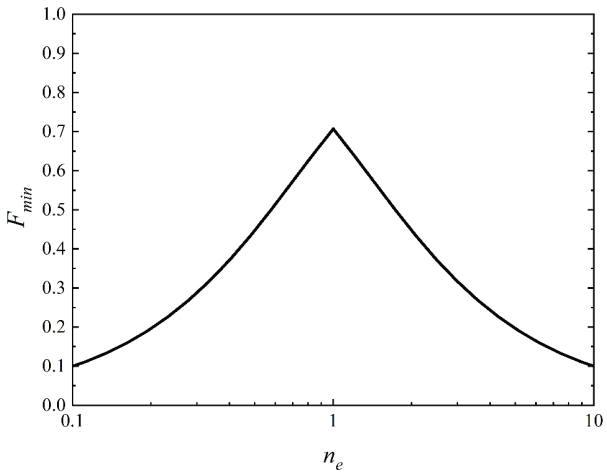


Fig. 6. Minimum value of normalized pattern of superposition of an electric dipole and a magnetic dipole.

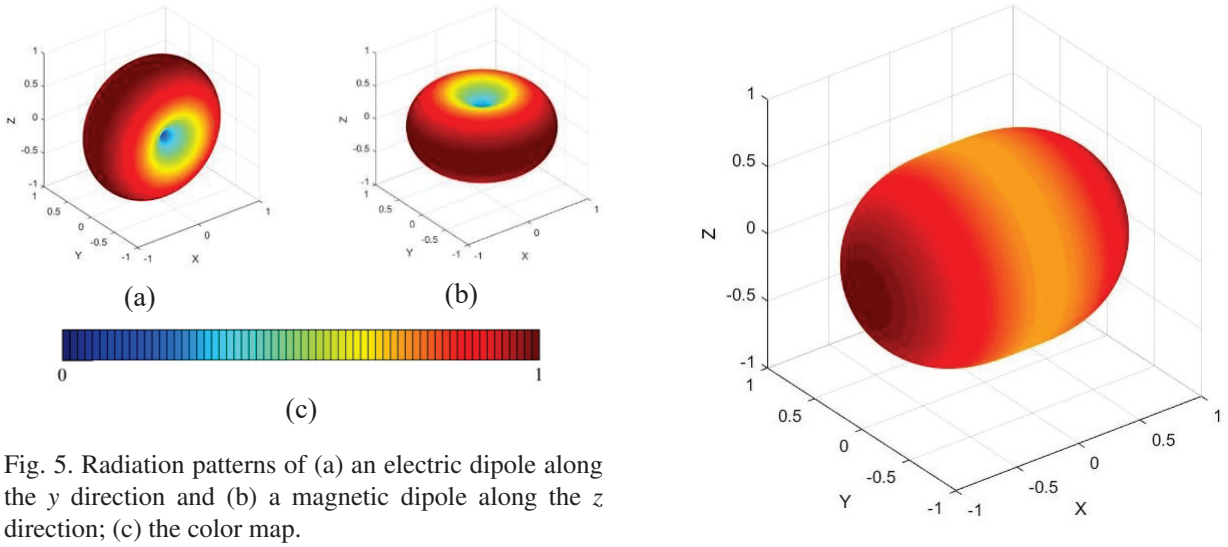


Fig. 7. When $I_e l/I_m l = 1/\eta$, the radiation pattern of the superposition of an electric dipole and a magnetic dipole. The color map is the same as that in Fig. 5 (c).

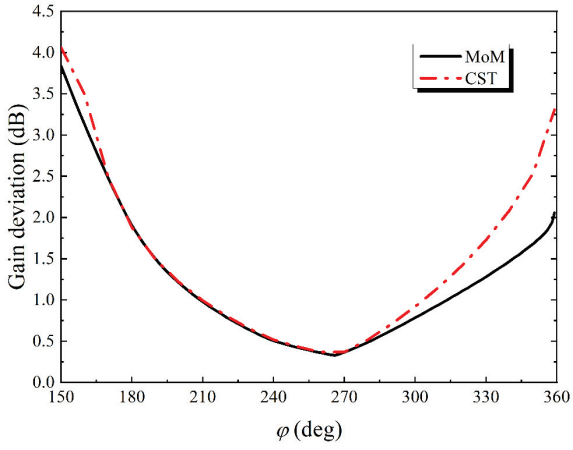


Fig. 8. Gain deviation of circular arc antenna with different angles when the arc antenna is the closest to 3D omnidirectional radiation.

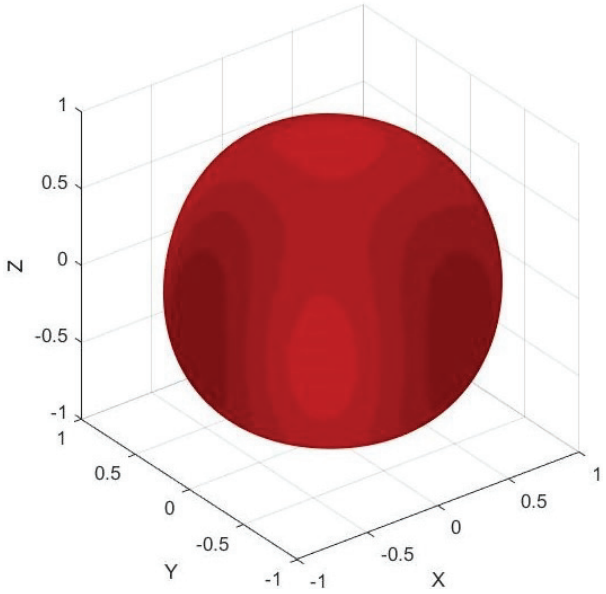


Fig. 9. The radiation pattern of 266-degree arc antenna which is the closest to isotropic radiation. The radius is 0.1542 λ. The color map is the same as that in Fig. 5 (c).

the results calculated by the MoM, the 266-degree arc antenna is the closest to 3D omnidirectional radiation, and the gain deviation is only 0.33 dB. The radiation pattern of the 266-degree arc antenna is shown in Fig. 9.

Figures 10 and 11 show the relationship between arc radius and arc antenna angle, and the relationship between arc length and arc antenna angle when the arc antenna is the closest to 3D omnidirectional radiation. It can be found that the arc radius and arc length decrease monotonically with the increase of the arc antenna angle.

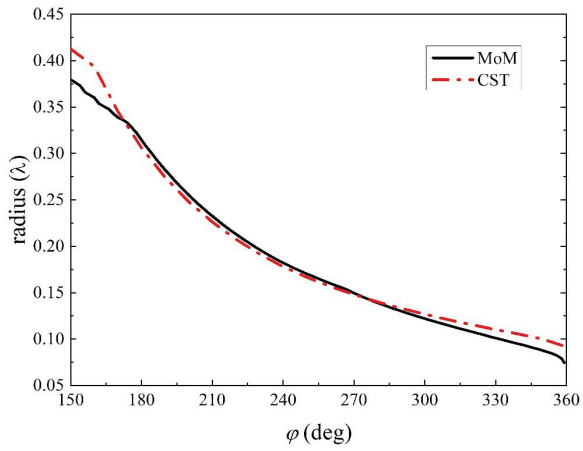


Fig. 10. Radius of circular arc antenna with different angles when the arc antenna is the closest to 3D omnidirectional radiation.

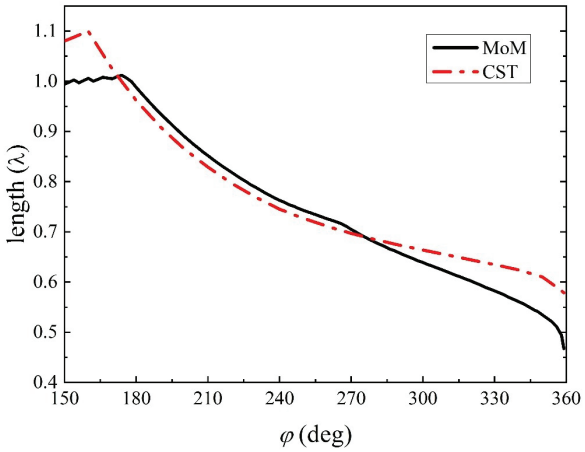


Fig. 11. Length of circular arc antenna with different angles when the arc antenna is the closest to near isotropic radiation.

That is to say, under this condition, when the arc is closer to a complete circle, the size is smaller. Although there are some differences in the simulation results of the two methods, the overall trend of change is consistent.

IV. SUMMARY

In this paper, the MoM is used to analyze the circular arc antenna. Firstly, the EFIE of the circular arc antenna is derived and verified by the known results of loop antennas. Secondly, the analytical approximate expression of the radiation field of the small circular arc antenna is derived by using the triangular current distribution. And by comparing the radiation field of circular arc antenna with that of electric dipole and magnetic dipole, it is concluded that the small circular arc antenna can be equivalent to a superposition of an electric dipole and a magnetic dipole. Thirdly, it is found that the superposition of an electric dipole and a magnetic dipole can realize quasi-3D omnidirectional radiation. Finally, by using MoM and CST software, it is shown that the small arc antenna can realize near isotropic radiation.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under Grant 61971384 and Grant 62071436.

REFERENCES

[1] S. M. Radha, M. Jung, P. Park, and I.-J. Yoon, "Design of an electrically small planar quasi-isotropic antenna for enhancement of wireless link reliability under NLOS channels," *Appl. Sci.*, vol. 10, no. 18, paper 6204, Sep. 2020.

[2] J. Kim, J. Park, A. A. Omar, and W. Hong, "A symmetrically stacked planar antenna concept exhibiting

quasi-isotropic radiation coverage," *IEEE Antennas Wireless Propag. Lett.*, vol. 19, no. 8, pp. 1390-1394, Aug. 2020.

[3] P. F. Hu, Y. M. Pan, X. Y. Zhang, and B. J. Hu, "A compact quasi-isotropic dielectric resonator antenna with filtering response," *IEEE Trans. Antennas Propag.*, vol. 67, no. 2, pp. 1294-1299, Feb. 2019.

[4] S. I. Hussain Shah, S. M. Radha, P. Park, and I.-J. Yoon, "Recent advancements in quasi-isotropic antennas: A review," *IEEE Access*, vol. 9, pp. 146296-146317, 2021.

[5] J. Ouyang, Y. M. Pan, S. Y. Zheng, and P. F. Hu, "An electrically small planar quasi-isotropic antenna," *IEEE Antennas Wireless Propag. Lett.*, vol. 17, no. 2, pp. 303-306, Feb. 2018.

[6] Q. Li, W.-J. Lu, S.-G. Wang, and L. Zhu, "Planar quasi-isotropic magnetic dipole antenna using fractional-order circular sector cavity resonant mode," *IEEE Access*, vol. 5, pp. 8515-8525, 2017.

[7] S. M. Radha, G. Shin, P. Park, and I.-J. Yoon, "Realization of electrically small low-profile quasi-isotropic antenna using 3D printing technology," *IEEE Access*, vol. 8, pp. 27067-27073, 2020.

[8] Y. Wang, M.-C. Tang, D. Li, K.-Z. Hu, M. Li, and X. Tan, "Low cost electrically small quasi-isotropic antenna based on split ring resonator," *Proc. Int. Appl. Comput. Electromagn. Soc. Symp.-China (ACES)*, pp. 1-2, Aug. 2019.

[9] S. Li, *System and Methodology of Look Ahead and Look Around LWD Tool*: U.S. Patent 10,605,073, 2020-3-31.

[10] R. F. Harrington, *Field Computation by Moment Methods*, MacMillan, New York, 1968.

[11] X. X. Nie, H. L. Liu, J. B. Liu, and J. M. Song, "Research on the small circular arc antenna based on the method of moments," *2022 International Applied Computational Electromagnetics Society (ACES-China) Symposium*, 2022.

[12] L. Li, L. C. Wang, X. L. Yin, and S. F. Li, "Design of an electrically small and near-3D omnidirectional loop antenna for UHF band RFID tag," 2013 *IEEE International Conference on Microwave Technology & Computational Electromagnetics*, pp. 246-248, 2013.

[13] C. A. Balanis, *Antenna Theory: Analysis and Design*, John Wiley & Sons, New York, 2015.

[14] W. L. Stutzman and G. A. Thiele, *Antenna Theory and Design*, John Wiley & Sons, New York, 2012.

[15] J. G. Van Bladel, *Electromagnetic Fields*, John Wiley & Sons, New York, 2007.

[16] J. M. Jin, *Theory and Computation of Electromagnetic Fields*, John Wiley & Sons, New York, 2010.

[17] CST Microwave Studio, ver. 2020, Computer Simulation Technology, Framingham, MA, 2020.

Hailong Liu received the B.S. degree in communication engineering and the M.S. degree in electronic information from Communication University of China, Beijing, China, in 2019 and 2023, respectively.

His research interests include computational electromagnetics and antenna design for RFID tags.

Jinbo Liu received the B.S. degree in electronic information engineering from Zhengzhou University, Zhengzhou, China, in 2010, and the Ph.D. degree in electronic science and technology from Beijing Institute of Technology, Beijing, China, in 2016.

He is currently an associate professor with the School of Information and Communication Engineering, Communication University of China, Beijing. His current research interests include computational electromagnetics and its applications, and frequency selective surfaces design.

Xiaoxia Nie received the B.S. degree in electronic information engineering from Taiyuan University of Technology, Taiyuan, China, in 2019, and the M.S. degree in electronic information from Communication University of China, Beijing, China, in 2022.

Her research interests include computational electromagnetics and antenna theory and design.

Jiming Song received the Ph.D. degree in electrical engineering from Michigan State University in 1993. From 1993 to 2000, he worked as a postdoctoral research associate, a research scientist and visiting assistant professor at the University of Illinois at Urbana-Champaign. From 1996 to 2000, he worked part-time as a research scientist at SAIC-DEMACO. Dr. Song was the principal author of the Fast Illinois Solver Code (FISC). He was a principal staff engineer/scientist at Semiconductor Products Sector of Motorola in Tempe, Arizona, before he joined the Department of Electrical and Computer Engineering at Iowa State University as an assistant professor in 2002.

Dr. Song currently is a professor at Iowa State University's Department of Electrical and Computer Engineering. His research has dealt with modeling and simulations of electromagnetic, acoustic and elastic wave propagation, scattering, and non-destructive evaluation, electromagnetic wave propagation in metamaterials and periodic structures and applications, interconnects on lossy silicon and radio frequency components, antenna radiation and electromagnetic wave scattering using fast algorithms, and transient electromagnetic fields. He received the NSF Career Award in 2006 and is an IEEE Fellow and ACES Fellow.

Zengrui Li received the B.S. degree in communication and information system from Beijing Jiaotong University, Beijing, China, in 1984, the M.S. degree in electrical engineering from Beijing Broadcast Institute, Beijing, China, in 1987, and the Ph.D. degree in electrical engineering from Beijing Jiaotong University, Beijing, China, in 2009.

He is currently a professor with the School of Information and Communication Engineering, Communication University of China, Beijing, China. He was a visiting scholar at Yokohama National University, Yokohama, Japan, from 2004 to January 2005. He also served as a senior visiting fellow at Pennsylvania State University from October 2010 to January 2011. His research interests include the areas of finite-difference time-domain (FDTD) methods, electromagnetic scattering, metamaterials and antennas. He is a senior member of the Chinese Institute of Electronics.

A Sector Ring Shape UWB Antenna by Tightly Coupling

Ziqin Wang[1], Zhihao Chen[1], Zhengming Tang[1], Lam Phav[3], and Fangyuan Chen[1,2]

[1]School of Electronic Information Engineering, China West Normal University
Shida Rd, Nanchong, Sichuan Province, 637000, China

[2]Jinyichang Science and Technology Co. Ltd
398 Zhenye Rd., Jiaxing, Zhejiang Province, 314000, China
fangyuanscu@gmail.com

[3]Ministry of Post and Telecommunications of Cambodia
Building 13, Monivong Blvd, Sangkat Srah Chak, Khan Daun Penh, Phnom Penh, 120210, Cambodia

***Abstract* –** In this paper, a sector ring shape UWB microstrip antenna based on the tight coupling is proposed. The continuous current and electric field distribution are formed by the coupling effect between microstrip patches. This design expanded the bandwidth of this antenna and realized a high-gain performance. To achieve the miniaturization design, the ring microstrip patches are staggered and the ground is removed to effectively reduce the size of the antenna. The ring shaped antenna has a 31 mm outside diameter and a 21 mm inside diameter, with a 2 mm thick substrate.

Besides the excellent mechanical features, such as being compact and easy to fabricate, the proposed antenna also shows good characteristics in radiation patterns and time-domain behaviors. The results show that the antenna achieves the effect of S_{11} less than -10 dB in the working frequency band of 2 GHz to 8 GHz, and the simulated results are in good agreement with the measured ones. The antenna realizes low profile and broadband in a novel approach, and this antenna can be used in V2X vehicle systems and IoT networks.

***Index Terms* –** tightly coupling, UWB.

I. INTRODUCTION

The concept of tight coupling was first proposed by Ben Munk in Wheeler's theory of Continuous Current Sheet Array (CCSA) [1], [2]. The novel interfingered dipoles were applied to the design of the tightly coupled array (TCA) antenna, the elements with coupling capacitors to extend the bandwidth to 4.5:1, which is widely used in antenna design. Volakis et al. used the rectangular spiral structure to realize the tightly coupled antenna with a bandwidth of 10:1; however, it has the disadvantage of complicated antenna unit design [3]. Li introduced a firmly coupled dipole structure with 1/10 wavelength to extend the frequency band [4]. In subsequent studies, it was found that when the distance between the array and the ground was half of the working wavelength, a short circuit effect would be caused, limiting the expansion of impedance bandwidth. Typically, a single or multi-layer dielectric plate is loaded above the array as a wide-angle matching layer, but this increases the profile and quantity of the antennas, which is not conducive to miniaturization design.

Traditional antennas transmit and receive electromagnetic waves through the radiation arm and the ground, the principle of which is to reflect electromagnetic waves through the ground. The general principle is to meet the performance of the antenna, such as gain, impedance matching, direction diagram, etc. As a novel way, the tight coupling can not only achieve effective radiation but also meet the effects of broadband antenna, low profile and high radiation efficiency. In order to avoid the coupling effect between the elements, the space between the elements of the traditional multi-element antenna is usually large. So the antenna size is large, and it is difficult to integrate and conform, greatly limiting the application scenarios of the antenna. In Luo's study, the size of the antenna reached a radius of 75 mm and a thickness of 3 mm. The antenna only operates in the dual-band Wi-Fi bands 2.37 GHz to 2.63 GHz and 5.52 GHz to 6.37 GHz [5]. Mohsen Gholamrezaei proposed a ring sector slot antenna that can achieve ultra-wideband characteristics of 3.48–5.5 GHz, 5.18–7.35 GHz, and 6.06–8.05 GHz on a floor of 32×35 mm^2 by etching a sector ring groove [6]. Moreover, a low-profile six-port circular patch antenna achieves 3:1 voltage standing wave ratio (VSWR) in the frequency band 5.8-7.5 GHz, with a floor diameter of 45 mm. [7]

However, the antenna profile can be greatly reduced while forming broadband by introducing the tightly coupling effect between antenna elements. Contrary to the

Submitted On: February 7, 2023
Accepted On: June 12, 2023

https://doi.org/10.13052/2023.ACES.J.380912

traditional antenna, the TCA uses the capacitive coupling effect between the elements to form a continuous current distribution and expand the antenna bandwidth. Therefore, the antenna structure is compact and the antenna size can be effectively reduced. It is more easily integrated into key systems such as automobiles and unmanned aerial vehicles. [8]

In this research, a novel design of a low-profile broadband antenna is presented. Several fan-ring microstrip patches are placed on the upper and lower sides of a ring-shaped dielectric substrate. These two layers of fan-ring microstrip patches are interspersed with each other, and radiation is carried out through the mutual coupling between the microstrip patches. Due to the special nature of the tight coupling, the ground is removed so that a circular cavity can be cut in the middle of the antenna to facilitate the integration of the antenna on different devices. The antenna proposed in this paper can reach the bandwidth of 2 GHz to 8 GHz under the condition that S_{11} is less than -10 dB.

In this design, the key parameters that influence the performance of the antenna are studied: the area of overlap between the top and bottom sides of the antenna, the radius of the ring, the spacing between the segments, and the equivalent circuit with inductances and capacitances. The proposed antenna presents advantages such as easy manufacturing and good radiation efficiency characteristics in the operating frequency band.

II. ANTENNA DESIGN AND ANALYSIS

A. UWB antenna design

According to the design of the annular antenna, a ring-shaped microstrip structure is placed on a medium substrate, and a ground is placed on the back. The ring microstrip antenna is concentrically divided into seven equal-angle fan rings, and the feeding position is placed between the two fan rings. The ground was further removed by coupling and replaced by five fan ring microstrip patches on the lower layer. In this design, two microstrip patch antennas connected to the feed port form an equivalent dipole antenna. The bandwidth is extended by coupling capacitors in the form of interleaved upper and lower radiation patches. By adjusting the different radius of the sector ring and the number of layers, results can be optimized below -10 dB. In this section, the evolution of the antenna and the optimization process of key parameters are discussed. ANSYS Electronics 2021R is used for the numerical simulation of antennas.

Figure 1 (a) shows the structure of the proposed antenna, and Fig. 1 (b) shows the geometric size of the proposed antenna. Figures 1 (b) and (c) show the dimensional parameters of the upper and lower elements of the antenna, respectively. The dielectric substrate material is

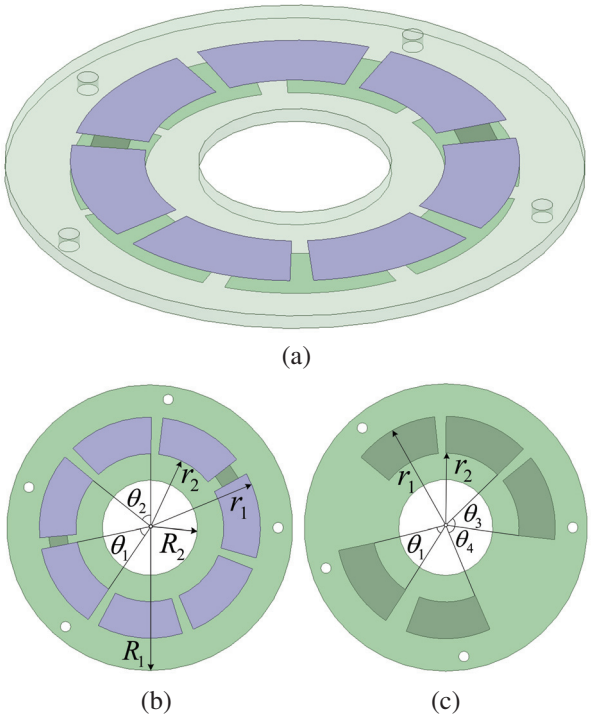


Fig. 1. Antenna geometry structure: (a) top view, (b) upper layer, and (c) bottom layer.

FR4 (with 2 mm thickness, $\varepsilon_r = 4.4$), with an outer diameter $R_1 = 39$ mm and an inner diameter $R_2 = 13$ mm. The thickness of the copper layer on the microstrip patch is 0.018 mm, the outer diameter r_1 of the upper fan ring is 31 mm, the inner diameter r_2 is 21 mm, and the θ_1 is 51 degrees. Seven concentric copper fan rings patches are printed on the top side of the FR4.

The inner and outer diameters of the five fan rings at the bottom side have the same geometry size as the upper layer. The five fan rings patches of the bottom layer are interspersed with those of the upper layer. The cross-placing layout of upper and lower copper layers is to maximize the coupling effect between the patches. Four asymmetric holes are made for positioning and adhesive fixation. The coaxial feed is selected to ensure the 50 Ω characteristic impedance at the antenna input. Recommended antenna parameters are shown in Figure 1, and the values are listed in Table 1.

Two adjacent microstrip patches on the upper surface form an excitation drive unit; the remaining microstrip patches are used as the coupled radiation

Table 1: Parameters of the proposed antenna

Parameter	R_1	R_2	r_1	r_2
Dimension (mm)	39	13	31	21
Parameter	θ_1	θ_2	θ_3	θ_4
Dimension	45°	51°	53°	50°

parasitic part of the excitation drive unit. The excitation drive unit excites other microstrip patches on the surrounding and bottom surface. The source of the excitation signal is established in the space between two adjacent microstrip patches that constitute the excitation unit. As a result, the high-frequency resonance of the microstrip patch is also constituting the excitation drive unit. The parasitic unit receives the electromagnetic wave generated by the high-frequency resonance of the excitation drive unit through coupling radiation. The microstrip patch that is not directly fed and does not generate the high-frequency resonance is excited by the adjacent patches. When the antenna is working, the excitation drive unit provided with feeding lines can be considered as the network port of the whole antenna.

The overall parameters of the tightly coupled UWB antenna unit are described by the normalized output to input voltage ratio of the antenna. For the radiated electromagnetic wave between the microstrip patches, the incident electromagnetic wave on the corresponding microstrip patch can be expressed as

$$a_i(z) = \frac{v_{io}^{+} e^{-yz}}{\sqrt{z_{io}}} = \frac{1}{2}\left[\frac{v_i(z)}{\sqrt{z_{io}}} + \sqrt{z_{io}} I_i(z)\right]. \quad (1)$$

The normalized reflected electromagnetic wave of the corresponding microstrip patch on the microstrip patch is expressed as

$$b_i(z) = \frac{v_{io}^{-} e^{-yz}}{\sqrt{z_{io}}} = \frac{1}{2}\left[\frac{v_i(z)}{\sqrt{z_{io}}} - \sqrt{z_{io}} I_i(z)\right]. \quad (2)$$

Therefore, the ratio of output and input voltage at the port of the matching network of tightly coupled UWB antennas is expressed as:

$$\frac{b_i(z)}{a_i(z)} = \frac{Z_i(z) - Z_{io}}{Z_i(z) + Z_{io}} = \Gamma_i(z) \quad (3)$$

The overall S-parameter matrix of the matching network of tightly coupled UWB antennas in the system is shown as follows:

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} S_{11} & \cdots & S_{1n} \\ \vdots & \ddots & \vdots \\ S_{n1} & \cdots & S_{nn} \end{pmatrix} \times \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \quad (4)$$

where a_1, a_2, ……, a_n is used to represent the incident wave of N ports, b_1, b_2, ……, b_n is used to represent its reflected wave, and S_{ij} is used to represent that when the incident wave a_j is used to connect the corresponding device port j, all ports except the device port are terminated with matching load. Therefore, as long as the reflected wave at the port i is measured, the corresponding S-parameters can be measured. To characterize the impedance matching characteristics of microstrip patches of different tightly coupled UWB antennas and the coupling characteristics between different microstrip patches the value range of i and j is $1 \sim n$, where n is the total number of several microstrip patches.

If all ports terminate in reference loads, there will be no reflections. A re-formation of the relationships allows for an observation of how S-parameters define all voltages reflected from any port i. The voltage port term U_{nh_i} is the sum of the incident voltage and the reflected voltage. The coupling effect between the excitation-driven microstrip patch and the parasitic microstrip patch in the tightly coupled UWB antenna unit can be equivalent to the mutual coupling effect between the multiple ports [9]. The corresponding voltage $U_{nh_i}$of the port i between different microstrip patches is

$$U_{nh_i} = Z_{inh_i} I_{nh_i} + \sum_{h_j} Z_{h_i h_j} I_{nh_j}, \quad (5)$$

where Z_{inh_i} is the input impedance of the microstrip patch h_i, $Z_{h_i h_j}$ is the mutual impedance between the microstrip patch i and microstrip patch j, I_{nh_i} is the corresponding port i current between different microstrip patches, I_{nh_j} and is the corresponding port j current between different microstrip patches.

As for each microstrip patch of tightly coupled antenna, its performance can be obtained.

The shape, size, and distance of each radiation patch are adjusted to form different degrees of coupling, to achieve inconsistent voltage, current, and peripheral field intensity on the antenna patch.

B. Parameters Analysis

In this design, the coupling effect between the same layer of fan ring microstrip patch units can be regarded as a capacitor, as well as between the microstrip patch units between different layers. Adjusting the distance between them can achieve the effect of adjusting the capacitance value. In order to achieve the desired coupling effect, it is necessary to design and adjust the sweep angle of each fan ring element, the gap between the same layer of fan ring elements, and the interlocking angle between the upper and lower layers.

Figure 2 shows the surface current at 2.45 GHz and 5.5 GHz. As can be seen from Fig. 3, adjusting the spacing between patches θ_3 shows that the coupling effect between patches has a vital effect on the full frequency band. After optimization, it is concluded that the antenna is well matched at θ_3 =53°, and the whole frequency band is below -10 dB. The difference θ_3 indicates that the spacing between antenna elements also means that the capacitance formed between antenna elements is not the same, resulting in different coupling effects. Adjusting its spacing can make reasonable use of the coupling effect between antenna elements to achieve good matching and bandwidth widening.

The width is r_1 minus r_2. As shown in Fig. 4, adjusting the width will affect the matching of the antenna. To achieve the target S_{11} and S_{22} at the same time, the length with width=10 mm is selected.

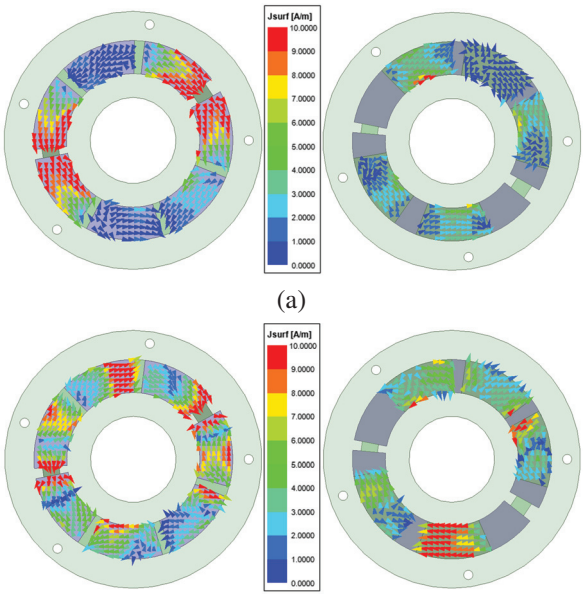


Fig. 2. Current distribution of the antenna: (a) 2.45 GHz and (b) 5.5 GHz.

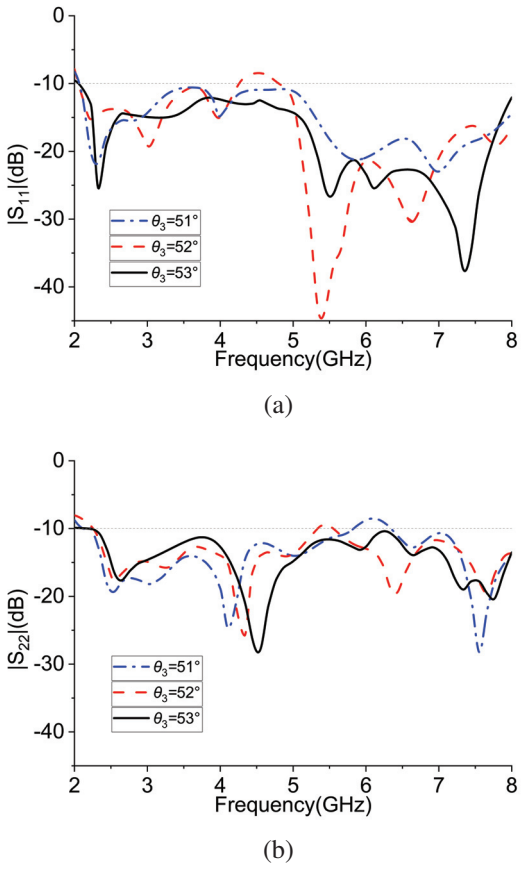


Fig. 3. The evolution of the proposed antenna: (a) S_{11} and (b) S_{22}.

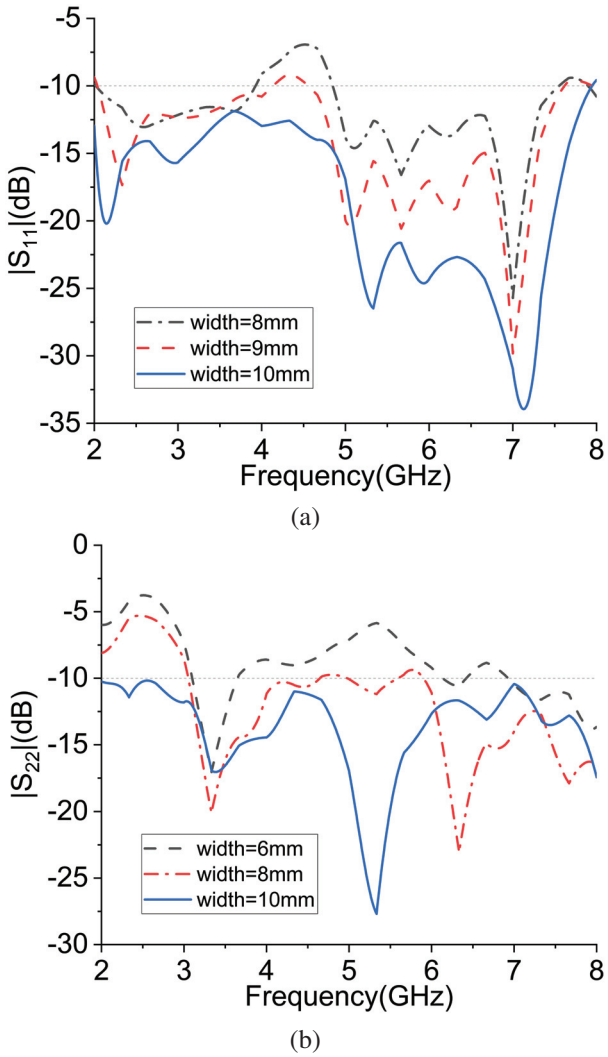


Fig. 4. The width evolution of the proposed antenna: (a) S_{11} and (b) S_{22}.

C. Equivalent Circuit

To demonstrate the working principle of the antenna from the point of view of the circuit, the simplified equivalent circuit model is used to analyze the antenna. The Fig. 5 shows the equivalent circuit model of the fan-ring-shaped antenna. It can be seen that a fan-ring microstrip patch can be considered as a set of inductors, capacitors and resistors in parallel. Therefore, the feed radiation patch is represented by L_r, C_r, R_r and the parasitic coupling patch is represented by L_p, C_p, R_p.

The equivalent circuit of a fan ring microstrip patch antenna is simplified into a module. The equivalent circuit of the antenna is presented as shown in Fig. 6, where A, B, C, D, E, F and G are the upper metal microstrip patch, and H, I, J, K and L are the metal microstrip patch staggered between the lower layer and the upper patch.

In this analysis, the strongly coupled antenna element is taken into consideration. The coupling between the patches is represented by a capacitor. According to the above analysis, to change θ_3, which means change the capacitance between the two modules, has a great impact on the coupling effect.

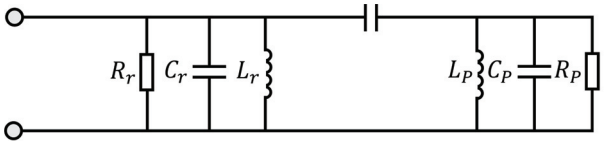


Fig. 5. Two loop antennas couple equivalent circuits.

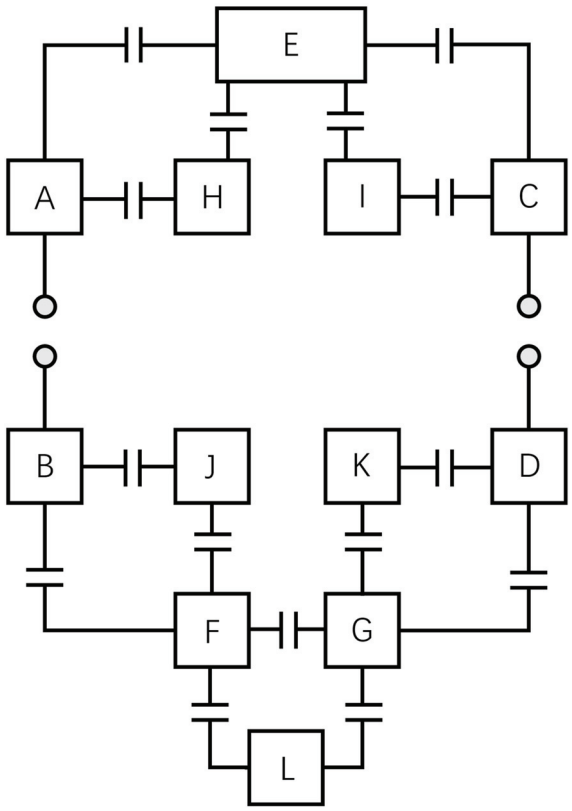
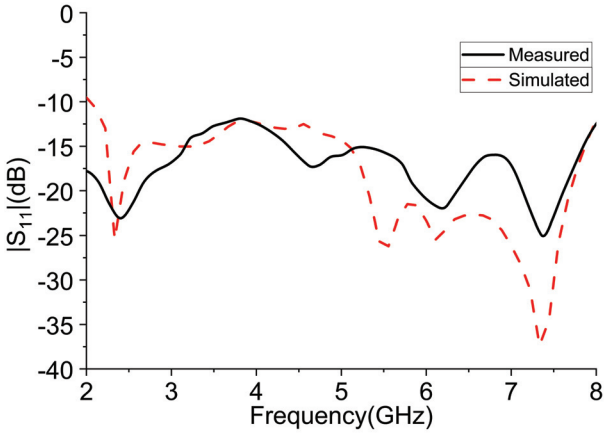


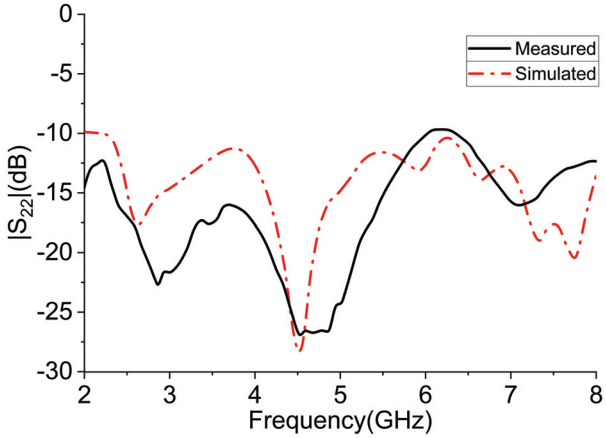
Fig. 6. Antenna equivalent circuit.

III. RESULTS AND ANALYSIS

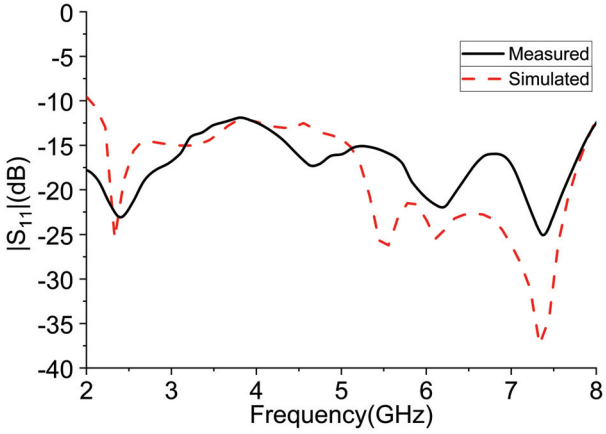
The antenna was manufactured and measured according to the design concepts and dimensions indicated above. As shown in Fig. 7, the scattering parameter S_{11} and S_{22} of the antenna was measured using the R&S ZNB40 vector network analyzer (the solid line), versus to the simulated results (the dashed line). The simulated curve and tested curved match well. It can be seen that both (a) S_{11} and (b) S_{22} are below -10 dB and



(a)



(b)

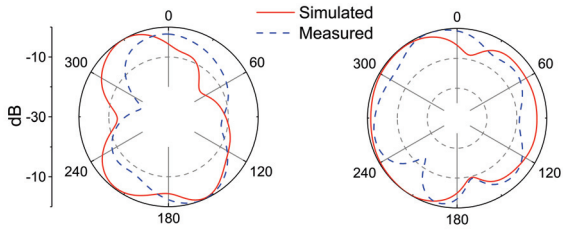


(c)

Fig. 7. Simulated and measured *S*-parameters of the proposed antenna: (a) S_{11}, (b) S_{22}, and (c)S_{12}.

the isolation degree (c) S_{12} is below -20 dB. The whole antenna has reached the ideal standard. The fabricated prototype shows a wide operating band as indicated in the simulation study.

Figure 8 (a) and 8 (b) show the simulated and measured far-field radiation patterns in the *xoz* plane

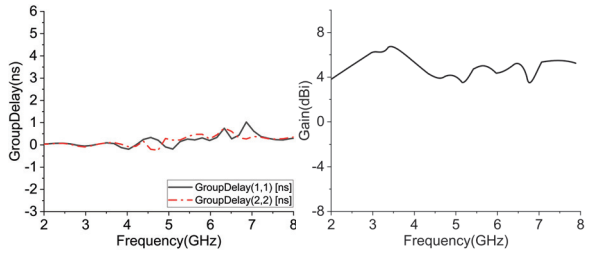


(a) *xoz* plane (E-plane) (b) *yoz* plane (H-plane)

Fig. 8. The radiation patterns of (a) *E*-plane and (b) *H*-plane at 5.5 GHz.

(E plane) and *yoz* plane (H plane) at a frequency of 5.5 GHz, respectively. Figure 8 shows that this antenna has a good omnidirectional radiation pattern in the E-plane and in the H-plane. In the E plane, the simulated and measured patterns are less consistent, this is due to the deformation of the 3-D pattern in the *xoz* plane.

To verify the time-domain performance of the proposed antenna, the group delay is also measured. In Fig. 9 (a), the measurements were made by placing two antennas facing each other at a distance of 30 cm.

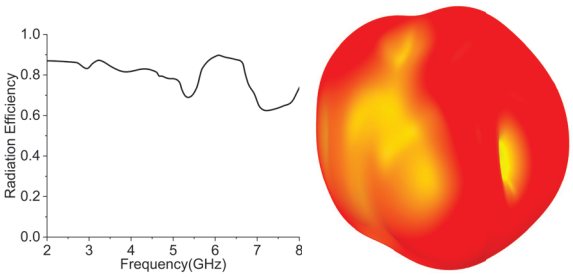


(a) (b)

Fig. 9. (a) Measured group delay and (b) measured peak gains.

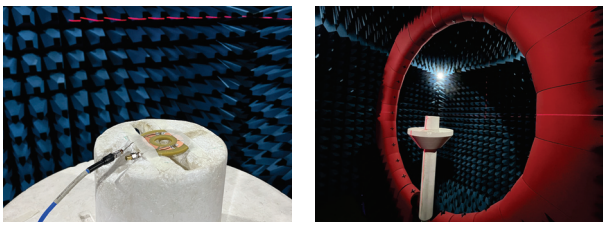
The antenna gain is greater than 4 dBi within the entire working frequency band, which shows good radiation performance as an omnidirectional antenna. The maximum gain of this antenna reaches 7.05 dBi at 3.5 GHz

Figure 10 shows the radiation efficiency and 3D pattern of this antenna. The average radiation efficiency reaches 79.5% from 2 GHz to 8 GHz, as tested in the anechoic chamber. The efficiency fluctuates as the frequency changes. This is due to the variation of the current distribution at different working frequencies. The current distribution led to a non-uniform electromagnetic field coupling effect. As can be seen from Fig. 10 (a), the working efficiency is higher at the starting working frequency at 2 GHz. The radiation efficiency decreases when the frequency increases. This is because the EM wave coupling effect is strong at low frequencies, and the effects are



(a) (b)

Fig. 10. (a) The radiation efficiency and (b) the 3D Pattern in 5 GHz.



(a) (b)

Fig. 11. The antenna in anechoic chamber.

Table 2: Performance comparison with other works

Reference Work	Bandwidth (GHz)	Gain (dBi)	Size (mm)	Efficiency
[5]	2.37-2.63 5.52-6.37	6	Ø75	/
[6]	3.48-5.5 5.18-7.35	4.25	32×35	90%
[7]	5.8-7.5	/	Ø45	63%
[8]	1.5-4.5	11.8	104×104	/
[10]	2.24-2.53 5.42-5.98	7.5	Ø170	80%
Proposed work	2-8	7.05	Ø78 (outside) Ø26 (inside)	79.5%

less strong when the wavelength is shortened in a higher frequency range.

IV. CONCLUSION

In this research, a low-profile wideband antenna tight coupling radiation is proposed. By placing the etched radiation patch on the upper and lower layers, a 2GHz to 8GHz radiation band is achieved. The ground structure is eliminated and the antenna miniaturization is realized. The equivalent circuit is provided and discussed, and the key parameters are studied. Then, the prototype of the proposed antenna is manu-

factured and its performance is tested. S_{11} simulated and measured. Radiation mode, group delay, and maximum gain were compared, showing good agreement. All these good properties make the proposed tightly coupling low-profile broadband antenna a good candidate for the communication system.

ACKNOWLEDGMENT

This work is supported by the Nanchong Key Laboratory of Electromagnetic Technology and Engineering. Jinyichang Science and Technology Co., Ltd. provides hardware support, technical assistance, and the test environment.

REFERENCES

[1] B. A. Munk, *Finite Antenna Arrays and FSS,* John Wiley & Sons, New York, 2007.

[2] H. Wheeler, "Simple relations derived from a phased-array antenna made of an infinite current sheet," *IEEE Transactions on Antennas and Propagation*, vol. 13, no. 4, pp. 506-514, July 1965.

[3] E. A. Alwan, K. Sertel, and J. L. Volakis. "A simple equivalent circuit model for ultrawideband coupled arrays," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 11, pp. 117-120, Jan. 2012.

[4] W. Li, S. Gao, L. Zhang, Q. Luo, and Y. Cai. "An ultra-wide-band tightly coupled dipole reflectarray antenna," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 2, pp. 553-540, Dec. 2017.

[5] Luo, Z. Kai, T. Su, and K. D. Xu. "A single-layer low-profile dual-wideband monopolar patch antenna with shorting vias and parasitic annular sectors," *IEEE Antennas and Wireless Propagation Letters*, Oct. 2022.

[6] M. Gholamrezaei, F. Geran, and R. A. Sadeghzadeh, "Completely in-dependent multi-ultrawideband and multi-dual-band frequency reconfigurable annular sector slot antenna (FR-ASSA)," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 2, pp. 893-898, Nov. 2016.

[7] K. L. Wong, C. J. Ho, and W. Y. Li, "Low-profile six-port circular patch antenna with six triple-shorted dual-resonant 60°-disk sectors to generate six uncorrelated waves for wideband mobile MIMO antennas," *IEEE Access*, pp. 80277-80288, Aug. 2022.

[8] Y. Fu, Y. Zhang, Q. Shi, M. Temiz, A. El-Makadema, and J. Shi, "A foldable tightly coupled crossed rings antenna array of ultrawide bandwidth and dual polarization," *IEEE Access*, pp. 86684-86695, Aug. 2022.

[9] T. R. Vogler, "Analysis of the radiation mechanisms in and design of tightly-coupled antenna arrays," Diss. Virginia Tech, 2010.

[10] S. Gao, L. Ge, D. Zhang, and W. Qin, "Low-profile dual-band stacked microstrip monopolar patch antenna for WLAN and car-to-car communications," *IEEE Access*, pp. 69575-695816, Oct. 2018.

[11] Z. F. Wu, W. J. Lu, J. Yu, and L. Zhu, "Wideband null frequency scanning circular sector patch antenna under triple resonance," *IEEE Transactions on Antennas and Propagation* vol. 66, no. 11, pp. 7266-7274, May 2020.

[12] S. M. Moghaddam, J. Yang, and A. U. Zaman, "Fully-planar ultrawideband tightly-coupled array (FPU-TCA) with integrated feed for wide-scanning millimeter-wave applications," *IEEE Transactions on Antennas and Propagation*, vol. 68, no. 9, pp. 6591-6601, June 2020.

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

[14] H. Zhang, S. Yang, S. Xiao, Y. Chen, and S. Qu, "Low-profile, lightweight, ultra-wideband tightly coupled dipole arrays loaded with split rings," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 6, pp. 4257-4262, Mar. 2019.

Ziqin Wang received the B.E. degree in applied electronic engineering from China West Normal University, Nanchong, China, in 2020. He is currently pursuing the master's degree in electronic information engineering at China West Normal University, Nanchong, China. He was an intern at the Jinyichang Science & Technology from 2022 to 2023. His research interests include microwave theory, numerical computation in electromagnetics, and antenna design.

Zhihao Chen received the B.E. degree in communication engineering from China Jiliang University, HangZhou, China, in 2021. He is currently pursuing the master's degree in electronic information engineering at China West Normal University, Nanchong, China. He was a teaching assistant at the school of electronic information engineering of China West Normal University. His research interests include microwave theory and technology, and antenna design.

Zhengming Tang received his M.S. degree in electrical engineering from Southwest Jiaotong University in 2012, and his Ph.D. degree from Sichuan University in 2016. Now he is a professor in the School of Electronics and Information Engineering, China West Normal University, Sichuan, China. His research interests are microwave heating, microwave theory, and microwave chemistry.

Fangyuan Chen received his Ph.D. degree in radio physics at Sichuan University, Chengdu, China, in 2016. From 2013 to 2015 he was a visiting doctoral student at Cornell University, New York, USA. Now he is the deputy chief engineer of Jinyichang Science & Technology Co., Ltd. Currently, he is responsible for strategic technology development in V2X network architecture design, multi-band frequency antenna design. His research interests are microwave theory and technology, and optimization algorithms.

Lam Phav received his B.E degree in computer science and engineering, from Royal University of Phnom Penh, Cambodia, in 2004. He received the M.Sc. degree in communication information systems, from Sichuan University, Chengdu, Sichuan, China, in 2013. Currently he is working at the General Department of Information and Communication Technology, Ministry of Post and Telecommunications (MPTC), Phnom Penh, Cambodia. His research intetrests are antenna theory and design, wireless communication technology.

Design of Multilayer Wideband Microwave Absorbers using Improved Grey Wolf Optimizer

Hao Nan Zhang, Zhi Fei Zhang, Yi Du, Wei Bin Kong*, Xiao Fang Yang, and Zhong Qing Fang

College of Information Engineering, Yancheng Optical Fiber Sensing and
Application Engineering Technology Research Center
Yancheng Institute of Technology, Jiangsu Yancheng, 224051, China
kongweibin2007@sina.com
*Corresponding author

Abstract – In this paper, an improved heuristic algorithm based on the disturbance and somersault foraging grey wolf optimizer (IDSFGWO) is proposed to optimize the design of multilayer wideband microwave absorbers for normal incidence. The multilayer absorber is designed to reduce maximum reflection coefficient by choosing suitable layers of materials from a predefined database. Three improvement strategies are given to enhance the performance of GWO, including tent map, nonlinear perturbation, and somersault foraging. The optimization results show that the reflection coefficients optimized by IDSFGWO are better than those of other algorithms for multilayer absorber design.

Index Terms – absorbing material, GWO, multilayer microwave absorbers, reflection coefficient.

I. INTRODUCTION

Microwave absorbers have important applications in stealth technology [1], wave anechoic chamber [2], imp-roving electromagnetic environment, and protecting daily safety [3–4]. Traditional single microwave absorbers are limited by their low electromagnetic parameters, narrow absorber band, and thick absorber structure. Multilayer microwave absorbers can complement and correlate the absorbing properties of each material and broaden the absorbing bandwidth to a certain extent [5]. Reasonable optimization of the absorbing structure can improve the absorptivity of a multilayer microwave absorber.

Over the years, primitive optimization algorithms like simplex algorithms [6], penalty function method [7], and multi-objective programming method [8] have been used. Simplex method is often used to find the optimal solution of objective function in linear constraint problems. Penalty function method is to transform the constrained optimization problem into a series of unconstrained optimization problems to solve. Multi-objective programming is a mathematical method for solving multi-objective decision-making problems based on linear programming. Recently, nature-inspired heuristic algorithms have been effectively applied to the design of absorbers [9–10], most notably the particle swarm optimization and the genetic algorithm [11–13]. They are accepted by researchers due to their higher probability of global optimal solution, higher robustness, and faster convergence speed.

This paper proposes the improved grey wolf optimizer (GWO) to design multilayer wideband microwave absorbers. Three enhancement strategies are provided to improve the performance of GWO. The tent map is used to initialize the populations. The nonlinear perturbation factor is adopted to balance mining and exploration capabilities. The somersault foraging strategy is introduced to prevent the algorithm from converging to local optima during the later stage. The reflection coefficient of incident electromagnetic waves on the surface of multilayer flat structures is calculated by the improved GWO through simulation experiments. It is shown that IDSFGWO has a certain advantage in designing multilayer microwave absorbers.

II. PHYSICAL MODEL OF MULTILAYER ABSORBER

The physical model of a multilayer microwave absorber is shown in Fig. 1. It is a multilayer system consisting of N layers of different materials backed by a perfect electric conductor (PEC). The total reflection coefficient of the multilayer microwave absorber can be calculated by using the equivalent transmission line method [14–15]. Each layer of media in Fig. 1 is considered as a transmission line. The equivalent transmission line model of the cascade of N different uniform transmission lines is shown in Fig. 2.

The input impedance of the Nth transmission line can be considered as the input impedance of the transmission line with a terminal short, which is expressed as

Submitted On: April 24, 2023
Accepted On: May 17, 2023

https://doi.org/10.13052/2023.ACES.J.380913

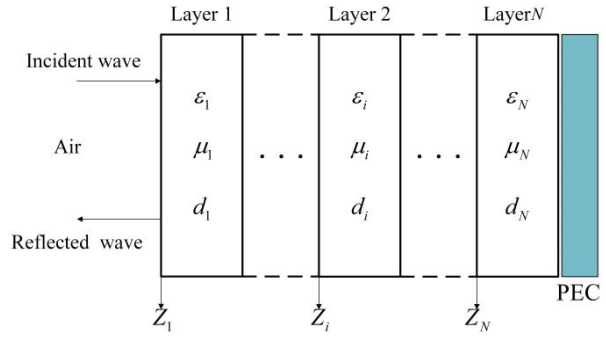


Fig. 1. Structure of multilayer microwave absorber.

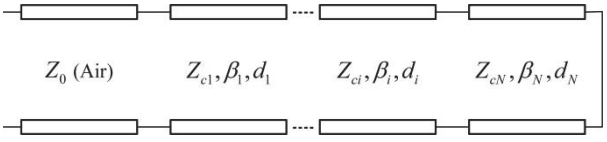


Fig. 2. Equivalent circuit of multilayer microwave absorber.

follows:

$$Z_i = jZ_c \tan(\beta d), \tag{1}$$

where, Z_c is the characteristic impedance, β is the propagation constant, d is the length of transmission line. Z_c and β are calculated in the TE mode as follows:

$$\beta = k_0\sqrt{\varepsilon_r\mu_r - \sin^2\theta}, \tag{2}$$

$$Z_c = \frac{\omega\mu}{\beta}, \tag{3}$$

while in TM mode Z_c and β are expressed as follows:

$$\beta = k_0\sqrt{\varepsilon_r\mu_r - \sin^2\theta}, \tag{4}$$

$$Z_c = \frac{\beta}{\omega\varepsilon}, \tag{5}$$

where, ε_r and μ_r are the relative permittivity and relative permeability of the materials, ε and μ are the permittivity and permeability of the materials.

The input impedance of the left adjacent transmission line can be expressed as follows:

$$Z_i = Z_c\frac{Z_1 + jZ_c\tan(\beta d)}{Z_c + jZ_1\tan(\beta d)}, \tag{6}$$

where, Z_1 is the terminal load of the left adjacent transmission line. It is also the input impedance of the Nth transmission line represented by Eq. (1).

Repeat the above process from right to left in turn. Then the reflection coefficient of the interface between air and medium is expressed as

$$R = \frac{Z_{10} - Z_{c0}}{Z_{10} + Z_{c0}}, \tag{7}$$

where, Z_{c0} is the characteristic impedance of the air corresponding to the transmission line, Z_{10} is the load impedance of the air corresponding to the transmission line.

For TE mode, Z_{c0} is defined as follows:

$$Z_{c0} = \frac{Z_0}{\cos\theta}. \tag{8}$$

For TM mode, Z_{c0} is defined as

$$Z_{c0} = Z_0\cos\theta, \tag{9}$$

where, Z_0 is the wave impedance in vacuum.

Finally, the total reflection coefficient of the multilayer microwave absorber is obtained:

$$RL = 20\log_{10}\left|\frac{Z_{10} - Z_{c0}}{Z_{10} + Z_{c0}}\right|. \tag{10}$$

III. GREY WOLF OPTIMIZER

A. Basic grey wolf optimizer

GWO is a new heuristic algorithm based on the predatory characteristics of grey wolves. It optimizes the search by imitating the leadership level and hunting mechanism of grey wolves. The leadership hierarchy of the grey wolf population divides the wolves into $\alpha\beta\delta\omega$, and the mathematical model includes stalking, encircling, and hunting [16]-[18].

1) Encircling prey

$$\overrightarrow{X}(t+1) = \overrightarrow{X_P}(t) - \overrightarrow{A}\cdot\left|\overrightarrow{C}\cdot\overrightarrow{X_P}(t) - \overrightarrow{X}(t)\right|, \tag{11}$$

where t indicates the current iteration, $\overrightarrow{A}$ and $\overrightarrow{C}$ are coefficient vectors, $\overrightarrow{X_P}(t)$ is the position vector of the prey, and $\overrightarrow{X}(t)$ indicates the position vector of a grey wolf.

The vectors $\overrightarrow{A}$ and $\overrightarrow{C}$ are calculated as follows:

$$\overrightarrow{A} = 2\overrightarrow{a}\cdot\overrightarrow{r}_1 - \overrightarrow{a}, \tag{12}$$

$$\overrightarrow{C} = 2\overrightarrow{r_2}, \tag{13}$$

where components of $\overrightarrow{a}$ are linearly decreased from 2 to 0 over the course of iterations, and $\overrightarrow{r_1}$, $\overrightarrow{r_2}$ are random vectors in [0,1].

The factor $\overrightarrow{a}$ is defined as follows:

$$a = 2(1 - t/t_{\max}), \tag{14}$$

where $t_{\max}$ is the maximum number of iterations.

2) Hunting

The first three best candidate solutions obtained so far oblige the other search agents to update their positions according to the position of the best search agents. So the updating equations for the wolves positions are proposed as follows:

$$\begin{cases} \overrightarrow{X_1} = \overrightarrow{X_\alpha} - \overrightarrow{A_1}\cdot\left|\overrightarrow{C_1}\cdot\overrightarrow{X_a} - \overrightarrow{X}\right| \\ \overrightarrow{X_2} = \overrightarrow{X_\beta} - \overrightarrow{A_2}\cdot\left|\overrightarrow{C_2}\cdot\overrightarrow{X_\beta} - \overrightarrow{X}\right| \\ \overrightarrow{X_3} = \overrightarrow{X_\delta} - \overrightarrow{A_3}\cdot\left|\overrightarrow{C_3}\cdot\overrightarrow{X_\delta} - \overrightarrow{X}\right| \end{cases}, \tag{15}$$

$$\overrightarrow{X}(t+1) = \frac{\overrightarrow{X_1} + \overrightarrow{X_2} + \overrightarrow{X_3}}{3}, \tag{16}$$

where, $\overrightarrow{X_\alpha}\overrightarrow{X_\beta}\overrightarrow{X_\delta}$ are the first three best solutions in the swarm at a given iteration t, $\overrightarrow{A_1}\overrightarrow{A_2}\overrightarrow{A_3}$ are defined as in Eq. (12), and $\overrightarrow{C_1}\overrightarrow{C_2}\overrightarrow{C_3}$ are defined using Eq. (13).

B. Improved grey wolf optimizer

The basic grey wolf optimizer uses randomly generated data in the initial population stage, resulting in poor population diversity and poor optimization results. A chaotic map is used to generate a random chaotic sequence generated by a simple deterministic system, which has the characteristics of nonlinearity, ergodicity, and randomness. When solving the optimization problem, chaotic maps can maintain the diversity of the population and improve the global search ability. The existing chaotic maps include Logistic map, Chebyshev map, Tent map, etc. [19–20]. Tent map has better ergodic uniformity than other maps. So Tent map is selected to initialize the grey wolf population. The Tent map expression is defined as follows:

$$X_{t+1} = \begin{cases} \frac{X_t}{u} & X_t \in [0,u) \\ \frac{1-X_t}{1-u} & X_t \in [0,u) \end{cases}, \quad (17)$$

where, the value of parameter u is 0 to 1. The obtained chaotic sequence has uniform distribution when u is 0.5.

Due to the linear change of the convergence factor a, the first half and the second half of the convergence of the algorithm have the same decline. Thus, the grey wolf optimizer cannot effectively complete both global search in the early stage and local search in the late stage. Adjusting linear to nonlinear is a general improvement [21]. The improved nonlinear perturbation factor can balance the capability of local and global search effectively [22]. The nonlinear perturbation factor E is defined as follows:

$$E = randn \cdot (\sin^k(\frac{\pi}{2} \cdot \frac{t}{t_{\max}}) + \cos(\frac{\pi}{2} \cdot \frac{t}{t_{\max}}) - 1), \quad (18)$$

$$\overrightarrow{A} = (2\overrightarrow{a} \cdot \overrightarrow{r}_1 - \overrightarrow{a}) + E, \quad (19)$$

where *randn* represents a random number subject to Gaussian normal distribution, k represents a constant that determines the peak value of the disturbance factor.

In order to avoid the situation that the basic GWO is prone to fall into the local optimum in the later stage of optimization, the somersault feeding strategy is introduced. This strategy is inspired by the behavior of manta rays that take the current optimal solution as the turning fulcrum and roll to the other side of the mirror relationship with their current position [23]. The mathematical expression of somersault for aging strategy is defined as follows:

$$X_i^d(t+1) = X_i^d(t) + S \cdot (r_1 X_{best}^d - r_2 X_i^d(t)) \quad i = 1, \cdots, N, \quad (20)$$

where, S is the flip factor, X_{best}^d is the location of prey, N is the population number, d is a dimension, and r_1,r_2 are random numbers in [0,1].

In each iteration of the GWO, the fitness value of the current grey wolf is compared with the fitness value of the grey wolf after jumping the fulcrum to determine whether to carry out the somersault feeding strategy. If the grey wolf has fallen into the local optimum at this time, the optimization result will be replaced. As the grey wolf optimizer iterates to the later stage, each wolf is getting closer to the optimal solution. The effect of jumping out of the local optimum is more obvious. The somersault foraging bebavior of grey wolves is shown in Fig. 3. The flow chart of improved GWO is shown in Fig. 4.

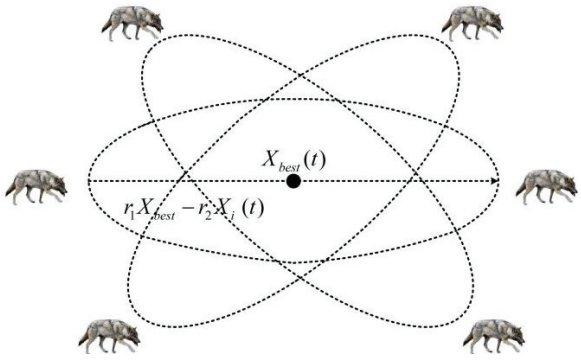


Fig. 3. Somersault foraging behavior of grey wolf.

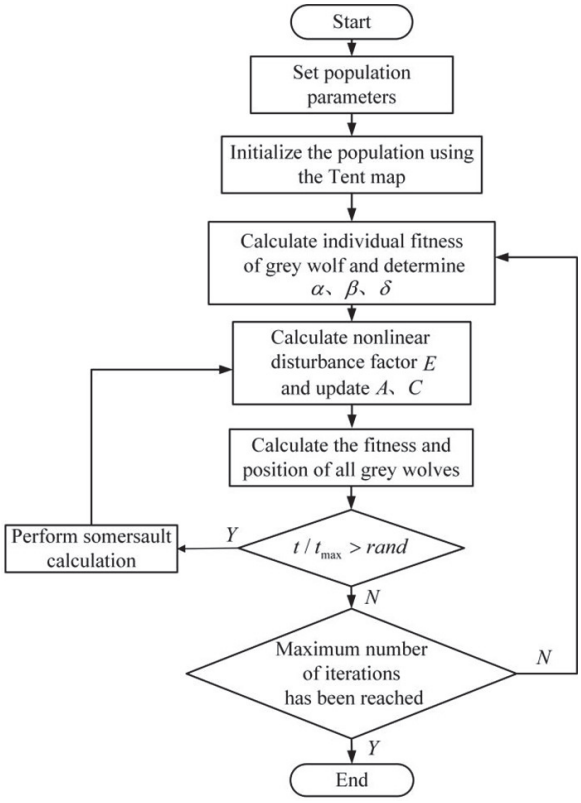


Fig. 4. Flow chart of improved grey wolf optimizer.

C. Objective function

Generally, the reflection coefficient is used to measure the absorbing effect. The reflection coefficient of the electromagnetic wave is closely related to the electromagnetic parameters of each layer of materials, the thickness of the layer, the arrangement of the layer and the incident frequency of the electromagnetic wave [24–25]. The optimization design is to optimize the selection and arrangement of layer thickness and materials type of each layer under the constraints of given number of layers, maximum thickness and band-width. The optimization goal is to minimize the maximum value of reflection coefficient.

Therefore, the maximum reflection coefficient in a specific frequency band is taken as the objective function. The objective function is represented as follows:

$$F_{obj} = 20\log_{10}(\max|R|) \tag{21}$$

IV. SIMULATION EXPERIMENT AND RESULT ANALYSIS

A. Database of absorbing materials

Absorbing materials convert the incident electromagnetic wave into heat energy or other forms of energy through various loss mechanisms, so as to absorb and attenuate the electromagnetic wave.

The single absorption principle database selected by many scholars limits the ability of the algorithm to search the optimal solution. There are 16 kinds of materials selected for this optimization design, which are divided into 4 categories: lossless dielectric materials, lossy magnetic materials, lossy dielectric materials, and relaxation magnetic materials, basically covering all types of electromagnetic materials. The database of absorbing materials is shown in Table 1.

Real and imaginary parts of permeability and permittivity for corresponding materials can be calculated using following equations.

For lossy magnetic materials:

$$\mu'(f) = \mu'(1GHz)/f^a, \tag{22}$$

$$\mu''(f) = \mu''(1GHz)/f^b. \tag{23}$$

For lossy dielectric materials:

$$\varepsilon'(f) = \varepsilon'(1GHz)/f^a, \tag{24}$$

$$\varepsilon''(f) = \varepsilon''(1GHz)/f^b. \tag{25}$$

For relaxation-type magnetic materials:

$$\mu'(f) = \frac{\mu_m f_m^2}{f^2 + f_m^2}, \tag{26}$$

$$\mu''(f) = \frac{\mu_m f_m f}{f^2 + f_m^2}. \tag{27}$$

Finally, the complex permittivity and permeability can be calculated as

$$\mu = \mu' - j\mu'', \tag{28}$$

$$\varepsilon = \varepsilon' - j\varepsilon''. \tag{29}$$

Table 1: Database of absorbing materials

Lossless Dielectric Materials ($\mu' = 1, \mu'' = 0$)				
#	ε'			
1	10			
2	50			
Lossy Magnetic Materials ($\varepsilon' = 15, \varepsilon'' = 0$)				
#	$\mu'(1GHz)$	a	$\mu''(1GHz)$	b
3	5	0.974	10	0.961
4	3	1.000	15	0.957
5	7	1.000	12	1.000
Lossy Dielectric Materials ($\mu' = 1, \mu'' = 0$)				
#	$\varepsilon'(1GHz)$	a	$\varepsilon''(1GHz)$	b
6	5	0.861	8	0.569
7	8	0.778	10	0.682
8	10	0.778	16	0.861
Relaxation-type Magnetic Materials ($\varepsilon' = 15, \varepsilon'' = 0$)				
#	μ_m	f_m		
9	35	0.8		
10	35	0.5		
11	30	1.0		
12	18	0.5		
13	20	1.5		
14	30	2.5		
15	30	2.0		
16	25	3.5		

B. Results and analysis

The application of the IDSFGWO to optimize multilayer microwave absorber is discussed and analyzed with the help of two design examples, including a 5-layer microwave absorber and a 7-layer microwave absorber. The simulation experiment considers the case of vertical incidence of electromagnetic waves.

Particle swarm optimization (PSO) is a classic heuristic intelligent algorithm which is based on the collaborative manner in which a swarm of insects, a herd of animals, a flock of birds, or a school of fish search for food [26]. The PSO algorithm is easy to execute and it has been seen to perform well on the design of broadband microwave absorbers [27].

WOA is a heuristic intelligent optimization algorithm that seeks the optimal solution by simulating the behavior of whale populations [28]. Both WOA and GWO can be used to solve nonlinear, nonconvex, and high-dimensional optimization problems. GWO performs well when dealing with large-scale problems or constrained optimization problems. WOA performs well when rapid convergence and high precision are required [29].

Example A: 5-layers microwave absorber

The design is a multilayer absorbing structure with 5 layers. The specified frequency range is 2-12 GHz,

Table 2: The best optimization result of 5-layer microwave absorber

Algorithm / Layers	IDSFGWO		GWO		PSO		WOA	
	Type and Thickness (mm)							
1	16	0.1962	16	0.2220	8	1.5941	16	0.2826
2	8	1.9755	16	0.0671	9	0.3662	1	0.1906
3	7	1.8286	6	1.9910	8	1.2310	7	2.0000
4	13	0.7471	5	1.6768	7	1.1402	3	2.0000
5	7	1.3248	12	2.0000	12	1.2028	5	2.0000
Total thickness (mm)	6.0722		5.9569		5.5343		6.4732	
Maximum reflection coefficient (dB)	-34.2194		-28.0019		-14.7287		-23.1191	

the frequency step is set to 0.1 GHz. The maximum total thickness of the microwave absorber is limited to 10 mm. Each algorithm is run 20 times independently in the experiment. The maximum number of iterations is 1000 times. The optimal design optimization results of the four algorithms after 20 runs are shown in Table 2.

The maximum reflection coefficient optimized by the IDSFGWO is significantly better than the other three algorithms. The maximum reflection coefficient result of GWO is equivalent to WOA and better than PSO. Although the total thickness optimized by the IDSFGWO is worse than the PSO algorithm, it is thinner than the WOA algorithm. In addition, relaxation materials appear most frequently in the optimization results in Table 2. It indicates that the absorbing properties of relaxation materials are more likely to meet the requirements of multilayer microwave absorbers. The results of the selection of materials show that the mixed composition of magnetic and dielectric multilayer structure is conducive to better absorption of electromagnetic waves. A reasonable combination of electromagnetic media is expected to achieve the performance requirements of strong absorption of microwave absorber.

The reflection coefficient curves corresponding to the optimization results in Table 2 are shown in Fig. 5. The optimized maximum reflection coefficient of the IDSFGWO is -34.2194 dB. The peak value of -43.6407 dB is reached at 3.8 *GHz*. The numerical results show that IDSFGWO achieves a better absorption effect. The reflection coefficients optimized by both GWO and IDSFGWO are below -28 dB in the frequency range of 2-12 *GHz*. It reflects a better absorption performance compared with the rest of the algorithms.

Figure 6 shows the convergence curves of the four algorithms for the absorption optimization. It can be found that the convergence accuracy of the IDSFGWO is significantly higher than the other three algorithms.

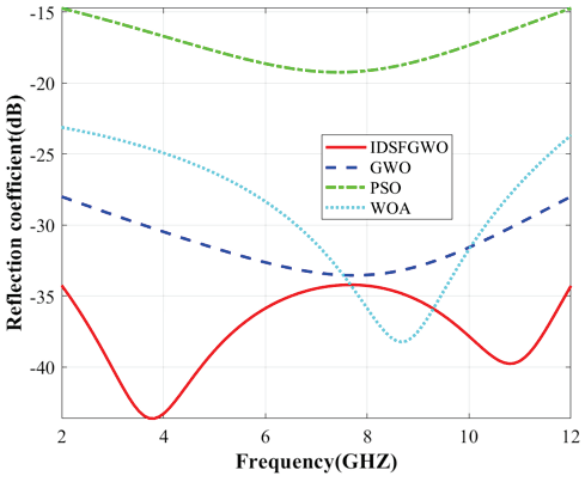


Fig. 5. Convergence curve for 5-layer design.

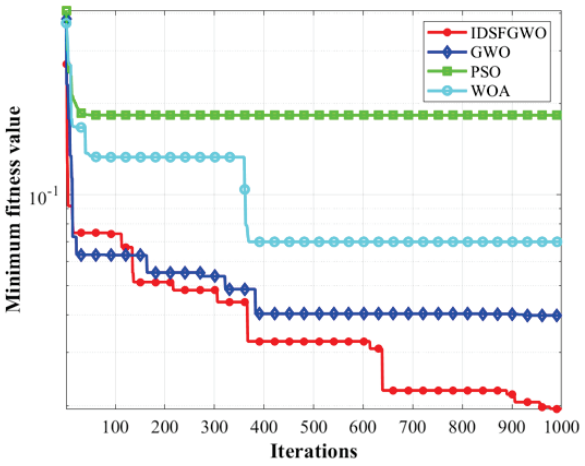


Fig. 6. Convergence curve for 5-layers design.

IDSFGWO has a strong convergence performance and eventually converges to a minimum after 1000 iterations.

Table 3: Comparison results of maximum reflection coefficient(in dB) of 5-layers design

Algorithm	Best	Worst	Mean	Standard Deviation
IDSFGWO	-43.6407	-34.2194	-37.4532	1.4507
GWO	-33.5502	-28.0019	-31.3174	1.7568
PSO	-19.2381	-14.7287	-17.3531	2.8072
WOA	-38.2401	-23.1191	-29.0179	4.6701

The statistical results of the optimal, worst, mean, and standard deviation values of the maximum reflection coefficient for each algorithm are given in Table 3. It can be seen that the IDSFGWO obtains a small optimal value of the reflection coefficient. The standard deviation value of IDSFGWO is also smaller than the other three algorithms. This indicates that the IDSFGWO obtains better absorption performance and also has better stability.

Example B: 7-layers microwave absorber

The four algorithms above are still used to optimize the design of the 7-layers microwave absorber. The maximum total thickness of microwave absorber is still limited to 10 mm. To observe the optimization results of four optimization algorithms for multilayer microwave absorber in a wider band range, the absorbing band width is widened to 0.1-20 GHZ.

The best design results obtained after 20 independent runs are also given, as well as the reflection coefficient curve, convergence curve, and the comparative results of maximum reflection coefficient. The optimization results of 7-layers are analyzed in comparison with the optimization results of 5-layers. According to the refle-ction coefficient curves, the absorption bandwidth of 7-layers is significantly better than 5-layers with the same reflection coefficient. It shows that more layers of microwave absorber has a good effect on broadening the absorbing bandwidth. However, it is also seen that the absorption peak of the optimized absorption of the IDSFGWO becomes smaller. The reflection coefficient curve is smoother than 5-layers. Therefore, multilayer microwave absorber broadening the absorbing bandwidth will also weaken the absorbing performance in part of the frequency band.

The best design results of the 7-layers microwave absorber are shown in Table 4. The reflection coefficient curve of the 7-layers microwave absorber is shown in Fig. 7. The convergence curve of the 7-layers microwave absorber is shown in Fig. 8. The comparative results of the maximum reflection coefficient of the 7-layers microwave absorber are shown in Table 5.

Table 4: The best design result of 7-layers design

Algorithm	IDSFGWO		GWO		PSO		WOA	
Layers	Type and Thickness (mm)							
1	16	0.1824	16	0.4331	16	0.3546	1	0.1420
2	8	1.3370	5	0.1995	1	0.3664	7	0.3073
3	6	1.2994	13	1.4285	7	0.3221	1	1.4285
4	4	0.0403	1	0.0130	16	1.4285	14	0.4135
5	16	0.4294	2	0.4031	12	0.8721	1	0.8741
6	1	0.6216	9	0.1387	7	0.4952	4	1.4285
7	3	0.3607	13	0.0322	6	1.2276	1	0.3201
Total thickness (mm)	4.2708		2.6481		5.0665		4.9140	
Maximum reflection coefficient (dB)	-30.4597		-21.5203		-19.7793		-17.7929	

Table 5: Comparison results of maximum reflection coefficient(in dB) of 7-layers design

Algorithm	Best	Worst	Mean	Standard Deviation
IDSFGWO	-35.0045	-30.4597	-33.1925	1.4287
GWO	-38.2451	-21.5203	-28.4008	5.1299
PSO	-23.2331	-19.7793	-21.9631	1.0417
WOA	-20.6594	-17.7929	-19.5723	0.9028

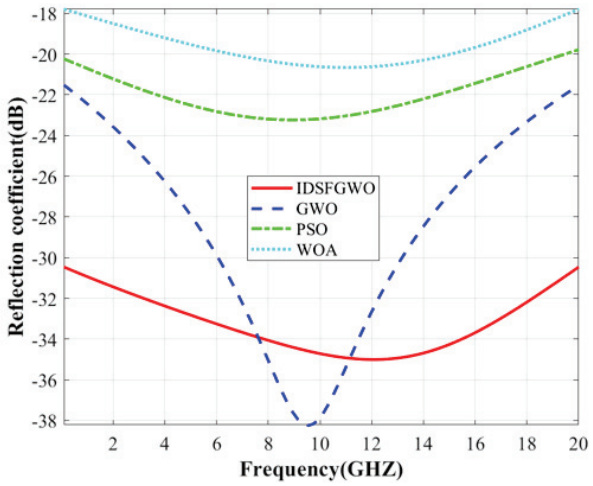


Fig. 7. Reflective coefficient curve of 7-layers design.

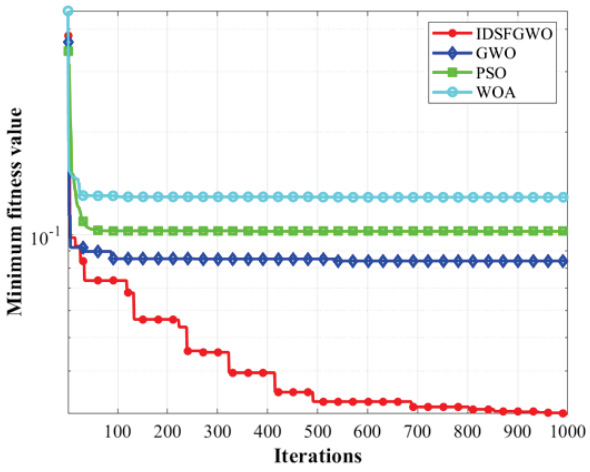


Fig. 8. Convergence curve for 7-layers design.

V. CONCLUSION

In this paper, an improved grey wolf optimizer combining three strategies is proposed to solve the problem of multilayer wideband microwave absorber design in the case of vertical incidence. Through this method, a set of coatings with the smallest reflection coefficient and thin thickness in a specific frequency range can be optimized. The simulation results show that IDSFGWO has better convergence accuracy, stronger optimization ability, and a more stable optimization process when solving the optimization problem of mul-tilayer wideband microwave absorbers.

ACKNOWLEDGMENT

This work was supported by the Jiangsu Graduate Practice and Innovation Program (Grant No. SJCX23-XY069), the Undergraduate Innovation and Entrepreneurship Training Program (Grant No. 2022464).

REFERENCES

[1] R. Panwar, S. Puthucheri, D. Singh, and V. Agarwala, "Design of ferrite-graphene-based thin broadband radar wave absorber for stealth application," *IEEE Transactions on Magnetics*, vol. 51, no. 11, pp. 1-4, Nov. 2015.

[2] T. Peng, C. K. Zhu, T. Y. Zhou, B. Zhang, D. X. Ye, X. J. Li, and L. X. Ran, "A compact microwave imager integrated with a miniaturized dual-angle anechoic chamber," *IEEE Transactions on Microwave Theory and Techniques*, vol. 69, no. 11, pp. 4831-4839, Nov. 2021.

[3] R. Kumar, H. K. Choudhary, S. P. Pawar, S. Bose, and B. Sahoo, "Carbon encapsulated nanoscale iron/iron-carbide/graphite particles for EMI shielding and microwave absorption," *Physical Chemistry Chemical Physics*, vol. 19, no. 34, pp. 23268-23279, Aug. 2017.

[4] M. Cao, X. Wang, W. Cao, X. Fang, B. Wen, and J. Yuan, "Thermally driven transport and relaxation switching self-powered electromagnetic energy conversion," *Small*, vol. 14, no. 29, pp. 1800987, July 2018.

[5] R. Panwar, and J. R. Lee, "Recent advances in thin and broadband layered microwave absorbing and shielding structures for commercial and defense applications," *Functional Composites and Structures*, vol. 1, no. 3, pp. 032022, July 2019.

[6] P. Tar, B. Stágel, and I. Maros, "Parallel search paths for the simplex algorithm," *Central European Journal of Operations Research*, vol. 25, no. 4, pp. 967-984, Dec. 2017.

[7] N. Xiao, X. Liu, and Y. X. Yuan, "A class of smooth exact penalty function methods for optimization problems with orthogonality constraints," *Optimization Methods and Software*, vol. 37, no. 4, pp. 1205-1241, Nov. 2020.

[8] A. Toktas, D. Ustun, and M. Tekbas, "Multi-objective design of multi-layer radar absorber using surrogate-based optimization," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 8, pp. 3318-3329, Aug. 2019.

[9] T. Alexandros, D. Georgios, "Nature inspired optimization algorithms related to physical phenomena and laws of science: A Survey," *International Journal on Artificial Intelligence Tools*, vol. 26, no. 6, pp. 1750022.1-1750022.25, Dec. 2017.

[10] X. S. Yang, S. Deb, S. Fong, X. He and Y. X. Zhao, "From swarm intelligence to metaheuristics: Nature-inspired optimization algorithms," *Computer*, vol. 49, no. 9, pp. 52-59, Sep. 2016.

[11] S. Roy, S. D. Roy, J. Tewary, A. Mahanti, and G. Mahanti, "Particle swarm optimization for optimal design of broadband multilayer microwave absorber for wide angle of incidence," *Progress in Electromagnetics Research B*, vol. 62, pp. 121-135, 2015.

[12] H. Mouna, V. Mekaladevi, and M. N. Devi, "Design of microwave absorbers using improvised particle swarm optimization algorithm," *Journal of Microwaves, Optoelectronics and Electromagnetic Applications*, vol. 17, no. 2, pp. 188-200, June 2018.

[13] X. Chen, X. X. Liu, X. J. Wang, and Y. Liu, "Optimized design for multi-layer absorbing materials based on genetic algorithm," *Advanced Materials Research*, vol. 681, pp. 324-328, Apr. 2013.

[14] L. Xia, Y. Feng, and B. Zhao, "Intrinsic mechanism and multiphysics analysis of electromagnetic wave absorbing materials: New horizons and breakthrough," *Journal of Materials Science and Technology*, vol. 130, no. 10, pp. 136-156, Dec. 2022.

[15] T. Wang, G. Chen, J. H. Zhu, H. Gong, L. M. Zhang, and H. J. Wu, "Deep understanding of impedance matching and quarter wavelength theory in electromagnetic wave absorption," *Journal of Colloid and Interface Science*, vol. 595, pp. 1-5, Aug. 2021.

[16] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46-61, Mar. 2014.

[17] H. Faris, I. Aljarah, M. A. Al-Betar and S. Mirjalili, "Grey wolf optimizer: A review of recent variants and applications," *Neural Computing and Applications*, vol. 30, pp. 413-435, Nov. 2017.

[18] S. Mirjalili, S. Saremi, S. M. Mirjalili, "Multi-objective grey wolf optimizer: A novel algorithm for multi-criterion optimization," Expert Systems with Applications, vol. 47, pp. 106-119, Apr. 2016.

[19] F. S. Gharehchopogh, I. Maleki, and Z. A. Dizaji, "Chaotic vortex search algorithm: Metaheuristic algorithm for feature selection," *Evolutionary Intelligence*, vol. 15, no. 3, pp. 1777-1808, Sep. 2022.

[20] Y. C. Li, M. X. Han, and Q. L. Guo, "Modified whale optimization algorithm based on tent chaotic mapping and its application in structural optimization," *KSCE Journal of Civil Engineering*, vol. 24, no. 12, pp. 3703-3713, Oct. 2020.

[21] C. X. Zhang, K. Q. Zhou, S. Q. Ye, and A. M. Zain, "An improved cuckoo search algorithm utilizing nonlinear inertia weight and differential evolution for function optimization problem," *IEEE Access*, vol. 9, pp. 161352-161373, Nov. 2021.

[22] H. Hakli and M. S. Kiran, "An improved artifcial bee colony algorithm for balancing local and global search behaviors in continuous optimization," *International Journal of Machine Learning and Cybernetics*, vol. 11, pp. 2051-2076, Feb. 2020.

[23] W. Zhao, Z. Zhang, and L. Wang, "Manta ray foraging optimization: An effective bio-inspired optimizer for engineering applications," *Engineering Applications of Artificial Intelligence*, vol. 87, pp. 103300.1-103300.25, Jan. 2020.

[24] E. Yigit and H. Duysak, "Determination of optimal layer sequence and thickness for broadband multilayer absorber design using double-stage artificial bee colony algorithm," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 8, pp. 3306-3317, Aug. 2019.

[25] P. Ranjan, A. Choubey, and S. K. Mahto, "Wide-angle polarization independent multilayer microwave absorber using wind driven optimization technique," *International Journal of Applied Engineering Research*, vol. 12, no. 19, pp. 8016-8025, 2017.

[26] T. M. Shami, A. A. El-Saleh, M. Alswaitti, Q. Al-Tashi, M. A. Summakieh, and S. Mirjalili, "Particle swarm optimization:A comprehensive survey," *IEEE Access*, vol. 10, pp. 10031-10061, 2022.

[27] S. Roy, S. D. Roy, J. Tewary, A. Mahanti, and G. Mahanti, "Particle swarm optimization for optimal design of broadband multilayer microwave absorber for wide angle of incidence," *Progress in Electromagnetic-s Research B*, vol. 62, pp. 121-135, 2015.

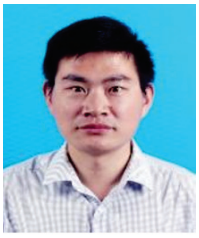
[28] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Advances in Engineering Software*, vol. 95, pp. 51-67, May 2016.

[29] F. S. Gharehchopogh and H. Gholizadeh, "A comprehensive survey: Whale optimization algorithm and its applications," *Swarm and Evolutionary Computation*, vol. 48, pp. 1-24, Aug. 2019.

Haonan Zhang received the B.S. degree from the Southeast University Chengxian College, Nanjing, China, in 2021, where he is currently pursuing the M.Eng. degree at Yancheng Institute of Technology. His current research interests include computational electromagnetics and wireless communications.

Zhifei Zhang received the B.S. degree in computer science and technology from Yancheng Institute of Techno-logy in 2022. He is currently pursuing the M.Eng. degree in mechanical engineering at Yancheng Institute of Technology. His main research interests focus on computational electromagnetics and artificial intelligence.

Yi Du received the B.S. degree in electrical information engineering from Suqian University, Suqian, China, in 2022. He is currently pursuing the M.Eng. degree in mechanical engineering at Yancheng Institute of Technology. His research interests include signal processing and artificial intelligence.



Weibin Kong received the B.S. degree in mathematics from Qufu Normal University, China, 2007, and the M.S. degree in mathematics from Southeast University, Nanjing, China, in 2010, and the Ph.D. degree in radio engineering from Southeast University, Nanjing, China, in 2015. Since 2020, he has been an associate professor with the College of Information Engineering, Yancheng Institute of Technology, Yancheng. His current research interests include computational electromagnetism, artificial intelligence, and wireless communication.

Xiaofang Yang received the B.S. and M.S. degrees from Jiangsu Normal University, Xuzhou, China, in 2009 and 2012, respectively, and the Ph.D. degree from Fudan University, Shanghai, China, in 2016. Since 2016, she has been a lecturer with the Coll-ege of Information Engineering, Yancheng Institute of Technology, Yancheng, China. Her current research interests include fiber sensors, computational optics, and computational electromagnetics.

Zhongqing Fang obtained the B.S degree in optical information science and technology from Anhui Uni-versity in 2014, and received his Ph.D. in optics from University of Science and Technology of China (USTC) in 2019. Since 2020, he has been a lecturer with the College of Information Engineering, Yancheng Institute of Technology, Yancheng. His current research interests focus on the fiber sensors and artificial intelligence.

Design of a Pulse Transformer for X-band Klystron

Yongfang Liu, Yonghua Wu, Xiaoxuan Zhou, and Jin Tong*

Shanghai Advanced Research Institute
Chinese Academy of Sciences, Shanghai, 201204, China
liuyongfang@sari.ac.cn, wuyh@sari.ac.cn, zhouxx@sari.ac.cn, tongj@sari.ac.cn
*Corresponding author

Abstract – Future linear accelerators require klystrons with higher radio frequency (RF) to drive higher gradient accelerating structure. An x-band accelerator structure was used to accelerate electrons at the Shanghai Soft X-ray Free Electron Laser Facility (SXFEL) in Shanghai Advanced Research Institute, Chinese Academy of Sciences (SARI-CAS). A pulse transformer is a crucial device in an RF system. This study presents a high-voltage pulse transformer used for a 50 MW x-band pulsed klystron in SXFEL. Typical specifications of the pulse transformer are peak pulse voltage 420 kV, peak pulse current 300 A, 50 Hz repetition rate and 1.5 μs flat-top pulse width. Design and optimization of pulse transformer are achieved by using equivalent circuit analytic methods and computational aided simulation. The relevant experiments show that this pulse transformer can meet the requirements of 50MW x-band klystron.

Index Terms – flat-top, klystron, pulse modulator, pulse transformer, rise time.

I. INTRODUCTION

X-ray free electron lasers (XFEL) are regarded as a new generation of advanced light sources. Shanghai Soft X-ray Free Electron Laser Facility, which is the first coherent x-ray light source in China, started user operation in 2023 and opened to scientists both from home and abroad [1]. The RF system of SXFEL's main accelerator adopts s-band, c-band and x-band technology, including five s-band RF units, fifteen c-band RF units and two x-band RF units. The RF unit is mainly composed of a klystron amplifier, a low-level RF (LLRF) system, a pulse modulator and a pulse transformer [2]. An X-band RF unit is used to achieve higher gradients (80 MV/m) and more compact footprints for SXFEL's main accelerator, which is driven by a pulsed 50MW x-band klystron. The pulse transformer which provides cathode voltage for the klystron connects the pulse modulator and klystron. Generally, a pulse transformer realizes voltage converting, dc isolation, matching impedances, polarity inversion and power delivery from the primary side to the secondary side. Figure 1 shows a typical schematic of a pulse transformer circuit. The x-band RF unit requires a high-power pulse transformer to drive the 50 MW x-band klystron [3–7]. A pulse transformer is a crucial device in an RF unit. The specifications of the high-voltage pulse transformer are listed in Table 1. Table 2 shows the performance of the x-band pulse transformers in other research institutions [8].

In this study, a high-power pulse transformer is developed. The design needs to guarantee the transformer with a fast leading and falling time, a minimum of overshoot and flat-top ringing. The optimized design method of electromagnetic devices like this pulse transformer can depend on an equivalent circuit model and computational aided simulation. The pulse transformer's design procedure, with all the basic stages, is given. Spice circuit simulation software (LTspice) and finite element analysis software (Comsol) are used to verify the design. Test data and waveforms are provided. The results indicate that the high-power pulse transformer meets requirements. The pulse transformer has been continuously operated for over two years without any failures.

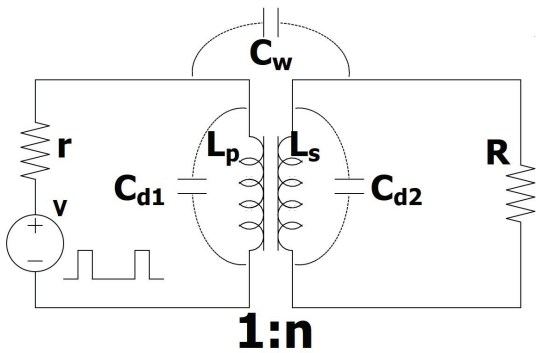


Fig. 1. A typical pulse transformer circuit. V is the source of the ideal pulse, r is resistance of pulse source, C_{d1} and C_{d2} are winding line distribution capacity, C_w is stray capacitance between primary and secondary windings, L_p and L_s are primary and secondary inductance, R is load.

Submitted On: February 21, 2023
Accepted On: May 22, 2023

https://doi.org/10.13052/2023.ACES.J.380914

Table 1: Specifications of the pulse transformer

Items	Value
Peak output power	126 MW
Primary nominal voltage	22 kV
Primary current	5700 A
Secondary voltage	420 kV
Secondary current	300 A
Pulse flat-top	1.5 μs
Pulse overshoot	¡ 1%
Pulse repetition rate	50 Hz
FWHM	3.5 μ
Insulating material breakdown field	10 kV/mm
Relative permittivity of insulating oil	3
Relative permeability of core material	3500

Table 2: Performance of the x-band pulse transformers in other research institution

Items	KEK	SLAC	CECT
Pulse voltage (kV)	500	464	446
Pulse current (A)	301	190	187
Repetition rate (Hz)	200	120	10
Flat-top (μs)	0.5	1.5	1.5

II. ELECTRICAL AND MECHANICAL DESIGN

The role and function of pulse transformer is to step-up pulse voltage from 22 kV to 420 kV with small flat-top drop, small overshoot and low oscillations. The transformer is of rectangular type core and is constructed with cool-rolled silicon steel sheets. In this section, electronical design and relevant geometrical parameters of the pulse transformer are provided and discussed.

A. Electrical circuit

A high-voltage pulse transformer is designed for x-band modulator and klystron. In this design, a filament heating power supply and a pump power supply are required for the x-band klystron cathode, so the pulse transformer adopts a scheme of four secondary windings. Figure 2 shows the practical schematic of the pulse transformer in the linear modulator. In pulse transformer, magnetic flux swings in one direction only, not fully utilizing the core. To avoid flux saturation, core reset circuits named bias power supply (PS)PS are used. The primary winding receives the output pulse waveform of pulse modulator. The filament heating power supply and pump power supply use two secondary windings separately. The secondary winding generates an induced high-voltage pulsed waveform.

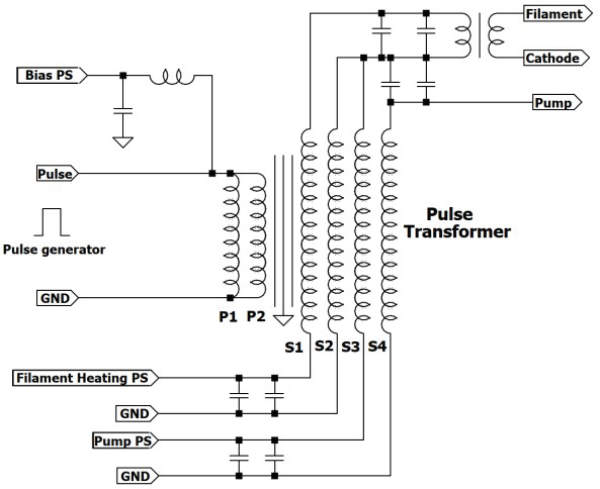


Fig. 2. Detailed circuit of pulse transformer application.

B. Ratio of pulse transformer

In this high-voltage pulse transformer, the secondary peak pulse voltage is 420 kV and the maximum primary peak pulse voltage is 25 kV, which is equal to half of the charging power supply. The minimum ratio of pulse transformer can be calculated as follows:

$$n = \frac{V_S}{V_p}, \tag{1}$$

where n is transformer ratio, V_S is the secondary peak pulse voltage, and V_p is the primary peak pulse voltage. The ratio of the pulse transformer is selected as 1:19 in consideration of suitable margins.

C. Coil Turns and core cross-sectional area

The primary winding turns, the secondary winding turns, and the minimum cross-sectional area of the core are determined by the following equations:

$$N_p = \frac{V_P\, T_P}{A_{core}\, \Delta B} \tag{2}$$

$$N_S = nN_P \tag{3}$$

$$\mathrm{A_{core}} = \frac{1}{\mathrm{N_p \Delta B}}(\mathrm{V_P T_P}) \tag{4}$$

where N_P and N_S are the primary and secondary winding turns, T_p is the pulse width at full width at half maximum (FWHM), ΔB is increment of magnetic induction intensity, A_{core} is the minimum cross-sectional area of the core to avoid saturation. The primary coil is designed as two 5-turn windings in parallel, and the secondary coil is designed as four 95-turn windings in parallel.

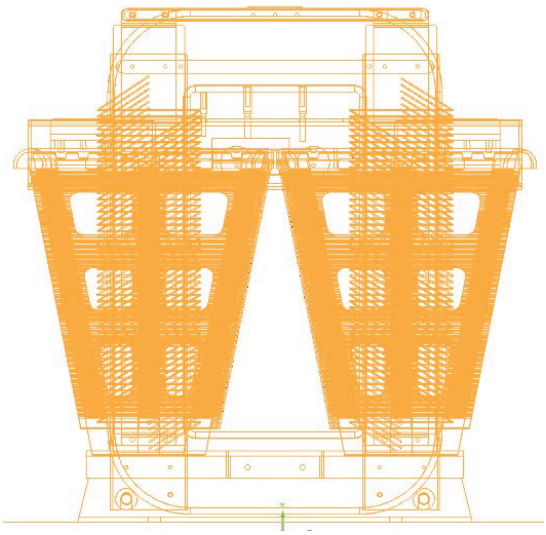


Fig. 3. The mechanical design of the pulse transformer.

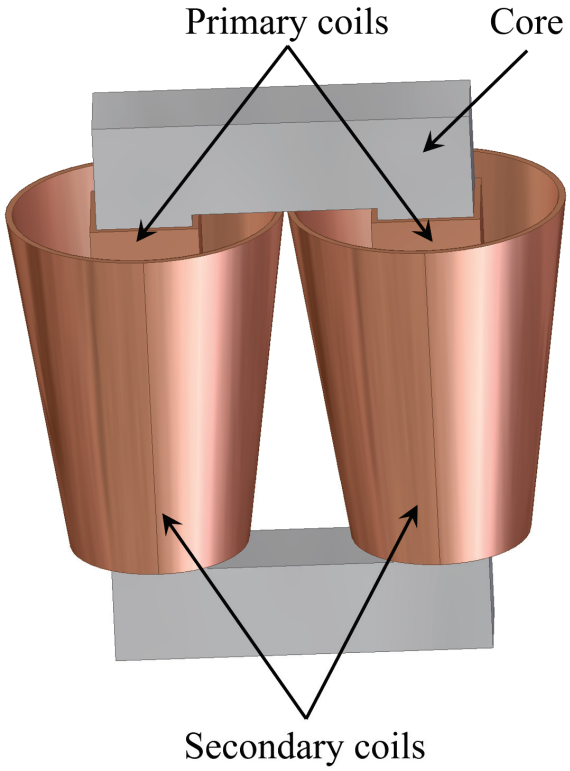


Fig. 4. 3D view of the pulse transformer structure.

The cross-sectional area of the core is designed as 112 cm^2. The geometric shape of the core is almost rectangular, except for the corners which are rounded. High-quality cool-rolled silicon steel sheets are chosen as the core's material.

D. Mechanical design

In order to ensure sufficient insulation distance, a wedge with an inverted trapezoidal cross section is designed as the secondary winding. The core is fixed to a ground plane covered with a special glass fiber layer. The secondary winding is supported by a fiberglass braced structure which surrounds the core. Creepage and clearance distances are the main considerations in the design. The transformer, together with the protective circuit and measuring circuit is placed in a metallic tank holding electric insulating oil which serves both as insulator and coolant. The primary winding is separated from secondary by a gap of 46 mm in case of electric spark due to insufficient insulation distance. In the transformer operation, heat is generated due to losses in windings caused by Joule heating and the core caused by hysteresis, eddy and anomalous effects. An eficient cooling system is necessary in order to minimize the risk of temperature rise. Figure 3 shows the mechanical design of the pulse transformer and Fig. 4 shows the 3D view of the pulse transformer structure.

III. MODEL ANALYSIS

A. Equivalent circuit of pulse transformer

In transformers, the time variation of the magnetic flux passing through the secondary coil depends on the current change of the primary coil [9–11]. The primary input of a pulse transformer is a pulse waveform generated by a pulse modulator. In fact, due to a finite frequency band of pulse transformers, output waveforms in the secondary side have a finite rise time, an overshoot, a flat-top droop and a backswing at fall time [12]. The principle of the pulse transformer is the same as an ordinary transformer. However, pulse transformer handles not a simple sine wave but a pulse waveform with complex spectrum. In the case of an ideal rectangular pulse, there is no rise time, no overshoot, a completely flat top and also no backswing.

The pulse transformer's equivalent circuit model including a pulse generator and a load is presented in Fig. 5.

B. Rise time

Generally, in the equivalent circuit of Fig. 4, it is possible to ignore R_e, C_d and C_L when analyzing the leading edge [13–15]. Then, the equivalent circuit for leading edge analysis can be given as shown in Fig. 6. We can obtain the leading edge by solving differential equation of second order circuit, the leading edge is calculated as follows:

$$T_r = 2\pi y(\sigma)\sqrt{L_e C_d} \tag{5}$$

where L_e denotes the leakage inductance, $C_d$denotes the distributed capacitance, R_L denotes the load

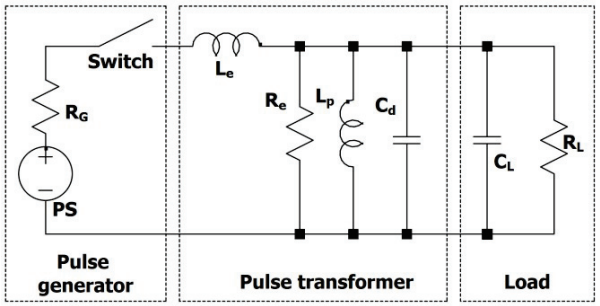


Fig. 5. Equivalent circuit model of the pulse transformer. R_G means the impedance of power supply, L_e indicates the leakage inductance, R_e means equivalent impedance of eddy current loss, L_P is the primary coil inductance, C_d is stray capacitance between coils and grand, $C_L$means the load stray capacitance, and R_L means impedance at primary side.

impedance. $\sigma = \frac{Z_T}{2R_L}$, $Z_T = \sqrt{\frac{L_e}{C_d}}$ and $y(\sigma)$ is a monotonically increasing function of the damping coefficient σ . It is obvious that leading edge depends on the pulse transformer distribution capacitance and leakage inductance. It is evident that small overshoot leads to short pulse rise time.

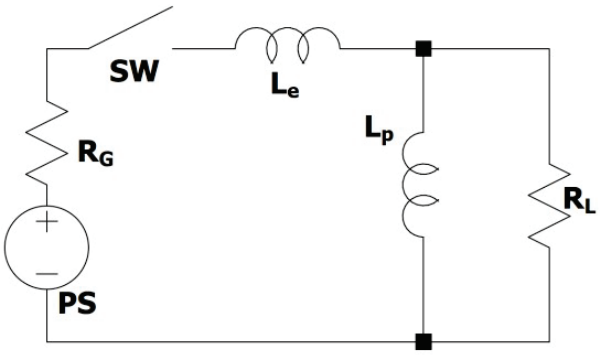


Fig. 6. Equivalent circuit for leading edge analysis.

In the pulse transformer design, leakage inductance is of considerable importance due to high insulation distance between the primary and secondary winding resulting in flux leakage [16]. For simulation of pulse transformer to evaluate leakage inductance, Comsol Multiphysics, which is a simulation tool for solving electromagnetic problems, is used. Magnetostatic simulation is used for the evaluation of leakage inductance of pulse transformer [17]. The model is specified in Comsol Multiphysics, together with material parameters, boundary conditions and other needed input values such as excitation current and frequency. Figure 7 shows the magnetic flux density of the designed pulse transformer. The leakage inductance simulation result is 1.78 μH. From leading edge calculation formula, it can be seen that the larger leakage inductance, the slower leading edge, and the smaller the distributed capacitance, leading to faster leading edge and smaller overcharge.

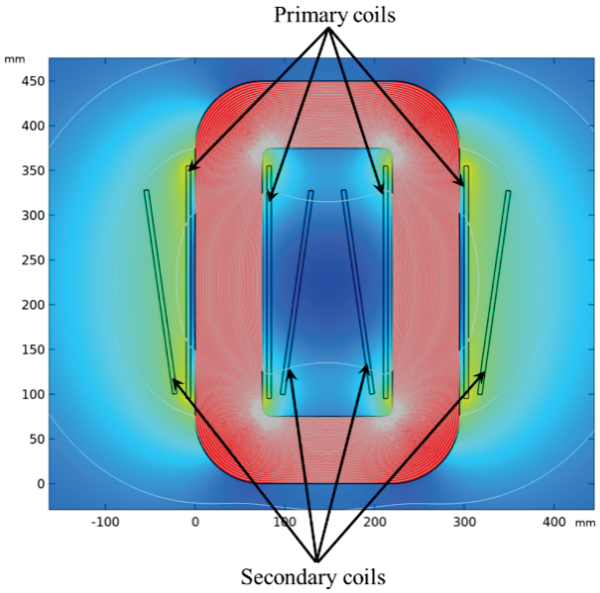


Fig. 7. Magnetic flux density of the designed pulse transformer.

C. Flat-top droop

In the flattop of the pulse waveform, pulse transformer behaves as a low-frequency equipment, due to the invariant voltage over time [18–19]. Thus, elements that are involved in the high-frequency component for L_e, C_d, C_L and R_e can be ignored. Then, the equivalent circuit of Fig. 4 may be rewritten as in Fig. 8. According to Kirchhoff's voltage law, the voltage on the load R_L can be expressed as follows:

$$U_L = \alpha E\, e^{-\frac{t}{T_d}} \tag{6}$$

$$\alpha = \frac{R_L}{R_L + R_{PFN}} \tag{7}$$

$$T_d = \frac{L_P(R_L + R_{PFN})}{R_L R_{PFN}} \tag{8}$$

where E is the voltage of the pulse, α is the pulse transformer primary to secondary transfer factor. T_d is the time constant of the pulse transformer. The flat-top droop is denoted by D, which is defined as follows:

$$D = 1 - e^{-\frac{t}{T_d}} \tag{9}$$

From Equations (6) to (9), the primary inductance of the pulse transformer is related to the top drop as follows:

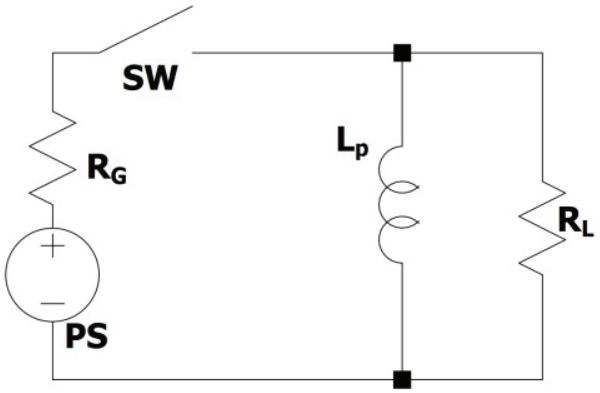


Fig. 8. Equivalent circuit for flat-top droop analysis.

$$D \approx ln\left(\frac{1}{1-D}\right) = \frac{t\ R_L\ R_{PFN}}{L_P\ (R_L + R_{PFN})}. \qquad (10)$$

The relationship between the primary inductance and the flat-top drop in the pulse transformer is shown as Fig. 9.

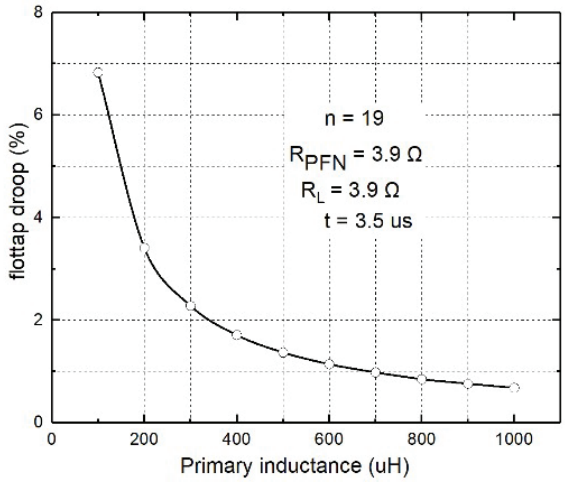


Fig. 9. The relationship between primary inductance and flat-top droop.

IV. RESULTS AND DISCUSSION

Figure 10 shows a picture of the high-voltage pulse transformer. The pulse transformer was installed inside a cylinder filled with electric insulating oil. A precise pulsed current sensor (0.1 V/A) installed in the secondary windings is used to convert pulse current to voltage for current measurement. A high-voltage divider (10850:1) close to the secondary side is used to measure the pulse voltage. Figure 11 shows output waveforms of pulse transformer. Table 3 summarizes the test parameters of the typical measured waveforms of the x-band high-power pulse transformer.

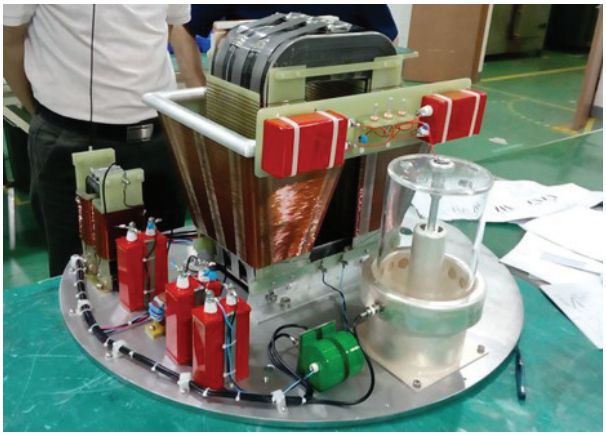


Fig. 10. Picture of the pulse transformer.

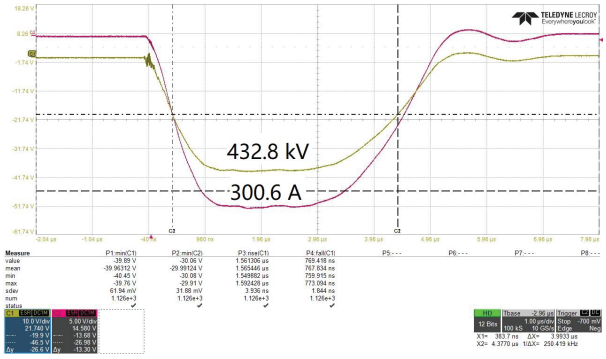


Fig. 11. Output waveform of the pulse transformer.

This paper presents the design process of the x-band klystron pulse transformer. Equivalent circuit modeling was used to analysis the leading edge and the fiat-top droop. It was indicated by relevant experimental data that the pulse transformer satisfied the requirements of the x-band 50 MW klystron.

Table 3: Test parameters of the waveform

Items	Value
Pulse voltage	432.8 kV
Pulse current	301 A
Leading edge	0.77 μs
Falling edge	1.56 μs
Flat-top	1.59 μs
FWHM	3.99 μs

ACKNOWLEDGMENT

This project was supported by the National Natural Science Foundation of China (No.12005282) and Youth Innovation Promotion Association of Chinese Academy of Sciences (No. 2021283).

REFERENCES

[1] Z. Zhao, "The SXFEL upgrade: From test facility to user facility," *Applied Sciences*, vol. 12, Article no. 176, 2021.

[2] A. Zavadtsev, D. Zavadtsev, O. Perevozchikova, and D. Churanov, "High-voltage pulse power supply system for klystron in transverse deflecting system of free-electron laser XFEL," *Journal of Physics: Conference Series*, vol. 1686, no. 1, Article no. 012070, 2020.

[3] J. Tan, "Design, RF measurement, tuning, and high-power test of an X-band deflector for Soft X-ray Free Electron Lasers (SXFEL) at SINAP," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 930, pp. 210-219, 2019.

[4] A. Zavadtsev, "High-voltage pulse power supply system for klystron in transverse deflecting system of free-electron laser XFEL," *Journal of Physics: Conference Series*, vol. 1686, no. 1, Article no. 012070, 2020.

[5] D. Aguglia, "Klystron modulator technology challenges for the Compact Linear Collider (CLIC)," *IEEE Pulsed Power Conference (PPC)*, Chicago, IL, USA pp. 1413-1421, 2011.

[6] Y. F. Liu, "Analysis and optimization of high-power pulse transformer for SXFEL," *Nuclear Science and Techniques* vol. 30, no. 7, Article no. 109, 2019.

[7] X. Deng, N. Liu, Y. Sun, Q. Guo, and M Zhang, "Design and analysis of a novel variable frequency transformer," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 33, no. 08, pp. 904-912, 2021.

[8] D. Sprehn, "X-band klystron development at the Stanford Linear Accelerator Center," *Proceedings of SPIE 4031*, Orlando, FL, USA, 2000.

[9] S. Candolfi, "Hybrid design optimization of high voltage pulse transformers for klystron modulators," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 22, pp. 3617-3624, 2016.

[10] X. Deng, K. Zheng, Y. Gu, A. Zhang, and Z. Jia, "Planar magnetic integration design based on LLC resonant converter," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 36, no. 11, pp. 1474-1483, 2021.

[11] S. Blume, "Design and optimization procedure for high voltage pulse power transformers," *IEEE Transactions on Plasma Science*, vol. 43, pp. 3385-3391, 2015.

[12] S. Kashyap, "Comparison of electromagnetic response in time and frequency domains," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 8, no. 2, pp. 17-43, 2022.

[13] Y. Wang, "Optimal design and experimental study of pulse transformers with fast rise time and large pulse duration," *IEEE Transactions on Plasma Science*, vol. 42, pp. 300-306, 2014.

[14] E. M. M. Costa, "Tesla transformer and its response with square wave and sinusoidal excitations," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 30, no. 9, pp. 1035-1040, 2021.

[15] X. Liang, "An analytical method for pulse transformer-based inductive pulsed power supply circuit," *IEEE Transactions on Applied Superconductivity*, vol. 31, no. 8, Article no. 0500704, 2021.

[16] T. Jalakas, "High-voltage pulse transformer for IOT modulators," *IET Electric Power Applications*, vol. 14, no. 12, pp. 2348-2354, 2020.

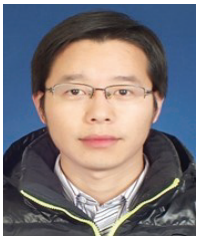
[17] X. Chu, "A new solid-state LC-Marx generator based on saturable pulse transformer," *Review of Scientific Instruments*, vol. 92, no. 5, Article no. 054712, 2021.

[18] F. Pan, "Design procedure of the leakage inductance for a pulse transformer considering winding structures," *IEEE Transactions on Plasma Science*, vol. 45, no. 9, pp. 2504-2510, 2017.

[19] Y. F. Liu, H. Matsumoto, M. Gu, G. Q. Li and S. Li, "Design of an Oil-Immersed Pulse Modulator for X-Band 50-MW Klystron," *IEEE Transactions on Plasma Science*, vol. 51, no. 3, pp. 802–807, 2023.

Yongfang Liu received the Ph.D. degree from University of Chinese Academy of Sciences, Beijing, China in 2020. He is an electrical engineer with the Shanghai Synchrotron Radiation Facility, Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai. His research interests include pulsed power supply and pulsed magnet technologies, such as pulsed klystron modulators, linear transformer drivers, accelerator pulsed magnet and power supply.

Yonghua Wu was born in Anhui, China, in 1972. He is a member of the Shanghai Synchrotron Radiation Facility in Shanghai Advanced Research Institute, Chinese Academy of Sciences. His current research interests include high-voltage switch power supply, pulsed modulator power supply, such as line-type pulse modulators and solid-state pulsed modulators.



Xiaoxuan Zhou was born in Henan, China, in 1984. He received the M.S. degree in instrument Science and technology from the China Jiliang University, Hangzhou, in 2010. His current research interests include linear high-voltage pulse modulators, solid-state pulse modulators, pulsed magnet power supply and control techniques.

Jin Tong graduated from the University of Chinese Academy of Sciences, majoring in Nuclear Technology and Application. Since graduation, he has been engaged in the design of pulse magnets and magnetic measurement systems over years, especially for special magnets used in electron beam injection and extraction systems, such as linear/nonlinear kicker magnets, eddy current septum magnets and Lambertson cutting magnets, which have been designed and researched in large scientific facilities.

A Miniaturized C-B and SIW Bandpass Filter based on LTCC

Wei Tang, Ruo-Feng Xu, and Lei Zhao

School of Information and Control Engineering, China University of Mining and Technology
Xuzhou, Jiangsu, 221116, China
weitangmax@163.com, xuruofeng@cumt.edu.cn, leizhao@cumt.edu.cn

***Abstract* –** In this paper, a miniaturized bandpass filter based on low temperature co-fired ceramic (LTCC) technology is proposed. The miniaturization is achieved by two double folded substrate integrated waveguide (DFSIW) resonant cavities. By interconnecting LTCC three-dimensional structure, the resonant cavities can be stacked vertically. Compared with the conventional SIW filter, the size is reduced by 75.6%. The experimental results show that the low insertion loss and good selectivity are achieved. The proposed miniaturized bandpass filter is promising for 5G application.

***Index Terms* –** Bandpass filter, double folded substrate integrated waveguide, folded waveguide resonator, low temperature co-fired ceramic (LTCC), resonant cavity.

I. INTRODUCTION

In modern satellite and mobile communication systems, there is a growing demand for low insertion loss, compact and easy-to-integrate bandpass filters. However, the conventional rectangular waveguide filter in C-Band has the dimensional limitation, which is difficult to integrate with planar circuits. A number of literatures have been studied on the nature of substrate integrated waveguide (SIW) cavities for the design of miniaturized filters [1–2], but the size reduction cannot satisfy the rapid development of modern wireless communication systems.

Nowadays, low temperature co-fired ceramic (LTCC) technology has been widely used in the design of compact components because of its high-density packaging and 3-D integration. It can be seen as the further miniaturization for compact filter processing on the basis of printed circuit board (PCB) technology [3]. In [4], a vertical LTCC rectangular waveguide bandpass filter with compact size and 2.1 dB insertion loss is proposed by using dielectric cavities. In [5], a multilayer SIW filter based on LTCC technology is designed, which uses a vertical structure to reduce the size. Although the LTCC technology with building the resonant cavities vertically can reduce its planar area, the size of the cross-section is still quite large and the insertion loss rises as the stacks increase.

To further reduce the cross-section of filters, the folded substrate integrated waveguide (FSIW) technology had been introduced [6]. The folding technology could effectively apply to design the compact structure in LTCC. Some miniaturized bandpass LTCC filters were proposed [7–8], which used the FSIW cavity to realize half the planar size of the structure. For earing more miniaturization spaces, a double folding substrate integrated waveguide (DFSIW) was applied to design three-order Chebyshev bandpass filters [9]. However, the insertion losses were quite large for realistic applications.

In the LTCC process, well-designed circuits and materials are important parts of key issues to cut down losses. For a single layer SIW cavity, the filtering circuits were optimized by etching different sizes of slots [10], which formed three resonance cavities and led to the minimum loss of 1.913 dB. In [11], a bandpass SIW filter using U-shape slots is proposed, and an electromagnetic band gap (EBG) was put into the structure to enhance the filtering performance and reduce the loss, which was only 1.42 dB at the center frequency. The one-layer structures are easily optimized to lower the loss, but hybrid coupling effects occur between cells, which may cause the loss difficult to control for a folding filter with multiple layers.

In [12], a new type of H-plane FSIW filter is presented. Multi-layer FSIW resonators and I/O sections were linked by the evanescent SIW sections, which eliminated the loss of extra vias. But the lowest insertion loss was 2.6 dB. In the meantime, silver as the conductor was used in the construction of SIW cavities, which had the practical impacts on the loss control [13]. Although these attempts on the loss reduction based on LTCC had a few progresses, the passband loss was still relatively high and there is few recent research combined with the folding technology. Thus, it is quite a challenge to design a FSIW bandpass filter using LTCC technology in C-Band with compact size and very low loss.

In this paper, a compact bandpass filter with low insertion loss is proposed. The filter is designed with DFSIW resonant cavity, and achieved greater miniaturization with the LTCC packaging process in C-Band.

Submitted On: April 3, 2023
Accepted On: July 8, 2023

https://doi.org/10.13052/2023.ACES.J.380915

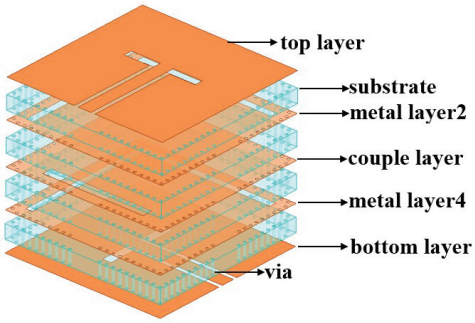
Based on the field distribution characteristics of SIW TE_{101} mode, the magnetic coupling is adopted between adjacent SIW resonators. In order to meet the test requirement, the ground coplanar waveguide (GCPW) is used to design the feed line. The fabricated and measured results are well matched.

II. FILTER DESIGN

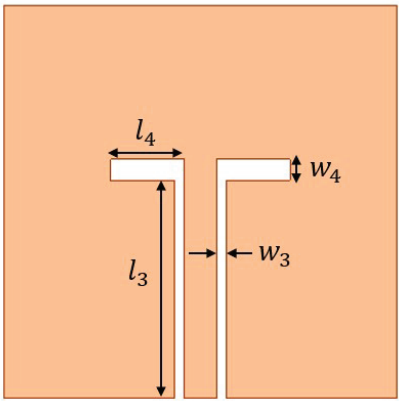
A. Filter geometry

The three-dimensional structure of the filter is shown in Fig. 1 (a). The filter consists of metal layers, substrate layers, and via arrays as vertical sidewalls. The top and bottom metal layers connect the GCPW as feeding lines of the filter, which is shown in Fig. 1 (b). The input and output GCPWs are orthogonal to avoid interactions. Two double-folded SIW resonators are vertically placed, which are coupled through a slot in couple layer. Figure 1 (c) shows the construction of this layer. The metal layer 2 and metal layer 4 are coupling layers of the DFSIW resonator, which are shown in Fig. 1 (d). There is a right-angle slot in each metal layer, which is close to the sidewall.

Ferro-A6 is used as substrates in the structure, which has the dielectric constant$\varepsilon_r = 5.9$and loss tangent $\tan\theta = 0.002$, respectively. The thickness of each dielectric layer is 0.2 mm. The parameters of the filter are shown in Table 1.

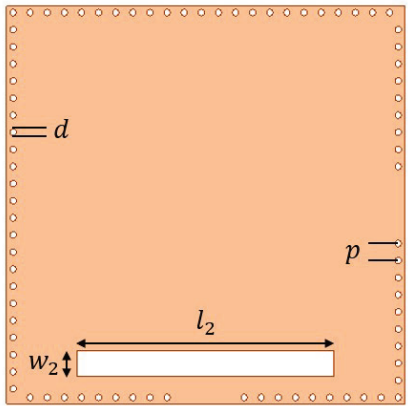


(a)

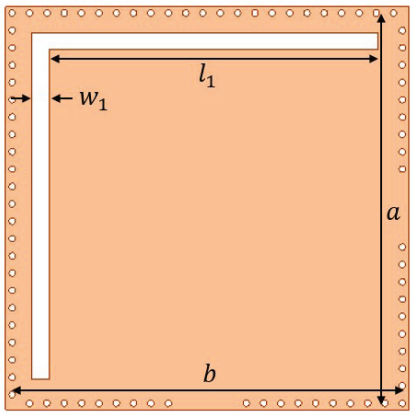


(b)

Fig. 1. *Continued*



(c)



(d)

Fig. 1. Design structure of the filter: (a) 3-D view, (b) the top view of feed part, (c) couple layer, and (d) the top view of a DFSIW resonator.

Table 1: Units for magnetic properties

Variable	Value (mm)	Variable	Value (mm)
d	0.15	p	0.3
l_1	8	w_1	0.4
l_2	6	w_2	0.6
l_3	5.15	w_3	0.2
l_4	1.74	w_4	0.5

The coupling matrix of the designed filter operating at 4.9 GHz is shown in Equation (1):

$$M = \begin{bmatrix} & S & 1 & 2 & L \\ S & 0.00 & 1.23 & 0.00 & 0.00 \\ 1 & 1.23 & 0.00 & 1.66 & 0.00 \\ 2 & 0.00 & 1.66 & 0.00 & 1.23 \\ L & 0.00 & 0.00 & 1.23 & 0.00 \end{bmatrix}. \quad (1)$$

B. DFSIW cavity design

The TE_{101} mode in the SIW cavity is similar to the corresponding mode of conventional rectangular metal waveguides. Therefore, the SIW cavity size can be determined by [14]

$$f_{m0n} = \frac{c_0}{2\pi\sqrt{\varepsilon_r}}\sqrt{\left(\frac{m\pi}{a_{eff}}\right)^2 + \left(\frac{n\pi}{b_{eff}}\right)^2}, \quad (2)$$

and

$$a_{eff} = a - \frac{d^2}{0.95p} b_{eff} = b - \frac{d^2}{0.95p}, \quad (3)$$

in which c_0 is the speed of light. The a_{eff} and b_{eff} are the effective width and length of the SIW cavity, and its value can be obtained by the formula (3). Here, d and p represent the diameter of the metallized vias and the distance of adjacent vias, a and b represent the actual width and length of the SIW cavity, respectively. When the resonant frequency of the SIW cavity is 4.9 GHz, the conventional SIW cavity size is $17.82\times17.82\times 0.2$ mm^3.

In the conventional TE_{101} mode of the SIW resonator, the strongest electric field is distributed in the center of cavity, and the magnetic field distribution reaches maximum at the edge. Figure 2 shows the E-field and M-field distributions in the TE_{101} mode of the designed DFSIW resonator. It can be observed that the electric field is the strongest along the right-angle slot, and the magnetic field is centrally distributed at the end of the slot. Through the folding technology, the positions of field distributions are changed, but the TE_{101} mode is still maintained by the DFSIW resonator. As the folding structure is achieved through the SIW folded twice along the right-angle slot, each side of the DFSIW resonator is reduced by half. The thickness of one DFSIW resonator is double that of the SIW cavity. Therefore, the size of designed cavities is theoretically $8.91\times8.91\times 0.4$ mm^3.

C. DFSIW filter coupling structure

The magnetic coupling mode is adopted between the folded resonant cavities. The slot in the coupling layer is placed near the sidewall. Due to the manufacturing limitations, the position of the coupling slot is 0.65 mm from the vias. By adjusting the position and size of the slot, the coupling strength can be controlled. In Fig. 3, the relationship between the length of slot and coupling coefficient is shown. When the length of the coupling slot increases, the coupling coefficient rises. To meet the requirement of the filter, the length of the coupling slot is chosen as 6 mm.

The insertion loss in the passband mainly consists of the dielectric and conductor loss. To reduce the conductor loss, silver is used to form the metal layer due to its low conductivity. The excellent LTCC processes ensure that each metal layer is only 0.008 mm, and the thickness of the whole dielectric layer is 0.8 mm.

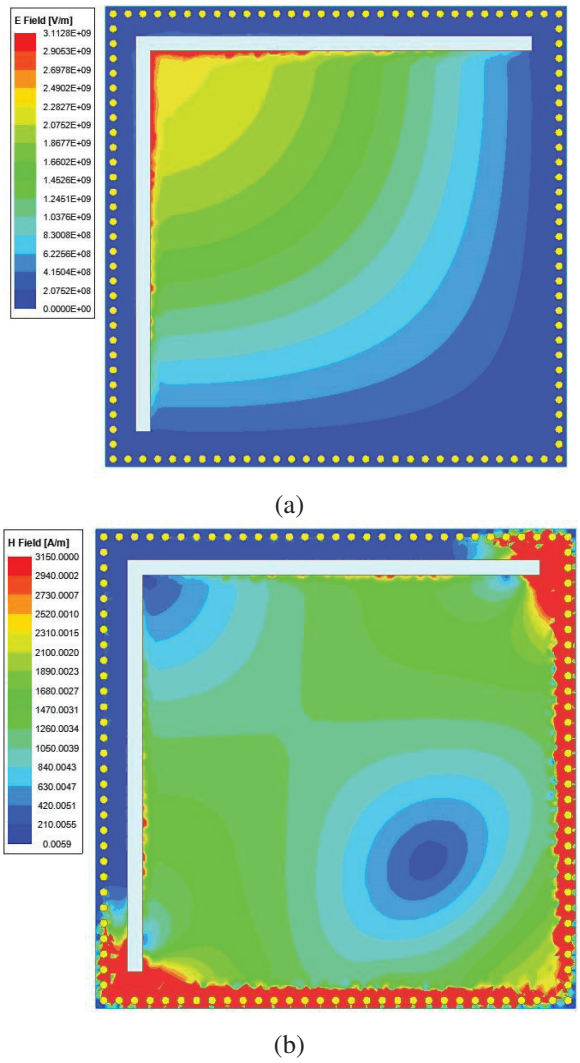


Fig. 2. Field plot: (a) Electric field distributions of a quarter DFSIW cavity and (b) magnetic field distributions of a quarter DFSIW cavity.

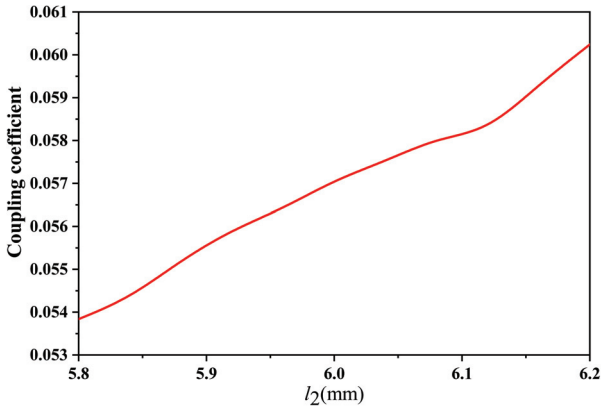
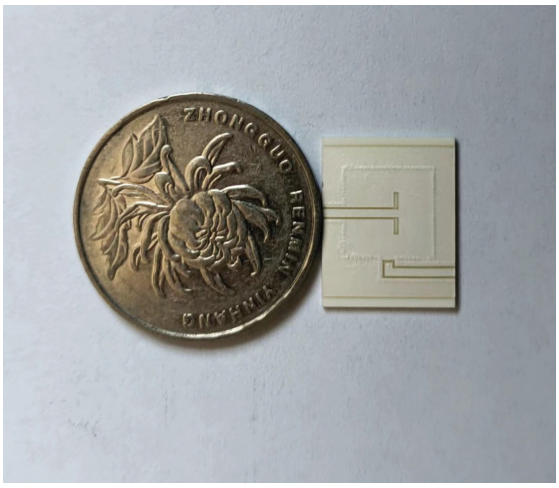


Fig. 3. Coupling coefficients and slot length.

III. SIMULATION AND TESTING

The designed filter is simulated by high frequency structure simulator (HFSS), and the actual size of the structure is $8.8\times8.8\times0.816$ mm^3. The simulated values of the DFSIW cavity are consistent with the theoretical ones. The filter is fabricated by multi-layer LTCC technology, and the production is shown in Fig. 4 (a). The experimental test was completed at the State Key Laboratory of Millimeter Wave in Southeast University. The sample was tested in an Agilent E8363B vector network analyzer, and the platform is shown in Fig. 4 (b).

Figure 5 shows that the simulated and fabricated results are well matched. The measured center frequency is 4.95 GHz, with insertion loss of 1.17 dB. The 3 dB passband ranges from 4.57 GHz to 5.12 GHz, with a relative bandwidth of 11%. In addition, the return loss is better than 10 dB over the range of 4.48-5.36 GHz. Compared with the simulation, the test results have 0.05 GHz center frequency shift, which may be caused by parasitic inductance.



(a)

(b)

Fig. 4. (a) Fabricated filter and (b) test platform.

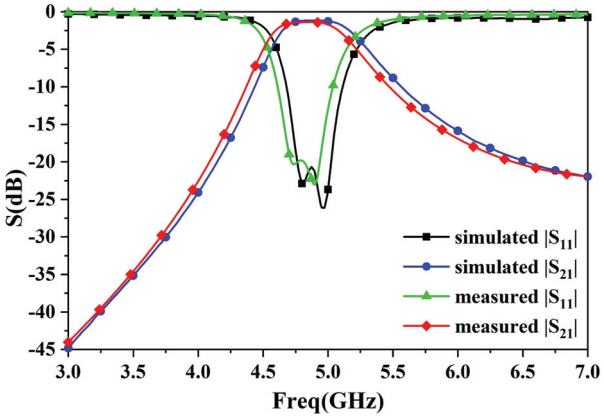


Fig. 5. Simulated and measured S parameters.

Table 2: Comparison of the proposed filter with other filters

Ref.	f_0 (GHz)	IL (dB)	Size (λ_g^2)
[8]	5	3.48	0.72×0.37
[15]	4.97	2.2	1.28×1.30
This work	4.9	1.17	0.35×0.35

In Table 2, the performance of the designed filter is compared with other SIW filters. From the comparison results, the filter designed in this paper has a lower insertion loss and a smaller size.

IV. CONCLUSION

In this paper, a novel DFSIW bandpass filter based on LTCC technology is demonstrated. The structure and the coupling principle of DFSIW resonators are well described, which achieve 75.6% size reduction comparing with the conventional SIW filters. Meanwhile, the LTCC filter is fabricated and measured at C-band with very low insertion loss. The great size reduction is realized by the double folding resonators, and low-loss bandpass performance is achieved by the well-designed coupling position through LTCC technology. The proposed filter has the extensive application in the sub-band of 5G systems.

REFERENCES

[1] Y. Cassivi, L. Perregrini, P. Arcioni, M. Bressan, K. Wu, and G. Conciauro, "Dispersion characteristics of substrate integrated rectangular waveguide," *IEEE Microw. Wireless Compon. Lett*., vol. 12, no. 9, pp. 333-335, Sep. 2002.

[2] Y. L. Zhang, W. Hong, F. Xu, K. Wu, and Tie Jun Cui, "Analysis of guided-wave problems in substrate integrated waveguides numerical simulations and experimental results," *IEEE MTT-S Int. Microwave Symp. Dig*., pp. 2049-2052, Jan. 2003.

[3] Y. Rao, H. J. Qian, J. Zhou, Y. Dong, and X. Luo, "Miniaturized 28-GHz packaged bandpass filter with high selectivity and wide stopband using multilayer PCB technology," *IEEE Microw. Wireless Compon., Lett*., vol. 32, no. 6, pp. 664-667, June 2022.

[4] A. Isapour and A. Kouki, "Vertical LTCC integrated rectangular waveguide and transitions for millimeter-wave applications," *IEEE Trans.Microw. Theory Techn*., vol. 67, no. 3, pp. 868-882, Mar. 2019.

[5] K. S. Chin, C. C. Chang, C. H. Chen, Z. Guo, D. Wang, and W. Che, "LTCC multilayered substrate-integrated waveguide filter with enhanced frequency selectivity for system-in-package applications," *IEEE Trans. Compon., Packag. Manuf. Technol*., vol. 4, no. 4, pp. 664-672, Apr. 2014.

[6] J. S. Hong, "Compact folded waveguide resonator," in *IEEE MTT-S Int. Microw. Symp. Dig*., pp. 213-216, June 2004.

[7] Y. Fei, L. Rui-zhu, Y. Hong-xi, H. Xin-yang, and Z. Ying, "Miniaturized folded substrate integrated waveguide filters in LTCC," *2011 IEEE International Conference on Microwave Technology & Computational Electromagnetics*, Beijing, China, 2011, pp. 171-173, 2011.

[8] L. F. Qiu, B. Xie, L. S. Wu, and J. F. Mao, "A flat-passband predistorted bandpass filter with folded substrate integrated waveguide," *IEEE Transactions on Circuits and Systems—II: Express Briefs.,* vol. 69, no. 2, pp. 324-328, Feb. 2022.

[9] H. Y Chien, T. M. Shen, T. Y. Huang, W. H. Wang, and R. B. Wu, "Miniaturized bandpass filters with double-folded substrate integrated waveguide resonators in LTCC," *IEEE Trans. Microw. Theory Techn.*, vol. 57, no. 7, pp. 1774-1782, July 2009.

[10] S. W. Wong, R. S. Chen, K. Wang, Z. N. Chen, and Q. X. Chu, "U-shape slots structure on substrate integrated waveguide for 40-GHz bandpass filter using LTCC technology," *IEEE Trans. Compon. Packag. Manuf. Technol.*, vol. 5, no. 1, pp. 128-134, Jan. 2015.

[11] S. W. Wong, K. Wang, Z. N. Chen, and Q. X. Chu, "Design of millimeter-wave bandpass filter using electric coupling of substrate integrated waveguide (SIW)," in *IEEE Microwave and Wireless Components Letters*, vol. 24, no. 1, pp. 26-28, Jan. 2014.

[12] L. S. Wu, X. L. Zhou, and W. Y. Yin, "A novel multilayer partial H-plane filter implemented with folded substrate integrated waveguide (FSIW)," in *IEEE Microwave and Wireless Components Letters.*, vol. 19, no. 8, pp. 494-496, Aug. 2009.

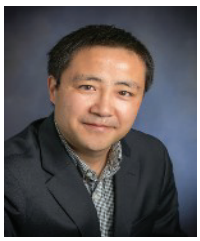
[13] Y. Li, L. A. Yang, L. Du, K. Z. Zhang, and Y. Hao, "Design of millimeter-wave resonant cavity and filter using 3-D substrate-integrated circular waveguide," *IEEE Microw. Wireless Compon., Lett*., vol. 27, no. 8, pp. 706-708, Aug. 2017.

[14] Y. Cassivi and K. Wu, "Low-cost microwave oscillator using substrate integrated waveguide cavity," *IEEE Microw. Wireless Compon. Lett*., vol. 13, no. 2, pp. 48-50, Feb. 2003.

[15] L. F. Qiu, L. S. Wu, B. Xie, W. Y. Yin, and J. F. Mao, "Substrate integrated waveguide filter with flat passband based on complex couplings," *IEEE Microwave and Wireless Components Letters*, vol. 28, no. 6, pp. 494-496, June 2018.

Wei Tang Master's degree student in School of Information and Control Engineering, China University of Mining and Technology, major research interests are SIW filter, active adjustable filter, etc.

Ruofeng Xu received the B.SC. (2004) degree from Suzhou University in electronics and Sscience and technology, and received M.Sc. (2006) and Ph.D. (2010) degrees in electronic engineering from the University of Kent, UK. He was a visiting scholar (2018) for one year in Rutgers University, USA. Now he is a lecturer in School of Information and Control Engineering, China University of Mining and Technology. His main research directions are microwave technology and application.



Lei Zhao (M'09–SM'18) received the B.S. degree in mathematics from Jiangsu Normal University, China, 1997, the M.S. degree in computational mathematics, and the Ph.D. degree in electromagnetic fields and microwave technology from Southeast University, Nanjing, China, in 2004 and 2007, respectively. Zhao serves as an associate editor for *IEEE Access,* an associate editor-in-chief for *ACES Journal*, and a reviewer for multiple journals and conferences including the *IEEE Trans. on Microwave Theory and Techniques*, *IEEE Trans. Antennas and Propagation*, *IEEE Access*, *IEEE Antennas and Wireless Propagation Letters*, *ACES Journal*, and other primary electromagnetics and microwave related journals.

A Compact Bandpass Filter with Active Switchable Passband Status

Ruofeng Xu, Wei Tang, Jun Wang, and Lei Zhao*

School of Information and Control Engineering
China University of Mining and Technology, Xuzhou, 221116, China
xuruofeng@cumt.edu.cn, 383028707@qq.com, jun-wang@cumt.edu.cn, leizhao@cumt.edu.cn*

Abstract – In order to meet the increasing demand for adjustable devices in the modern wireless communication system, a compact bandpass filter with active switchable passband is proposed. The filter is based on the half mode substrate integrated waveguide (HMSIW), and uses its cut-off characteristics to form the lower stopband. The resonant characteristics of a quarter-wavelength shorted stub generates a transmission zero (TZ) at higher frequency, which forms the upper stopband. The active control of DC power supply determines the biased states of PIN diodes to change the electrical size of the filtering parts. When the diodes are reverse biased, the filter works in C band with the passband of 3.9-5.2 GHz; when the diodes are forward biased, the passband shifts to S band of 2.7-3.6 GHz. The 28.6% relative bandwidth in both frequency bands remains constant. The simple design realizes the active switching between bands, providing a promising idea for active adjustable devices.

Index Terms – Active, bandpass filter, half mode substrate integrated waveguide, switchable.

I. INTRODUCTION

The bandpass filter is one of the most important parts in satellite radar communication systems, microwave communication systems, wireless transceiver systems, and other electromagnetic devices, which can effectively ensure the signal transmission in the target frequency band and shield against interference. While traditional filters can provide excellent bandpass performance in the fixed frequency band with advantages of simple structures and steady characteristics, they lack the ability of the multiband selection and adjustment. As the growing demand of high-integration and reconfigurable devices in wireless communication systems, the technical need rises rapidly for compact active tunable filters with multifunctional features, like wideband tuning, easy implementation, low cost, etc. To meet these requirements, DC controlled units are integrated into the structure to dynamically change the distribution characteristics of the electromagnetic field during propagation, which can enable a single filter to efficiently process multiple signals in different frequency bands [1–3].

Currently, most tuning studies mainly focus on microstrip filters. By utilizing the on/off characteristics of the classic PIN diode, the planar microstrip lines can be easily accessed or disconnected from the filtering part, which changes the resonant size. Thus, adjustment of the center frequency, working frequency, or bandwidth can be achieved. A reconfigurable bandpass filter was introduced that can generate passbands in multiple frequency bands by controlling different combinations of the microstrip stubs [4]. A switchable filter was constructed by a set of microstrip lines. The switch between the ultra-wideband and wideband was determined by the connection of high impedance quarter-wavelength shorted stubs or a shorted stepped impedance resonator [5]. A reconfigurable dual-band bandpass filter was designed, which can alter three bandwidths through the control of PIN diodes without changing the center frequency [6]. However, due to the multiple harmonic effect of the microstrip line, it is difficult to achieve high isolation and broadband regulation. Moreover, the active feed network is relatively complex. Therefore, the studies turn to the active tuning of the substrate integrated waveguide (SIW) filter.

The SIW structure has the advantages of low loss, easy integration, and low cost [7]. The PIN diode works as a microwave switch to shift between different components in the guide-wave structure. The electrical control unit changes the combination or the values of the distributed impedances in the cavity to alter the propagation characteristics [8, 9]. An adjustable SIW filter was demonstrated by using the PIN diode to connect the different lengths of slot lines in the cavity. The center frequency can be changed with a 30 MHz bandwidth expansion [10]. But the overlap of each passband was a bit high, and the introduction of additional parasitic elements led to the narrow adjustable band and bandwidth. A switchable SIW filter realized the band and bandwidth switching among three center frequencies through the connections of corresponding metal vias in different positions of the cavity [11]. However, the links and bias

Submitted On: March 2, 2023
Accepted On: October 27, 2023

https://doi.org/10.13052/2023.ACES.J.380916

networks for inner vias increased the processing complexity, and the parasitic inductance brought additional insertion loss. Due to the high integration of the SIW structure and the limitation of available parameter adjustment, there is little research on the filter of the wide-range and cross-band active switch by using PIN diodes.

This paper presents a compact switchable half mode substrate integrated waveguide (HMSIW) filter, which realizes the bandpass feature through the coupling between the fundamental mode of HMSIW and the resonance of a quarter-wavelength shorted stub. The active switching changes the electrical size of the filtering part by utilizing the on/off status of PIN diodes in order to switch the passband between the S band and C band with constant relative bandwidth. When the diodes are reverse biased, the passband is from 3.9 GHz to 5.2 GHz in the C band, centers at 4.55 GHz with a relative bandwidth of 28.6%. When the diodes are forward biased, the passband shifts to 2.7-3.6 GHz in the S-band, centers at 3.15 GHz, without changing the relative bandwidth. The measured and simulated results show good agreement and validate the feasibility of this technology for achieving cross-band switching in SIW filters.

II. FILTER DESIGN

A. Structures

The structure of the proposed bandpass filter is shown in Fig. 1, and the relative parameters are listed in Table 1. The structure consists of an upper metal layer, a middle dielectric layer, and a bottom ground plane. The feedline section adopts the microstrip line to coplanar waveguide (CPW) construction and introduces transverse slots at the terminal of the CPW to further reduce port reflections. The functions of the bandpass filtering and band switching are primarily accomplished by altering the configuration of the upper metal layer. The open side of the middle substrate acts as an equivalent magnetic wall of HMSIW. A line of metal vias spaced much

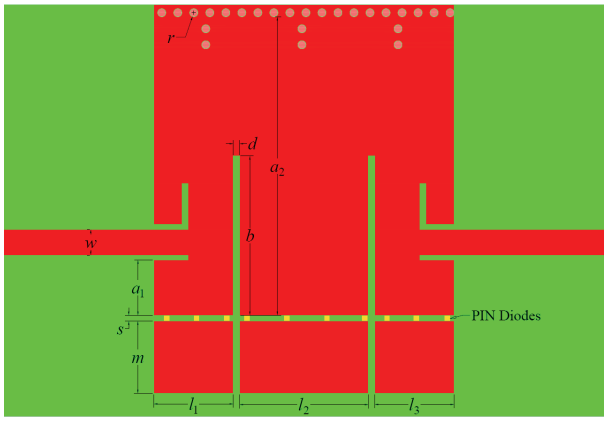


Fig. 1. Top view of the filter.

Table 1: Parameters of the structure

Symbol	a_1	a_2	l_1	l_2	l_3
(**mm**)	2.75	14.9	3.97	6.44	3.97
Symbol	b	s	d	m	w
(**mm**)	7.95	0.3	0.3	3.6	1.28

less than quarter-wavelength act as an equivalent electric wall on the other sidewall of the substrate; an extra six metal vias are horizontally oriented very near to the electric wall to enhance out-of-band suppression. Two surface slots are arranged at a distance from the side of the metal vias and connected longitudinally on the upper metal layer, creating a quarter-wavelength shorted stub. This part constitutes the filtering structure in the C band.

Three rectangular metal pads with the same width are placed longitudinally at a very short distance from the open end of the HMSIW. The gaps between the metal sheets correspond to the slots of the shorted stub. These three metal pads are connected to the upper metal layer of the C-band structure via several MADP-000907-14020X PIN diodes, and their welding polarities keep the same directions. The filtering performance will have significant interference by the number of diodes if the two ends of pads are shorted. Therefore, the number of diodes and their spacing have been optimized to ensure minimum parasitic effects to the propagation and surface current. This part constitutes the switching structure of the S band. The DC power supply drives the PIN diodes through ferrite chokes, which can effectively filter high-frequency signals with minimal interference to DC signals. The forward or reverse biased states of the PIN diodes determine whether the switching parts are connected to the filtering structure. The Rogers RT/Duroid 5880 substrate has dielectric thickness of 0.5 mm, copper foil thickness of 0.035 mm, $\varepsilon_r = 2.2$, and $\tan\delta = 0.0005$ is used as the middle dielectric layer.

B. Active switch realization

The bandpass performance is mainly realized by lower and upper stopbands. The lower stopband is created by the high-pass characteristics of the HMSIW fundamental mode, and the upper stopband is created by the resonance of the quarter-wavelength shorted stub, which generates a transmission zero (TZ) at higher-frequency [12]. The cutoff frequency f_c of HMSIW $TE_{1/2,0}$ mode and the resonance frequency f_r of the shorted stub are

$$f_c = \frac{c}{2a_{eff}\sqrt{\varepsilon_r}}, f_r = \frac{c}{\lambda_p}. \quad (1)$$

Here c is the speed of light in vacuum, ε_r is the relative dielectric constant of the medium substrate, a_{eff} is the equivalent width of HMSIW, and λ_p is the waveguide wavelength of the shorted stub in the substrate.

When the PIN diodes are reverse biased, the equivalent width of the waveguide a_1 is 14.6 mm, and

the cut-off frequency is 3.5 GHz; the length of the shorted stub b is 7.9 mm, and the resonant frequency is 7.3 GHz, where the upper TZ appears. Considering a few transition bandwidths, the filter is under the C-band bandpass working status.

When the PIN diodes are forward biased, three metal sheets are connected to the HMSIW structure to change the electrical size, internal impedance distribution, and the resonance characteristics of the filter. Therefore, the equivalent width of the waveguide (a_1+s+m) increases to 18.5 mm, and the cut-off frequency moves to 2.7 GHz; the length of the shorted stub (b_1+s+m) increases to 11.7 mm, and the upper TZ is generated at 5 GHz. The actual passband shifts to the S band. Three activated metal sheets enhance the transverse electrical lengths of the waveguide and the shorted stub simultaneously. It leads to a reduction of f_c and f_r respectively, which switches the bandpass to S band. By controlling the states of PIN diodes through the DC circuits, the working status of the bandpass filter can be effectively switched between S and C bands.

The transverse E field of HMSIW $TE_{1/2,0}$ mode can be expressed as [13]

$$E_y = \frac{-j\omega\mu a_{eff}}{\pi} A \sin\frac{\pi x}{a_{eff}} e^{-j\beta z}. \qquad (2)$$

Here, A is the random amplitude of the wave and μ is the permeability of the dielectric. As a_{eff} is changed by the bias of diodes, the maximum area of E-field will shift in the $TE_{1/2,0}$ mode. When the filter works at S band, the extra capacitance created by the air gap is paralleled with the inductance of diodes, which could generate harmonic resonances in the passband. The behavior of the E field can be analyzed by using FEM methods.

Figure 2 shows the comparison of the electric field distributions at different phases (0^o and 90^o) while working in C band and S band. The PIN diodes are off at C band; thus, the surface electric field propagates as the $TE_{1/2,0}$ mode of HMSIW, which allows the electromagnetic wave transmitting through the quarter-wavelength shorted stub to produce a bandpass effect. There is little energy coupling to three metal sheets, which have very weak impacts on the performance. When the PIN diodes are activated, the surface current spreads to the metal sheets through the short circuits. Obviously, the electromagnetic distribution still follows the rule of the fundamental mode of HMSIW, but the altering structures of the waveguide and shorted stub form a new bandpass filtering in a lower frequency band. Although the excited position of the feedline is moved to the middle of the waveguide when diodes are active, the fundamental field distributions have very little affection.

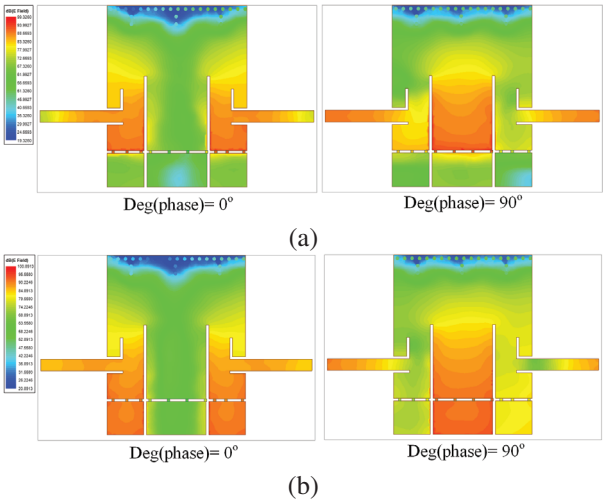


Fig. 2. Distributions of the electric field at different phases: (a) Propagation in C band and (b) propagation in S band.

III. SIMULATION AND TESTING

The filter's structure was simulated using the high frequency structure simulator (HFSS), and the comparison of the simulated and measured results are shown in Fig. 3. The simulated results show that excellent bandpass filtering is achieved in the 4-5.4 GHz band when PIN diodes are reverse biased. The in-band insertion loss is less than 0.3 dB, and the return loss is lower than -15 dB. The upper TZ appears at 7.3 GHz, which greatly improves the selectivity. When the PIN diodes are forward biased, the passband of the filter is switched to the 2.9-3.9 GHz band. The in-band insertion loss is less than 0.5 dB, and the return loss is lower than -13 dB. The upper TZ is moved to 5 GHz, and the out-of-band roll-off is very steep. The relative bandwidth of each status is about 29%, and the simulated results are consistent with the theoretical analysis.

The filter was fabricated by using traditional single-layer print circuit board (PCB) processing technology. Two SMA-KFD0851 connectors work as input/output ports to connect the E5063A network vector analyzer for full-wave simulation testing. The solderless connectors eliminate additional welding losses with only 0.2 dB in-band loss. The fabricated product is shown in Fig. 4 (a). To reduce the ferrite choke's influence on the top metal layer, this choke was welded on the bottom ground plane, which links the positive pole of the DC power supply. Each metal sheet was welded to one ferrite choke which links the negative pole. Figures 4 (b) and (c) show the testing environment and two working statuses of the filter when switching the biases.

The measured results in Fig. 3 show that the filter has a passband in C band from 3.9GHz to 5.2GHz

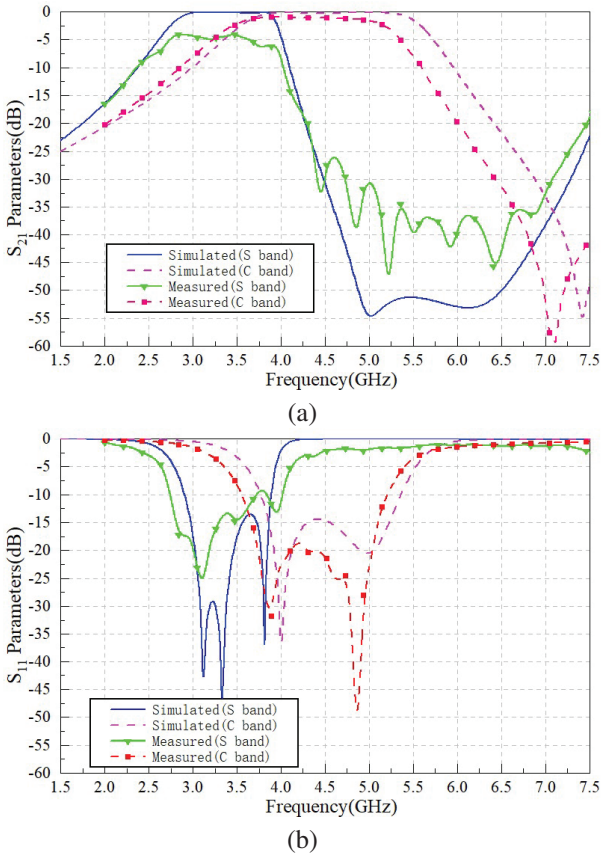


Fig. 3. Comparisons between simulated and measured results: (a) S_{21} parameters and (b) S_{11} parameters.

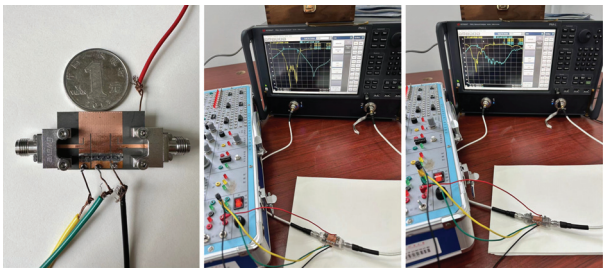


Fig. 4. Fabricated filter and test environment: (a) Product, (b) passband in C band, and (c) passband in S band.

when PIN diodes are off. The in-band insertion loss is less than 1 dB and return loss is lower than -18dB. When PIN diodes are forward biased, the passband is switched to S band of 2.7-3.6 GHz, with an insertion loss less than 5 dB and a return loss lower than -10 dB. The relative bandwidth of these two passbands is 28.6%. The active network has a 2-3 ns switching speed. However, the parasitic inductance made two passbands shift to the lower frequency at about 0.3 GHz. Although the in-band insertion loss is relatively larger than simulated ones, the return loss keeps the good performance. It means the additional measured losses are mainly caused by the energy storage elements in the feeding network. More diodes may reduce the in-band loss because better impedance matching will be achieved. Moreover, the welding process increases the surface roughness and brings more losses. These drawbacks can be improved by optimizing the processing technology and isolating circuits. It illustrates that the designed filter can achieve real-time bandpass switching between S and C bands under active control, The passband response and overall trends in the measurement match well with simulated results.

IV. CONCLUSION

This paper presents a HMSIW bandpass filter capable of switching passbands. Detailed theoretical analysis, simulations, and the testing are accomplished. The measured passband can be switched between S band (2.7-3.6 GHz) and C band (3.9-5.2 GHz) under the active control while maintaining a constant relative bandwidth of 28.6%. The designed filter has the advantages of simple structure, low cost, easy realization, and compact size, which is potentially applicable to the reconfigurable filtering devices in microwave communication systems.

REFERENCES

[1] J. S. Hong, "Reconfigurable planar filters," *IEEE Microw. Mag.*, vol. 10, no. 6, pp. 73-83, Oct. 2009.

[2] C. Lugo and J. Papapolymerou, "Six-state reconfigurable filter structure for antenna based systems," *IEEE Trans. Antennas Propag.*, vol. 54, no. 2, pp. 479-483, Feb. 2006.

[3] I. Hashinur, D. Saumya, B. Tanushree, and A. Tanweer, "Diode based reconfigurable microwave filters for cognitive radio applications: A review," *IEEE Access*, vol. 8, pp. 185429-185444, Oct. 2020.

[4] S. Ramkumar and R. Boopathi, "Compact reconfigurable bandpass filter using quarter wavelength stubs for ultra-wideband applications," *International Journal of Electronics and Communications,* vol. 151, 2022.

[5] A. Bandyopadhyay, P. Sarkar, and R. Ghatak, "A bandwidth reconfigurable bandpass filter for ultrawideband and wideband applications," *IEEE Transactions on Circuits and Systems-II: Express Briefs*, vol. 69, no. 6, June 2022.

[6] X. K. Bi, X. Zhang, S. W. Wong, S. H. Guo, and T. Yuan, "Reconfigurable-bandwidth DWB BPF with fixed operation frequency and controllable stopband," *IEEE Transactions on Circuits and*

Systems—II: Express Briefs, vol. 68, no. 1, Jan. 2021.

[7] X. P. Chen and K. Wu, "Substrate integrated waveguide filters: design techniques and structure innovations," *IEEE Microwave Magazine*, vol. 15, no. 6, pp. 121-133, 2014.

[8] R. F. Xu, B. S. Izquierdo, and P. R. Young "Switchable substrate integrated waveguide," *IEEE Microwave and Wireless Communication Components Letters,* vol. 21, no. 4, pp. 194-196, 2011.

[9] H. D. Chen, W. Q. Che, T. Y. Zhang, Y. Cao, and W. J. Feng, "HMSIW-based switchable units using super compact loaded shunt stubs and its applications on SIW/HMSIW switches," *International Journal of Electronics*, vol. 105, no. 6, pp. 1036-1050, 2018.

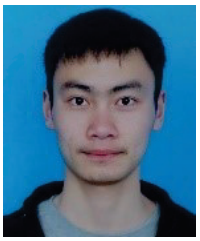
[10] F. Cheng, P. Lu, and K. Huang, "Center frequency and bandwidth switchable substrate integrated waveguide filters," *International Journal of Microwave and Wireless Technologies,* vol. 12, no. 4, pp. 282-287, 2020.

[11] Y. Yi and Y. R. Zhang, "A bandpass filter with switchable frequency and bandwidth on substrate integrated waveguide," 2021 International Conference on Microwave and Millimeter Wave Technology (ICMMT), pp. 23-26, May 2021.

[12] Y. Q. Wang, W. Hong, Y. D. Dong, B. Liu, H. J. Tang, J. X. Chen, X. X. Yin, and K. Wu, "Half mode substrate integrated waveguide (HMSIW) bandpass filter," *IEEE Microwave and Wireless Communication Components Letters*, vol. 17, no. 4, Apr. 2007.

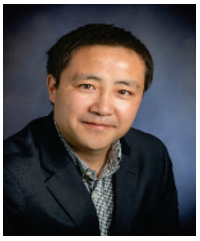
[13] David M. Pozar, *Microwave Engineering*, Fourth Edition. Beijing: Publishing House of Electronics Industry, 2019.

Ruofeng XU received the B.Sc. (2004) degree from Suzhou University in electronics and science and technology and the M.Sc. (2006) and Ph.D. (2010) degrees in electronic engineering from the University of Kent, UK. He had been the visiting scholar (2018) for one year in Rutgers University, USA. Now he has been a lecturer in School of Information and Control Engineering, China University of Mining and Technology. His main research directions are microwave technology and application.



Wei Tang has been a master student at the School of Information and Control Engineering, China University of Mining and Technology. His main researches are on SIW filters and active adjustable filters.

Jun Wang has been a lecturer in School of Information and Control Engineering, China University of Mining and Technology. His main researches are on radio frequency/microwave devices, antennas and transmission, and SPP components.



Lei Zhao has been a professor and doctoral supervisor in School of Information and Control Engineering, China University of Mining and Technology. He has been a senior member of IEEE, member of the Antenna Branch of the China Electronics Society, and chairman of IEEE AP-S Chapter Xuzhou. His main researches are on RF microwave devices, new electromagnetic materials, vortex wave communication, and computational electromagnetics.

Magnetic Field Analysis and Measurement of Pulsed Septum Magnet

Jin Tong and Yongfang Liu*

Shanghai Advanced Research Institute
Chinese Academy of Sciences, Shanghai, 201204, China
*liuyongfang@sari.ac.cn

Abstract – Traditional off-axis injection becomes inadequate in diffraction limited storage ring (DLSR) for its small dynamic aperture (DA). On-axis injection with thin septum could solve the problem. This paper focuses on the theoretical calculation, field simulation, and measurement of thin septum magnets. The scheme of eddy-current type thin septum magnets (the thinnest portion is with the thickness of 0.9 mm) was adopted with laminated silicon steel sheets as the magnet core. The simulation of main field, stray field along the beam trajectory, the leakage field decayed over 1 millisecond time was carried out within Opera 2D/3D. Field measurement and analysis of thin septum magnets also has been conducted comprehensively. The results meet requirements, and the work laid a foundation for injection technology of an advanced light source.

Index Terms – Diffraction limited storage ring (DLSR), eddy current, injection and extraction, magnetic field simulation, septum magnet, thermal analysis.

I. INTRODUCTION

Diffraction limited storage ring (DLSR) injection by traditional off-axis methods is inadequate for its small emittance aperture. Strip-line combined with a Lambertson or thin septum could offer an on-axis injection scheme, as shown in Fig. 1 [1]. This article focuses on the magnetic field simulation, measurement, and analysis of the thin septum. The eddy current type of septum magnet had been adopted for its relatively low power consumption and simple structure [2].

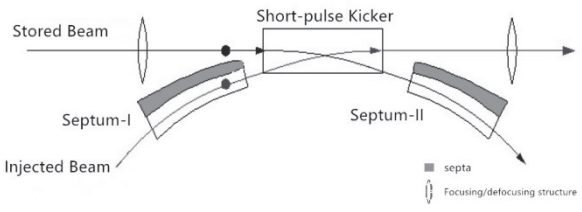


Fig. 1. Diagram of on-axis injection schematic.

II. CALCULATION OF KEY PARAMETERS

The eddy-current type of thin septum magnet is mainly constructed from 4 parts: exciting coil, c-type silicon steel core, septa, and supporting frame. The exciting coil is just one turn oxygen-free copper, which could be led out of the vacuum through feedthrough. The core comprises more than 13,000 lamination silicon steel sheets [3]. The septa part, which is the synthesis of oxygen-free copper and shielding layer, is the most significant assembly unit for the whole septum, and the supporting frame is constructed by oxygen-free copper [4].

$$I_p = B * (G_m + {}^{l_c}/_{u})/(u_0 * N). \quad (1)$$

In equation (1), G_m is the magnet gap height, which is 12 mm and l_c is the length of core path, which is approximately 150 mm. Besides, N indicates the number of turns on magnet. The peak of exciting current is 5322 A.

$$L = \mu_0 * A_m * w_m * N^2/(G_m + {}^{l_c}/_{u}). \quad (2)$$

Similarly, A_m and W_m represent the magnet gap width 40 mm and the magnet gap height 12 mm respectively. The equivalent inductance of the septum magnet is 2.5 μH.

According to the requirements of physical design, key parameters of the thin septum may be seen in Table 1.

Table 1: Key parameters of the eddy-current septum

Parameters	Value	Unit
Beam energy	2	GeV
Deflection angle	50	mrad
Integral field	0.32	T*m
Peak field	5560	Gauss
Good field region	28*10	mm*mm
Mini. septum Thickness	0.9	mm
Peak current	5322	A
Magnet inductance	2.5	μH
Leakage field	0.1%	–

Submitted On: February 26, 2023
Accepted On: November 22, 2023

https://doi.org/10.13052/2023.ACES.J.380917

III. MODELING ANALYSIS

A. 2D transient optimization

According to the magnet size setting, two-dimensional static analysis could be conducted in OPERA finite element analysis software by modeling the magnet with clear boundary conditions. Specific analysis shows that the magnetic field between the center of the magnetic gap is approximately 5600 gauss.

The basic operation principle of thin septum is that the pulsed magnetic field could induct eddy current on oxygen-free copper septa, which could generate a field to effectively offset the leakage part instantly. High-permeability materials could deal with almost all of the remaining overflow magnetic field.

Figure 2 shows some dynamic analysis results. Transverse homogeneity of the main field in central area is±0.45%, and the leakage field under different thicknesses of septa, which was decaying over 1 ms, was thoroughly recorded.

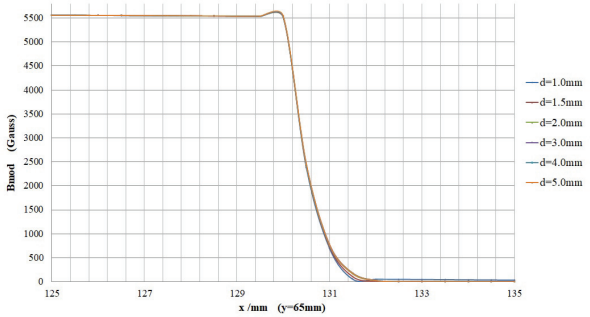


Fig. 2. Leakage magnet field with different septa.

B. 3D transient analysis

The key point of 3D transient analysis of the thin eddy-current septum with OPERA lies in the technique of 3D mesh generation [5]. Poor meshing could lead to misconvergence output. In this project, the mechanical length of the magnet is 660 mm, and the length of the silicon steel core (laminated coefficient is 0.98) is 600 mm. Meanwhile, the thickness of the septa is as thin as 0.9 mm (0.6 mm oxygen-free copper plus 0.3 mm shield layer). In order to achieve a balance between precision outcomes and appropriate computational volume, the ratio between the small-scale grids of the key parts and the large scale grids of other parts should not be too disparate during the process of meshing (Fig. 3). The maximum angle between elements and maximum deviation from surface were set as 90°and 0.2 mm. The maximum element sizes are 1 mm, 0.8 mm, 0.5 mm, and 0.3 mm for the silicon steel core, oxygen-free copper framework, oxygen-free copper septa, and the shield layer, respectively. Finally, more than a million regular tetrahedrons

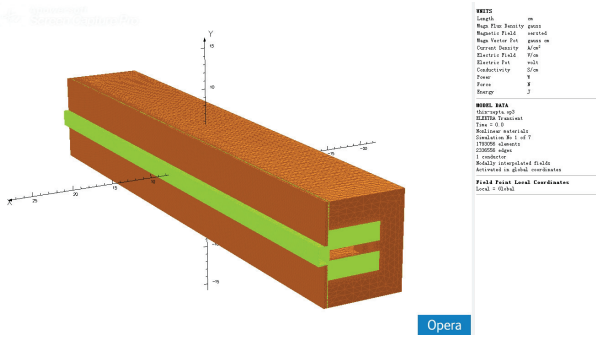


Fig. 3. 3D model of the thin septum magnet.

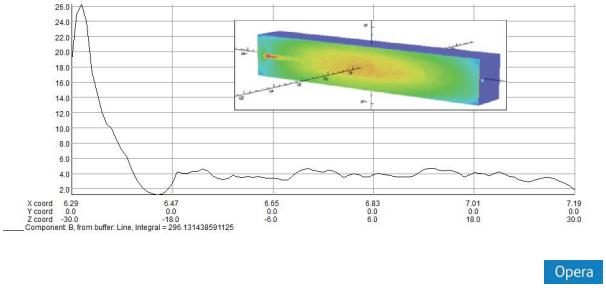


Fig. 4. Results of three dimensional analysis.

had been divided with the opera-3D modeler. The simulation was conducted on an Intel(R) Core^{TM} i5@3.2 GHz, 4GB RAM desktop computer and took almost 30 hours for the result.

The line integrals of the main magnetic field and the leakage field along the beam direction, just 2 mm away from the septa, are shown in Fig. 4. The final result after the ends shielding optimization is 0.09%, which could satisfy the requirement of 0.1%.

Furthermore, the dynamic analysis of the single thin sheet silicon steel of the magnet septum was conducted, and the distribution of induced current could be clearly seen, as shown in Fig. 5.

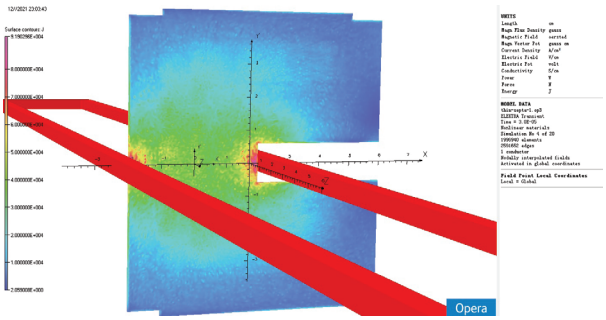


Fig. 5. Eddy current on single slice of silicon steel.

C. Thermal analysis

Because the eddy current septum is designed as an in-vacuum magnet, special attention should be paid to its own heating power consumption and temperature distribution. Heat conduction analysis of the septum magnet should be conducted strictly.

According to the basic theory of steady-state heat conduction [6], The temperature gradient $grad(t)$ is defined as the limit of the temperature increment Δt and the ratio of the normal distance Δn along the isothermal surface,

$$grad(t) = \lim_{\Delta n \to 0} \frac{\Delta t}{\Delta n} = \frac{\partial t}{\partial n} \vec{n}. \tag{3}$$

The heat flux $\vec{q}$ is proportional to the temperature gradient grad.

$$\vec{q} = -\lambda \text{grad}(t) = -\lambda \frac{\partial t}{\partial n} \vec{n} \tag{4}$$

where λ is the thermal conductivity of W/(mK).

According to the analysis, the sources of heating for the septum are copper exciting coil, laminated silicon steels, and oxygen free copper septa. After setting the thermal conductivity of related materials (see Table 2), the steady-state thermal analysis was carried out.

Table 2: Thermal conductivity of septum components

	Material	Thermal conductivity
Coil	Ox- free copper	387.6 W/(m·K)
Insulating	Polyphenylene sulfide	0.28 W/(m·K)
Core	Non-oriented silicon steel	0.3 W/(m·K)
Septa	Ox- free copper	387.6 W/(m·K)
Shielding	Oriented silicon steel	0.3 W/(m·K)
Frame	Ox-free copper	387.6 W/(m·K)

The initial temperature of all the components is 24 degrees Celsius . According to the power density results at 0.5 Hz repetition rate, the highest temperature of the whole magnet is about 28°C (Fig. 6).

Figure 7 shows the temperature distribution on the outer surface of the septa plate with eddy current septum, which is 1∼2°C higher than the base temperature. Therefore, the excitation heating problem is not serious, hence no need for water cooling and other special measures. The magnet base and stainless steel adjustment platform could conduct the heat to the outside of the vacuum box for heat dissipation.

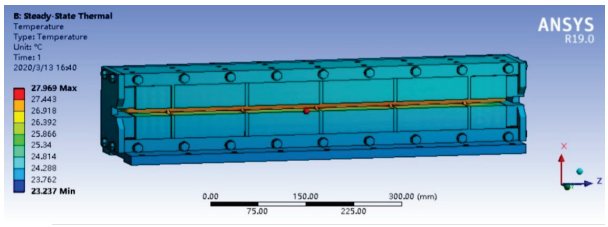


Fig. 6. Internal temperature distribution of the septum.

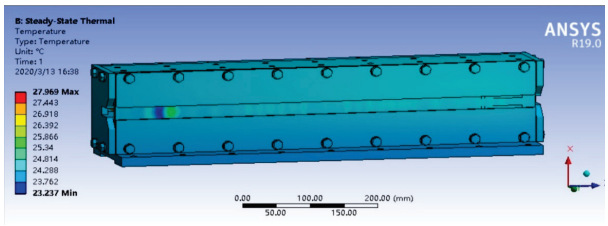


Fig. 7. Temperature distribution of the septa.

IV. Magnetic field measurement and analysis

After the thin septum magnet was assembled, the capacitor charging and LC resonance discharge pulse power (Fig. 8), which can output a stable peak of 5500 A with bottom width of 50 μS half-sine wave, could be used to energize the magnet coil. The convenient and accurate magnetic field measurement platform was built for the analysis of magnetic field data related to the septum magnet, too (Fig. 9). The automatic magnetic measurement system, using carbon fiber to carry the magnetic coil, works effectively in the lab.

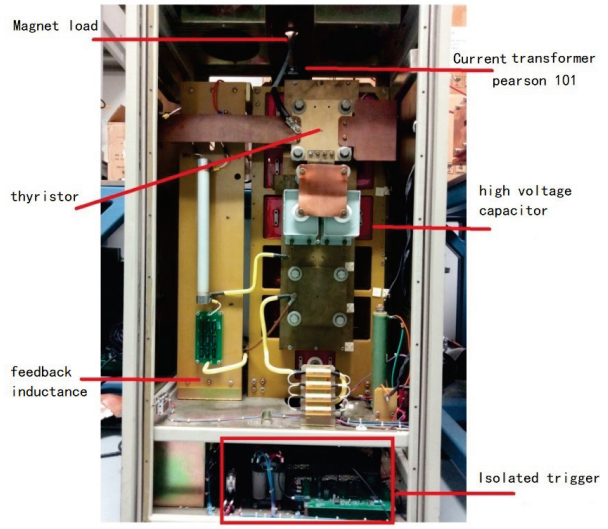


Fig. 8. Layout of excitation pulse power supply.

Fig. 9. Point coil for measuring the end field.

Figure 10 shows the measurement of distribution of the main magnetic field along the beam direction using a point coil with a diameter of 4 mm and 20 turns. Due to the limitation of the length of the measurement probe, only half of the length of the magnet has been tested. The other half of the magnetic field can be measured in the same way. The accuracy of the 12-bit oscilloscope HDO 4104 with digital integrator is so high that the magnetic field test accuracy of half-length magnet is mainly determined by the position accuracy of the moving platform, and the position accuracy is 0.1 millimeter. The other half of the field is measured by rotating the magnet and then testing.

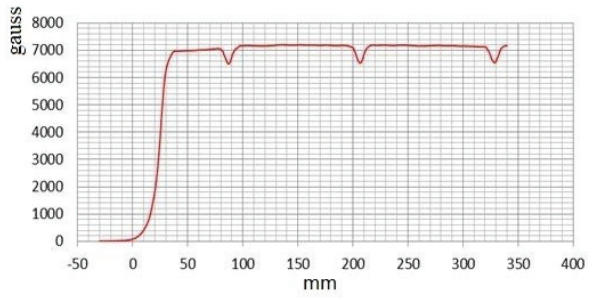


Fig. 10. Field distribution along beam direction.

Furthermore, the point coil (20 turns of φ4 mm) can be used to test the data related to the magnet end field, including the spatial distribution of the end field (Fig. 11) and the decay of the leakage field outside the septa within time and space (Fig. 12) [7]. The *y*-axis in Fig. 12 represents the spatial distribution of the leakage magnetic field at the end region of the septum with 0,1 to 10.13 mm away from the septa .

In Fig. 13, the red waveform shows the excitation current, which reaches 6000A at the peak. The green waveform shows the induced voltage, which is mixed with a certain amount of interference. Finally, the yellow waveform shows the integral of the induced voltage, which is equal to leakage field.

Using the same long coil with a width of 6 mm and a length of 900 mm, the integral value of the main magnetic field against time and the integral value of the leakage magnetic field against time at 3 mm outside the septa have been tested: 2.5 mVs and 1.985 μVs, respectively. The integral ratio of the leakage against the main magnetic field is estimated to be 0.079%, better than the technical requirement of 0.1%.

Finally, the measured temperature distribution is in good agreement with the theoretical calculation. as shown in Fig. 14.

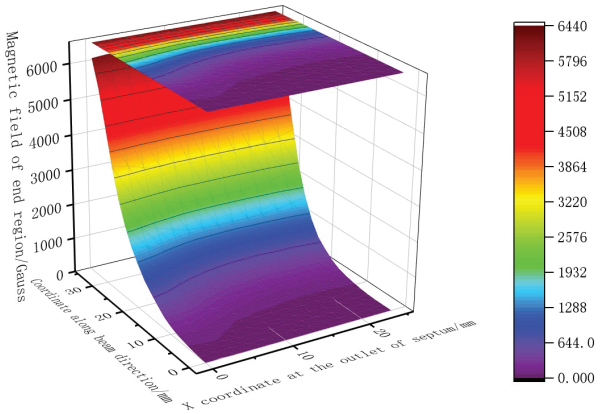


Fig. 11. Color diagram of end field distribution.

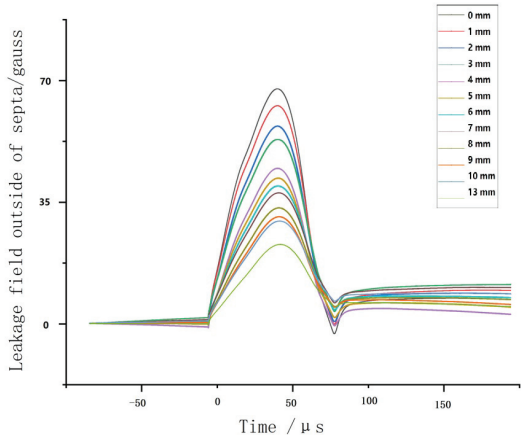


Fig. 12. Leakage field varying with time and space.

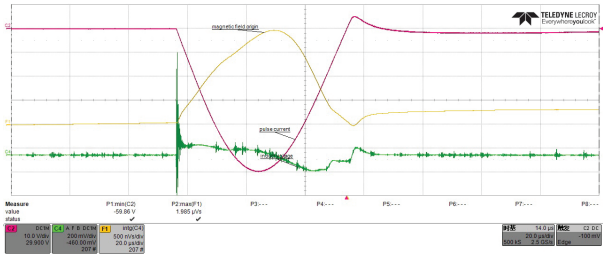


Fig. 13. Induced voltage and integration of leakage field.

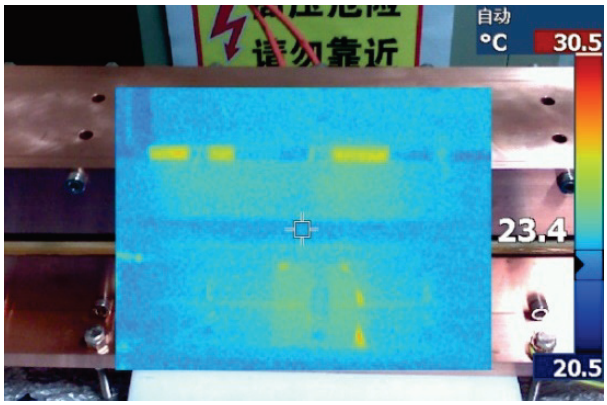


Fig. 14. Temperature distribution at 0.5Hz.

V. CONCLUSION

This paper aims at 3D transient analysis and processing optimization of the thin eddy-current septum for beam injection of a diffraction limited storage ring, especially about simulation and field measurement analysis of the septa. The magnetic excitation power source and the magnetic measurement platform had also been conducted. The results show that the key performance measures of the magnet could all meet the requirements. Furthermore, the theoretical calculations are in good agreement with the measured data in terms of magnetic field and temperature distribution. The work lays a foundation for the next generation of light-source injection technology.

ACKNOWLEDGMENT

Project supported by the National Natural Science Foundation of China (No.12005282) and Youth Innovation Promotion Association of Chinese Academy of Sciences (No. 2021283).

REFERENCES

[1] C. Swenson, D. Arbelaez, J.-Y. Jung, J. Osborn, S. Prestemon, E. Wallen, R. Schlueter, C. Steier, C. Sun, and D. Robin, "Conceptual design of storage ring magnets for a diffraction limited light source upgrade of ALS, ALS-U," in *Proceedings of IPAC2016, Busan, Korea*, pp. 1161-1163, May 2016.

[2] L. H. Ouyang, M. Gu, B. Liu, and R. Chen, "Eddy current septum magnets for injection and extraction at SSRF," *Chinese Physics C*, vol. 34, no. 3, pp. 389-39, 2010.

[3] W. Kang, X. P. Deng, Q. Han, J. B. Pan, and S. Y. Chen, "Theoretical analysis and computer simulation of eddy-current septum magnets," *High Energy Physics and Nuclear Physics*, vol. 25, no. 3, pp. 264-270, 2011.

[4] C. Y. Kuo, C. S. Yang, F. Y. Lin, and Y. L. Chu, "Design and measurement of the septum magnet for the Taiwan Photon Source," *IEEE Transactions on Applied Superconductivity*, vol. 22, no. 3, 4101704, 2012.

[5] F. Zhang, X. Yan, X. Zhang, and F. Wang, "Numerical analysis of magnetic field measurement based on Faraday rotation in a no-core tellurite fiber," *Optical Fiber Technology*, vol. 63, 2021.

[6] L. Christensen, R. Celestina, S. Sperling, R. Mathison, H. Aksoy, and J. Liu, "Infrared temperature measurements of the blade tip for a turbine operating at corrected engine conditions," *Journal of Turbomachinery*, pp. 1-35, 2021.

[7] A. A. Kirilenko, D. Y. Kulik, L. A. Rud, V. I. Tkachenko, and N. Herscovici, "Electromagnetic modeling and design of dual-band septum polarizers, *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 21, no. 2, pp. 155-163, 2022.

Jin Tong graduated from the University of Chinese Academy of Sciences, majoring in nuclear technology and application. Since graduation, he has been engaged in the design of pulse magnets and magnetic measurement systems over years, especially for special magnets used in electron beam injection and extraction systems, such as linear/nonlinear kicker magnets, eddy current septum magnets and Lambertson cutting magnets, which had been designed and researched in large scientific facilities.

Yongfang Liu received the Ph.D. degree in Shanghai Institute of Applied Physics (SINAP) from the University of Chinese Academy of Sciences (UCAS), Beijing, China, in 2020. He is a member of the Shanghai Synchrotron Radiation Facility in Shanghai Advanced Research Institute, Chinese Academy of Sciences. His research interest includes pulsed power and pulsed magnet technologies, such as pulsed klystron modulators, linear transformer drivers (LTD), accelerator pulsed magnets, and power supply.